

## Customer Equity: Measurement, Management and Research Opportunities

Julian Villanueva<sup>1</sup> and Dominique M. Hanssens<sup>2</sup>

<sup>1</sup> *IESE Business School, Camino del Cerro del Águila, 3, 28023 Madrid, Spain, villanueva@iese.edu*

<sup>2</sup> *UCLA, Anderson School of Management, 110 Westwood Plaza, Los Angeles, CA 90095-1481 USA, dominique.hanssens@anderson.ucla.edu*

### Abstract

Despite the recent academic interest in the study of customer equity (CE henceforth), a comprehensive discussion of the prevailing research issues has not been provided. There is a shift in the interest of managers and researchers from a traditional focus on product management to a more recent focus on customer relationship management (CRM). We believe that research on CE could provide the necessary tools to link CRM to long-term financial performance. In this paper, we (a) discuss the academic and strategic importance of CE, (b) provide an extensive literature review, and (c) prioritize future research. We argue that there are two major agendas for future research in CE. The first is to provide better measures (e.g. the measurement of customer lifetime value), and the second is to identify the strategies that lead to CE maximization. We emphasize modeling approaches that have been used or could be used to tackle the suggested research questions. A special focus is given to statistical models that are capable of incorporating long-run dynamics.

# 1

---

## Introduction

---

The Customer Equity paradigm (CE henceforth) recognizes customers as the primary source of both current and future cash-flows. In this framework, the firm is interested in maximizing the net present value of both current and future pools of customers, which is considered a good proxy for the value of a firm (Gupta et al., 2002). Thus, CE models emerge as powerful tools to maximize the return on marketing investments, and to guide the allocation of the marketing budget (Blattberg and Deighton, 1996, Rust et al., 2004, Reinartz et al., 2005).

The roots of the current research in CE can be found on several overlapping research streams: direct marketing, service quality, relationship marketing, and brand equity (Hogan et al., 2002c). Although relationship marketing has become widely popular during the 1990s, it was used long ago by some direct marketers, small enough to address their customers individually (Petrisson et al., 1997). Like direct marketing, the research streams of relationship marketing, service quality, and brand equity have focused on customer retention as one of their primary objectives. Not surprisingly, some researchers have warned against an inadequate implementation of some of these approaches. For example, it has been argued that many firms lack a clear understanding of what

## 2 Introduction

consumers want, and annoy customers by trying to build long-lasting relationships when a transaction approach might be more adequate. Thus, these firms might be obtaining exactly the opposite of what they intended (Fournier et al., 1998). Researchers have also criticized firms obsessed with providing increasing levels of service with the objective of satisfying their customers well beyond what is economically reasonable, falling into a “satisfaction trap” (e.g., Reichheld and Teal, 1996). It has also been pointed out that, after spending huge amounts of money in Customer Relationship Management (CRM henceforth) technologies, some firms do not know how to manage customer relationships with these new databases, and have therefore achieved negative returns for these investments (Rigby et al., 2002).

In an attempt to complement the previous research streams, the literature on CE has two well-defined objectives: (i) the economic measurement of customer relationships; and (ii) the identification of strategies that build profitable relationships. Thus, CE models are about guiding resource allocation with the objective of maximizing the value of a firm.

Early work on the economic measurement of customer relationships introduced the Customer Lifetime Value model (CLV henceforth), which measures the discounted stream of cash flows of an existing customer. Nevertheless, direct marketers and firms in financial industries had implemented these techniques much earlier (Jackson 1989a, 1989b, 1989c). Recently, however, more interesting models have been developed. For example, to guide optimal resource allocation between the acquisition of new customers and the retention of existing ones (Blattberg and Deighton, 1996), or to measure the value of a firm through the value of both its current and future relationships (Gupta et al., 2002).

Eventually, the maximization of CE can be decomposed into smaller problems through the optimization of the acquisition, retention and add-on selling processes (Blattberg et al., 2001). Yet, these processes tend to be interrelated among each other and models that do not incorporate the nature of those links may be biased (Thomas, 2001).

We believe that research on Customer Equity and Customer Lifetime Value is particularly promising because it can help management

practice in the following areas:

- (i) allocating marketing spending for long-term profitability,
- (ii) understanding the connection between marketing spending, marketing metrics and financial performance,
- (iii) providing a customer focused approach for measuring firm value,
- (iv) providing much needed frameworks, tools and metrics for enhancing the productivity of CRM platforms.

Put differently, managers who do not embrace a CE view of the firm, are at risk in several ways: (i) allocating resources (e.g., management focus, money) to marketing actions that produce larger short-term gains at the expense of long-term performance (e.g., investing in acquisition channels that generate lots of new customers who subsequently defect at the first opportunity), (ii) spending on actions or monitoring metrics that do not meaningfully impact customers' behavioral change, (iii) investing in firms that may seem attractive from a standard financial perspective, but whose customer metrics (such as the monthly evolution of acquisition costs, new customers, retention rates, margins per customer) tell a different story, and (iv) investing in expensive CRM platforms without careful consideration of how these platforms will be used to grow CE.

With this paper, we try to achieve several objectives. First, we are interested in reviewing current models of CE and in providing a typology of them. Second, we are interested in reviewing models that can increase CE by optimizing each of its drivers: acquisition, retention, and add-on selling. Third, we want to posit the more fundamental questions of whether a CE orientation of a firm is beneficial when competing with other firms or not, and if firms should discriminate loyals with respect to switchers when it is possible to do so. Finally, we want to provide directions for future research in this important stream of research.

# 2

---

## Models to Compute Customer Equity

---

Direct marketers were the first to calculate the expected lifetime value of a group of customers. Armed with large databases on customer transactions, and always interested in fine tuning their marketing intervention strategies, they looked at the CLV metric as a powerful tool to guide resource allocation for long-term profitability, yet in the late eighties very few companies were known to use this metric (Jackson, 1989a, 1989b, 1989c).

Unlike company valuation models, which have a long research history and enjoy a high degree of acceptance at all levels of management and for different types of firms, customer valuation models have received relatively little attention by researchers and managers until very recently. One possible reason for the apparent lack of implementation of CE models is that research on this area is still very recent and relatively few models have been publicized to managers (with the exception of Blattberg and Deighton's and Rust et al.'s work).

Our objective in this section is to review current models of CE and to provide some guidelines for future research in this area. Specifically,

we attempt to address the following questions:

- Which definitions of the CE concept are available?
- What are the differences and similarities between the concepts of brand equity and CE?
- Which are the characteristics that CE models should have to be complete on important issues?
- Which types of models have been provided so far?

Finally, we provide ideas for further research.

## 2.1 Definitions of Customer Equity

To avoid confusion when we describe these models, we will first define three different concepts that can be used when looking at the future cash flows of customer relationships over time.

- (a) *Customer Lifetime Value* (CLV): CLV is the discounted sum of cash flows generated over the lifetime of an individual customer, or of a segment of customers within the firm. Therefore, as we shall emphasize later on, the CLV metric does not account for indirect effects that an acquired customer has on other customer segments (e.g., generating referrals). Some early authors have used the term NPV (Net Present Value) instead of CLV since the rationale is the same: some investments (acquisition costs) generate future revenues (that decay over time with a retention rate). Furthermore, some authors have discounted only revenues to compare them later with acquisition costs, while others have included both in their CLV equations.
- (b) *Static Customer Equity* (SCE): Static CE is the sum of the CLVs of a specific cohort of customers. For instance, a firm might be interested to discount the expected future cash-flows of the pool of customers acquired at time  $t$ . Models that estimate the CLV of each individual in a cohort, can be used to estimate the SCE of that cohort. On the contrary, there are SCE models that cannot estimate the individual CLV but just the “average” CLV.

- (c) *Dynamic Customer Equity* (DCE): Dynamic CE can be defined as the discounted sum of both current and future cohorts' CE. It has been suggested that this metric is a good proxy for the value of a firm because it accounts for both current and future relationships (Gupta et al., 2002). DCE treats the customers as *renewable resources* and hence it is useful for firms interested in the long-term equilibrium of their strategies (Drèze and Bonfrer, 2003)

The literature on CE has criticized strategies that maximize short-term metrics such as market-share. But it is also possible to criticize early models of CLV in that they do not maximize DCE but instead the CLV of a group of customers or of a specific cohort. In that way, those models might lead to suboptimal decisions.

Other interesting metrics proposed in the literature are: (i) *Customer Equity Contribution* (CEC), which is defined as the increase in the DCE of a firm from a newly acquired customer; and (ii) *Customer Equity Elasticity*, which measures the percentage increase in CE from a 1% change in the marketing mix (e.g., advertising spending) or in any parameter of the CE specification (e.g., retention rate). These metrics might be more helpful than the previous ones in guiding resource allocation and in establishing casual links between marketing spending and financial performance.

## 2.2 Customer Equity Versus Brand Equity

Compared to brand equity, customer equity is a relatively novel framework in the marketing literature. As it can be seen in Table 2.1, 39 papers have been published in five major journals on the subject of brand equity, while only 3 have been published on CE. Similarly, only 7 papers appeared on “customer lifetime value,” 12 on “lifetime value” and 14 on “customer value.” Moreover, unlike the concept of brand equity, concepts such as “customer value” could be used by researchers with different meanings.

The literature on brand equity is more prolific because it was developed earlier, while the literature on CE is still in its infancy. Some researchers see many similarities between these two research streams.

Table 2.1 Published research on brand equity and customer equity

	<i>Number of papers whose abstract contains the word. . .</i>				
	Brand equity	Customer equity	Customer lifetime value	Lifetime value	Customer value
<i>Journal of marketing research</i>	12	1	2	3	1
<i>Marketing science</i>	11	2	3	6	9
<i>Journal of consumer research</i>	3	0	0	0	0
<i>International journal of research in marketing</i>	10	0	0	0	1
<i>Management science</i>	3	0	2	3	3
<i>Total</i>	39	3	7	12	14

Source: *The search was done with the EBSCO database using publication name and abstract*  
Date of search: 19th Dec, 2006

For example, Kumar et al. (2006) consider it is important to study the relationship between brand equity and CE. Specifically, they pose questions on the similarities, the differences and the links between them.

*Similarities* Both concepts are long term in nature, in that they measure the intangible value of marketing assets (customers versus brands) and in that both rely on the loyalty of the customer as a fundamental construct.

*Differences* BE and CE are significantly different in several ways (see Table 2.2): (i) Definition: While the definition of the CE metric or the methods to value customers are quite standard in the literature, there are many models that compute the value of a brand, several of them developed by practitioners (Fernández, 2001), (ii) Unit of analysis: product versus customer, (iii) Level of analysis: while the BE concept usually measures consumer attitudes, the CE concept measures observed behaviors, (iv) Methodologies: the models developed to value brands are mainly descriptive, whereas models to value customers tend to be analytical and statistical in nature, (v) Metric: brand value and CE (or CLV) are different in the way in which they are impacted by marketing and, in turn, in the way in which they drive financial performance, (vi) Drivers and Sub-drivers: CE drivers are easily defined and often observable, while the definition of BE drivers is more complex and grounded in consumer psychology, and (vii) Type of relationships and industries: in some industries the CE framework is difficult to implement or may not even be meaningful,

## 8 Models to Compute Customer Equity

Table 2.2 Differences between brand equity and customer equity frameworks

	Brand equity	Customer equity
Definition	Several definitions	The sum of the discounted stream of cash flows generated from a company's pool of customers
Unit of analysis	Product	Customer
Level of analysis	Attitudinal: The "mind" of the customer	Behavioral: Observed customer purchasing behavior
Methodologies	Mainly heuristic	Deterministic Non parametric Statistical Models
Metric – consensus on relationship between drivers and metric	Low	High
Metric – financial interpretation	Low	High
Components	Brand awareness Brand association Brand attachment Brand experience	Acquisition Retention Add-on selling
Marketing drivers	Advertising & other communication Promotions Word-of-mouth	Customer Satisfaction Loyalty Programs Product offerings Channels and tactics of acquisition Word-of-mouth Brand Equity Competition
Suitable customer-firm relationships	Non addressable customers	Addressable customers Targetability is high
Industries with high applicability (examples)	CPG Luxury goods Automotive	Financial services Insurance B2B
# published papers	39	3

while the opposite may be true for BE. Nevertheless, for some firms or industries both concepts could be useful.

*The link between CE and BE* If both concepts can be useful for some companies (e.g., a hotel chain using a loyalty card), can they be linked? Indeed, one of the first papers on CE established BE as one of the primary drivers of CE (Rust et al., 2004). In this model, it is possible to calculate the Return on CE derived from further investments on some subdrivers affecting BE. More recently, Leone et al. (2006) have

suggested a framework to connect BE to CE. We believe that this is an under-researched area that merits further attention.

## **2.3 Building Complete Customer Equity Models**

Little (1970) provides some general criteria that decision calculus models should have. Models should be simple, robust, easy to control, adaptative, complete on important issues, and easy to communicate with. Although completeness may sometimes compete with simplicity, a model should try to capture all the relevant elements of the problem, and the researcher should consider that completeness is a relative concept to the specific situation (Leeflang et al., 2000).

Researchers developing customer equity models must first understand which are the relevant elements of the problem. We provide here some guidelines about the characteristics that should be included if the model is to be complete on important issues. However, sometimes circumstances such as data availability or a higher need for simplicity may limit the incorporation of some of these characteristics into the model specification.

### **2.3.1 The lifetime of a customer**

Inherent in all models that try to value the long-run financial contribution of a customer is the expected length of the relationship. The first models of CLV incorporate a retention parameter  $r$ , a time horizon of the study, or both. Since retention rates are usually much smaller than one, some researchers have argued that the time horizon should be infinity (Gupta et al., 2002). Estimating the retention rate might be an especially daunting task for non-contractual businesses in which it is not directly observable whether a customer is active or not. We come back to the modeling details of this issue later in the article. Some researchers have also developed models for which the expected lifetime value of a customer depends on the current state of the relationship. These models can accommodate the migration of customers from one state to another (Dwyer, 1997, Bitran and Mondschein, 1996, Berger and Nasr, 1998, Pfeifer and Carraway, 2000) and are therefore especially

appropriate for always-a-share customers.<sup>1</sup> Another approach has been that of measuring brand switching behavior through a Markov process (Rust et al., 2004). Finally, a VAR modeling approach has also been suggested as a way to capture the declining pattern of a customer relationship (Yoo and Hanssens, 2005, Villanueva et al., 2006b).

Most of the previous approaches assume that once a customer is lost, she will not come back again. Or if she comes back, she will behave as any other prospect (with the same response to acquisition and retention efforts). Thus, these models measure the cash flows generated from the first period of a relationship (i.e., acquisition period) until the customer leaves the firm (i.e., defection). In some cases, this approach might underestimate the true CLV of a customer. Consider, for instance, the case of a newspaper for which it is much cheaper to acquire customers who were subscribers in the past. In this scenario, not accounting for the probability of a customer coming back would underestimate the impact of customer acquisition and might lead to underspending on acquisition and overspending on retention.

### 2.3.2 Profit expansion through direct effects

The acquisition of a new customer contributes to the customer equity of the firm through several effects. We distinguish between direct and indirect effects. Direct effects capture the overtime value of a customer generated from the product or services that she buys as long as the relationship exists. The direct effect of a particular customer can be measured using traditional CLV models. As we will explain in Section 4.2, some researchers have suggested that long-life customers are more profitable than first-time customers (e.g., Reichheld and Sasser, 1990). This occurs, they hypothesize, because long-life customers pay higher prices, are cheaper to serve, bring customers through word-of-mouth, and buy more. All but word-of-mouth are direct effects that result in higher customer profitability over time and therefore suggest that, if

---

<sup>1</sup> Jackson (1985) defines always-a-share customers as customers that purchase repeatedly from the same product category and in which they usually buy from different vendors at any given point in time. On the opposite side, she defines lost-for-good customers as those that at any given time buy from only one vendor.

not accounted for, the manager would overspend on acquisition and underspend on retention. Recently, however, using data from a catalog retailer, Reinartz and Kumar (2000, 2003) have shown that long-life customers are not necessarily more profitable than first-time customers.

In any case, if these effects exist, positive or negative, they should be included in the model. Many CLV models assume that individual profits are constant over time. In general, we would expect these to be changing and nonlinear. Statistical models of Customer Equity provide a more flexible structure in that they can more easily capture nonlinearities from the data generation process.

### 2.3.3 Profit expansion through indirect effects

If a customer is lost, the company not only loses the expected CLV of that customer, but also some other effects that drive future revenue. We identify the following indirect effects: cross-effects, feedback effects, adoptions, word-of-mouth, and network externalities. (a) *cross-effects* measure the impact that the behavior of a specific consumer segment has on the lifetime value of another segment. For instance, when customers buying top-of-the-line products influence other existing customers in their willingness to upgrade their products (e.g., Palm users trading their old models for new ones when they talk to users of the new models); (b) *feedback effects* measure the relationship between the firm's performance and future customer relationships. For example, when improvements in profitability derived from existing customers help in future customer acquisitions because of a higher acquisition budget that is closer to the optimal acquisition spending; (c) *adoptions*. It is a well-known fact that the *adoption* of a product category depends on the number of existing customers because some consumers imitate previous adopters (Bass, 1969, Rogers, 1983). Especially in the early stages of a life cycle, the effect that a customer has on future adoptions might represent an important amount of the future expected profits. Hence, they should be included in the model in order to assess the value of a lost customer (Hogan et al., 2002b); (d) *word-of-mouth* to the firm when existing customers bring customers who were buying from a competitor. These effects are especially important to be modeled

when the firm is allocating marketing resources among consumer segments and it is known that some of these segments are better than others at generating referrals (Villanueva et al., 2006b); (e) *network externalities*. Finally, in some industries *network externalities* might exist when the willingness to buy depends on the number of customers that are already using the product or service (e.g., VHS versus Beta-max in the early life cycle of VCRs). Needless to say, the above effects could also be negative (e.g., negative word-of-mouth).

### **2.3.4 Competition**

Even though competition directly affects the expected customer equity of a firm, most models have not explicitly included it. It can be argued, however, that if the previous data generation process from which the model is estimated, is representative of the future, competition is implicitly captured. These models, however, cannot capture what would happen if, for instance, a competitor reduces its prices by 10%. Competition has been modeled in previous CE models using questionnaires (Rust et al., 2004), Markov processes (Rust et al., 2004), and time series modeling using panel data (Yoo and Hanssens, 2005), yet more work needs to be done on this area. For example, previous models have not explicitly studied how competition affects the retention and acquisition processes.

### **2.3.5 Endogenous parameters**

If we look at any mathematical model of CE, it is difficult to accept the assumption that parameters that capture the previously described phenomena are completely exogenous. For example, it has been argued that current efforts in acquisition impact future retention probability and hence this interrelationship should be included in the model (Thomas, 2001). Indeed, we could argue that efforts in acquisition impact everything: future acquisitions, future prices, marketing costs, and so on. We could similarly hypothesize that retention efforts impact other variables. For example, by endogenizing the acquisition and the retention rate, Yoo and Hanssens (2005) show that for some firms an increase in the acquisition rate is followed by an increase in the reten-

tion rate, while for others the impact is negative. It has also been suggested that inter-communication timing through permission-based marketing impact the CLV of a customer (Drèze and Bonfrer, 2001). Hence, the frequency of contacts cannot be taken as exogenous to the CLV equation, but instead it should be optimized.

