

---

## Modelling customer lifetime value in contractual settings

---

Christoph Heitz\*, Marcel Dettling and  
Andreas Ruckstuhl

Institute of Data Analysis and Process Design,  
Zurich University of Applied Sciences,  
Rosenstrasse 3, CH-8401 Winterthur, Switzerland

E-mail: [heit@zhaw.ch](mailto:heit@zhaw.ch)

E-mail: [dtli@zhaw.ch](mailto:dtli@zhaw.ch)

E-mail: [rkst@zhaw.ch](mailto:rkst@zhaw.ch)

\*Corresponding author

**Abstract:** Service provision is often governed by a contract (e.g., newspaper subscriptions, phone contracts, and credit agreements). Typically, such a contract includes rules that influence the dynamics of the customer in the marketplace. Typical examples are minimum contract durations, or fixed time instants for contract termination. The goal of these rules is to increase the future total profit gained from the customer, which is usually denoted with the term customer lifetime value (CLV).

We analyse the problem of calculating the CLV under general contract structures. We show that classical Markov models for describing the customer dynamics are not appropriate and may lead to huge errors in the CLV. We propose a semi-Markov formulation which leads to substantially better results. We apply the framework to data of newspaper subscription.

**Keywords:** marketing; customer lifetime value; CLV; Markov chain model; Semi-Markov model; direct marketing; mathematical modelling; contractual settings; modelling; customer dynamics; newspaper subscription.

**Reference** to this paper should be made as follows: Heitz, C., Dettling, M. and Ruckstuhl, A. (2011) 'Modelling customer lifetime value in contractual settings', *Int. J. Services Technology and Management*, Vol. 16, No. 2, pp.172–190.

**Biographical notes:** Christoph Heitz received his Diploma in Physics in 1992 and his PhD in Theoretical Physics in 1996 from the University of Freiburg i.Br., Germany. After an employment in two companies, he joined Zurich University of Applied Sciences (ZHAW) in 2000 as a Professor for Operations Management. He is a founding member and current President of the Swiss Institute of Service Science, an interdisciplinary research centre for service. His research interests include service science, in particular from a marketing perspective, operations management, and stochastic modelling and simulation.

Marcel Dettling graduated with an MSc in Mathematics from ETH Zurich in 2000. Subsequently, he received his PhD in Applied Statistics at the Seminar for Statistics, also at ETH Zurich. He then worked until 2006 as a Postdoc/Research Associate at John Hopkins Medical School at Baltimore, USA. In 2006, he joined the Institute for Data Analysis and Process Design at ZHAW where he works as a Senior Researcher in Data Analysis, mainly focusing on customer intelligence and business analytics, as well as health and

transportation data. He also works as a Lecturer in Statistical Data Analysis and Time Series Analysis at both ZHAW and ETH Zurich. His research interests mainly focus on development and application of methods for the analysis and prediction of complex and high-dimensional, and the corresponding software implementations.

Andreas Ruckstuhl is a graduate in Mathematics of the Swiss Federal Institute of Technology (ETH) Zurich where he also wrote a PhD thesis in Applied Statistics. From 1996 to 1999, he was a Lecturer at the Department of Statistics and Econometrics and a Research Fellow at the Centre for Mathematics and its Applications, simultaneously, both of the Australian National University. In 1999, he joined the Zurich University of Applied Sciences (ZHAW) in Winterthur. Since 2001, he is a Professor of Statistical Data Analysis and currently the Associate Head of the Institute of Data Analysis and Process Design (IDP). His main research fields are in statistical data analysis and in the development of robust inference methods. Applications concern problems from environmental sciences, business analytics as well as from the hedge fund industry.

This paper is a revised and expanded version of a paper entitled 'Customers as an intrinsic part of the service system: customer dynamics in service contracts, customer lifetime value, and marketing' presented at European Workshop on Service Systems (EWOSS), Eindhoven, The Netherlands, 9 November 2009.

---

## 1 Introduction

The concept of customer lifetime value (CLV) has attracted increasing attention in the last decades, especially in the field of marketing. Valuating individual customers with respect to their future expected profit is a key input for many marketing decisions and has become central to relationship management (Berger and Nasr, 1998; Malthouse and Blattberg, 2004). CLV is usually defined as the sum of the future revenues of a given customer, minus the costs of attracting, selling and servicing this customer (Berger and Nasr, 1998; Jain and Singh, 2002).

The sum of the CLVs of all customers is often called the customer equity, and is widely considered as a proxy of a firm's value (Gupta et al., 2006; Berger et al., 2006). This is particularly true in service contexts, where the customer equity is often much larger than the other assets of the company.

Since the CLV is defined by the future profits gained from a customer, the CLV depends on the future behaviour of the customer which is not known in advance. The better this behaviour is described, the more the calculated CLV value is a reasonable metrics for the value of a customer. Considerable effort has been undertaken to develop good models for calculation of the CLV [see e.g., Gupta et al. (2006) for a review].

Often, the relationship between firm and client is governed by contracts. Examples are telecom, service subscriptions such as TV, newspapers, or video rentals, but also service contracts for auto repair, credit contracts, and many others. Viewed from a business perspective, such contracts aim at influencing the future behaviour of the customers. For example, a customer signing a minimum duration clause in a contract is effectively bound to the firm for a certain amount of time, thus suppressing contract

termination and churn. Another example is a yearly renewal policy: by restricting the possibility of cancelling to a short period of a year, the churn probability is reduced, and an increase of the contract duration can be expected.

Thus it can be expected that the contract design is a relevant component of the CLV of a signing customer. This is particularly relevant for dynamic markets with low switching costs, for example communication providers (telecom, and internet), contract-based services such as video rentals or the like.

The focus of this paper is to clarify the relationship between the CLV of an individual client and the contract design with respect to contractual elements that influence the future dynamics of the customer. We show that classical CLV models do not take these dynamics-changing elements of contracts into account. In particular, the popular Markov chain models (MCM) are not able to cope with such contract structures.

We derive a new modelling framework based on Semi-Markov models (SMM) for the calculation of CLV which is able to describe contract rules influencing the customer dynamics for many typical applications. We furthermore show that the usage of classical MCMs may lead to significant errors in the calculated CLV values and thus may lead to wrong marketing decisions.