Endogenizing parameters is extremely difficult to specify with traditional deterministic CLV models, but can be captured by some stochastic models more easily.

### **2.3.6 Including different sources of risk**

Hogan et al. (2002a) have pointed out that by using a single discount rate in the CE equation the firm might be undervaluing the CE contribution of its long-life customers and overvaluing it for its new customers or prospects. These authors argue that payoffs derived from long-life customers have lower levels of uncertainty and risk, and should be discounted with a different rate than the one applied to new customers or prospects. If the company does not adjust for different risks, the firm might be overspending on acquisition relative to retention.

### **2.3.7 Measuring the forecasting accuracy of CLV models**

The output of a CLV model is a forecasted value the company assigns to a customer, and this value is usually estimated using past data. Estimating CLV requires many inputs and assumptions (e.g., retention rate, profit expansion, etc.) and is subject to prediction error despite the use of sophisticated forecasting techniques such as ARIMA models (Mulhern, 1999). Malthouse and Blattberg (2005) use databases from several firms and divide these databases in two in order to compare the CLV predicted from a base period to the actual CLV in a target period. They find that, of the top 20% customers, approximately 55% are misclassified, while of the 80% bottom customers, approximately 15% are misclassified. Thus in several companies in different industries, using past data to estimate CLV brings a substantial amount of error, as customers' behavior is difficult to predict.

Marketing researchers should be aware of these potential misclassifications, especially if they are planning to develop different marketing plans for different customer segments, based on their past behavior.

## **2.4 A Typology of Customer Equity Models**

We propose a typology of customer equity models based on the data source. We support this criterion for our typology due to the following considerations. First, the type of data available for a particular company usually depends on the types of relationships with their customers. For example, when customer relationships are governed by a contract, internal databases might capture the needed information. On the other hand, when customers switch among brands very frequently, panel data – if available – is desirable. Second, database limitations for some firms make particularly important to develop models that use alternative data sources (e.g., survey data). Finally, the objectives of the researcher may require different data sources. For example, when the objective is company valuation, data from company reports may be sufficient and when it is ease of use, managerial judgment might be more attractive. Using this typology and reviewing previous and current research on CE we propose the following data sources: internal databases, surveys, company reports, panel data, and managerial judgment. We provide in Table 2.3 a summary of the models presented here.

### **2.4.1 Internal databases**

Customer valuation models that use internal databases are the most common. The widespread interest in CRM has persuaded marketers to store a wide array of variables measuring the development of the relationships between the firm and each individual customer. These data can be very rich in that it is both cross-sectional and time-series in nature and therefore can be used for measurement and predictions of the CLV of individual customers as well as consumer segments or the whole customer base. Firms can estimate the CLV of a particular customer, and may use this information to target customers individually. The main limitation of these data is that they usually lack the

Table 2.3 A typology of Customer Equity Models based on the data source

Data and references	Model	Description	Level of aggregation	Competition is included	Application
Jackson (1989a, 1989b, 1989c)	Deterministic	Application of the CLV concept to the insurance industry	Product	No	Insurance
Keane and Wang (1995)	Deterministic	Application of the CE concept to a newspaper	Geographic zones	No	Newspaper
Bitran and Mondschein (1996)	Stochastic (Markov)	Optimalcatalogmailing policy with an algorithm to maximize CLV	Firm/RFM cells	No	Catalogs
Dwyer (1997)	Deterministic/ CMM <sup>a</sup>	A CLV model for lost-for-good customers and another for always-a-share	Firm/Recencycell	No	Magazine/ Catalogs
Berger and Nasr (1998)	Deterministic	Provides different mathematical models to calculate CLV	Any	No	Illustrative example
Stauss and Friege (1999)	Deterministic	Introduces the Second Life time Value (STLV) metric	Any	No	Illustrative example
Blattberg et al. (2001)	Deterministic	Comprehensive CE equation that captures key marketing mix phenomena	Consumer segment	No	No
Pfeifer and Carraway (2000)	Stochastic (CMM)	Develops a CLV model using Markov chains, based on demographic info	Consumer segment	No	Illustrative example
Reinartz and Kumar (2000, 2003)	Deterministic	Implements Equation (2.2) of Berger and Nasr (1998)	Individual	No	Several firms
Libai et al. (2002)	Stochastic (CMM)	Incorporates customer switching among different consumer segments	Consumer segment	No	No
Drèze and Bonfrer (2001)	Stochastic	Maximizes CLV by optimizing inter-communication timing	Firm	No	Entertainment
Venkatesan and Kumar (2004)	Stochastic	Allocates contacts to more CLV responsive customers	Individual	No	Computer manufacturer
Thomas et al. (2004)	Stochastic	Expected STLV based on the offer and history of a previous customer	Individual	No	Newspaper

Table 2-3 (Continued)

Data and references	Model	Description	Level of aggregation	Competition is included	Application
Lewis (2005a)	Dynamic program	Studies the effect of marketing policy on CLV	Individual	No	Newspaper
Fader et al. (2005a)	Stochastic	Estimates CLV using Pare to/NBD and links RFM to CLV	Individual	No	Internet
Villanueva et al. (2006b)	Stochastic (VAR)	Estimates the CE contribution of different cohorts of customers	Consumer segment	No	Internet
Simester et al. (2006)	Stochastic	Model to optimize mailing decisions considering long-run profitability	Individual	No	Catalog
Lewis (2006)	Stochastic	Connects acquisition efforts to expected CLV	Individual	No	Grocer/ Newspaper
<i>Survey</i>					
Rust et al. (2000)	Deterministic	Identifies three CE drivers: Value, Brand and Relationship Equity. Shows which drivers are the most important at building CE	Firm	Yes	Illustrative examples
Rust et al. (2004)	Stochastic (Markov)	Can measure the CE impact of different strategies and incorporates switching among brands	Firm	Yes	Several
<i>Company reports</i>					
Gupta and Lehman (2003)	Deterministic	Explains how to use CE modeling for managerial decision making	Firm	No	Public firms
Gupta et al. (2002)	Deterministic	Infinite-period model to estimate CE based on publicly available data.	Firm	No	Public firms
Wiesel and Skiera (2004)	Deterministic	Extends Gupta et al. (2002) to connect CE to shareholder value	Firm	No	ISP

Table 2.3 (Continued)

Data and references	Model	Description	Level of aggregation	Competition is included	Application
<i>Panel with competition</i>					
Yoo and Hanssens (2005)	Stochastic (VAR)	Measures how the firm's marketing mix efforts impact CE	Brand	Yes	Automobile
Reinartz et al. (2005)	Stochastic	Optimizes acquisition and retention spending to maximize LT cust profits	Firm	Yes	High-tech manufacturer
<i>Managerial judgment</i>					
Blattberg and Deighton (1996)	Deterministic	Finds optimal acquisition and retention spending	Firm	No	Illustrative example
Berger and Nasr (2001)	Deterministic	Extends BD to the allocation among promotional vehicles/markets	Firm	No	Illustrative example
Ryals (2005)	Deterministic	Calculates the CLC of key accounts using KAM's judgments	Individual	No	Insurer

<sup>a</sup> CMM: Customer migration model.

observation of the customer's behavior outside of the firm. Hence, models that use these cannot explicitly model competitive behavior.

#### 2.4.1.1 Deterministic models

The first modeling attempts use deterministic equations in which some inputs are entered into the equation in order to calculate CLV (see Berger and Nasr (1998) for a review of CLV models). Some researchers have used these models, but have measured some of their inputs (e.g., retention rate) stochastically. A basic CLV model is described by Jain and Singh (2002),<sup>2</sup>

$$\text{CLV} = \sum_{t=1}^T \frac{R_t - C_t}{(1 + d)^{t-0.5}}, \quad (2.1)$$

where  $t$  is the time period of the relationship,  $T$  is the last period of the relationship,  $R_t$  is the revenue from the customer in period  $t$ ,  $C_t$  is the cost per customer to generate the revenue in that period, and  $d$  is the discount rate.<sup>3</sup> Some models assume a constant gross contribution margin (GC) and marketing costs ( $M$ ), such as Eq. (1) in Berger and Nasr (1998),<sup>4</sup>

$$\text{CLV} = \text{GC} \sum_{t=0}^T \frac{r^t}{(1 + d)^t} - M \sum_{t=1}^T \frac{r^{t-1}}{(1 + d)^{t-0.5}}, \quad (2.2)$$

where  $r$  is the retention rate per period, which is assumed to be constant over time and  $T$  is the time horizon of the study. Many extensions of these basic models exist (e.g., Jackson, 1989a, 1989b, 1989c, Keane and Wang, 1995, Dwyer, 1997), but perhaps the most comprehensive of all

<sup>2</sup>We use the same notation throughout the paper to help the reader in comparing these models. Therefore, we modify the original equations to our particular notation.

<sup>3</sup>Note that this model does not account for acquisition costs.

<sup>4</sup>Equations (2.1) and (2.2) follow the same logic but each makes different assumptions that are noteworthy: (a) Attrition: In Eq. (2.1) revenues and costs are given for each time period  $t$  and therefore there is no need for a retention rate to account for the probability of the customer being alive at time  $t$ , while in Eq. (2.2) contribution and costs are conditional on the customer being alive, thus the need for the retention rate, (b) The timing of the cash flows: In Eq. (2.1) both revenues and costs are accounted for in the middle of the time period while in Eq. (2.2) the gross contribution arises at the beginning of the period (e.g., as in the case of a newspaper subscription), while marketing costs are accounted for in the middle of the time period.

is the equation to calculate CE provided by Blattberg et al. (2001, p. 23) which incorporates number of prospects, acquisition spending and consumer segments.

The applications of these models have been varied. For example, it has been used to estimate the CLV per territorial zone for a newspaper subscription business (Keane and Wang, 1995), to estimate the average CLV of a new insurance policy subscriber (Jackson, 1989a, 1989b, 1989c), or for estimating the individual CLV based on previous purchase history (Reinartz and Kumar, 2000, 2003).

Deterministic models of CE present several limitations when compared to other CE models (see Table 2.3). First, they require substantial data at the individual level. Second, when used at the aggregate level, practitioners estimate each parameter separately and do not take into account any relationship between the parameters of interest (e.g., between acquisition and retention). Third, these models are generally merely descriptive and do not help in guiding action. Fourth, competitive behavior is usually not included. Finally, it is not possible to model latent behavior (e.g., cross-effects among consumer segments). The main advantage is that they are relatively easy to implement, if data are available.

#### 2.4.1.2 Customer Migration Model

A particularly interesting CLV model is what has been called the Customer Migration Model (Dwyer, 1997). These particular CLV models are better suited for scenarios in which customers follow an always-a-share scenario and where it is difficult to observe whether a customer is alive or not. They predict customer behavior based on historical probabilities of purchase depending on recency (i.e., number of periods since last purchase) and the current recency state in which the customer is located (see Berger and Nasr, 1998, case 5, and also Pfeifer and Carraway, 2000). The model has recently been generalized to include more complete segmentation variables, such as RFM, or any other demographic variable (Libai et al., 2002). In this latter paper, CE is specified as,

$$\text{SCE} = \sum_{t=0}^T \frac{MM_t C_t P_t}{(1+d)^t}, \quad (2.3)$$

where  $MM_t$  is a matrix that contains the probabilities of customers moving from one segment to another<sup>5</sup> at time  $t$ ,  $C_t$  is a vector containing the number of customers in each segment at time  $t$ , and  $P_t$  is the profit from each segment at time  $t$ .

### 2.4.1.3 Other stochastic models using internal data

More recently, researchers have proposed a VAR modeling approach to estimate the CE contribution of different segments of customers, depending on the acquisition channel from which they were acquired (Villanueva et al., 2006b). For ease of exposition, assume a three-variable system that captures the dynamic interrelationships among the number of customers acquired through mass advertising at time  $t$  ( $AD_t$ ), the number of customers acquired through word-of-mouth at time  $t$  ( $WOM_t$ ), and a proxy variable for the firm's performance (e.g., profit) at time  $t$  ( $V_t$ ). The VAR(p) model would be specified as,<sup>6</sup>

$$\begin{pmatrix} AD_t \\ WOM_t \\ V_t \end{pmatrix} = \begin{pmatrix} a_{10} \\ a_{20} \\ a_{30} \end{pmatrix} + \sum_{l=1}^p \begin{pmatrix} a_{11}^l & a_{12}^l & a_{13}^l \\ a_{21}^l & a_{22}^l & a_{23}^l \\ a_{31}^l & a_{32}^l & a_{33}^l \end{pmatrix} \begin{pmatrix} AD_{t-l} \\ WOM_{t-l} \\ V_{t-l} \end{pmatrix} + \begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix}. \quad (2.4)$$

Assuming stationarity, we can rewrite the VAR model in standard form given by Eq. (2.4) as a moving average representation (see Enders, 1994),

$$\begin{pmatrix} AD_t \\ WOM_t \\ V_t \end{pmatrix} = \begin{pmatrix} \overline{AD} \\ \overline{WOM} \\ \overline{V} \end{pmatrix} + \sum_{i=0}^{\infty} \begin{pmatrix} \phi_{11}(i) & \phi_{12}(i) & \phi_{13}(i) \\ \phi_{21}(i) & \phi_{22}(i) & \phi_{23}(i) \\ \phi_{31}(i) & \phi_{32}(i) & \phi_{33}(i) \end{pmatrix} \begin{pmatrix} \varepsilon_{1t-i} \\ \varepsilon_{2t-i} \\ \varepsilon_{3t-i} \end{pmatrix}. \quad (2.5)$$

The coefficients  $\phi_{jk}(i)$  are called impact multipliers and measure the impact of a one-unit change in  $\varepsilon_{kt-i}$  (i.e., one more customer from

<sup>5</sup>These models are subject to the Markov property. That is, the expected future behavior depends on the current state and not on the path taken up to that state.

<sup>6</sup>For this VAR model of order  $p$ , where  $(e_1, e_2, e_3)'$  are white-noise disturbances following  $N(0, \Sigma)$ , the direct effects are captured by  $a_{31}, a_{32}$ , cross effects by  $a_{12}, a_{21}$ , feedback effects by  $a_{13}, a_{23}$  and finally, reinforcement effects by  $a_{11}, a_{22}, a_{33}$ . A deterministic trend, seasonal dummies, and exogenous variables can also be included in this VAR. Instantaneous effects are not included directly in this VAR, but they are reflected in the variance-covariance matrix of the residuals ( $\Sigma$ ).

segment  $k$ ) on the  $j$ th variable. We can calculate the CE contribution of an average customer from segment  $k$  as follows,

$$\text{CEC}_k = \sum_{i=0}^m \frac{1}{(1+d)^i} \phi_{vk}(i). \quad (2.6)$$

The main advantage of this metric is that it calculates the total customer equity contribution from a newly acquired customer and therefore not only accounts for the CLV of that customer (*direct effects*) but also for many other *indirect effects*. This modeling approach does not require very complex data and can accommodate profit expansion over time. One of the limitations is that it cannot estimate the CLV of individual customers, since it is estimated using aggregated data. VAR models can also include exogenous variables.

Another interesting approach to estimate the CLV of a customer using internal data has been that of Drèze and Bonfrer (2001). They define CLV as a function of the interval of time between email contacts sent to a customer,

$$\text{CLV}(\tau) = \frac{(1+d)^\tau}{(1+d)^\tau - p(\tau)} A(\tau), \quad (2.7)$$

where  $\tau$  is a fixed time interval between contacts,  $A(\tau)$  is the expected surplus from communications following that interval, and  $p(\tau)$  is the probability of retention given that interval. Note that this equation is similar to Eq. (2.10) when  $\tau = 1$ , except that the latter assumes that cash flows are generated at the beginning of time 0. Using data from a firm in the entertainment industry, they estimate the relationship between the time interval and the customer lifetime value. Their objective is to find the optimal interval time for permission-based emails to a customer base.

An interesting model is the one developed by Reinartz et al. (2005). Although their model does not explicitly compute CLV or CE, it actually models customer profitability and relationship duration at the individual level, and allows the firm to optimally allocate acquisition and retention resources for long-run profitability. The model incorporates competition through an operationalization of share-of-wallet, because in a B2B setting, it is easier to observe a customers' current budget in the category of study.

A stochastic CLV model that works for always-a-share scenarios is the one developed by Venkatesan and Kumar (2004). In this model contribution margin and relationship duration are modeled separately and as a function of the number of contacts through several channels. This model requires panel data and allows to estimate the CLV at the individual level based on her previous purchase history and the number and nature of marketing actions targeted to her. In a lost-for-good scenario, Lewis (2005b) develops a CLV model using structural dynamic programming to replicate consumers' dynamic decision making process. In this way, the model incorporates the effect that marketing variables (e.g., a promotion) has on customer behavior, and allows the researcher to study the effects of marketing policy on CLV.

Fader et al. (2005b) develop a model where the traditional RFM model is linked to CLV by using the Pareto/NBD model to estimate future purchases and a gamma-gamma submodel for spend per transaction. In this model, iso-value (CLV) curves of frequency and recency are estimated and show how for low recency values, customers with high frequency present less CLV than other customers with lower frequency, suggesting iso-value curves are highly nonlinear. The model is calibrated and tested using a hold-out sample.

#### **2.4.2 Survey data**

Rust et al. (2000), and Rust et al. (2004) develop a model that uses survey data to estimate the customer equity of a firm. One of the advantages of this model is that it does not require sophisticated databases and complicated modeling techniques. This might be especially relevant for small firms that do not have the resources to invest in database technologies. Also, this model can handle competition very well. Additionally, the model allows managers to identify which subdrivers of CE are the most important (value equity, brand equity and relationship equity) and where should they concentrate their efforts. The disadvantages are that it assumes purchase volume and interpurchase time to be exogenous, and that it might be difficult to update frequently. Their CLV equation uses markov switching probabilities, and it is therefore subject to the markov property.

### 2.4.3 Company reports

Recognizing customers as important assets of the firm, Gupta et al. (2002) suggest that the value of a firm could be very well approximated using the expected lifetime value of both its current and future customers. They develop a model that uses publicly available data and illustrate it empirically. This model has three main advantages. First, it does not require sophisticated databases. Second, the model is conceptually simple and relatively easy to implement. Finally, it is accessible to people outside of the organization (e.g., analysts). They provide the following equation to measure the value of the firm's customer base,

$$\text{DCE} = \sum_{k=0}^{\infty} \frac{n_k}{(1+d)^k} \sum_{t=k}^{\infty} m_{t-k} \frac{r^{t-k}}{(1+d)^{t-k}} - \sum_{k=0}^{\infty} \frac{n_k c_k}{(1+d)^k} \quad (2.8)$$

and the continuous version of the same equation is,

$$\text{DCE} = \int_{k=0}^{\infty} \int_{t=k}^{\infty} n_k m_{t-k} e^{-d \cdot k} e^{-\left(\frac{1+d-r}{r}\right)(t-k)} dt dk - \int_{k=0}^{\infty} n_k c_k e^{-d \cdot k} dk, \quad (2.9)$$

where  $m$  is the net margin per customer,  $n$  is the number of customers,  $t$  is time and  $k$  is the customer cohort (each year a new cohort is acquired).