Estimating CLVs of individual customers or customer segments of a company, and usage of this information for controlling the marketing activity of a firm, may lead to a significant competitive advantage, in particular in competitive markets. On the other hand, exploiting existing and available customer data by using a dynamical customer model as described in this paper, and using this as a basis for predicting an organisation's future profit, is a highly sophisticated and knowledge-intensive task. This task is usually performed by analytics specialists, either as an internal service of a business analytics department for the firm's decision makers, or as a B2B service, offered by a specialised marketing analytics company.

Thus, the described formal approach is, by itself, a good example for a knowledge-intensive service with a high value-adding potential for every firm whose profits depend significantly on dynamical customer relationships.

## 2 Literature review

The notion of CLV of customers has become a popular research topic during the last decade (see Reinartz and Kumar, 2003; Rust et al., 2000). It has been shown to be an important and useful concept in marketing, especially in micro marketing and direct marketing. The CLV is usually defined as the estimated present value of an individual customer's future cash flows, including future revenues as well as future costs (Jain and Singh, 2002). Historically, CLV has been used primarily in the realm of direct marketing due to the availability of individual data on the customers. In the advent of internet-based communication, however, many more opportunities have appeared. This has significantly renewed the interest in CLV among marketing researchers (Jain and Singh, 2002).

Conceptually, the notion of CLV is defined on an individual customer level. However, integrating all customers of a firm, a natural extension is to define the customer equity as the sum of the future profits of all customers (Blattberg and Deighton, 1996; Rust et al., 2004; Gupta et al., 2005; Villanueva and Hanssens, 2007; Berger et al., 2006). Customer equity is usually considered as a good proxy for the value of a firm (Gupta et al., 2005).

The concepts of CLV and CE has been used for solving relevant management questions on, e.g., allocation of marketing spending, understanding the connection between market spending and financial performance, and optimising customer relationship management (Rust et al., 2000; Gupta and Lehmann, 2003; Rust et al., 2004; Kumar et al., 2006; Villanueva and Hanssens, 2007). Especially in relationship marketing, CLV has become central (Berger and Nasr, 1998; Malthouse and Blattberg, 2004).

A crucial point when using any CLV approach is the method how to determine the CLV of individual customers, based on the available data. The first CLV models have been introduced by Dwyer (1989) in the context of direct marketing. An overview of CLV models is given in Berger and Nasr (1998) and Jain and Singh (2002). In the original paper of Dwyer (1989), he proposed that customers can be broadly divided into two groups: lost-for-good customers, and always-a-share customers. Lost-for-good customers are assumed to be lost forever once they leave the company. Such a case can be expected in situations with high switching costs. In the always-a-share model, customers may switch frequently between different providers.

Classical retention models usually assume lost-for-good situations. In these models, a retention rate (or, equivalently, a churn rate) is estimated for each customer. Based on this rate, the residual lifetime can be calculated. Coupled with financial assumptions this may be used to calculate the CLV.

It has been argued (Rust et al., 2004) that the lost-for-good assumption systematically understates CLV because a lost customer may return, and thus an always-a-share assumption is more realistic. This leads to the so-called migration models (Berger and Nasr, 1998; Dwyer, 1989) where a customer is assumed to be in one of several states, and his/her dynamics is governed by transition probabilities between these states.

The lost-for-good model can be interpreted as a two-state migration model with the states 'alive' and 'dead', and the additional assumption that there is no migration from 'dead' to 'alive' (see Pfeifer and Carraway, 2000). Thus, the always-a-share model is more general and includes the lost-for-good case as a special case.

An excellent overview on CLV models in a wider context of customer-based analysis is given by Fader and Hardie (2009). The authors stress that fact that for predicting the future behaviour of customers, probabilistic models have to be used due to incomplete knowledge.

The modelling of future customer behaviour seems to be governed by two streams of research. The first stream focuses on the lost-for-good model. All classical retention models belong to this category. In this context, duration dependent churn or retention effects have been studied, and models have been developed for coping with these effects (see e.g., Reichheld and Teal, 2001). Fader and Hardie (2009) pointed out that duration-dependent effects may be rather strong and should not be neglected. As an example, they cite Fielding (2005) who reports that "renewal rates at regional magazines vary; generally 30% of subscribers renew at the end of their original subscription, but that figure jumps to 50% for the second-time renewals and all the way to 75% for long-time readers". However, in the literature, modelling of duration dependent transition rates seems to be restricted to this lost-for-good case.

The second stream focuses on the always-a-share idea where the dynamics of a customer manifests itself as migrations between different states. Pfeifer and Carraway (2000) and Rust et al. (2000) propose the general framework of Markov chains for

modelling these dynamics. The states of a customer may be defined differently. For example, Pfeifer and Carraway (2000) define the state of a customer as his/her recency in one application. In another application, the states are defined as ‘prospect customer’, ‘customer’, and ‘former customer’. Rust et al. (2000) attach the states to specific brands – the fact that a customer is in state  $i$  at period  $t$  means that he is buying brand  $i$  at this period.

To our knowledge, all migration models cited in literature are based on a Markov assumption. This has the advantage of leading to a simple model where the dynamics of a customer are fully described by a single switching matrix. As a consequence, however, the switching probability may not depend on the duration of the sojourn time in the state. While this may be a good approximation in some settings, this may not be the case for many others as empirical evidence from the first stream of research suggests. In particular, it is not a realistic assumption in contractual settings where the contract usually has a direct impact on the user behaviour.

In our contribution, we combine the two research streams and develop a model which is:

- a able to cope with duration-dependent transition probabilities
- b can be applied to the more general always-a-share model, in particular can be used for an arbitrary number of states.

We provide a formal definition and a calculus to determine the CLV from the probabilistic model. The new model is based on a Semi-Markov approach which is a generalisation of the MCM, and thus contains the MCM as a special case. This allows a direct comparison of the results of the MCM and the new model.

Our model is based on a customer individual level, and the model parameters are estimated for each customer or each customer segment separately. This is a similar approach as in Rust et al. (2004). It allows using the model for direct marketing purposes.