The model, however, cannot measure the lifetime value of individual customers or of a group of customers. It also assumes: (a) constant average margin, (b) constant retention rate, (c) the acquisition cost is calculated as the marketing cost divided by the number of new customers, which may be a very strong assumption for mature firms that spend heavily on retention. An interesting feature of this model is the way they estimate the number of new customers, using the Technological Substitution Model.

Wiesel and Skiera (2004) extend the philosophy of the Gupta et al. (2002) model by computing different definitions of Customer Equity and by showing how CE approximates the value of the firm. Specifically, they add indirect customer related expenditures, non-operating assets, and non-equity claims to link CE to shareholder value.

Additionally, Gupta and Lehman (2003) show how to calculate the lifetime value of an average customer using the following formula, which

also relies on publicly available data,

$$\text{CLV} = \sum_{t=1}^{\infty} \frac{m r^t}{(1+d)^t} = m \left( \frac{r}{1+d-r} \right). \quad (2.10)$$

They call the factor  $r/(1+d-r)$  the “margin multiple,” which in their examples ranges from 1.07 to 4.50. The margin multiple is high when the discount rate is low or the retention rate is high. This multiple provides an easy way to estimate the lifetime value of an average customer if the margin is known and the assumptions of the model hold.

Despite the necessary assumptions that models using these data should have, they are very interesting because of their likelihood to impact the business community.

#### 2.4.4 Panel data

The use of panel data in CE models offers a great challenge for researchers and opportunities for future work. Two main advantages come from models that use these data. First, it is easy to explicitly capture the competition; and second, it is possible to model the interaction between the marketing mix and CE. Hence, the use of these data allows the researcher to estimate elasticities of CE to the change in the marketing mix (e.g., price). Using data from the automobile market, Yoo and Hanssens (2005) develop a VAR model in which it is possible to measure the contemporaneous and total effects of prices, sales to existing customers and sales to new customers, on the firm’s CE. They also show how acquisition and retention rates are interrelated, thanks to the fact that VAR models treat all variables as endogenous.

Other stochastic models of CE using panel data could be developed in the future. It could also be an interesting idea to combine these data with the firm’s internal databases and develop models to estimate more precise measures of individual CLV.

#### 2.4.5 Managerial judgment

One problem of marketing science models, and in this CE models are no exception, is that managers rarely use them (Little, 1970). Even though

in the last decades, firms have implemented some marketing models, models that evaluate the lifetime value of customers are rarely used in practice. This occurs because they are relatively new and most require very rich databases and sophisticated modeling techniques. Decision calculus emerged as a powerful tool to help managers make decisions using formal models that use as input their own judgment and usually no hard data (Little and Lodish, 1969, Little, 1970, Lodish, 1971, Montgomery et al., 1971, Little, 1975). Some researchers have warned against the pitfalls of decision calculus (e.g., Chakravarti et al., 1981) while others have supported them (Little and Lodish, 1981).

Decision calculus has therefore potential in the development of CE models. Indeed, the first CE model is based on decision calculus (Blattberg and Deighton, 1996). The objective of their model is to help managers find the optimal level of acquisition and retention spending. They do it by following two steps:

First, the manager responds to two simple questions: (i) How much she spent on acquisition last year ( $A$ ) and what proportion of prospects she got ( $a$ ); and (ii) what is the maximum percentage of prospects she could acquire with unlimited spending (ceiling rate). Assuming the function that translates acquisition spending into acquisition rate is exponential, and the intercept is at 0, they derive the following equation,

$$a = \text{ceiling rate}[1 - \exp(-k_1A)], \quad (2.11)$$

where  $k_1$  measures the steepness of the curve and it is found by solving the previous equation. The net contribution from acquiring a prospect in the first year is  $am - A$ , where  $m$  is the margin per transaction. The optimal acquisition level is obtained by finding the maximum.

Second, the manager responds to the following questions: (i) How much she spent on retention last year ( $R$ ) and what percentage of customers she retained ( $r$ ); and (ii) what is the maximum percentage of customers she could hope to retain if retention spending was unlimited (ceiling rate). Following the same assumptions, they derive the following equation,

$$r = \text{ceiling rate}[1 - \exp(-k_2R)], \quad (2.12)$$

where  $k_2$  is also obtained as a solution to the above equation. The value of a customer in any given year  $t$  is  $r^t(m - R/r)$ . Finally, the average CE per customer can be expressed as,

$$\text{SCE} = am - A + a(m - R/r) \left( \frac{r}{1 + d - r} \right) \quad (2.13)$$

and it is easy to find the value of  $R$  that yields the maximum SCE. This equation measures static CE because it does not include future cohorts of customers.

Their model is elegant and conceptually simple, although it has several limitations. First, it does not simultaneously solve for optimal retention and acquisition. Second, it assumes that the intercept of both the retention and the acquisition curves is at zero (that is, zero spending leads to zero acquisitions or retentions). Third, it is only valid at the aggregate level of the firm. Nevertheless, the model has had a large impact in the management community and many of its assumptions are easy to relax.

Berger and Nasr (2001) extend the previous model to the problem of allocating a promotion budget between acquisition and retention, and among different promotional vehicles and/or market segments. As in Blattberg and Deighton (1996), they offer several illustrative examples.

Managerial judgment can also be used to calculate CLV at the individual level. For example, deterministic models of CLV can be difficult to determine for segments with very few customers, in which historical data may not be informative enough to estimate retention or profitability at the customer level. Ryals (2005) uses managerial judgment from key account managers of an insurer to estimate relationship duration, customer revenue and service costs for each key account. With these data, CLV at each account is estimated.

## 2.5 Future Directions

A close look at the current research in CE models suggests that this area is in its early stages of development. We provide here directions that we think are priorities for future model development.

### 2.5.1 Develop stochastic models of CE

Early CLV models are deterministic in nature. They require many inputs that are sometimes difficult to obtain and their usefulness is limited in that they are merely descriptive, although theoretically could also be used for predictive purposes. But they are not very useful to guide resource allocation. Stochastic models of CE have several advantages over them. First, they can pick up much better the declining and nonlinear pattern of customer lifetimes. Second, they can capture evolution in profits of repeating customers without a very heavy parameterization. Third, they can accommodate much more easily indirect effects, as described above. Finally, it is easier to make some parameters endogenous, especially in VAR models.

### 2.5.2 Incorporate competition

Table 2.3 shows that very few of the current models of CE explicitly incorporate competition, yet heightened competition is known to reduce the CLV of a customer in many different ways: shortening the expected lifetime, decreasing prices, and increasing acquisition costs. Panel data is a very promising data source for some firms, and surveys can also be very useful in capturing the effect of competition.

### 2.5.3 Individual CLV models

Although many of the deterministic models could be applied to calculate the CLV of an individual customer, more research needs to be done when data availability is scarce. For example, bayesian estimation methods could help to estimate the predicted CLV of customers with insufficient purchase history. In the age of one-to-one marketing, the CLV of an individual customer emerges as a powerful metric to guide market resource allocation. Nevertheless, and as pointed out in Section 2.3.7, researchers should bear in mind the possibly low forecasting accuracy of CLV models, and develop models that improve upon it.

### 2.5.4 Develop models that can estimate CE elasticities

Yet in order to allocate optimally, the manager should not just measure CE but instead how CE reacts to a change in the marketing mix.

This could be done at different aggregation levels. With the exception of Yoo and Hanssens (2005), current CE models cannot measure CE elasticities. Some models such as Blattberg and Deighton (1996) or Villanueva et al. (2006b) can measure how CE reacts to an additional dollar spent on acquisition, but not to a 1% decrease in prices or a 1% increase in advertising, because they do not include information on the marketing-mix.

### **2.5.5 Develop more user-friendly CE models**

Decision calculus and models using publicly available data are candidates for rapid adoption. These models are easy to understand and implement by managers and can disseminate the CE concept more rapidly in the management community. If managers “buy” the CE concept, it will be more likely that more sophisticated models of customer valuation will be implemented in the future.

# 3

---

## Drivers of Customer Equity: The Acquisition Effort

---

Firms are enticed to grow by acquiring customers that are new to the market and by stealing customers from their competitors. Furthermore, since there is always a proportion of defectors, a company that does not acquire new customers will see its market share gradually declining over time. Depending on the industry characteristics and the product life cycle, decisions on customer acquisition have a larger or smaller impact on the firm's CE. The following factors are suggested as enhancers of the importance of customer acquisition:

(a) *High switching costs.* When firms realize that customers will present switching costs in the future, they compete more fiercely for acquisition before the customer has attached herself to a supplier (see our discussion about switching costs in Section 4.1 and Blattberg et al. 2001). This happens because, once a customer has been acquired, the company can benefit from the switching cost in later periods (e.g., Klemperer, 1987a, McGahan and Ghemawat, 1994).

(b) *Low switching costs with undifferentiated products.* When switching costs are low and products are undifferentiated, it is relatively easy to induce customers to switch by small price cuts. In those scenarios, we

would expect firms to compete aggressively both for customer acquisition and customer retention.

(c) *Early stages of the life cycle.* There are two reasons why a firm may be motivated to emphasize customer acquisition in the early stages of its life cycle. First, the marginal CE contribution of a dollar spent on acquisition is usually higher than the same dollar spent on retention of existing customers (see discussion in Section 4.2). Second, for some businesses, category adoption depends heavily on imitation effects (Bass, 1969), and therefore the speed with which the firm acquires customers in its early stages of the life cycle will affect its future diffusion process.

(d) *Infrequently purchased products.* For some products the purchase cycle is so large and the retention rates are so difficult to increase (e.g., housing market) that most of the marketing budget goes into customer acquisition, without making a distinction between new and existing customers (Blattberg et al., 2001).

(e) *New entrants.* New entrants into a market have to steal customers from the incumbents. They have to manage their costs of acquisition very well and their future probability of retention, while at the same time considering the incumbents' competitive reaction.

Regardless of the relative importance of customer acquisition, a firm should not acquire just any customer, but the *right* kind of customers (e.g., Reichheld, 1993, Hansotia and Wang, 1997). That is, firms should acquire customers who contribute positively to the firm's CE, after accounting for their acquisition costs. Moreover, the manager has to decide how much is spent on acquisition relative to how much is to be spent on retention or add-on selling (Blattberg and Deighton, 1996, Blattberg et al., 2001), and how to allocate that spending among different acquisition channels.

In this section, we consider models that can help managers answer the following three questions:

- How much should be spent on acquisition?
- What are the characteristics of the best prospects?

- How to allocate the acquisition budget among different acquisition channels?

### 3.1 Optimal Acquisition Spending

In relationship businesses, one of the most important decisions managers should make is setting the acquisition budget. As we showed above, for some businesses this question is more important than for others. In any case, for most firms a large proportion of the marketing budget is allocated into acquisition. How much is spent on acquisition will affect the firm's short and long-run market share, competitive reaction and, as a consequence, its customer equity. Hence, setting an optimal acquisition budget becomes most critical.

#### 3.1.1 The breakeven approach

A first heuristic approach has been to spend on acquiring a customer as long as the cost of acquisition is lower than the expected CLV (e.g., Jackson, 1989a, 1989b, 1989c, Hansotia and Wang, 1997). The point in which the cost of acquisition equals the CLV reflects the breakeven at which the firm is indifferent between acquiring the customer or not. For example, Hansotia and Wang (1997) illustrate this approach with an example from a motor club that wants to acquire new customers by mailing a solicitation package to a pool of prospects. They first develop a response model that estimates a prospect's probability of acquisition as a function of some individual predictors. Assuming an average CLV per customer and a constant acquisition cost, they calculate the breakeven response rate as,

$$r^* = \frac{c}{\text{CLV}}, \quad (3.1)$$

where  $c$  is the cost of contacting a prospect. Customers with a predicted probability of acquisition lower than  $r^*$  will not be contacted. These authors extend their approach to situations in which it is possible to target consumers with different packages and where customers have different expected CLVs.

### 3.1.2 The acquisition response function approach

The previous breakeven approach is valid whenever individual customers are addressable and when the firm has enough individual information so that their response likelihood is estimated with no much error, two conditions that do not exist in many business situations. For example, a cable company must decide on a “subsidized” introductory price and the amount of advertising for a specific region. In those cases, the breakeven point only tells the manager the maximum she can spend to avoid long-run losses but not the optimal spending amount. In these situations, an objective profit function that depends on the acquisition rate and spending level should be built and maximized. An interesting approach has been that of Blattberg and Deighton (1996) in which an acquisition response function is estimated using managerial judgment, and then the first-year value of a customer is maximized (see Section 2.4.5 for a description of this model). This approach, however, underestimates the value generated from customer acquisition and results in underspending. However, their CE specification shown in Eq. (2.13) can easily be maximized to simultaneously find the optimal level of acquisition and retention spending. It turns out that in their model the optimal level of retention spending per customer is independent of the level of customer acquisition spending even when both retention and acquisition are jointly optimized. This happens in their model because of the following two conditions: (i) the retention response function assumes all customers are equal. In other words, once a customer is acquired, it is assumed to behave similar to the existing pool. This can be easily seen in Eq. (2.12) in that the retention response function does not depend on the acquisition response function; and (ii) the optimal budget for acquisition and retention is lower than the budget limit.

The first condition is rarely met in practice. Instead, one should expect that as the saturation point in the acquisition response function approaches, the quality of the acquired customer diminishes and so it becomes harder (i.e., more expensive) to retain that particular customer. The second condition will vary across firms. In many firms the budget for each business unit is set from the top based on simple

proportionate rules instead of using response functions (e.g., Mantrala et al., 1992). Hence, the manager might have a small budget that is far from optimal. When the budget for acquisition and retention is constrained, the optimal acquisition and retention spending per customer will depend on the size of the pool of existing customers and of prospects. That is, depending on how many customers the firm has and how many prospects are in the market, the limited budget will be allocated differently. If the firm has few existing customers compared to the prospect pool, it is likely that the marginal benefit of spending in acquiring customers is much higher than the marginal benefit on retention spending. This is especially true when the firm needs a critical mass of customers to cover its fixed costs. If the budget were unlimited, optimal retention spending per customer in this period would be independent of the sizes of the customer pools.

Finally, it is important to bear in mind that the shape of the response functions (which reflect consumers' responsiveness to spending) will have an important influence on the optimal acquisition and retention spending. For example, an acquisition response function with a large intercept that is close to the saturation level, will result in a low marginal benefit of additional acquisition expenditures. Using concave or S-shaped response functions will also impact the results.

### **3.2 Identifying the Best Prospects**

The best prospect should be the one that brings the highest CE contribution to the firm, not just the one that is less likely to defect or even the one that has the highest expected CLV. Yet the expected CLV might be a good proxy for her CE contribution; e.g., when the size of the indirect effects derived from each customer is linearly proportional to her CLV. In many circumstances this might not be the case. For example, using data from an Internet firm Villanueva et al. (2006b) estimate that customers who were acquired from marketing generate on average a total of 1.77 referrals (i.e., customers who sign-up because of word-of-mouth), while customers acquired through word-of-mouth are expected to generate about 3.64 customers. This means that even at equal levels of CLV, the latter customers are much more attractive

than the former. However, sometimes building a CE contribution metric is very hard given the available data, and other dependent variables are used (e.g., CLV, duration, sales).

We will briefly review two ways of identifying the best prospects:

(a) *Profiling the best customers.* Once the dependent measure has been chosen, a firm should identify which characteristics are candidates to explain the “quality” of a customer. Using data from the current pool of customers, the firm can compare the dependent variable (e.g., CLV) to personal characteristics. This approach has been called *profiling* (e.g., Blattberg et al., 2001). Profiling customers can be implemented by simply ranking the best customers and looking at their average characteristics (duration, sales, household size), or by using a statistical model such as the ones described in Section 4.3. Many companies use database marketing to profile their best customers. For example, Merrill Lynch developed a model that substantially improved its prospecting efforts, using discriminant analysis. Using data from their existing customers, they studied which characteristics can correctly classify customers as high quality. The model implemented outperformed a baseline model by capturing 167% higher assets, 39% higher revenues, and a 43% higher conversion rate (Labe R. P. Jr., 1994). One of the main problems with profiling is that the characteristics of current customers depend on the previous targeting strategy of the firm. Hence, if there were a profile of customers that has great potential but has never been targeted before, no model would detect it.

(b) *Measuring response likelihood.* Another approach to identify the best prospects is to target a small sample from which you have some demographic information, record their response to the offer (yes or no), and estimate a simple logistic regression or discriminant analysis to predict response based on the demographic information. Later, target only those that, based on the estimated model, have a response likelihood that is large enough (see discussion in Section 3.1).

Both of the previous two approaches help the firm identify the best prospects. The first approach helps in understanding which

characteristics are predictive of key customers, once they have been acquired. The second one helps in predicting which prospects are more likely to be acquired, and hence help us in finding out customers that can be acquired in a cost effective way. Both approaches should complement each other.

### 3.3 Acquisition Budget Allocation

In deciding the best allocation of the acquisition spending, managers make decisions at two levels: (i) which major classes or types of media (which we define here as acquisition channel), and (ii) which specific vehicles within each type (see Sissors and Petray, 1976). At the same time, the creativity for a specific campaign plays a significant role, as previous research in advertising has emphasized. We have focused here on the first level, because the second level has extensively been studied by previous research on media modeling (e.g., Gensch, 1973, Rust, 1986).

When a firm only uses direct marketing to target its customers, it can follow the breakeven approach using an individual response model (as described in Sections 3.1 and 3.2) to decide whether or not to target a specific prospect. Nevertheless, most firms use a variety of channels to acquire their customers: all types of advertising (e.g., TV, radio, outdoor), and of direct marketing (e.g., telemarketing, catalogs), sales representatives, public relations, mass coupons, and even word-of-mouth. Some firms use a wide array of channels while others concentrate on a few of them. Each of these channels has a different audience group that can be more or less close to the profile of the best customers. Furthermore, the cost of acquiring a customer varies among these channels. Hence, we can expect significant heterogeneity in the cost effectiveness of these channels.

Given that there could be significant differences in the cost effectiveness of different acquisition channels, we need to develop models that can: (a) measure the cost effectiveness of each channel, and (b) allocate the acquisition budget among acquisition channels in order to maximize CE. Without these models, managers and media planners will allocate suboptimally. A suboptimal allocation can have a great impact into

the firm's profit, even greater than having a suboptimal budget. This will happen when the firm's profit is relatively flat around the optimal budget (Mantrala et al., 1992). Furthermore, these authors argue that when the response function is s-shaped – instead of concave – departure from optimality will be larger.

Some evidence suggesting that this allocation is far from optimal comes from a survey conducted by Forrester (2001). In this survey, Internet managers were asked two questions: (i) which media they believed most effective to drive web site traffic, and (ii) which media they predominantly used. The results show that managers do not predominantly use the media that they believe is the most effective. For example, affiliate programs was said to be very effective, but they were rarely used.