### 3 Modelling CLV with sojourn time dependent customer dynamics

#### 3.1 *Introductory example*

The concept of CLV is based on a calculation of future profits of a given customer where profit is defined as revenue minus costs. Within this paper, we use the following expression for the CLV of a given customer  $k$ :

$$R_k = \sum_{t=0}^{\infty} c_k(t) \alpha^t \quad (1)$$

where  $R_k$  is the total future profit,  $c_k(t)$  is the net cash flow in time step  $t$  (revenue minus costs in period  $t$ ), and  $\alpha$  is a discount factor with  $\alpha < 1$ . Time is given in discrete steps where a time step may denote a day, a month, or any reasonable time unit. We use an infinite time horizon, but our approach can be easily adapted to a finite time horizon.

Since the future cash flows  $c_k(t)$  are subject to uncertainty, the future profit is a random variable with both an expectation value as well as variability. Since we focus on

B2C applications, in particular direct marketing purposes with many customers, we ignore the variability and define the CLV as the expectation value of  $R_k$ :

$$CLV_k = E(R_k) = E\left(\sum_{t=0}^{\infty} c_k(t) \alpha^t\right) \quad (2)$$

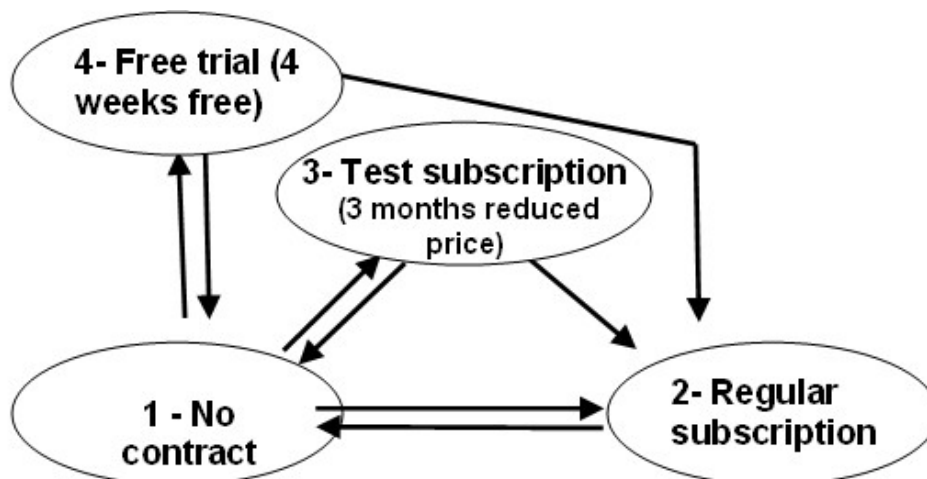
We assume that the customer has a choice of different contracts  $i$  ( $i = 1, \dots, N$ ) and may change his contract over time. At each time, only one contract is possible. For example, different subscription options for a telecom provider may be available to the customer. The change of contract obviously has an impact on the profit. For the sake of simplicity, we assume that the profit is a constant profit rate (profit per time period) which only depends on the contract type but not on the specific behaviour of the customer. This is the case for subscription contracts with a flat rate, e.g., many telecom and mobile contracts. Other popular examples are many video-on-demand subscription, or newspaper and journal subscriptions. It should be noted, however, that our approach can be extended to customer-specific revenue patterns as well as variable revenue under each contract  $i$ .

The customer dynamics itself is a stochastic process with a discrete state space, each state  $i$  being defined by the corresponding contract.

As an example, we consider the case of a newspaper subscription. State 1 corresponds to the situation where a customer has no contract at all. In this case, the profit per time period is  $c_1 = 0$ . State 2 is a non-limited subscription contract with a weekly profit of  $c_2 > 0$ .

Let us assume that the provider offers a test subscription for three months at a reduced price. This is a third state  $i = 3$  with corresponding revenue  $c_3 < c_2$ . In addition, we assume that the provider offers a free four-week test subscription by direct phone marketing. This corresponds to a state  $i = 4$  with  $c_4 < 0$ , because there are no revenues but still delivery and production costs arise.

**Figure 1** State diagram of contract states for newspaper subscription



In Figure 1, the state space diagram is shown. The arrows show the possible transitions. The dynamics of a customer is influenced by different factors. Some transitions are related to autonomous behaviour of customers, e.g., all transitions to a lower contract. For example, the transition  $2 \rightarrow 1$  means a termination of the regular subscription by the customer. Depending on the contract, this might happen at any point of time (sometimes after an initial non-cancellation period), or on a yearly basis, or at any other pattern.

In contrast to the autonomous transitions, the transition  $1 \rightarrow 2$  is triggered by a marketing activity. The transitions  $3 \rightarrow 2$  and  $4 \rightarrow 2$  are mutations of the temporal contract to a regular subscription – it is exactly these transitions that the provider wants to generate by offering a test subscription or the free trial.

Calculating the CLV of a customer in such a framework requires a probabilistic dynamical model for the customer's dynamics within the state space, and a combination of this model with the profit structure of the different states.

### 3.2 Customer dynamics

In this section, the general model for the description of the customer dynamics is described. We will show that a MCM is not able to cope with a sojourn time dependent dynamics, and we propose a new modelling approach based on a SMM.

At every time instant  $t$  ( $t = 1, 2, 3, \dots$ ), a customer is in a state  $X(t)$ :

$$X(t) = i, \text{ with } i = 1, 2, \dots, N$$

where  $N$  marks the number of states. The initial condition is denoted by  $X(0)$ . The temporal development of  $X(t)$  for  $t > 0$  is a stochastic process whose characteristic has to be described by a suitable model.

The simplest model is a MCM which assumes that a customer in state  $i$  has a time-independent probability  $p_{ij}$  of changing the state to  $j$  in the next time step (see also Pfeifer and Carraway, 2000). With this model, the expectation value of equation (2) can be calculated analytically. A central property of the Markov model is that the process is assumed to be 'memoryless', i.e., the dynamics in a given state only depends on the actual state, but neither of what happened before nor how long the customer is already in this state. An immediate consequence is that sojourn times are always random and geometrically distributed (Ross, 1995).

As discussed above, this central assumption is not realistic in many cases, and sojourn time dependent effects should be taken into account. This is possible with the so-called Semi-Markov model (SMM, see Ross, 1995). A SMM is defined by the following property: The dynamics may depend on the sojourn time in the present state, but not on what happened *before* the last state change. Thus, a part of the memoryless property is preserved, which makes the model analytically tractable and motivates the name Semi-Markov. Hence, SMM is a natural extension of the MCM when considering sojourn time dependent dynamics.

In SMM, the model consists of two parts: the first part models the sojourn time characteristics, i.e., *when* the current state is left. The second part models *how* the state change happens, i.e., to which other state the process jumps once it leaves the current state. This is worked out in the following.

Assume that the process is in state  $i$  with a given sojourn time  $T$ . There are two formal model elements that describe the dynamics in the next time step:

- a The hazard function  $h_i(T)$  specifies if the customer leaves the current state in the next time step. More specifically,  $h_i(T)$  is the probability of leaving the current state  $i$ , given the sojourn time  $T$ . Note that because we are in a discrete time setting with  $T = 0, 1, 2, \dots$ ,  $h_i(T)$  is a discrete function in contrast to the usual definition of a hazard rate which is used for continuous time models. Each state may have a specific hazard function  $h_i(T)$ .
- b The  $N \times N$  transition matrix  $P(T)$  describes to which state the transition occurs if the current state is left at sojourn time  $T$ . More specifically, the matrix element  $p_{ij}(T)$  is the probability that the process jumps into state  $j$ , if it leaves state  $i$  at time  $t$ . Since  $P(T)$  may depend on the sojourn time  $T$ , this is a function rather than a constant matrix.

Using these modelling elements, any Semi-Markov process can be parameterised.

Note that MCM are special cases of SMM, characterised by the fact that the hazard function as well as the transition matrix are constant, i.e., do not depend on the sojourn time  $T$ . The usual basic building block of an MCM, the probability of jumping from state  $i$  to state  $j$  in the next time step, is given by  $h_i$  and  $p_{ij}$  which again are sojourn time independent.

### 3.3 Examples of hazard functions

In Figure 2, some prototypical hazard functions for contractual settings are shown. Figure 2(a) is the reference case for a Markovian dynamics: the hazard function is constant independent of the sojourn time  $T$ . Figure 2(b) shows the hazard function for a case of a contract with a minimum contract duration, leading to a zero hazard function during this period, where the termination rate is assumed to be constant after this initial period.

There might be customers who decide to cancel within the minimum duration period, and cancel at the first possible time instant. This would lead to a cancellation peak directly at the end of the minimum duration period, and would result to a hazard function as shown in Figure 2(c). Figure 2(d) shows a hazard function which is typical for a contract with yearly renewal policy: cancellations can only be made at a multiple of one year.

While the deviation from Markovian behaviour in Figure 2(b) to Figure 2(d) is strongly determined by contract restrictions, Figure 2(e) shows an example where the shape of the hazard function is determined not by a contract rule but solely by the loyalty effect in the autonomous behaviour of the customers: long-term customers usually have a lower churn rate than new customers, and this is reflected in a decreasing hazard rate (Reichheld and Teal, 2001).

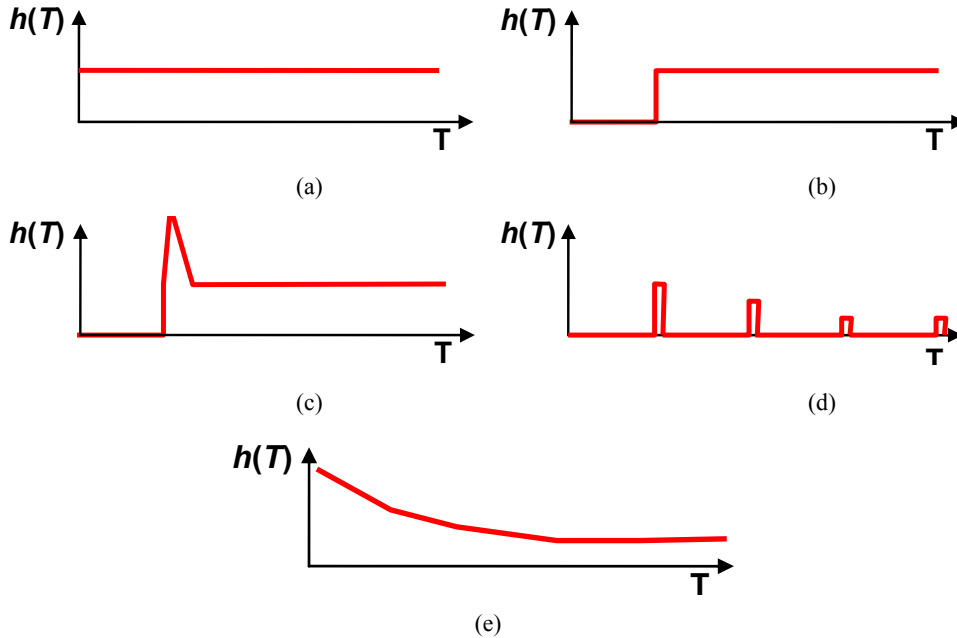
From the examples of Figure 2, we can conclude that in many common situations we would expect a strong deviation from Markovian dynamics. In particular, many typical contract restrictions have a strong influence on the hazard function.

Note that the hazard function is always a superposition of both contract rules and autonomous customer behaviour, and the interaction of both these elements may be complex. This is the case, for example, with the cancellation peak in Figure 2(c).

Empirically, the hazard function can be estimated by tracking customers over time. This estimated hazard function thus reflects both dynamical elements.



**Figure 2** Different hazard functions  $h(T)$  for contractual settings, (a) no contract restrictions, Markovian dynamics (b) minimum contract duration (c) minimum contract duration with cancellation peak (d) periodic withdrawal dates (e) no contract restrictions, long-time customers are more loyal (see online version for colours)



Note: In (a), the reference case of a Markovian dynamics is shown.

### 3.4 Calculation of CLV with SMM

In this section, the calculation of the CLV is derived under a Semi-Markov dynamics.

We assume the following:

- A customer is defined by an individual set of hazard functions  $h_i(T)$ ,  $i = 1, \dots, N$ , and a transition matrix function  $P(T)$ . His state at  $t = 0$  is  $X(0) = k$ , with a current sojourn time  $T_0$ .
- The profit per period for this customer in state  $i$  is given by  $c_i$  ( $i = 1, \dots, N$ ). The vector of profits is denoted by  $\vec{c}$ .
- Future profits are discounted with a discount factor  $\alpha$ .