Some of the reasons why managers are expected to allocate sub-optimally are the following: First, many companies who outsource the creativity of their campaigns completely delegate the media planning. Hence, their agencies decide which channels should be used. For example, an advertising agency that focuses almost exclusively on mass advertising will not recommend alternative channels such as direct marketing. Second, managers tend to use “soft” metrics of communication effectiveness (e.g., brand image) and not “hard” metrics of revenues able to capture the Return On Investment (ROI) (Greyser and Root, 1999). Finally, media mix choices are usually decided by the fit between the creativity and the media characteristics (Silk et al., 1997), and therefore not by the long-run profit potential of the media.

## **3.4 Future Directions**

### **3.4.1 Channels of customer acquisition**

Most of the academic literature dealing with channels or vehicles of acquiring customers has focused on advertising, direct marketing, and sales force management. However, some other channels present opportunities for further research. For example, word-of-mouth plays a significant role in many categories, bringing many customers to the firm without any cost. Especially interesting is the study of “incentivized”

word-of-mouth, where the firm pays the existing customer, the new customer, or both, when the prospect signs up as a new customer. Additionally, public relations might be an effective channel to increase the reputation of the firm and attract a significant portion of customers.

### **3.4.2 Normative models of customer acquisition**

Managers need models that help them find the optimal acquisition spending, and the optimal allocation among acquisition channels or among individual customers. Firms use a combination of acquisition channels without a clear understanding of the long-run consequences of their allocations. As we explained above, they delegate decision-making in media allocation to firms that may not allocate optimally across channels, but only within a specific channel or major class of media (e.g., Advertising or TV advertising). Finally, in the field of direct marketing, more research is needed on how to target individual customers in order to maximize CE, not simply immediate market-share, as most current models do.

# 4

---

## Drivers of Customer Equity: Customer Retention

---

The duration of a customer relationship is definitely an important factor strongly correlated with CLV. Hence, it is understandable why researchers and practitioners have been interested in understanding how customer retention (or brand loyalty) can be increased. Most of the previous literature on customer retention has been focused on understanding the determinants of customer retention, and the consequences of higher retention rates.

We first review current literature on the determinants and consequences of customer retention, to later review modeling efforts in this area. We conclude with recommendations for future research. Specifically, we want to address the following questions:

- What are the determinants of customer retention?
- What is the relationship between customer retention and profitability?
- Which models are available to predict customer retention (i.e., churn)?
- What is the cost of losing a customer?
- Why do customers leave?

- How relevant is it to account for the link between acquisition and retention, when developing models of customer retention?

## 4.1 Determinants of Customer Retention

The reasons for which customers decide to be loyal to a specific company or brand over time are not easy to determine. Some of these reasons may be intrinsic to the customer (e.g., propensity to switch or price sensitivity) while others may be extrinsic (e.g., competitors' actions). Some may be easy to be affected by a company, while others may not. It has been shown that in highly competitive markets (e.g., automobiles) only very high levels of satisfaction lead to loyalty. In non-competitive markets, even companies with low levels of satisfaction have high levels of loyalty (Jones and Sasser, 1995). This suggests there are many factors that affect the decision of a customer to continue on a relationship with a firm. We review here the most important factors that previous literature has considered as determinants of customer retention: switching costs, satisfaction and future considerations.

### 4.1.1 Switching costs

There are at least three different types of switching costs (Klemperer, 1987a): (a) transaction costs (e.g., changing banks), (b) learning costs (e.g., moving from Mac to PC), and (c) artificial or contractual costs (e.g., frequent-flyer program). One of the most popular strategies that firms have followed to artificially build costs of switching is the introduction of reward programs (e.g., frequent-flyer programs, 1% cash back when using a specific credit card). Nevertheless, it has been argued that these programs cost a lot of money to the company and are not always beneficial (Dowling and Uncles, 1997), but only when the customer changes her habit and becomes loyal as a consequence of the program (O'Brien and Jones, 1995). Lewis (2004) developed a model to evaluate the long-term effects of a loyalty program through simulation and policy experiments and applied it to an online grocery retailer. Companies have built switching costs by following many other strategies (e.g.,

incompatibility among software programs, providing different levels of service), yet a more fundamental question remains as to whether switching costs help create customer equity for the firm or not.

Switching costs affect the probability of retention and also the level of competition in the market. If a customer has a switching cost of  $\theta$  and is currently buying from firm  $A$ , assuming the products are homogenous, a second firm  $B$  has to offer a price  $P_B < P_A - \theta$  in order to induce the customer to switch. Switching costs have therefore the potential effect of making demand inelastic and reduce rivalry among firms in the market (e.g., Klemperer, 1987a, 1987b). Moreover, in the age of information, when firms have perfect information on each individual's switching cost, they can target individual customers with customized prices, making switching much more difficult. But, recognizing the value of locking-in customers might increase competition in the early stages of a market, when firms compete to gain market share (Klemperer, 1987a). This could potentially have negative effects in the long-run (e.g., reducing consumers' reference price and switching costs). Moreover, if customers behave strategically, they could recognize the implication of attaching themselves to a supplier, and request lower prices when starting a relationship for the first time. This could increase price competition for initial market share. Additionally, consumers' switching costs are not static and may depend on their experience with a product or in a category. For example, consumers starting to buy on the Internet would be expected to experience higher switching costs than consumers who have a longer purchase history in this channel. In summary, even though intuition suggests that switching costs will benefit the firms in the market, it is not clear this will happen in all markets and for all firms.

Consumers' switching costs have also been recently studied with the advent of the Internet. It has been argued, for instance, that costs of switching will be radically diminished on the Internet because comparing prices and visiting stores is much easier than in a brick-and-mortar world (Alba et al., 1997, Bakos, 1991, 1997). Nevertheless, loyalty on the Internet has been found to be higher than in the offline world (Brynjolfsson and Smith, 2000) and it has also been argued that consumers present lock-in effects on the Internet and do not search

as much as expected (Johnson et al., 2004, Zauberan, 2003). For example, Johnson et al. (2004) using Mediametrix data have modeled visits behavior of a pool of customers in the music category. They find that consumers search less than it was expected, presenting evidence of “lock-in” effects. Zauberan (2003) using an experiment that studied consumer behavior in the travel agency category, also presents evidence of “lock-in” effects and what he defines as “consumer’s myopia”: consumers fail to predict future switching costs. Hence, even though theory suggests that the characteristics of the Internet as a purchasing channel would reduce consumers’ switching costs, this has not been empirically proven.

We can conclude that much needs to be studied on the relationship between switching costs and long-run financial performance, particularly when competing firms build switching costs through expensive reward programs and customers behave strategically. Sometimes retention increases because of higher switching costs, could come at the expense of lower profitability for the firm, because of the costs of building switching costs.

#### **4.1.2 Customer satisfaction**

The literature on customer satisfaction and its link with customer retention is important to mention here. Some researchers have shown empirical evidence of a positive correlation between customer satisfaction and profitability and/or the duration of the relationship (e.g., Anderson et al., 1994, Bolton, 1998, Bolton and Lemon, 1999, Garbarino and Johnson, 1999, Gustafsson et al., 2005). The nature of the links between satisfaction, retention and profitability are not necessarily immediate, symmetric and linear (Anderson and Mittal, 2000), and hence some firms might be mistaken if cutting investments on increasing satisfaction as a consequence of not being able to observe a direct and immediate impact on retention and profitability. More recent research has suggested as an empirical generalization the relationship between improvements in customer satisfaction and firm’s financial performance (Gupta and Zeithaml, forthcoming).

The relationship between satisfaction and retention has been studied extensively before. For example, it has been argued that customer characteristics affect the relationship between satisfaction, repurchase intent and repurchase behavior (Mittal and Kamakura, 2001). More specifically, these authors show that customer characteristics affect satisfaction thresholds and response bias. Furthermore, they find that the relationship between satisfaction and repurchase intent is highly nonlinear, although the nonlinearity does not depend on customer characteristics. Bolton and Lemon (1999) use a dynamic model of customer usage of services that identifies causal links between customers' prior usage levels, satisfaction evaluations, and subsequent service usage. They also introduce the concept of payment equity, which helps to determine how customers use price and usage over time to update their evaluations of the fairness of the exchange. They find that evaluation affects overall satisfaction and future usage. Garbarino and Johnson (1999) argue that satisfaction is a good predictor of future intentions for low relational customers, whereas for high relational customers trust and commitment, rather than satisfaction, are the important constructs. They use structural equation analysis on data of customers of a New York off-Broadway repertory theater company. Finally, Seiders, Voss, Grewal and Godfrey find that satisfaction has a direct effect on repurchase intentions but not on repurchase behavior, where customer and marketplace characteristics play important moderating roles.

The relationship between satisfaction and profitability has also been studied extensively. For example, Anderson et al. (1994) use a database that consists of a national customer satisfaction index and measures of economic return (e.g., ROI) to test a series of hypotheses and conclude that satisfaction is positively related with profitability. Bolton (1998) shows that customer satisfaction can have an important effect on the value of a firm. She finds that satisfaction is positively related to lifetime revenues, and more specifically, to the duration of the relationship and the dollar amount of purchase across billing cycles. Bolton uses a left-truncated, proportional hazards regression model with cross-sectional and time-series data in the cellular category. Anderson et al. (2004) also find a positive association between customer satisfaction and shareholder value.

In conclusion, even though there are other determinants of customer retention, all else equal, satisfaction has been shown to impact the relationship's duration positively, and as a consequence, profitability. Different findings in this area suggest that at least the link between satisfaction and retention could be established as an empirical generalization.

#### **4.1.3 Expected future usage and expected regret**

It has been argued that when consumers decide whether to continue a relationship with a firm, they take into account future considerations regarding the product or service, such as anticipated future benefits (e.g., usage) and anticipated future regret (Lemon et al., 2002). These authors argue that customer retention models that do not account for customer's future orientation will be incorrectly specified. This is especially important for ongoing services. As a consequence, marketing decisions that try to maximize retention will not be optimal if based on these models. Recognizing the importance of consumer's expected future benefits and regret can help marketing managers integrate these concepts in their communications in order to reduce churn.

In summary, several determinants of customer retention exist and can be managed for maximizing retention. Firms usually invest to reduce churn rates using different programs, and understanding the key determinants of customer retention is extremely important to increase retention and possibly customer equity. Models should be developed that help the firm make decisions on how much and where to spend in order to increase retention up to a desired level. Maximizing retention does not always lead to maximum profits, as we shall explain in the next section.

## **4.2 The Relationship Between Customer Retention and Profitability**

This section first reviews the possible benefits of customer retention in terms of its contribution to a firm's CE, and then concludes with

some insights on how the shape of the response functions affects these benefits.

#### 4.2.1 The benefits of customer retention

The literature on CRM has argued that there are several benefits attached to long-life customers (e.g., Reichheld and Sasser, 1990, Reichheld and Teal, 1996), which are basically contained in the following five propositions:

- PROPOSITION 1. *It is cheaper to retain customers than to acquire them*
- PROPOSITION 2. *The costs of serving long-life customers are less than those of serving new customers*
- PROPOSITION 3. *Long-life customers increase the reputation of the company and attract new customers through word-of-mouth (WOM)*
- PROPOSITION 4. *Long-life customers are less price sensitive than new customers and therefore pay higher prices*
- PROPOSITION 5. *Long-life customers are more likely to buy more from the company, so that the company can increase their share-of-wallet through up-selling and cross-selling*

These five propositions can be summarized in two:

- PROPOSITION 6. *Customer lifetime is positively related to profitability.*
- PROPOSITION 7. *Profits for retained customers increase over time.*

Despite the apparent intuition of these propositions, little empirical evidence has been published. An exception is the work of Reinartz and Kumar (2000) where it is shown that long-life customers are not necessarily profitable in a noncontractual setting. They use three years of data of a catalog company, and estimate the probability of each customer being “alive,” using the NBD/Pareto model suggested by

Schmittlein et al. (1987) and Schmittlein and Peterson (1994). They calculate the CLV of each customer and finally compare it to their estimated duration to test Propositions 2, 4, 6 and 7. Reinartz and Kumar (2002, 2003) extend their findings to the customer base of four different companies. They find that some of the long-life customers are much less profitable than transactional customers. Moreover, they find that (i) in none of the four companies, long-life customers were cheaper to serve, and (ii) in one company, long-life customers paid lower prices and no significant difference was found for the other firms. They argue that the violation of the previous propositions will be more likely to happen in noncontractual situations, in which there is a significant cost to entice customers to repeat buying.

We provide in Table 4.1 both arguments in favor and against these propositions. We think Proposition 1 is too strong and only valid for certain situations. Propositions 2 to 5 can be very context specific. Propositions 2 and 4 have been shown not to be true for several firms in the published studies that have formally tested them (Reinartz and Kumar 2000, 2002, 2003). The same studies question Propositions 6 and 7. The bottom line is that a firm should expect heterogeneity in the customers' behavior, and not necessarily all customers will exhibit these effects. Believing Propositions 1 to 7 as true in all scenarios might lead the firm to suboptimal spending decisions. Instead, continuous measurement and management of that heterogeneity is key for the long-run success of the firm. Models that find the optimal retention spending – especially under a limited budget – should correctly specify the acquisition and retention response functions, as we will show next.

#### 4.2.2 The shape of the response functions

The shape of the response functions helps resolve some of the contradictory findings explained above. These insights should be incorporated in models that capture the link between retention spending and CE.

- (a) *The relationship between retention spending and retention rate may follow a concave or an s-shaped pattern.* The relationship between retention spending and retention rate cannot be linear, nor can it be convex since the function has

Table 4.1 The relationship between customer retention and profitability

Proposition	Arguments in favor	Arguments against
P <sub>1</sub> : <i>It is cheaper to retain customers than to acquire them</i>	Whenever there are switching costs, the company that already has a relationship with a customer has an advantage over its competitor. Recognizing the CLV of a customer might lead the company to offer lower prices in order to acquire a customer for the first time.	There must be an interval in which the return on CE from acquisition spending is higher than the return on CE from retention spending. For instance, when the firm is in its early stage and needs to grow fast to reach a critical mass of customers.
P <sub>2</sub> : <i>The costs of serving long-life customers are less than those of serving new customers</i>	Customers learn over time and have less questions or problems. Higher marketing efficiency.	Reward programs may be expensive and increase the costs of serving long-life customers. Long-life customers may require better service in reward for their loyalty.
P <sub>3</sub> : <i>Long-life customers increase the reputation of the company and attract new customers through word-of-mouth (WOM)</i>	Long-life customers are usually more satisfied and hence generate referrals. The longer the relationship, the higher the probability of spreading WOM.	Long-life customers might be heavy users and buy at the same time from different vendors. So their referrals are spread among different firms. Some customers even with low satisfaction levels tend to exhibit high levels of loyalty (e.g. if they have high switching costs).
P <sub>4</sub> : <i>Long-life customers are less price sensitive and therefore pay higher prices</i>	Product preference: Because consumers value the product more than that of the competitors, they are willing to pay a price premium. Switching costs: retained customers are more likely to have high SC. SC could even increase over time.	Long-life customers might be heavy users who have better price information about the competition.

Table 4.1 (Continued)

Proposition	Arguments in favor	Arguments against
P <sub>5</sub> : <i>Long-life customers are more likely to buy more from the company, so that the company can increase their share-of-wallet through up-selling and cross-selling</i>	As they repeat with the firm, customers become aware of more products and, if satisfied, buy them.	As the experience with that product category increases, consumers become aware of other players in the market and might start to buy from more than one vendor.
P <sub>6</sub> : <i>Customer lifetime is positively related with profitability</i>	If P <sub>1</sub> – P <sub>5</sub> are true, the higher the average retention, the higher the profitability for the company.	Some transactional customers might buy more and at higher prices than long-life customers.
P <sub>7</sub> : <i>Profits for retained customers increase over time</i>	Customers use the product slowly at first, then start to exhibit the effects described in P <sub>2</sub> – P <sub>5</sub> Reichheld and Sasser (1990) argue they found that trend in 100 analyzed firms in 24 industries.	Not if the firm is spending too heavily on retention. The marginal benefit from customer retention might be lower than the cost of increasing retention.

to be bounded by 1.<sup>1</sup> The simplest functional form is a concave one, with its origin at zero (see Blattberg and Deighton, 1996). Nevertheless, imposing this form might be too restrictive in some scenarios. First, it is more than likely that spending zero in retention will lead to positive retention rates. Second, a more complex functional form might be imposed (e.g., s-shaped). This could occur, for instance, when the likelihood of an individual continuing a relationship is significantly increased only after a certain spending threshold, or point of inflection. Or when the distribution of the switching costs is not uniformly distributed but denser in a certain interval. Hence, slightly increasing retention spending in that interval will increase the average retention rate significantly.

- (b) *There exists a point of diminishing returns from which further spending on retention would lower the firm's customer equity.* Previous literature recognizes that some customers may not be profitably served, and therefore the zero defections argument should be followed only for those customers that are profitable to the company (e.g., Reichheld and Sasser, 1990). This insight follows from the previous one in that, when the retention rate gets sufficiently high, further increases in retention become very expensive and there has to be a point from which the marginal cost of an increase in retention is higher than the marginal benefit. This can be seen for instance in Blattberg and Deighton's (1996) work in which the model finds precisely the level of spending that maximizes the CLV of a customer.
- (c) There exists at least one interval in which the return on customer equity from one dollar spent on increasing retention is lower than the return from the same dollar spent on acquisition. Intuitively, firms starting their businesses first need to acquire a customer base that is sufficiently

---

<sup>1</sup>In the case in which retention spending is related to number of retained customers instead of retention rate, the bound should be the number of current customers.

large, if they want to become profitable (e.g., they need to cover their fixed costs). If the customer pool were still small, further investments in retention would derive less customer equity contribution than investments in acquisition. Similarly, when customer retention rates are very high, it might be more profitable to acquire customers with high CE potential than to retain customers with low interest in the firm. And this could happen even before reaching the point of diminishing returns. There might be intervals for intermediate retention spending where the CE elasticity of acquisition spending is higher than that of retention, especially if the relationship between spending and retention is s-shaped. This will also depend on how mature the market is, the number of competitors, product differentiation, and the magnitude and distribution of the switching costs.

### 4.3 Models to Estimate the Probability of Retention

A particularly interesting question is the measurement of how “alive” a customer really is. We deal here with models that can answer all or some of the following questions:

- How many customers does the firm have?
- How likely is a particular customer to leave?
- How can marketing spending increase average retention rates?
- How can marketing spending increase the retention probability of an individual customer?