The algorithm has two steps: In a first step, a vector  $\vec{J} = (J_1, J_2, \dots, J_N)$  is calculated which corresponds to the CLV-values of the customer for all possible initial states  $i$ , under the additional assumption that the customer has just moved into the state  $i$  (i.e.,  $T_0 = 0$ ). In a second step, the CLV is calculated for the current initial state  $k$  and the current  $T_0$ . For details on the calculation we refer to Heitz et al. (2010).

It can be shown that

$$J_i = \frac{c_i}{1-\alpha} \cdot (1-x_i) + \sum_{j \neq i} y_{ij} \cdot J_j, \quad i=1, \dots, N \quad (3)$$

with

$$x_i = \sum_{T=1}^{\infty} f_i(T) \alpha^T \quad (4)$$

$$y_{ij} = \sum_{T=1}^{\infty} p_{ij}(T) f_i(T) \alpha^T$$

The function  $f_i(T)$  is the probability distribution of the sojourn time which can be calculated from the hazard function by

$$f_i(1) = h_i(0)$$

$$f_i(2) = (1-h_i(0)) \cdot h_i(1)$$

$$f_i(3) = (1-h_i(0)) \cdot (1-h_i(1)) \cdot h_i(2)$$

.....

Equation (3) is a system of coupled linear equations that can be solved easily leading to the values  $J_i$ ,  $i = 1, \dots, N$ .

From this result, the CLV of a customer in state  $k$  with sojourn time  $T_0$  is given by

$$CLV_k(T_0) = \frac{c_k}{1-\alpha} \cdot (1-\tilde{x}_k) + \sum_{j \neq k} \tilde{y}_{kj} \cdot J_j \quad (5)$$

with

$$\tilde{x}_i = \sum_{T=1}^{\infty} \tilde{f}_i(T) \alpha^T \quad (6)$$

$$\tilde{y}_{ij} = \sum_{T=1}^{\infty} p_{ij}(T) \tilde{f}_i(T) \alpha^T$$

and the modified probability distribution  $\tilde{f}_i(T)$

$$\tilde{f}_i(T) = \frac{1}{\sum_{T=1}^{\infty} f_i(T+T_0)} f_i(T+T_0), \quad T > 0 \quad (7)$$

So, as an important result, we can state that the CLV of a given customer can be calculated analytically according to equation (5), given the model parameters  $h_i(T)$ ,  $i = 1, \dots, N$ , and the transition matrix function  $P(T)$ . As can be seen by comparing equation (5) with equation (3), the quantities  $J_i$  correspond to the CLV of state  $i$  with  $T_0 = 0$ .

### 3.5 Comparison of SMM and Markov model

In the last section, we have shown how the CLV can be calculated using an SMM. However, using SMM leads to both more efforts in estimating the model parameters from data, and more complicated formulae for the CLV calculation, compared to the classical case of the MCM. In this section, we analyse the improvement of the CLV calculation when using the SMM approach instead of a simpler MCM approach.

For doing this, we start from the assumption that the process is a Semi-Markov process with given model parameters. The true CLV would therefore be given by equation (5). What would be the estimated CLV if we applied a MCM instead as an approximation?

In order to analyse this questions, we assume that the MCM is modelled by both hazard function and transition matrix as with the SMM, but require that both elements are constant rather than dependent on  $T$ . We denote them by the symbols  $\hat{h}_i$  and  $\hat{P}_i$ . What are the best values for  $\hat{h}_i$  and  $\hat{P}_i$  if we try to mimic the observed dynamical behaviour of a Semi-Markov process approximately? For simplicity, we restrict ourselves to the case of a SMM with a constant transition matrix  $P$ .

An intuitive approach is to require that the approximate MCM yields

- a the correct expected total sojourn time for a single state sojourn. It can be shown that this means that the  $x_i$  in equation (4) have the same numerical values as for the original SMM:  $\hat{x}_i = x_i$ , for all  $i$ .
- b the correct proportion of state changes from each state  $i$  to the possible next states  $j \neq i$ . This means that the corresponding transition matrix of the approximate MCM must be identical to the original one:  $\hat{P} = P$ . (Note that we have assumed that  $P$  is constant).

It is easy to see that the quantities  $J_i$  of equation (3), calculated via the approximate MCM, remain unchanged. This means, that the CLV of  $T_0 = 0$  is calculated correctly even when using the simplified MCM approach!

However, this is no more true for  $T_0 > 0$ . Since the Markov model is memoryless by definition, the MCM-CLV is the same for all possible  $T_0$ , where in reality the CLV depends on  $T_0$  which is reflected in equation (5) due to the T-dependence of  $\tilde{x}_k$  and  $\tilde{y}_{kj}$ .

Hence, we can conclude: for a customer in state  $i$  with sojourn time  $T_0$ , the approximate MCM yields a CLV which is identical to SMM CLV, evaluated at  $T_0 = 0$ .

### 3.6 Example

As an example for comparing the difference between Markov modelling and Semi-Markov modelling, we consider the hypothetical case of a contract with minimum contract duration of one year. We assume that a specific customer has a probability of 50% of terminating the contract after the first year. If he does not cancel the contract at this time, he will stay in the contract with a constant hazard function and an average sojourn time of five additional years, thus having a total contract duration of six years. The resulting hazard function is similar to the one displayed in Figure 2(c). We assume furthermore that the customer never comes back after termination of the contract, and we

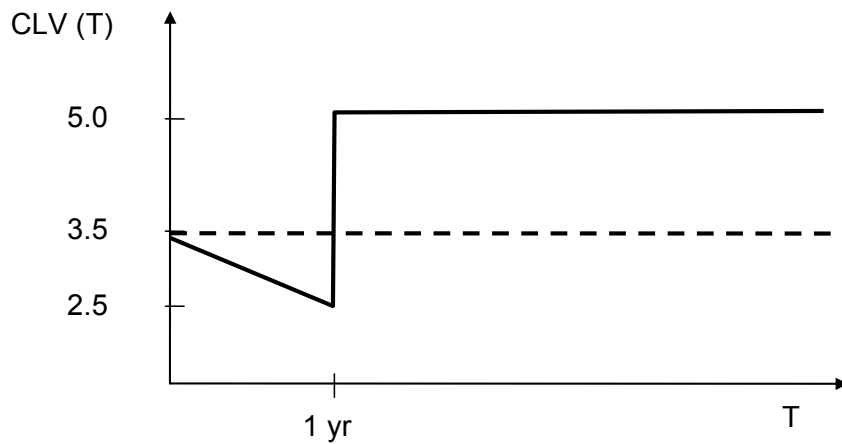
disregard discounting, setting  $\alpha = 1$ . The revenue is one per year during the contract period, and 0 afterwards.

The expected lifetime for a customer who just has subscribed to the contract ( $T_0 = 0$ ) is 3.5 years. That means that  $CLV(T_0 = 0) = 3.5$ . Corresponding to the discussion above, the customer would be assigned a CLV of 3.5 for all  $T_0$ , if a MCM were used.

However, in reality the CLV changes substantially during time. Shortly before the end of the first year, the CLV may be 0 (if the customer is going to terminate at  $T = 1$  yr), or 5 (if the customer stays within the contract). Thus, the expected CLV for  $T_0 = 1$  yr is 2.5.

Directly after  $T = 1$  yr, the average remaining contract time is five years, if the customer is still in the contract. That means that  $CLV(T_0 > 1 \text{ yr}) = 5$ . This value will not change anymore since the hazard function is assumed to be constant for  $T > 1$  yr. In Figure 3, the CLV is displayed both calculated with the best MCM (i.e., the Markov model with the correct *average* contract duration) and the SMM. It can be seen that the CLV for both models are identical for  $T_0 = 0$ . However, during the first year, the MCM overestimates the CLV up to 40% (3.5 instead of 2.5 at  $T = 1$  yr). After the minimum duration period, the Markov model underestimates the CLV by 30% (3.5 instead of 5).

**Figure 3** Calculated CLV as a function of sojourn time  $T$ , calculated with Markov model (dashed line) and SMM (solid line), respectively



Note: A hazard function with cancellation peak after minimum contract duration according to Figure 2(c) has been assumed.

This simple example shows that the estimated CLV may differ substantially from the correct value if assuming a Markovian dynamics and using a Markov model. The difference may easily be in the range of more than 30% – in fact it is easy to construct much more dramatic examples with still realistic assumptions.

This demonstrates that care has to be taken to model the dynamics correctly. In many cases, and in particular in contractual settings, the dynamics are far from being Markovian. In such cases, CLV estimations based on a Markov model can be expected to be substantially wrong.

Furthermore, when using a Markov model for estimating CLV as a basis for making business decisions in a non-Markovian dynamical environment, there is a high risk that the resulting decisions are wrong. For example, selecting customers for a marketing

campaign based on their CLV may result in completely wrong selections and dramatically deteriorate the efficiency of the campaign.

### 3.7 *Conclusions*

In this section, we have derived a new model for calculating the CLV of a customer in contractual settings based on a Semi-Markov approach. In this model, the dynamics of a customer are described by a set of hazard functions and a transition matrix that may depend on the sojourn time. With this model, both contract restrictions and the sojourn dependent autonomous behaviour of customers can be described in the same framework.

It has been shown that the CLV may be heavily influenced by the shape of the hazard functions. Using the simpler Markov approach by assuming a constant hazard function may lead to substantially wrong CLV estimates.

## 4 **Case study: estimating CLV for newspaper subscription customers**

### 4.1 *Introduction*

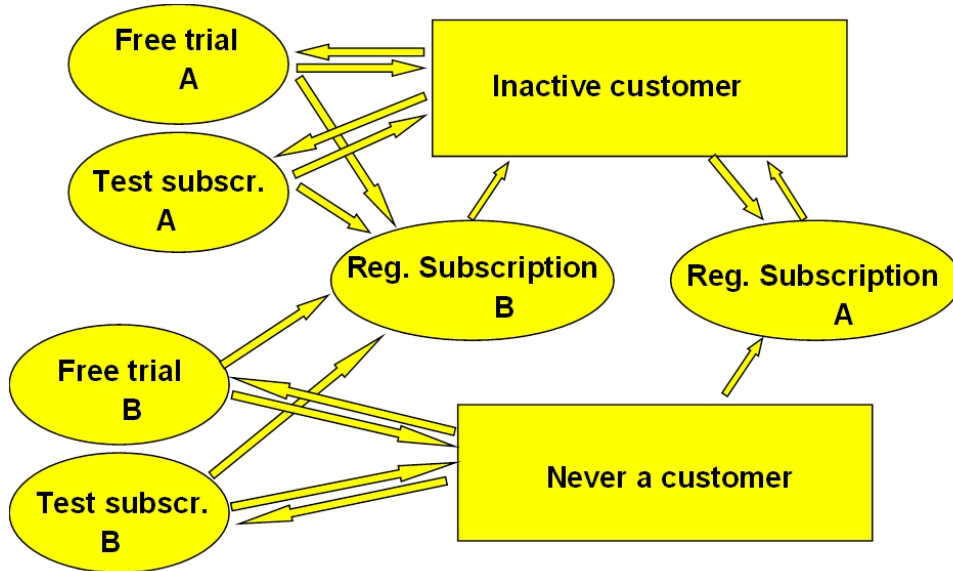
In this section, we describe the application of the SMM to subscription data of a major Swiss newspaper. We obtained the complete contract history of 444,778 customers over a time span ranging from January 2002 to April 2008, which consists of nearly a million different contracts. For each person, some demographic data such as sex, age, region of residence, etc. were available, too. The goal of the study is to evaluate the differences between a classical Markov modelling approach and our SMM approach.

### 4.2 *State space model*

The state space, as given by the original data, is specified in Figure 1 and consists of only four states. Through direct marketing, the company tries to sell 12-week test subscriptions at a reduced rate. Sometimes, also four-week free trials are offered. Both the test subscription and the free trial end with an upselling campaign to a regular contract. In addition, both customers and non-customers may subscribe autonomously to the regular contract at any time. There is, however, no direct marketing aimed at selling regular contracts to non-customers. The regular contract is on a yearly or (more expensive) half-yearly renewal basis, i.e., terminations are only possible after 6 or 12 months. Exceptions are only accepted in special cases, such as when customers move to another town, etc.