All these questions are extremely relevant for the firm, and they emphasize the importance of this area for future modeling work. The most suitable model will depend on the characteristics of the customer-firm relationship. More specifically, it will depend on the following five factors: (a) *contractual versus non-contractual relationships*. Especially in noncontractual situations (e.g., catalog retailer), firms often do not know how many active customers they have and, for a particular customer, what is the probability that she will buy again from the firm;

(b) *always-a-share versus lost-for-good*. When the customer buys from more than one firm at the same time (e.g., credit cards), a substantial drop in the quantity of products bought might signal the customer is about to leave, but in lost-for-good situations (e.g., health insurance) the information used to predict switching might be harder to observe because there is no behavioral information other than the premiums paid; (c) *frequency of purchases*. For infrequently purchased products (e.g., PDA) it might be very difficult to predict when a customer is about to make a decision to either repeat with the same brand or switch, whereas for frequently purchased goods there is usually more historical information from previous purchases; (d) *addressability*. Even when consumers purchase a product frequently, in many situations the company cannot address customers individually and hence information on individual retention rates is worthless; (e) *targetability*. Some firms have much richer information about consumers' preferences than others, which could allow the firm to estimate individual retention probability more precisely.

In this section, we briefly review current modeling efforts in this area. Not all models are helpful for all situations.

#### **4.3.1 Heuristic methods**

Managers often establish decision rules based on simple but reasonable heuristics in order to detect consumer segments with high (or low) probabilities of churn. For example, the fact that a customer wrote a letter of complaint indicates she is more likely to defect than other customers. This can even be validated using past information, not only managers' intuition. A very popular model that is frequently used by firms to detect quality degrees of customers is the RFM model (Recency, Frequency, and Monetary value). We categorize the RFM model as heuristic because it simply uses past information on recency, frequency and monetary value to rank order customers (see Roberts and Berger, 1989). If the objective is to predict probability of retention, the RS matrix has been suggested as more appropriate (see Blattberg et al., 2001, p. 91). The RS (Recency, Sales) matrix calculates for each customer her recency (time since last purchase) and the average interpurchase

time. The higher the recency is than the interpurchase time, the more likely is that customer has been lost. The advantage of these methods is that they are cheap, easy to implement, and easy to understand by managers. Yet they are overly simplistic and in most cases highly inaccurate. Therefore, formal statistical models can do much better in optimizing the firm's retention spending, such as the one developed by Fader et al. (2005b).

### 4.3.2 Binary logit model and discriminant analysis

Both of these models can be used to predict the probability of churn (or retention) of individual customers, based on some covariates. The dependent variable is binary (churn or retention) but can be ordinal with more than two outcomes (in which case we have a multinomial logit instead of a binary logit model). Discriminant analysis is a linear model and may not work as well as the logit in some situations.

A particular problem is when the proportion of 1's and 0's is highly unbalanced. In those cases, neither model would work well because they would predict everybody being either 0 or 1 (the one with a larger proportion of customers) and still have one of the groups perfectly classified, while the other group is not well classified at all. In these situations researchers may opt for balancing the sample used for estimation (see Ben-Akiva and Lerman (1985) for a discussion on choice-based sampling), but that would have an effect on a bias of the intercept if traditional estimation techniques are used (e.g., MLE).<sup>2</sup>

Another problem of these models is that they are hard to use for always-a-share customers in both non-contractual (e.g., catalog retailer) and contractual (e.g., credit card) situations. In the former case because it is not clear at any point in time whether a customer will buy again in the future or not, and in the latter case because the relevant information is not only whether the customer still has the credit card or not but also to what extent he is using it as much as he did in the past.

<sup>2</sup>Two alternative solutions have been suggested in the literature. One is to use what is called the weighted endogenous sampling maximum likelihood (WESML) estimator (Manski and Lerman, 1977), and the other is adding a constant to the estimated intercept (Manski and Lerman, 1977, McFadden).

### 4.3.3 The SMC model

Perhaps the most complete model to estimate the probability of a customer being alive, this model was first developed by Schmittlein et al. (1987) and later extended by Schmittlein and Peterson (1994). In order to compute the probability of a customer being alive, this model only uses information on the purchasing or transaction events. More specifically, it only needs the following information for each customer: the observation period, number of purchases in that period, and the time at which the last purchase was made. The model used is the NBD/Pareto, in which the following assumptions are made: (i) for individual customers: Poisson purchases, and exponential lifetime; and (ii) to account for heterogeneity across customers: individuals' purchasing rates are distributed gamma, death rates are distributed gamma, and purchasing rates and death rates are independent. Once estimated, the model can compute the probability of being alive for each customer at each point in time. Unlike some other models explained in this section, the SMC model takes into account the individual interpurchase time of customers in order to predict her probability of being alive. Schmittlein and Peterson (1994) extend it to account for purchase volume. The model is especially suitable for situations in which there is not a contractual relationship, because in these situations, the firm does not observe directly how many "active" customers it has.

Despite the powerful insights of this model, very little research has used it, with the exception of Reinartz and Kumar (2000, 2003) and Fader et al. (2005b). One possible reason is its difficult estimation. The likelihood function of this model is difficult to tract, and with the exception of Reinartz and Kumar (2003) the only estimation procedure used is based on a two-step method of moments suggested by Schmittlein and Peterson (1994). The method of moments, although consistent, does not guarantee efficiency of the parameter estimates (Greene, 2000).

Fader et al. (2005a) modify the SMC model to a beta-geometric/NBD and show how the parameters can be obtained easily even in Microsoft Excel. They show how the two models yield very similar results in different data settings.

Using stochastic models such as the Pareto/NBD and the BG/NBD presents a clear advantage over scoring methods because previous customer behavior is better modeled. Also, nonlinearities of iso-value curves between recency and frequency can be much better modeled (Fader et al., 2005b).

#### 4.3.4 Hazard models

Hazard models have been extensively used in marketing research to study interpurchase time (for a review of hazard models in marketing see Leeflang et al., 2000, Helsen and Schmittlein, 1993). One of the main advantages of these models is that they can handle very well right censoring in the data and are usually easy to estimate. Another advantage is that they can also include covariates, such as individual demographic information. One such example is the Cox proportional hazard model (Cox, 1975),

$$\lambda(t|x(t)) = \lambda_0(t) \exp[x(t)'\beta],$$

where  $\lambda_0(t)$  is the baseline hazard, and  $x(t)$  a vector of covariates. The second term shifts the baseline hazard depending on the covariates included on  $x(t)$ . Both continuous-time and discrete-time specifications have been applied in the marketing literature before. One problem of these models is when the probability of churn increases as the contract end approaches. In these cases, one should expect the hazard function to be non-smooth, with some kicks at the contract end date. The previous model cannot account for that effect, so it should be modified in those scenarios. These models are also limited to business situations in which the duration of the relationship is observed, since otherwise the data is not appropriate for the model.

#### 4.3.5 Artificial Neural Networks (ANN)

A neural network translates some inputs to a prediction of an outcome, using the weights that, once applied to the inputs have the highest predictive ability.<sup>3</sup> It could therefore be used, for instance, to predict how

<sup>3</sup>For an introductory description of ANN, look at Berry and Linoff (1997).

likely a customer is to churn. A training set is used to calculate the internal weights. Different combination functions can be used to apply weights to the inputs, such as a weighted sum, or even logical operators (e.g., AND or OR). Finally, a transfer function transfers the value of the combination function to the predicted output. The most widely used transfer functions are the sigmoid, linear, and hyperbolic. Different algorithms can be used to train a neural network (e.g., genetic algorithms). Logistic regression or even discriminant analysis can be considered special cases of ANNs. The main criticism of a neural network is that it is considered a “black box,” because it is usually impossible to understand how the model works internally, and the prediction is the only output of the model.

#### **4.3.6 Decision trees**

In a decision tree, a set of rules is used to classify individuals. For example, if we are interested in classifying people depending on whether they can be considered as defectors or not, we could use a set of questions such that: (i) Did the customer buy in the last three months?; (ii) Is the recency/sales ratio larger than X?; (iii) Is the customer older than a certain age?; and so on. The questions, or rules, can be developed by the experts in that industry and the general idea is that the tree be developed so that questions that are better at discriminating among classes or segments should be asked first. The answer to the first question determines the subsequent question. The accuracy of a tree in classifying individuals can be easily studied using a sample of records and measuring the percentage of people correctly classified. Different algorithms can be used for designing a decision tree, like CART, or CHAID (see Berry and Linoff, 1997). Like in the ANN, a training set is also needed in order to design the tree.

#### **4.3.7 Combined models**

An interesting approach to model churn is that of combining different models to increase the predictive accuracy of the resulting model. Indeed, the winner of a churn tournament organized by the Teradata Center for CRM at Duke University was a model that used a

combination of decision trees (Neslin et al., 2006). Another example is that of estimating a sequence of binary choice models from resampled versions of a calibration sample. A final choice model can be obtained by simple aggregation. Lemmens and Croux (2006) show that when these models are used they can significantly improve the predictive accuracy of a simple binary choice model.

In summary, the researcher can use a wide array of models to predict the probability of churn at an individual or a segment level. Some of these models could also be used to “count” customers, when the number of active customers is unknown (e.g., in noncontractual situations). Not all of these models can be used in all situations, and the decision of the model to be used will depend on several factors such as data availability, complexity, and suitability for the particular customer-firm relationship.

## **4.4 Customer Defections**

The importance of managing customer defections has been recognized early in the customer relationship literature (e.g., Reichheld and Sasser, 1990). Managers should be able to identify defectors and predict those with a higher likelihood to churn. Once those “vulnerable” customers have been identified, marketing intervention strategies should be developed in order to minimize defections. Marketing science can help in the development of models that identify “vulnerable” customers and in designing models that can calculate the net effect of marketing strategies aimed at minimizing churn. But before discussing these two topics in more detail, first we should discuss what is the cost of losing a customer.

### **4.4.1 The cost of losing a customer**

We argued in Section 2.3 that the value a new customer adds to the company should be measured not only by its own CLV but also by any indirect effects that the firm experiences through that particular customer acquisition. For example, word-of-mouth generation (Reichheld and Sasser, 1990) and category adoptions (Hogan et al.,

2002b) are considered as “extra” effects of an active customer. Hence, when a firm loses a customer, it not only loses her expected CLV, but also all the indirect effects expected from that relationship had the customer continued. Defectors can also negatively affect future customer acquisitions through negative word-of-mouth. The value of a lost customer (VLC) has been suggested by Hogan et al. (2002b) as,

$$\text{VLC} = \alpha \text{VLC}_{\text{disadopter}} + (1 - \alpha) \text{VLC}_{\text{defector}}. \quad (4.1)$$

These authors argue that in some industries  $\alpha \rightarrow 0$  (e.g., industries that are not technologically intensive), while in others this parameter could be higher. As we argued above, a good metric to calculate the value of a lost customer could be her customer equity contribution. Nevertheless, it would be a mistake to assume that the VLC is the CE contribution of an average customer, because customers who defect do so by different reasons and in many cases they will be expected to have lower CE contributions than active customers. Hence, models that estimate the cost of losing a customer should recognize that customers are heterogeneous and that some of the lost customers could not be profitably retained for further periods.

#### 4.4.2 How and why do customers leave?

Before developing a model to predict probability of churn, companies should first pay close attention to both the transactional pattern of customers about to leave, and the reasons for defection. The answer to these two important questions will help in developing models to predict churn. For example, in some situations, the information contained in a previous history of purchases might not be strong enough to predict churn and models such as the SMC will be very inaccurate. This could happen, for instance, when most of the customers who defect do so because of a complaint that was not very well handled by the firm. Including in the model not only transactional but also other kinds of information (e.g., complaints) will help the firm to accurately estimate probability of churn. In these cases, data mining models (e.g., ANN or decision trees) might work better than some stochastic models.

### 4.4.3 Intervention models to reduce defections

A particularly interesting, yet under researched area in marketing is the development of models that can relate marketing spending to retention probability. These models are necessary to assess the potential profitability of different strategies aimed at minimizing churn, even at the individual level. It is important that these models consider both the costs of such strategies and the long-run change in customers' behavior. To give an example, there is evidence that price promotions have a high elasticity of demand in the short-run, but almost negligible in the long-run (e.g., Nijs et al., 2001). Another interesting area for future research is the development of models that measure market response of customers who already defected.

## 4.5 The Acquisition-Retention Interface

As we discussed in Section 3, the acquisition effort has a clear impact on the quality of the customer pool. Hence, it has been suggested that retention models that do not capture this relationship will generally be biased. For example, we explained in Section 3.1 that the reason why the optimal retention spending per customer in the Blattberg and Deighton (1996) model does not depend on the acquisition spending is that their retention response function (Eq. (2.12)) is independent of the acquisition process. If these functions were indeed dependent, the optimal allocation between acquisition and retention would depend on that relationship.

In an interesting empirical research, Thomas (2001) proposes a new methodology that studies customer retention, while accounting for the impact of acquisition on retention, even when data on prospects are not available. Her model addresses both truncation and censoring, and shows that models that do not account for acquisition will give biased estimates of the duration of a relationship. This, of course, will therefore bias any estimate of CE or CLV.

The offer with which a customer is acquired has an important impact on future CLV. Lewis (2006) finds that promotionally acquired customers have lower repurchase rates and smaller CLVs. He offers empirical illustrations of this phenomenon for the customer base of a

newspaper and an online grocer. The same author develops a dynamic programming approach to optimal pricing that maximizes customer value. In his empirical illustration he suggests it is better to offer a series of price discounts throughout the customer's tenure rather than a deep promotion discount to first time buyers (Lewis, 2005a).

An interesting line of research that links acquisition and retention is that of re-acquiring lost customers. Stauss and Friege (1999) offer a model for winback strategies, and Thomas et al. (2004) build on that research to study the optimal pricing strategy to lapsed customers. They explicitly model the Second Lifetime Value (SLTV), which is the expected CLV of a re-acquired customer. These authors warn that re-acquired customers tend to have lower lifetimes than new customers and hence their expected CLV should be considered while setting prices for acquisition and retention.

The bottom line is that the retention process cannot be completely isolated from the acquisition process. The researcher should take good care that her retention model is not biased if it does not explicitly account for this relationship.

## **4.6 Future Directions**

We provide the following list of research opportunities in the area of customer retention and its impact to CE management.

### **4.6.1 On the relationship between retention and profits**

Most of the literature on relationship marketing has taken Propositions 1–7 as true, yet little empirical evidence has been published. As we explained above, Reinartz and Kumar (2000, 2003) have challenged some of these Propositions testing them using purchase behavior in different firms. Research should examine which laws govern this relationship and which factors affect them. For instance, it could be the case that customer characteristics might explain profitability more than lifetime duration does. Models should be developed that could capture these effects, as we have explained in Section 2.3.

### **4.6.2 A better understanding of available models to predict retention**

As we have explained above, there are a variety of models that can help managers in predicting customer retention (or churn). There is a need for a better understanding of which of these models are better and for which situations. Firm specific factors and database limitations might incline the researcher for the use of one model or another.

### **4.6.3 Developing normative models of customer retention**

There is a need for models that can guide action with the objective of maximizing CE through customer retention. That is, we need models that maximize long-run profits through increases in customer retention, while taking into account the corresponding marketing costs. For example, the relationship between switching costs and customer retention has been extensively researched. But the manager should know how much to spend on increasing switching costs, based on the CE derived from such strategy. The same applies to the management of customer satisfaction. It is clear that, as we explained above, there must be a point from which further spending on customer satisfaction will increase retention levels, but at a marginal benefit that is lower than the marginal cost. Models could also study how to spend on customer retention on different segments of customers. For that it is necessary to build different retention response functions, one for each segment.

### **4.6.4 Specifying the acquisition-retention interface**

There are few papers that have created models of customer retention that explicitly incorporate this interface, as we have covered above. There is an urgent need for more research on this area.

### **4.6.5 Intervention models of customer defection**

Surprisingly, there is a lack of research on normative models that can help manage customer defections. It would be interesting to measure how customers that are about to leave respond to different targeting strategies (e.g., a price cut or a month for free). For example, AOL

often offers a customer who calls to cancel her subscription some periods for free without any obligation. This strategy increases the probability that the customer keeps the service for a longer period, but has some costs associated with it, and may trigger strategic behavior from some opportunistic customers. Additionally, normative models to guide re-acquisition of lost customers are very important for a company with high addressability and low costs of targeting. In those scenarios, targeting previous customers with customized offers might represent a source of competitive advantage.

# 5

---

## Drivers of Customer Equity: Add-on Selling

---

As we explained in Proposition 5 some researchers have argued that long-life customers are expected to buy more from the company than new customers (e.g., Reichheld, 1993). This process, known as “add-on selling” (Blattberg et al., 2001) consists on increasing sales due to the following effects: (i) cross-selling; (ii) up-selling; and (iii) higher quantity of the same product or service. Clearly, add-on selling is one of the most important factors to maximize CE in that it increases the baseline CLV of an acquired customer. For example, Jackson (1989a, 1989b, 1989c) illustrates how by offering six products (cross-selling and upgrades) to existing insurance policy owners during a 12 month period, a firm increased the baseline CLV of an average customer by 40%. This means that the acquisition allowance can be substantially increased once the total CLV of a potential customer is accounted for. The add-on selling effect can occur as a natural process (e.g., long-life customers, if satisfied, will be more likely to try and buy new products from the firm). But most multi-product firms adopt an aggressive strategy targeting specific customers to induce cross- and up-selling. For example, Rust and Verhoef (2005) show how to optimize marketing interventions at the individual level across several contact media with the objective

of increasing next period customer profitability. We will focus our discussion on modeling developments aimed at optimizing the firm's proactive strategies to increase add-on selling.

The economic value of an add-on selling offering depends on the following factors (Blattberg et al., 2001): (i) number of feasible add-on offers per period per customer ( $I$ ); (ii) response rate of each offer ( $s$ ); (iii) sales per add-on offer ( $S$ ); (iv) margins of the offer ( $m$ ); (v) number of customers targeted ( $H$ ); and (vi) per capita cost of targeting the product offering ( $c$ ). We add a seventh factor, which is: (vii) the subsequent retention rate for that offer ( $r$ ). The resulting net present value (or CE contribution) for the add-on selling strategy could be calculated as,

$$\text{VAOS} = \sum_{h=1}^H \left[ \sum_{i=1}^I \left( \sum_{t=1}^T S_i s_{ih} m_i \frac{r_h^t}{(1+d)^t} \right) - c_{ih} \right]. \quad (5.1)$$

The retention should only be included when the additional products or services sold are expected to generate repeat purchases in the future. Additionally, when there is not enough information or when a group of homogeneous customers is targeted with the same product offerings, the equation could be simplified by using average responses and costs per offering. Equation (5.1) highlights some of the most important questions that models for add-on selling should address:

(a) *Product offering.*

- Which products should be offered?
- How many products should be offered?
- What is the best timing of product offerings?

(b) *Market response*

- What is the average expected response of a product offering?
- What is the expected response from a specific customer to a specific product offering?
- What is the expected response from a specific customer to product offerings made through a specific channel (e.g., telemarketing, direct mail)?

In the following section, we review current modeling efforts in this area and we later propose future modeling directions.

## 5.1 Models for Product Offering Selection

The questions of which products to offer to whom, how many, and eventually the timing of the different offerings are extremely important in order to optimize the add-on selling strategy. Knott et al. (2002) provide some guidelines on these modeling issues: which statistical model, which data, and which sample. We review here some modeling approaches to deal with this issue.