A preliminary analysis of the hazard functions shows that there is a substantial difference in the upselling dynamics of customers who have had a regular subscription before, and customers who did never have a regular subscription, respectively. Furthermore, customers who signed autonomously for a regular subscription had a markedly different sojourn dynamics compared to the ones who attained this state by an upselling campaign. Both these findings indicate that the behaviour of a customer is influenced not only by his current state but also by prior states. This type of ‘memory’ cannot be modelled with the SMM: while the dynamics in the SMM may depend on the sojourn time, it is assumed to be independent on prior states.

**Figure 4** Enlarged state space allowing to distinguish between customers who had a regular subscription before (inactive customers), and customers who never had a regular subscription (see online version for colours)



Hence, we cannot continue with the simple four-state space from Figure 1, but have to shape up an enlarged space as displayed in Figure 4. Two different states for non-customers, namely ‘never a customer’ and ‘inactive customer’ are used, where the latter had a regular subscription at an earlier time. Also the regular subscriptions are divided in two states, i.e., ‘regular subscription A’ which means that an order was placed autonomously, and ‘regular subscription B’, where the state was reached via upselling from a test subscription or a free trial. Furthermore, we also split the states ‘free trial’ and ‘test subscription’ in two versions A and B each, depending on if the customer has had a regular subscription before (A) or not (B).

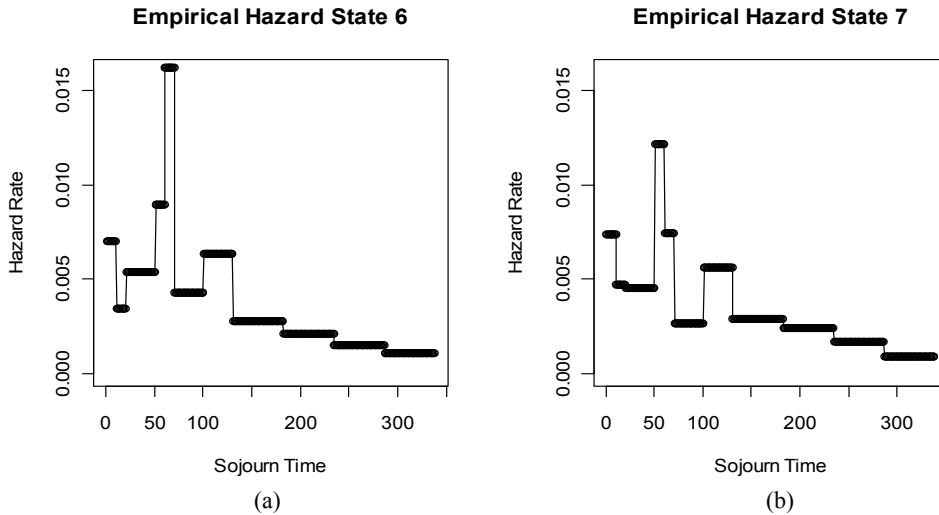
When using this state space, the above mentioned differences in customers’ dynamics can be modelled reasonably with the SMM. The possible transitions between the states are indicated in Figure 4 as arrows.

#### 4.3 Estimation of the model parameters

The hazard functions of the test subscription and the free trial do not need to be estimated from the data since they are completely defined by the contract. For example, for the test subscription,  $h(T) = 0$  for  $T < 3$  months, and  $h(T) = 1$  for  $T = 3$  months.

In Figure 5, we show the estimated hazard functions for states 6 and 7, which are the regular contracts that were either achieved by upselling (state 6), or autonomously (state 7). It is apparent that contracts that were signed due to direct marketing as in state 6 are much more likely to be terminated after 12 months, or more generally, during the first two years of the subscriptions. Once a so-called ‘upselling’-customer has been a subscriber for more than two years, the dynamics does not seem to differ much from an autonomous customer any more. However, before it does.

**Figure 5** Example of the empirical hazard functions for states 6 and 7, i.e., the regular contracts that were signed by upselling (a) and autonomously (b) (time is given in weeks)



Note: If the regular contract was attained via an upselling campaign, the hazard function is significantly higher than if customers signed autonomously.

#### 4.4 Calculation of the CLV

The CLV for all customers are calculated according to equation (5), where evaluating the expression for  $T_0 = 0$  leads to the CLV estimate according to the MCM. The weekly profits are indicated in the following table in arbitrary units:

State no	State	Weekly net profit $c_i$
1	Never a customer	0
2	Test subscription B	1
3	Free trial B	-5
4	Test subscription A	1
5	Free trial A	-5
6	Regular subscription B	8
7	Regular subscription A	8
8	Inactive customer	0

The regular subscriptions (states 6 and 7) have weekly profits of eight, where the test subscriptions at reduced price have only small profits of one per week. The free trials have negative net profits: Since the offer is free, no revenue is earned from the customers, but costs still arise due to the delivery. The states without a contract have weekly net profits of 0.

The weekly discount factor is set to 0.996, leading to a yearly discounting of approximately 20%.

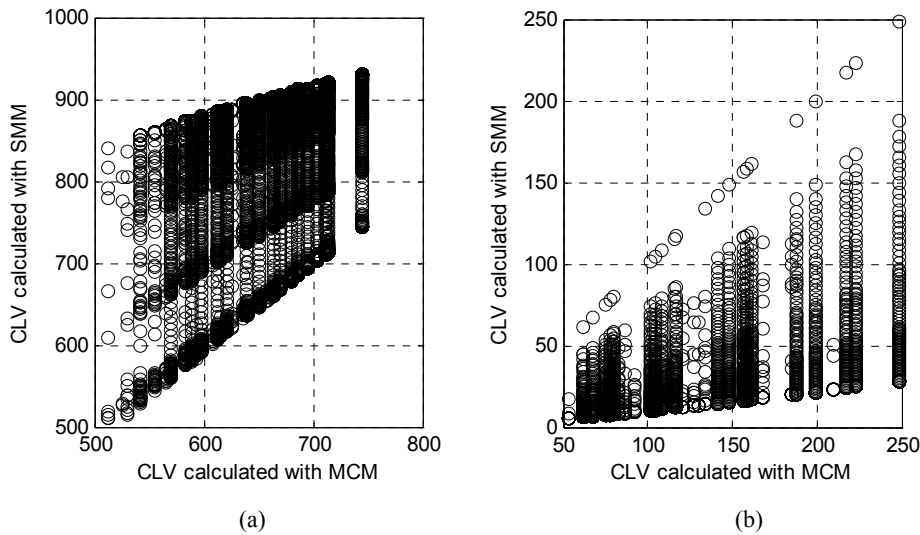
In Figure 6(a), a scatter plot of the CLVs of all customers in state 6 (regular subscription B) is shown, where the two axes are the two CLVs calculated with MCM and SMM, respectively. The CLVs of the MCM vary between 500 and 750, the CLVs of

the SMM between 500 and 900. The different points align vertically. A set of vertical points correspond to a set of customers who have the same dynamical properties but different sojourn times in the state<sup>1</sup>.