### 5.1.1 Cross-buying analysis

One very simple metric is called cross-buying analysis (Blattberg et al., 2001). This metric uses information on the percentage of times products  $i$  and  $j$  are bought together ( $Z_{ij}$ ), and the percentage of times that these products are bought without the other, which is denoted by  $X_i$  and  $X_j$ . These authors provide the following metric,

$$CB_{ij} = \frac{Z_{ij}}{X_i X_j}.$$

The larger this metric is, the higher the cross-buying effect between these two products. This metric could be used, for instance, to decide which product to offer to customers who have bought one product but not another one that has a high cross-buying likelihood. This metric is aggregated and it does not take into account individuals' behavioral and demographic information. Incorporating that information in a model could significantly improve the success of cross-buying strategies.

### 5.1.2 Collaborative filtering

This technique, developed in the early nineties, has recently attracted considerable research interest (e.g., Resnick and Varian, 1997). The central idea of collaborative filtering or recommender systems is to predict the utility of an individual for a given product or service based on the preferences of similar people. For example, a firm selling a large number of products (e.g., books) can use the information from

previous purchases of a particular customer and similar customers to recommend specific books to that customer. This technique can use two different kinds of information from consumers, which are called *explicit votes* and *implicit votes*. Explicit votes refer to the situation in which a particular consumer rates individual products (e.g., Netflix, an online DVD rental firm asks its customers to rate previously seen movies). Implicit votes refer to less salient information about consumers' utilities for the products (e.g., Amazon uses purchase history as a predictor of consumer's tastes).

Once the votes have been gathered, an algorithm is needed to predict consumers' utilities for specific products. There are two general classes of collaborative filtering algorithms (Breese et al., 1998): memory-based and model-based. Memory-based algorithms use the entire consumer's database to predict, while model-based algorithms estimate the expected utility for a certain product.

Assume a firm has  $n$  customers, each providing votes  $z_{i,j}$  to a certain number of products, where  $J_i$  represents the products for which consumer  $i$  has voted. Hence, consumer  $i$  average vote is calculated as  $(1/J_i)\sum_{j\in J_i} z_{i,j}$ . An example of a memory-based specification is the following (see Breese et al., 1998),

$$V_{a,j} = \bar{z}_a + \kappa \sum_{i=1}^n w(a,i)(z_{i,j} - \bar{z}_i), \quad (5.2)$$

where  $V_{a,j}$  is the predicted vote of consumer  $a$  to the product  $j$ , the subscript  $a$  represents the consumer of interest (i.e., active user) and the weight  $w(a,i)$  is some measure of similarity between the active user and consumer  $i$ . There are different specifications other than Eq. (5.2), and also different measures of similarity. One particular problem of these techniques is when there are few observations per customer. In those cases it is quite difficult to predict with sufficient accuracy. Breese et al. (1998) review several of these models and apply them to some data in order to compare their predicted accuracy.

### 5.1.3 Other models

Probabilistic models that predict whether a customer would use a particular product or service based on the ownership of other products or

services are one modeling stream that could be used (Kamakura et al., 1991). Kamakura et al. (2003) extend previous factor analysis procedures to predict consumption of new or current products by current customers who do not use them yet. They combine transaction data with survey data for a sample of the population to identify the best prospects for each cross-selling offer.

## **5.2 Predicting Individual Response to Add-on Offerings (Customer Selection)**

Another interesting area is the development of models to identify individuals who are more responsive to marketing contacts. One way is to model the CLV elasticity to marketing at the customer or segment level, and another way is to predict individual's response to cross-buying and up-selling initiatives. Although there are models that sample target customers from a prospect population for a direct marketing campaign, few models have explicitly considered CLV, cross- and up-selling. We emphasize two research approaches: market response models and data mining techniques.

### **5.2.1 Market response models**

Market response models to estimate the probability of an individual buying a cross- or up-selling offer or their expected CLV if contacted present a high potential. For example, using data from previous product offerings to individual customers together with their observed responses, demographic and behavioral information, a simple binary logit model could be estimated. Models could explain the individual's likelihood of buying a specific product offering as a function of the products that the customer has already purchased, the household size, age, etc.

An interesting model to select target customers from a population, and at the same time maximize profitability, has been developed by Bult and Wansbeek (1995). In their model it is possible to find the cut-off point from which mailings are sent to individual customers. The firm should mail as long as the marginal benefit is larger than the marginal cost of mailing. The CLV metric has been used in a model for customer

selection and resource allocation (Venkatesan and Kumar, 2004). These authors develop a stochastic model using panel data that predicts purchase frequency and contribution margin at the customer level. They illustrate the model using data from a computer hardware and software manufacturer in a B2B setting. When customers are selected for contacts using this model, the company generates 83% more long-run profitability than when using its previous allocation method. Zhang and Krishnamurthi (2004) develop a model to provide recommendations on when to promote, how much and to whom. They develop a joint purchase incidence – brand choice – purchase quantity model to derive the optimal price discount to each household on each shopping trip, and apply it to an online retailer. Given that customers present different variety seeking behavior and also present different responsiveness to price cuts, a recommendation system to implement customized promotions can significantly improve the ROI of price promotions. Li et al. (2005) develop a cross-selling model that can incorporate the sequential acquisition pattern of some products across different households, as it occurs, for example in financial services. Incorporating this natural sequential process can enhance the cross-selling opportunities of the firm.

One problem of selecting customers for a cross-selling offer is that of adverse selection. The most interesting customers may not respond and those who respond may even be turned down because of credit risk. Cao and Gruca (2005) develop a bi-variate probit model that simultaneously accounts for response and approval probability and show how using this model rather than a simple binary probit for response probability increases the profitability of the cross-selling campaign.

### **5.2.2 Data mining techniques**

Decision trees have been applied to cross-selling databases to classify individuals as either buyers of multiple products or as single buyers (i.e., only one product) (Giuffrida et al., 2000). These authors develop a new algorithm that applies a multivariate, instead of a univariate search, as most of the previous algorithms have used. They show how for complex databases such as those of a firm studying cross-selling opportunities

using many customers and information per customer, their algorithm can capture the relevant information to correctly classify individuals. In their application these authors get around 90% of customers correctly classified. Other algorithms, when applied to this database, either did not complete their task or did not find any useful information.

Alternative data mining technologies such as Artificial Neural Networks (see Section 4.3) could also be developed here. Finally, heuristic methods such as the RFM model (recency, frequency and monetary value) are applied by direct marketing firms to rank order consumers by their probability of response.

### **5.3 Estimating the Potential Value of a Customer**

Estimating the potential value of a customer is important because it gives the company an idea of how much more can they get from that customer (the difference between wallet and share of wallet) and it can also serve as a basis for customer segmentation. If the business the firm gets from a particular customer is close to its potential value, it will not be optimal to target that customer with new product offerings, but instead to incentivize retention. For example, Harrah's Entertainment, a casino company, has attributed its superior financial performance to its CRM initiatives whose objectives are competing for share of wallet, rather than rewarding loyalty.

Customer potential can be defined in two different ways. The first one is the total value of products or services currently bought by a customer and that could be offered to that customer from the firm. The second adds to the first definition the potential of new purchases not currently made from a customer (e.g., up-selling).

The problem of measuring the current volume of purchases from a customer is that the firm usually does not observe purchases with its competitors. Moreover, the purchase levels observed from a company may be poorly correlated with the purchase amounts made by its customer base outside the firm (Du et al., forthcoming). In other words, the company's best customers may be close to their potential and some mediocre customers may have substantial untapped potential. One way to overcome this data limitations is to survey a representative sample

of customers. Verhoef and Donkers (2001) follow this approach and use both a linear regression and a multivariate probit model to explain customer potential as a function of some demographic information. The probit model estimates the probabilities of each customer owning a specific product. In that case, the potential value of customer  $i$  would be calculated as,

$$PV_i = \sum_{j=1}^J \Pr(\text{customer } i \text{ owns product } j) \text{Profit}_j, \quad (5.3)$$

where  $\text{Profit}_j$  is the expected profit from product  $j$ . Using a linear regression model could be helpful to study which demographics influence potential value, but not the probability of purchasing each particular product. Hence, the multivariate probit can be used not only for estimating the potential value but also for cross-selling strategies.

Du et al. (forthcoming) apply a “list augmentation” approach to a bank database, in order to calculate customers’ potential values. This approach consists of four steps (see Kamakura and Wedel, 2003): (i) the firm surveys a random sample of customers and collects information needed to estimate its potential, (ii) the survey data are linked to the internal data, (iii) several predictive models are fitted on these sample data containing both survey and internal data, (iv) the best performing model is applied to the entire internal database in order to estimate survey results for that group. The authors conduct a series of targeting simulations and find substantial lifts in targeting efficiency when using estimates of share of wallet and total wallet.

An interesting model that has been proposed to estimate the up-selling potential of individual customers is the stochastic frontier model (Kim and Kim, 1999). These authors applied the model to an insurance firm, in which the policy premiums were explained as a function of customers’ demographics, marketing activity, competition, and macroeconomic environment. The model is specified as,

$$Y_i = \beta_0 + \beta_1 X_1 + \cdots + \beta_k X_{ki} + \varepsilon_i \quad \varepsilon_i = v_i - u_i. \quad (5.4)$$

As it is noted in Eq. (5.4), the error term is decomposed in two components. The first term is the standard error of the classical regression,

in which  $v_i : N(0, \sigma_v^2)$ , and it captures the statistical noise, measurement error, and random shocks outside the firm's control. The second term captures management inefficiency under the firms' control (i.e., lost up-selling potential). The authors assume that  $u_i : N(0, \sigma_u^2)$ ,  $u_i \geq 0$ , and both error terms are independent of each other. Hence, the premium paid by a customer  $i$  will always be equal or lower than her frontier (i.e.,  $\beta_0 + \beta_1 X_1 + \dots + \beta_k X_{ki} + v_i$ ) and the frontier can be defined as the maximum premium the customer can pay. The authors show how the company could have obtained a 25% increase in business with more than half of its customers.

Another interesting question is how to maximize the value of a customer through permission-based marketing. When firms have low cost vehicles to target their customers with offers aimed at increasing their CLV, there must be an optimal communication timing that maximizes their CLV (Drèze and Bonfrer, 2001). These authors argue that a firm sending too many emails to their customers risk annoying them and increasing the probability of churn. On the other hand, if a firm does not target its customers, it will miss selling opportunities. They develop an interesting model (see Eq. (2.7)) and estimate it using data from a firm in the entertainment industry. They find that the effect inter-communication timing has on the CLV of a customer is asymmetric, and managers are better-off by erring toward longer inter-communication times than toward shorter ones.

## 5.4 Antecedents of Add-on Selling

A relevant question for the firm is to understand why do customers buy more over time at an aggregate level. Satisfaction and payment equity (i.e., fairness of the exchange) have been hypothesized as possibly affecting future cross-buying (Verhoef et al., 2001). These authors estimate a multivariate probit model, in which the dependent variable is the number of policies bought from a customer at a certain point in time, minus the same figure one year earlier. They find that the difference in payment equity of the firm with its competitors is an antecedent of cross-buying. But they do not find support for the hypothesis that

satisfaction affects cross-buying. Additional research with data from more companies is necessary to test these hypotheses.

## **5.5 Future Directions**

### **5.5.1 The relationship between add-on selling and retention**

It has been argued that the larger the number of products or services a customer buys from a certain firm, the higher her probability of retention (Blattberg et al., 2001). This could occur when the customer has large switching costs, because she might, for example, find some synergies by buying several products from the same firm. For example, consumers might attach some value to paying their Internet, telephone and cable fees under the same bill. This relationship between add-on selling and future retention probability is important as it might account for a significant portion of the CE contribution of a customer. If this effect is large and is not accounted for, firms would sub spend in customer acquisition and add-on selling generation. To the best of our knowledge, there are no published models that account for this.

### **5.5.2 Models for add-on selling**

There are very few models of add-on selling that have been published in marketing. We believe this modeling area represents opportunities for future research. The development of collaborative filtering seems to have a huge potential for cross-selling applications. Similarly, response models can help in measuring those with higher probabilities of responding to a specific product offering. Also, models are needed to estimate the potential value of customers. In general, these models should support managers' actions by recommending one-to-one marketing decisions on product offerings, media used to deliver those offerings, and customer selection. Finally, identifying the antecedents of add-on selling (e.g., satisfaction and payment equity) can help firms increase their future sales by investing at an aggregate level in those factors that generate the largest future customer equity through add-on selling.

# 6

---

## Marketing Customization and Customer Equity Maximization

---

We have so far introduced models that can help managers maximize CE. We said that customers are heterogeneous and that firms should target them with the objective of maximizing their CE. Hence, firms could be able to discriminate among their customers and target them with different offers so that the long-run financial performance of the firm (to which CE was considered a good proxy) is maximized.

### 6.1 A Framework for Marketing Customization

Figure 6.1 presents a framework for marketing customization. First, a firm willing to implement a customization strategy will decide which actions to use (e.g., price, product, communication, service or a combination of them). Second, the targeting ability of the firm will be determined by the type of customer heterogeneity, how this heterogeneity is distributed across the customer pool, the quantity and quality of available data, and the model used for customizing the chosen marketing action. Third, the output of the model and the strategy used to customize will result in specific recommendations in terms of the type and the depth of the offer, the channel used to address the customer and

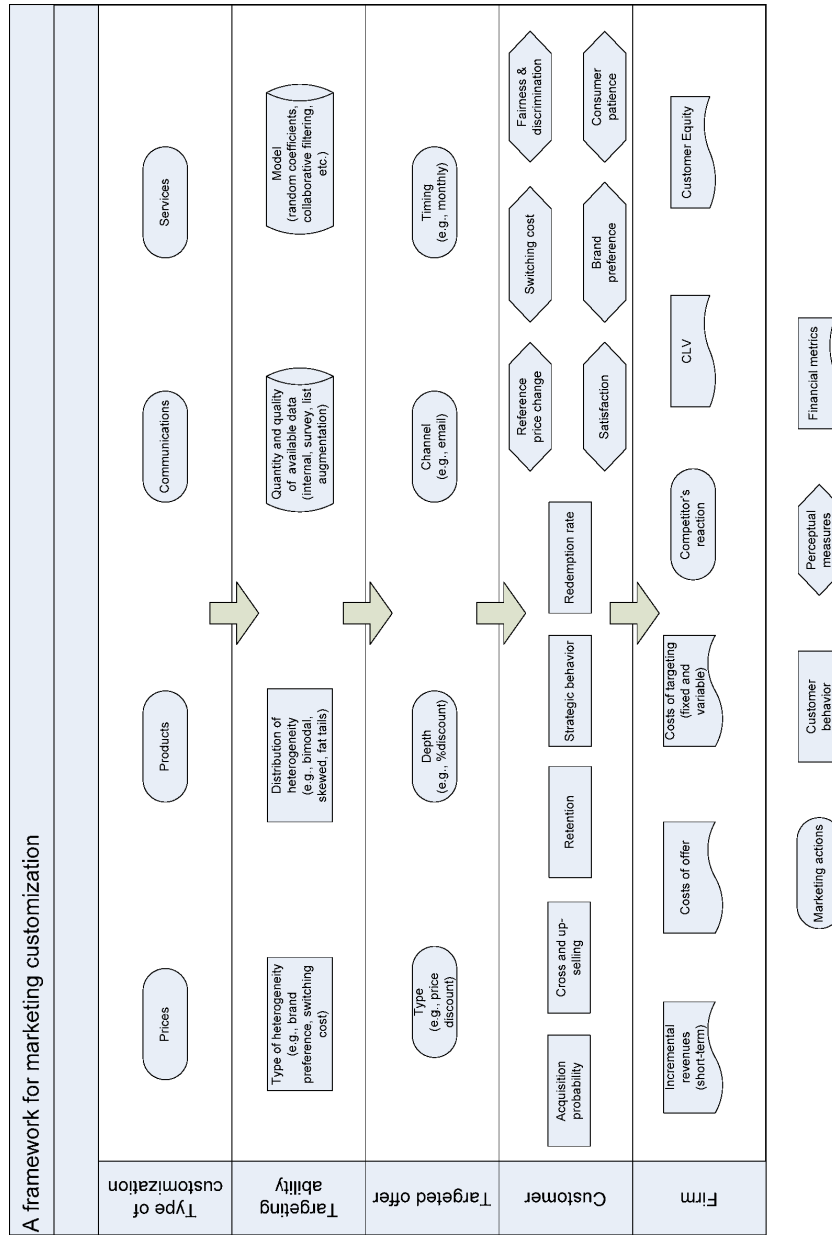


Fig. 6.1 A framework for marketing customization

the number of offers in a certain period. For example, a supermarket implementing a customized couponing strategy may test several offer types (e.g., a discount on a current or on a future purchase occasion), depth levels (a 10%, 15%, or 20% discount), delivery channels (coupons printed at the check-out counter or an SMS coupon) and timing choices (coupons printed on a monthly basis or per shopping trip). Fourth, customers will react to a specific offer or a combination of offers. This will result in behavioral changes (e.g., a higher retention rate) and attitudinal changes (e.g., higher satisfaction). Finally, the firm will obtain some additional revenues and costs associated with its customization campaign. Some of these costs will be variable (such as the cost of printing a coupon) and some will be fixed (e.g., the cost of a data mining department). In order to test the long-run consequences of a customization strategy, competition as well as future customer behavior (e.g., reference price changes) should be taken into account.

## 6.2 The Objective Function to Maximize

Measuring the ROI of a customization action is difficult. In practice, many firms compare the results to what they consider a baseline (e.g., sales of the same product the week before the action), and others like Harrah's Entertainment use experiments and always test against a control sample. This approach makes ROI calculation easy in principle. However, it could be quite challenging when implementing hundreds of offers at the same time. In these approaches the objective function to maximize is the short-term return that is truly incremental.

A short-term objective function could be appropriate to evaluate a marketing action that is tactical in scope (e.g., a specific supermarket coupon campaign). However, this could be quite inappropriate to evaluate an entire customization strategy. As shown in Figure 6.1, long-run changes in customers and competitors' behaviors could be important (e.g., an escalation of marketing costs in the form of more expensive loyalty programs). We argue that CE and CLV are two powerful metrics to evaluate the ROI of a customization strategy.

A major problem of using CLV as an objective function for customization decisions (e.g., printing a coupon only if it is predicted

CLV will increase thanks to this coupon) is that of forecasted error. As explained above, and as suggested by Malthouse and Blattberg (2005), forecasting CLV requires using historical data, so if customers are unpredictable we may end up using an objective function (CLV) that is good in theory but inaccurate in implementation.

If the decision at hand is at the firm level, CE is a useful metric to use as an objective function to maximize. As explained above, dynamic CE should be equal to long-run profitability, which is consistent with shareholder value. The advantage of using CE is that, due to its mathematical derivation, it is easier to link marketing customization actions to customer responses and firm value (e.g., one could model how acquisition spending impacts customer acquisition rate and, as a consequence, long-run profitability).

In this section, we focus only on the strategic implications of customized marketing actions. We discuss three fundamental questions about CE management:

- Should a firm discriminate loyals with less attractive offers (e.g., higher prices) than those for switchers?
- Is targeted pricing beneficial to the firm?
- Does a marketing customization strategy where the objective function to maximize is CE or CLV bring higher profits in the long-run than a myopic (i.e., short-term) strategy?

This section is organized as follows. First, we review the literature on targeted pricing. Here, we will discuss the first two questions raised above. Even though some authors have developed models that deal with other types of customization, such as that of communications (e.g., Ansari and Mela, 2003), we will focus our attention on targeted pricing given that this research area is more developed. Second, we discuss the optimality of using dynamic CE (i.e., long-run profitability) as the objective function to maximize. Finally, we propose future research directions.

### **6.3 Targeted Pricing**

Consumers are heterogeneous in their willingness to pay for the same product or service. In general, the reservation price of existing

customers is higher than that of new customers, because existing customers tend to exhibit higher switching costs and also higher brand preference for that product. When firms offer a uniform price to all customers, they leave some surplus to consumers with higher willingness to pay, while forsaking sales to customers with lower willingness to pay. In an ideal world, in which firms were able to perfectly identify the reservation price of each customer in the market and offer individual prices to each, they would price at the reservation price and capture the entire consumer's surplus. Previous literature has identified at least four types of price discrimination: first, second and third-degree (Pigou, 1920, Tirole, 1988) and behavior-based (Shaffer and Zhang, 1995, Villas-Boas, 1999, Fudenberg and Tirole, 2000). Price discrimination can take many forms (e.g., coupons, bonuses, targeted price reductions, and any other type of discounts), and can be based on different levels of information (e.g. last purchase, history of purchases, price elasticity, volume, current versus competitor's customer).

### 6.3.1 Who should be discriminated?

The question of whether managers should penalize loyals or reward them (with respect to switchers) is still open to debate. In most industries firms discriminate against loyals offering coupons or even temporary price reductions, because consumers with low levels of loyalty are more likely to buy at a discount price. Most of the analytical literature on price discrimination has found that it is optimal to penalize loyals with higher prices than switchers (Narasimhan, 1988, Villas-Boas, 1999, Fudenberg and Tirole, 2000). With the exception of Shaffer and Zhang (2000) who find that when demand is asymmetric, charging a lower price to existing customers might be optimal. In an interesting study, Anderson and Simester (2004) using data from a mail order company find that price promotions to existing customers increase short-term demand but have a negative impact in the long-run. However, the same price promotion to prospects increases both short- and long-run demand. However, the literature on relationship-marketing has questioned this practice arguing the higher profitability of loyal customers. Even though the literature admits that in general long-life customers

tend to pay higher prices than new customers, it is also suggesting that when customers are evaluated with respect to their lifetime value they are worth much more than when looking only at the short-term, and therefore the common practice of rewarding switchers is more than questionable: losing a loyal customer due to overcharging is extremely costly for the firm.

The emphasis on price discrimination strategies that attract switchers with lower prices has at least four negative effects: (i) attracting customers with low prices lowers the quality of the customer base, reducing the firm's ability to deliver value to its customers, with the consequence of making more difficult the retention of the most interesting customers (Reichheld, 1996); (ii) frequent price cuts educate customers to switch and increase their price sensitivity (Papatla and Krishnamurthi, 1996, Mela et al., 1997, Jedidi et al., 1999); (iii) discriminatory prices make loyals unhappy when they observe that they are penalized for being loyals (Reichheld, 1996); and (iv) consumers can even become strategic by forgoing a purchase today to get a lower price tomorrow (Villas-Boas, 1999, Chen and Zhang, 2001), and also stockpile product that is bought only when there is a promotion (e.g., Gupta, 1988, Bell et al., 1999).

In summary, discriminatory prices to loyals seem to be supported by the majority of the analytical models, yet common wisdom and some empirical studies have warned against the pitfalls of such strategy.

### **6.3.2 Is targeted pricing beneficial to the firm?**

In theory, targeted pricing helps firms extract consumers' surplus. Nevertheless, in the presence of competitors also offering targeted prices, this question does not have a trivial answer. For example, it has been argued that coupon targeting gives rise to a prisoners' dilemma in which both companies are worse-off than if they were to set the same price for all customers (Shaffer and Zhang, 1995, Chen et al., 2001). This happens when targetability is perfect or sufficiently high. Other studies have also found that firms are worse-off by following targeted pricing (e.g., Chen, 1997, Villas-Boas, 1999, Fudenberg and Tirole, 2000).

On the contrary, competitive targeted prices have sometimes been shown to be beneficial for the firms. For example, in industries in which individual's information is poor, improvements in targetability could lead to win-win situations in which firms are better-off (Chen et al., 2001). Even when average prices are reduced as a consequence of targeted prices, a company with more loyals than its competitor (markets are asymmetric) could benefit from targeted pricing (Shaffer and Zhang, 2002). This will occur when the market share effect of targeted pricing offsets the reduction in prices. One could also argue that when only few companies are able to offer targeted prices, they could have a competition advantage over other firms in the market.

### 6.3.3 Elements of targeted pricing models

We will refer here to analytical models of price discrimination and the different assumptions that have been or could be made in these models. Changing some of these assumptions in future modeling efforts could help in a better understanding of the previously raised questions.

(a) *Consumer's heterogeneity.* Inherent in all models of price discrimination is the specification of consumer's heterogeneity. Previous models have captured this heterogeneity through brand preference á la Hotelling (1929), switching costs (e.g., Shaffer and Zhang, 2000), or both (e.g., Klemperer, 1987a, 1987b). When there are more than two periods in the model, switching costs increase competition in the first periods because firms recognize that by acquiring customers now they may retain them more easily in the future (Klemperer, 1987a, 1987b). On the contrary, when consumers' heterogeneity is reflected in their brand preference, a firm will be less incentivized to compete fiercely for a customer who prefers another brand, because that customer will be very vulnerable in the future. This occurs because a switching cost benefits the firm to which the customer is currently buying, while brand preference does benefit a specific brand.

(b) *Addressability.* Most of the previous research has assumed that all customers are addressable. Nevertheless, in some circumstances, this assumption might not hold. Many firms can address its own existing customers but cannot address all of their prospects. That scenario

would mean that the firm is in the position to send individual offers to its own customers but not to its prospects. Privacy laws in many countries require firms to have opt-out lists and some laws even require the firm to have previous authorization from a customer before targeting her with an offer. Finally, in some markets or countries there might not exist accurate prospect lists, so firms have to rely on mass marketing communications to acquire new customers.

(c) *Targetability*. It is defined as the ability to predict preferences and purchase behaviors of individual customers for the purpose of customizing its price or product offer (Chen and Zhang, 2001). When targetability is high, firms can engage in individual pricing (Shaffer and Zhang, 2002). Other models have assumed the firm only offers two prices: one to its existing customers and another to new customers (or switchers) (Narasimhan, 1988, Shaffer and Zhang, 1995, 2000, Chen et al., 2001). It could also be possible for firms to have different levels of targetability based on whether the customer is currently buying from the firm or not or even based on her previous purchase history (Shaffer and Zhang, 1995, Villas-Boas, 1999).

(d) *Cost of targeting*. Most models have not included a cost of targeting but in some industries this could be an important factor to consider. Shaffer and Zhang (1995) introduce a model of coupon targeting in which targeting customers with coupons has a cost. They show how when the cost of targeting is low, firms should engage in more defensive targeting. They later develop another model (Shaffer and Zhang, 2002) in which firms compete for a population of switchers and it is shown how there are some targeting regions for which one firm is better-off not offering individual promotions to the individuals in that region. This occurs when the probability of acquiring a customer in a region is sufficiently low, and the expected profit of targeted pricing to that region is not high enough, given the targeting cost.

(e) *Number of periods*. Previous models have either one period (e.g., Narasimhan, 1988, Shaffer and Zhang, 1995, Chen et al., 2001), two periods (e.g., Klemperer, 1987a, 1987b, McGahan and Ghemawat, 1994, Fudenberg and Tirole, 2000, or infinite periods with overlapping

generations of customers (Villas-Boas, 1999). Models that capture the CLV or the CE concept should at least include two periods.

(f) *The profitability of long-life customers.* An interesting concept that firms should consider when setting targeted prices is the expected CE contribution of a customer or, at least, her expected CLV. We explained earlier in this paper the hypothesized higher profitability of long-life customers. This could imply that firms should take into account these effects and avoid setting prices while maximizing only the short-term.

(g) *Individual's preference or switching cost changes over time.* It would be interesting to study what would be the effect on the competition among two rival firms when consumers are expected to change their brand preference or switching costs over time. For example, a new customer that enters the market might be expected to reduce its switching cost as it acquires experience in the product category. Alternatively, switching costs or brand preference could depend on previous switching behavior. That is, customers who keep switching because of price promotions would be expected to exhibit lower switching costs over time, while customers who keep repeating would be expected to reinforce their brand preference or increase their switching cost. To the best of our knowledge no model has been developed that captures this behavior.

(h) *Strategic customers.* Another interesting element of some models is allowing customers to behave strategically. Customers may behave strategically when they learn over time that their current decisions will affect the prices they are offered in the future. Thus they could forgo a purchase to secure a lower price in the future (Villas-Boas, 1999, 2004, Chen and Zhang, 2001). While Villas-Boas (2004) shows that strategic customers can even make monopolists worse-off when they implement targeted pricing, Chen and Zhang (2001) show that the negative effects of customer recognition may not offset the benefits of targeted pricing. It could also be possible that loyal customers are less strategic than switchers because switchers have generally more price information, while loyals tend to be less sensitive to prices. We believe customers will be more likely to behave strategically when at least one of the following conditions exists: (i) the price difference for similar

products is large (e.g., air fares); (ii) transparency in prices is high (e.g., Internet); (iii) the importance of the product in the consumer's basket is high (e.g., automobiles).

(i) *Asymmetry in brand preference, or switching costs.* Although most models assume that both firms have the same proportion of customers of each type (e.g., same distribution of switching costs), this could not hold in many situations. Shaffer and Zhang (2000) allow for asymmetry in brand preferences (i.e., ex-ante switching costs) and show that when this occurs it might be profitable for a company to charge a lower price to its existing customers.

In summary, analytical models of targeted pricing can have different elements that can help to shed some light on the important question of whether to discriminate loyals against switchers or not. The decision of which elements to include in the model will depend on the particular industry and the way consumers behave. We think future models that look at this question should definitely include at least two periods and should take into account the profit expansion effects of repeating customers.

#### **6.4 On the Optimality of Customer Equity Maximization**

The CE literature takes for granted the assumption that a strategy aimed at maximizing CE is optimal, and definitely superior to a myopic (short-term) strategy. This is definitely the case when the firm allocates its resources more efficiently, and everything else remains unchanged. But, in a competitive scenario, when all firms set their marketing spending to maximize CE, this could give rise to a prisoner's dilemma, in which all firms are worse-off. Villanueva et al. (2006a) study this research question and show that firms might be better-off by being myopic, when they target customers depending on their previous purchase history (which carries information on their switching costs). This happens because taking into account the CLV when setting targeted prices could make firms compete more fiercely for retention and acquisition.

In many markets it is not rare to see managers following myopic strategies. Short-term metrics (e.g., market-share) are usually reliable and have strong credibility at all levels of management (Keil et al.,

2001), while the CE concept is more difficult to explain and even harder to measure with some accuracy. Inaccurate decision support models of CE might lead managers to deviate substantially from optimality, while short-term metrics such as market-share might be sufficiently correlated with long-run performance. In general, maximizing the short-term has been criticized in marketing (e.g., Dekimpe and Hanssens, 1995, 1999, Keil et al., 2001). We think that three important questions are whether CE models are good enough to guide resource allocation, whether managers understand them, and whether they will have credibility at other levels of management or not.

We can conclude that it is not that clear whether a generalized shift by all players in an industry from a myopic to a CE objective will benefit them or not. The answer to the previous question will depend on the following factors: (i) how much faster than its competitors a firm moves from a myopic to a CE maximization strategy; (ii) market shares of the different players; and (iii) the nature and quantity of the profit expansion effects. In summary, much needs to be done in developing analytical models that scrutinize the optimality of a CE maximization strategy.

## **6.5 Future Directions**

Both analytical and empirical models can be used to increase our understanding of the previously discussed questions. We identify the following key directions for further research.

### **6.5.1 Analytical models of CLV maximization**

Analytical models of targeted offers (e.g., targeted pricing and price discrimination) have not explicitly considered the strategic consequences on maximizing CLV. Dynamic models of targeted pricing have maximized the total profit for the firm while solving for equilibrium prices. Much research needs to be done, however, changing the assumptions made in the main characteristics of these models (as described in Section 6.3.3) to reflect more realistic market conditions, and to study how each of these characteristics affects targeted pricing in a CLV framework. Another interesting area for future work is in developing

models in which firms learn about consumers' preferences or switching costs over time and can use that information to price discriminate even further.

### **6.5.2 The optimality of CE maximization**

As we have pointed out above, it is unclear whether firms competing for the same pool of customers will be better-off by maximizing CE. If the answer is no, then we could conclude that myopic managers might not be so irrational after all. Some market characteristics, such as the ones described in Section 6.3.3 could also have an effect on the optimality of CE maximization. We might also expect advantages for the first firm to maximize CE, or for the larger firm. It could also be interesting to study how should managers be compensated to focus more on the long-run when CE Maximization is superior.

# 7

---

## Conclusions

---

We have covered the main objectives of this paper. First, we have reviewed current models to compute CE. We presented the most important elements that CE models need in order to be complete on important issues. For the first time in the CE literature, we have provided an extensive review and a typology of current modeling efforts. After reviewing current models of CE it is clear that there are still huge opportunities for model development in this area. More specifically, we believe that competitive behavior and endogenous parameters are two key aspects that few models have incorporated so far. Stochastic models of CE present also opportunities, as they may capture latent behavior difficult to be modeled using earlier deterministic models.

Second, we have provided key strategic questions on the drivers of CE: acquisition, retention, and add-on selling, and reviewed current modeling efforts in those questions. Some models seem to be well suited for direct marketers, or companies with extensive customer databases, while others apply also to traditional mass marketers with less knowledge about individual customers. We emphasized the importance of normative models of customer acquisition, retention, and

add-on selling, which could guide resource allocation with the objective of CE maximization.

Finally, we have reviewed current literature dealing with two questions at a more strategic level: (i) Is CE maximization an optimal strategy, under a competitive scenario?; and (ii) Should firms discriminate loyals with higher prices than those for switchers? We suggested that both empirical and analytical modeling can shed some light on the answers to these critical questions.

## References

---

- Alba, J., J. Lynch, B. Weitz, C. Janiszewski, and others (1997), 'Interactive home shopping: Consumer, retailer, and manufacturer incentives to participate in electronic marketplaces'. *Journal of Marketing* **61**(3), 38–53.
- Anderson, E. and D. Simester (2004), 'Long-run effects of promotion depth on new versus established customers: Three field studies'. *Marketing Science* **23**(1), 4–20.
- Anderson, E. W., C. Fornell, and D. R. Lehmann (1994), 'Customer satisfaction, market share, and profitability: Findings from Sweden'. *Journal of Marketing* **58**, 53–66.
- Anderson, E. W., C. Fornell, and S. K. Mazvacheryl (2004), 'Customer satisfaction and shareholder value'. *Journal of Marketing* **68**(4), 172–185.
- Anderson, E. W. and V. Mittal (2000), 'Strengthening the satisfaction-profit chain'. *Journal of Service Research* **3**(2), 107–120.
- Ansari, A. and C. F. Mela (2003), 'E-customization'. *Journal of Marketing Research* **40**(2), 131–145.
- Bakos, J. Y. (1991), 'A strategic analysis of electronic marketplaces'. *Mis Quarterly* **15**(3), 295–310.

- Bakos, J. Y. (1997), 'Reducing buyer search costs: Implications for electronic marketplaces'. *Management Science* **43**(12), 1676–1692.
- Bass, F. M. (1969), 'A new product growth for model consumer durables'. *Management Science* **15**(5), 215–227.
- Bell, D. R., J. Chiang, and V. Padmanabhan (1999), 'The decomposition of promotional response: An empirical generalization'. *Marketing Science* **18**(4), 504–526.
- Ben-Akiva, M. and S. R. Lerman (1985), 'Discrete Choice Models'. MIT Press.
- Berger, P. D. and N. N. Bechwati (2001), 'The allocation of promotion budget to maximize customer equity'. *Omega* **29**(1), 49–62.
- Berger, P. D. and N. I. Nasr (1998), 'Customer lifetime value: Marketing models and applications'. *Journal of Interactive Marketing* **12**(1), 17–30.
- Berry, M. J. A. and G. Linoff (1997), *Data Mining Techniques for Marketing, Sales and Customer Support*. New York: John Wiley & Sons.
- Bitran, G. and S. Mondschein (1996), 'Mailing decisions in the catalog sales industry'. *Management Science* **42**(9), 1364–1381.
- Blattberg, R. C. and J. Deighton (1996), 'Manage marketing by the customer equity test'. *Harvard Business Review* **74**(4), 136–144.
- Blattberg, R. C., G. Getz, and J. S. Thomas (2001), *Customer Equity: Building and Managing Relationships as Valued Assets*. Boston, Massachusetts: Harvard Business School Press.
- Bolton, R. N. (1998), 'A dynamic model of the duration of the customer's relationship with a continuous service provider: The role of satisfaction'. *Marketing Science* **17**(1), 45–65.
- Bolton, R. N. and K. N. Lemon (1999), 'A dynamic model of customers' usage of services: Usage as an antecedent and consequence of satisfaction'. *Journal of Marketing Research* **36**(2), 171–186.
- Breese, J. S., D. Heckerman, and C. Kadie (1998), 'Empirical analysis of predictive algorithms for collaborative filtering'. *Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence*.
- Brynjolfsson, E. and M. D. Smith (2000), 'Frictionless commerce? A comparison of internet and conventional retailers'. *Management Science* **46**(4), 563–585.

- Bult, J. R. and T. Wansbeek (1995), ‘Optimal selection for direct mail’. *Marketing Science* **14**(4), 378–394.
- Cao, Y. and T. S. Gruca (2005), ‘Reducing adverse selection through customer relationship management’. *Journal of Marketing* **69**(4), 219–229.
- Chakravarti, D., A. Mitchell, and R. Staelin (1981), ‘Judgment-based marketing decision models: Problems and possible solutions/commentary on “Judgment-Based Marketing Decision Models”’. *Journal of Marketing* **45**(4), 13–40.
- Chen, Y. (1997), ‘Paying customers to switch’. *Journal of Economics and Management Strategy* **6**, 877–897.
- Chen, Y., C. Narasimhan, and Z. J. Zhang (2001), ‘Individual marketing with imperfect targetability’. *Marketing Science* **20**(1), 23–41.
- Chen, Y. and Z. J. Zhang (2001), ‘Price-for-information effect and benefit of behavior-based targeted pricing’. *Working Paper*.
- Cox, D. R. (1975), ‘Partial likelihood’. *Biometrika* **62**, 269–276.
- Dekimpe, M. G. and D. M. Hanssens (1995), ‘The persistence of marketing effects on sales’. *Marketing Science* **14**(1), 1–21.
- Dekimpe, M. G. and D. M. Hanssens (1999), ‘Sustained spending and persistent response: A new look at long-term marketing profitability’. *Jmr, Journal of Marketing Research* **36**(4), 397–412.
- Dowling, G. R. and M. Uncles (1997), ‘Do customer loyalty programs really work?’. *Sloan Management Review* **38**(4), 71–82.
- Drèze, X. and A. Bonfrer (2001), ‘To pester or leave alone: Lifetime value maximization through optimal communication timing’. Working Paper. Wharton-SMU Research Center: .
- Drèze, X. and A. Bonfrer (2003), ‘A renewable-resource approach to database valuation’. Working Paper. Wharton-SMU Research Center.
- Du, R. Y., W. A. Kamakura, and C. F. Mela (forthcoming), ‘Size and share of customer wallet’. *Journal of Marketing*.
- Dwyer, F. R. (1997), ‘Customer lifetime valuation to support marketing decision making’. *Journal of Direct Marketing* **11**(4), 6–13.
- Enders, W. (1994), *Applied Econometric Times Series*. New York: John Wiley & Sons.

- Fader, P. S., B. G. S. Hardie, and K. L. Lee (2005a), “Counting your customers” the easy way: An alternative to the pareto/NBD model’. *Marketing Science* **24**(2), 275–284.
- Fader, P. S., B. G. S. Hardie, and K. L. Lee (2005b), ‘RFM and CLV: Using iso-value curves for customer base analysis’. *Journal of Marketing Research* **XLII**, 415–430.
- Fernández, P. (2001), ‘Valuation of brands and intellectual capital’. Working Paper.
- Forrester (2001). *Effective Email Marketing*.
- Fournier, S., S. Dobscha, and D. G. Mick (1998), ‘Preventing the premature death of relationship marketing’. *Harvard Business Review* **76**(1), 42–44.
- Fudenberg, D. and J. Tirole (2000), ‘Customer poaching and brand switching’. *RAND Journal of Economics* **31**(4), 634–657.
- Garbarino, E. and M. S. Johnson (1999), ‘The different roles of satisfaction, trust, and commitment in customer relationships’. *Journal of Marketing* **63**(2), 70–87.
- Gensch, D. (1973), *Advertising Planning Mathematical Models in Advertising Media Planning*. New York: Elsevier Scientific Publishing Company.
- Giuffrida, G., W. Chu, and D. M. Hanssens (2000), ‘Mining classification rules from datasets with large number of many-valued attributes’. *Lecture Notes in Computer Science 1777*. Springer-Verlag, pp. 335–349.
- Greene, W. H. (2000), *Econometric Analysis*. Prentice Hall.
- Greyser, S. A. and H. P. Root (1999), ‘Improving advertising budgeting’. *MSI Working Paper Report No. 00-126*.
- Gupta, S. (1988), ‘Impact of sales promotions on when, what, and how much to buy’. *Journal of Marketing Research* **25**(4), 342–355.
- Gupta, S. and D. R. Lehman (2003), ‘Customers as assets’. *Journal of Interactive Marketing* **17**(1), 9–24.
- Gupta, S., D. R. Lehman, and J. A. Stuart (2002), ‘Valuing customers’. *Journal of Marketing Research* **XLI**, 7–18.
- Gupta, S. and V. Zeithaml (forthcoming), ‘Customer metrics and their impact on financial performance’. *Marketing Science*.

- Gustafsson, A., M. D. Johnson, and I. Roos (2005), 'The effects of customer satisfaction, relationship commitment dimensions, and triggers on customer retention'. *Journal of Marketing* **69**, 210–218.
- Hansotia, B. J. and P. Wang (1997), 'Analytical challenges in customer acquisition'. *Journal of Direct Marketing* **11**(2), 7–19.
- Helsen, K. and D. C. Schmittlein (1993), 'Analyzing duration times in marketing: Evidence for the effectiveness of hazard rate models'. *Marketing Science* **11**(4), 395–413.
- Hogan, J. E., D. R. Lehmann, M. Merino, R. K. Srivasta, J. S. Thomas, and P. C. Verhoef (2002a), 'Linking customer assets to financial performance'. *Journal of Service Research* **5**(1), 36–38.
- Hogan, J. E., K. N. Lemon, and B. Libai (2002b), 'The true value of a lost customer'. *MSI Working Paper Report No. 02-108*.
- Hogan, J. E., K. N. Lemon, and R. R. Rust (2002c), 'Customer equity management: Charting new directions for the future of marketing'. *Journal of Service Research* **5**(1), 4–12.
- Hotelling, H. (1929), 'Stability in competition'. *The Economic Journal* **39**(153), 41–45.
- Jackson, B. B. (1985), *Winning and Keeping Industrial Customers the Dynamics of Customer Relationships*. Lexington, Mass: Lexington Books.
- Jackson, D. (1989a), 'Determining a customer's lifetime value'. *Direct Marketing* **51**(11), 60–62, 123.
- Jackson, D. (1989b), 'Determining a customer's lifetime value (Part 2)'. *Direct Marketing* **52**(1), 24–32.
- Jackson, D. (1989c), 'Insurance marketing: Determining a customer's lifetime value (Part 3)'. *Direct Marketing* **52**(4), 28, 30.
- Jain, D. and S. Singh (2002), 'Customer lifetime value research in marketing: A review and future directions'. *Journal of Interactive Marketing* **16**(2), 34–46.
- Jedidi, K., C. F. Mela, and S. Gupta (1999), 'Managing advertising and promotion for long-run profitability'. *Marketing Science* **18**(1), 1–22.
- Johnson, E. J., W. W. Moe, P. S. Fader, S. Bellman, and G. L. Lohse (2004), 'On the depth and dynamics of online search behavior'. *Management Science* **50**(3), 299–308.

- Jones, T. O. and W. E. Jr. Sasser (1995), 'Why satisfied customers defect'. *Harvard Business Review* **73**(6), 88–99.
- Kamakura, W., S. Ramaswami, and R. Srivasta (1991), 'Applying latent trait analysis in the evaluation of prospects for cross-selling of financial services'. *International Journal of Research in Marketing* **8**, 329–349.
- Kamakura, W. A. and M. Wedel (2003), 'List augmentation with model based multiple imputation: A case study using a mixed-income factor model'. *Statistica Neerlandica* **57**(1), 46–57.
- Kamakura, W. A., M. Wedel, F. d. Rosa, and J. A. Mazzon (2003), 'Cross-selling through database marketing: A mixed data factor analyzer for data augmentation and prediction'. *International Journal of Research in Marketing* **20**, 45–65.
- Keane, T. J. and P. Wang (1995), 'Applications for the lifetime value model in modern newspaper publishing'. *Journal of Direct Marketing* **9**(2), 59–66.
- Keil, S. K., D. Reibstein, and D. R. Wittink (2001), 'The impact of business objectives and the time horizon of performance evaluation on pricing behavior'. *International Journal of Research in Marketing* (18), 67–81.
- Kim, B.-D. and S.-O. Kim (1999), 'Measuring upselling potential of life insurance customers: Application of a stochastic frontier model'. *Journal of Interactive Marketing* **13**(4), 2–9.
- Klemperer, P. (1987a), 'The competitiveness of markets with switching costs'. *Rand Journal of Economics* **18**(1), 138–150.
- Klemperer, P. (1987b), 'Markets with consumer switching costs'. *Quarterly Journal of Economics* **102**(2), 375–394.
- Knott, A., A. Hayes, and S. Neslin (2002), 'Next-product-to-buy models for cross-selling applications'. *Journal of Interactive Marketing* **16**(3), 59–75.
- Kumar, V., N. K. Leman, and A. Parasuraman (2006), 'Managing customers for value: An overview and research agenda'. *Journal of Service Research* **9**(2), 87–94.
- Labe R. P. Jr. (1994), 'Database marketing increases prospecting effectiveness at merrill lynch'. *Interfaces* **24**(5), 1–12.

- Leeflang, P. S. H., D. R. Wittink, M. Wedel, and P. A. Naert (2000), *Building Models for Marketing Decisions*. Boston/Dordrecht/London: Kluwer Academic Publishers.
- Lemmens, A. and C. Croux (2006), 'Bagging and boosting classification trees to predict churn'. *Journal of Marketing Research* **43**(2), 276–286.
- Lemon, K. N., T. B. White, and R. S. Winer (2002), 'Dynamic customer relationship management: Incorporating future considerations into the service retention decision'. *Journal of Marketing* **66**, 1–14.
- Leone, R. P., V. R. Rao, K. L. Keller, A. M. Luo, L. McAlister, and R. Srivasta (2006), 'Linking brand equity to customer equity'. *Journal of Service Research* **9**(2), 125–138.
- Lewis, M. (2004), 'The influence of loyalty programs and short-term promotions on customer retention'. *Journal of Marketing Research* **XLI**, 281–292.
- Lewis, M. (2005a), 'A dynamic programming approach to customer relationship pricing'. *Management Science* **51**(6), 986–994.
- Lewis, M. (2005b), 'Incorporating strategic consumer behavior into customer valuation'. *Journal of Marketing* **69**, 230–238.
- Lewis, M. (2006), 'Customer acquisition promotions and customer asset value'. *Journal of Marketing Research* **XLIII**(2), 195–203.
- Li, S., B. Sun, and R. T. Wilcox (2005), 'Cross-selling sequentially ordered products: An application to consumer banking services'. *Journal of Marketing Research* **XLII**, 233–239.
- Libai, B., D. Narayandas, and C. Humby (2002), 'Toward an individual customer profitability model: A segment-based approach'. *Journal of Service Research* **5**(1), 69–76.
- Little, J. D. C. (1970), 'Models and managers: The concept of a decision calculus'. *Management Science* **16**(8), B466–B485.
- Little, J. D. C. (1975), 'BRANDAID: A marketing-mix model'. *Operations Research* **23**(4), 628–673.
- Little, J. D. C. and L. M. Lodish (1969), 'A media planning calculus'. *Operations Research* **17**(1), 1–35.
- Little, J. D. C. and L. M. Lodish (1981), 'Judgment-based marketing decision models: Problems and possible solutions/commentary

- on “Judgment-Based Marketing Decision Models”. *Journal of Marketing* **45**(4), 13–40.
- Lodish, L. M. (1971), ‘CALLPLAN: An interactive salesman’s call planning system’. *Management Science* **18**(4), 25–40.
- Malthouse, E. C. and R. C. Blattberg (2005), ‘Can we predict customer lifetime value?’. *Journal of Interactive Marketing* **19**(1), 2–16.
- Manski, C. and S. Lerman (1977), ‘The estimation of choice probabilities from choice-based samples’. *Econometrica* **45**, 1977–1988.
- Mantrala, M. K., P. Sinha, and A. A. Zoltners (1992), ‘Impact of resource allocation rules on marketing investment-level decisions and profitability’. *Journal of Marketing Research* **29**(2), 162–176.
- McGahan, A. M. and P. Ghemawat (1994), ‘Competition to retain customers’. *Marketing Science* **13**(2), 165–176.
- Mela, C. F., S. Gupta, and D. R. Lehmann (1997), ‘The long-term impact of promotion and advertising on consumer brand choice’. *Journal of Marketing Research* **34**(2), 248–261.
- Mittal, V. and W. A. Kamakura (2001), ‘Satisfaction, repurchase intent, and repurchase behavior: Investigating the moderating effect of customer characteristics’. *Journal of Marketing Research* **37**, 131–142.
- Montgomery, D. B., A. J. Silk, and C. J. Zaragoza (1971), ‘A multiple-product sales force allocation model’. *Management Science* **18**(4), 3–24.
- Mulhern, F. J. (1999), ‘Customer profitability analysis: Measurement, concentration, and research directions’. *Journal of Interactive Marketing* **13**(1), 25–40.
- Narasimhan, C. (1988), ‘Competitive promotional strategies’. *Journal of Business* **61**(4), 427–449.
- Neslin, S., S. Gupta, W. Kamakura, J. Lu, and C. Mason (2006), ‘Defection detection: Measuring and understanding the predictive accuracy of customer churn models’. *Journal of Marketing Research* **46**(2), 204–211.
- Nijs, V. R., M. G. Dekimpe, J.-B. E. M. Steenkamp, and D. H. Hanssens (2001), ‘The category-demand effects of price promotions’. *Marketing Science* **20**(1), 1–22.
- O’Brien, L., , and C. Jones (1995), ‘Do rewards really create loyalty?’. *Harvard Business Review* **73**(3), 75–82.

- Papatla, P. and L. Krishnamurthi (1996), 'Measuring the dynamic effects of promotions on brand choice'. *Journal of Marketing Research* **33**(1), 20–35.
- Peterson, L. A., R. C. Blattberg, and P. Wang (1997), 'Database marketing – Past, present, and future'. *Journal of Direct Marketing* **11**(4), 109–125.
- Pfeifer, P. E. and R. L. Carraway (2000), 'Modeling customer relationships as markov chains'. *Journal of Interactive Marketing* **14**(2), 43–55.
- Pigou, A. C. (1920), *The Economics of Welfare*. London: Macmillan and co.
- Reichheld, F. F. (1993), 'Loyalty-based management'. *Harvard Business Review* **71**(2), 64–73.
- Reichheld, F. F. (1996), 'Learning from customer defections'. *Harvard Business Review* **74**(2), 56–61.
- Reichheld, F. F. and W. E. Sasser (1990), 'Zero defections: Quality comes to services'. *Harvard Business Review* **68**(5), 105–111.
- Reichheld, F. F. and T. Teal (1996), *The Loyalty Effect the Hidden Force Behind Growth, Profits, and Lasting Value*. Boston, Mass: Harvard Business School Press.
- Reinartz, W. and V. Kumar (2002), 'The mismanagement of customer loyalty'. *Harvard Business Review* **80**(7), 86–94.
- Reinartz, W. and V. Kumar (2003), 'The impact of customer relationship characteristics on profitable lifetime duration'. *Journal of Marketing* **67**(1), 77–99.
- Reinartz, W., J. S. Thomas, and V. Kumar (2005), 'Balancing acquisition and retention resources to maximize customer profitability'. *Journal of Marketing* **69**, 63–79.
- Reinartz, W. J. and V. Kumar (2000), 'On the profitability of long-life customers in a noncontractual setting: An empirical investigation and implications for marketing'. *Journal of Marketing* **64**(4), 17–35.
- Resnick, P. and H. R. Varian (1997), 'Recommender systems'. *Association for Computing Machinery. Communications of the ACM* **40**(3), 56–58.
- Rigby, D. K., F. F. Reichheld, and P. Schefter (2002), 'Avoid the four perils of CRM'. *Harvard Business Review*, February.

- Roberts, M. L. and P. D. Berger (1989), 'Direct Marketing Management'. Prentice Hall Collage Div.
- Rogers, E. M. (1983), *Diffusion of Innovations*. New York: Free Press.
- Rust, R. T. (1986), *Advertising Media Models*. Lexington Books.
- Rust, R. T., K. N. Lemon, and V. A. Zeithaml (2004), 'Return on marketing: Using customer equity to focus marketing strategy'. *Journal of Marketing* **68**, 109–127.
- Rust, R. T. and P. C. Verhoef (2005), 'Optimizing the marketing interventions mix in intermediate-term CRM'. *Marketing Science* **24**(3), 477–489.
- Rust, R. T., V. A. Zeithaml, and K. N. Lemon (2000), *Driving Customer Equity How Customer Lifetime Value is Reshaping Corporate Strategy*. New York: Free Press.
- Ryals, L. (2005), 'Making customer relationship management work: The measurement and profitable management of customer relationships'. *Journal of Marketing* **69**, 252–261.
- Schmittlein, D. C., D. G. Morrison, and R. Colombo (1987), 'Counting your customers: Who are they and what will they do next?'. *Management Science* **33**(1), 1–24.
- Schmittlein, D. C. and R. A. Peterson (1994), 'Customer base analysis: An industrial purchase process application'. *Marketing Science* **13**(1), 41–67.
- Shaffer, G. and Z. J. Zhang (1995), 'Competitive coupon targeting'. *Marketing Science* **14**(4), 395–416.
- Shaffer, G. and Z. J. Zhang (2000), 'Pay to switch or pay to stay: Preference-based price discrimination in markets with switching costs'. *Journal of Economics & Management Strategy* **9**(3), 397–424.
- Shaffer, G. and Z. J. Zhang (2002), 'Competitive one-to-one promotions'. *Management Science* **48**(9), 1143–1160.
- Silk, A. J., L. R. Klein, and E. R. Berndt (1997), 'Intermedia substitutability in the U.S. national advertising market'. *Working Paper*.
- Simester, D. I., P. Sun, and J. N. Tsitsiklis (2006), 'Dynamic catalog mailing policies'. *Management Science* **52**(5), 683–696.
- Sissors, J. Z. and E. R. Petray (1976), *Advertising Media Planning*. Chicago, IL: Crain Books.

- Stauss, B. and C. Friege (1999), 'Regaining service customers'. *Journal of Service Research* **1**(4), 347–361.
- Thomas, J., R. Blattberg, and E. Fox (2004), 'Recapturing lost customers'. *Journal of Marketing Research* **38**(2), 31–45.
- Thomas, J. S. (2001), 'A methodology for linking customer acquisition to customer retention'. *Journal of Marketing Research* **38**(2), 262–268.
- Tirole, J. (1988), *The Theory of Industrial Organization*. Cambridge, Mass: MIT Press.
- Venkatesan, R. and V. Kumar (2004), 'A customer lifetime value framework for customer selection and resource allocation strategy'. *Journal of Marketing* **68**, 106–125.
- Verhoef, P. C. and B. Donkers (2001), 'Predicting customer potential value an application in the insurance industry'. *Decision Support Systems* **32**, 189–199.
- Verhoef, P. C., P. H. Franses, and J. C. Hoekstra (2001), 'The impact of satisfaction and payment equity on cross-buying: A dynamic model for a multi-service provider'. *Journal of Retailing* **77**, 359–378.
- Villanueva, J., P. Bhardwaj, Y. Chen, and S. Balasubramanian (2006a), 'Managing customer relationships: Should managers focus on the long term?'. *Working Paper*. University of California, Los Angeles.
- Villanueva, J., S. Yoo, and D. M. Hanssens (2006b), 'The impact of marketing-induced versus word-of-mouth customer acquisition on customer equity'. *Working Paper*.
- Villas-Boas, J. M. (1999), 'Dynamic competition with customer recognition'. *RAND Journal of Economics* **30**(4), 604–631.
- Villas-Boas, J. M. (2004), 'Price cycles in markets with customer recognition'. *RAND Journal of Economics* **35**(3), 486–501.
- Wiesel, T. and B. Skiera (2004), 'Enterprise valuation by using customer lifetime values'. *Working Paper*.
- Yoo, S. and D. M. Hanssens (2005), 'Modeling the sales and customer equity effects of the marketing mix'. *Working Paper*.
- Zauberman, G. (2003), 'The intertemporal dynamics of consumer lock-in'. *Journal of Consumer Research* **30**(3), 405–419.
- Zhang, J. and L. Krishnamurthi (2004), 'Customizing promotions in online stores'. *Marketing Science* **23**(4), 561–578.