It can be seen that the SMM-CLV in Figure 6(a) is larger than or equal to the Markov-CLV. The reason is that with increasing sojourn time, the hazard function decreases, and the expected additional lifetime within the contract increases, see below.

In Figure 6(b), the CLVs of customers in state 8 (inactive customers) are shown. Here, the situation is different: The SMM-CLVs are always lower than the Markov-CLVs. In this case, the CLV of a customer is solely created by the probability that he could subscribe again. The longer a customer stays an inactive customer, the lower this probability is, which in turn reduces the CLV.

**Figure 6** CLV for all customers in (a) state 6 – regular subscription B and (b) state 8 – inactive customer, calculated with MCM and SMM, respectively



Note that in Figure 6(a), the Markov CLV may be up to 40% too low, while in Figure 6(b), the Markov-CLV may be several times too high. Thus, the huge discrepancies between using the MCM and the SMM can be seen in this case, too. Furthermore, since the average difference between MCM and SMM is in the same order of magnitude as the variability of the CLV, the error created by using MCM instead of SMM is similar to the average difference of the CLV among different customers! This makes the results of the CLV calculation with the simpler MCM practically worthless.

The influence of the sojourn time on the CLV can be demonstrated by selecting a set of customers with the same MCM-CLV and plotting their sojourn time against the SMM-CLV. In Figure 7(a), the CLV as a function of the sojourn time is plotted for all customers in state 6 with a Markov-CLV of 570 [compare to Figure 6(a)]. The dependence of the CLV on the sojourn time is non-trivial: during the first year, the CLV decreases slightly because there is a certain probability of contract termination at the end of the first year. After this cancellation peak which is similar to the one described in Section 3.6, the CLV increases rapidly, reflecting the decreased churn probability of



customers who did not cancel at the first opportunity (compare the corresponding hazard function in Figure 5).

In Figure 7(b), the CLV as a function of sojourn time is shown for state 2 (test subscription B) and customers with a MCM-CLV of 37. In this case, too, the CLV increases with increasing sojourn time. However, the reason is different: The value of the CLV at the end of the test subscription is around 58, reflecting the fact that there is a certain probability of switching to the regular subscription. One week before the end of the test subscription, the negative profit of  $-5$  has to be added to the CLV which leads to a reduction. Consequently, the CLV of a customer with  $T_0 = 0$  is around 20 lower than at the end of the test subscription period.

**Figure 7** Influence of sojourn time on CLV, (a) state 6 (regular subscription A), customers with Markov CLV=570 (b) state 2 (test subscription B), customers with a Markov CLV = 37

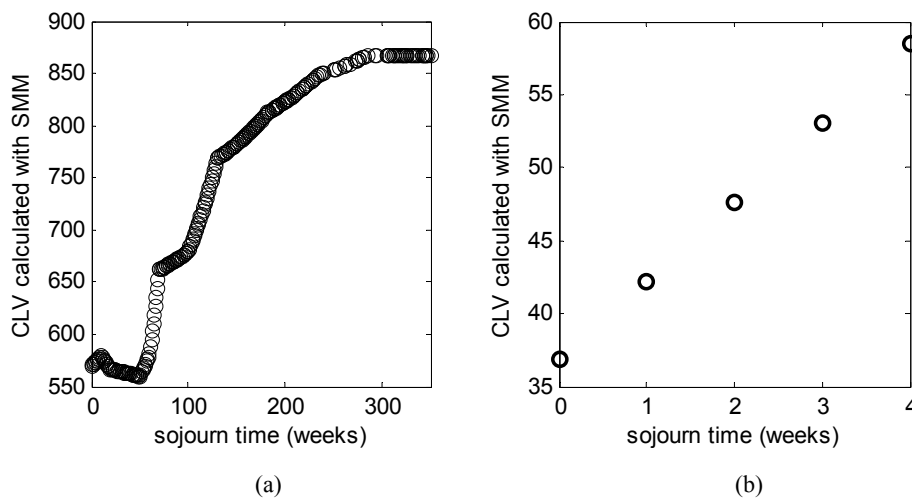


Figure 6 and Figure 7 show impressively that there is a huge difference between modelling with a SMM and the traditional Markov approach. For several states, the variability of the CLVs of all customers within this state is of the same magnitude as the difference between the MCM-CLV and the SMM-CLV. This demonstrates that the dynamical information contained in the hazard functions and the transition matrix is extremely relevant for the CLV. Disregarding this information by using the simpler MCM approach may lead to substantially wrong CLVs in many cases.

## 5 Conclusions

We have developed a general theory for calculating the CLV of customers with a SMM, extending the established MCM. The SMM allows the appropriate modelling of different non-Markovian dynamical elements which are typical for contractual settings, such as specific contract termination times, or minimum duration rules. In addition, the same framework allows including sojourn time dependent churn rates which are frequently observed empirically but have to be ignored in the traditional MCM approaches.

We have shown that the CLV may indeed be drastically influenced by such dynamics, and using a simplified Markov model may lead to substantially wrong estimates of the CLV. Consequently, business decisions based on these numbers tend to be seriously perturbed by using the wrong model.

The framework of the SMMs is an extremely flexible tool. It is slightly more complicated than MCMs but much more adequate for describing the customer dynamics in many markets, in particular in contractual settings. The dynamical model consists of the two elements of a set of hazard functions and a transition matrix that both may depend on the sojourn time. Both elements can be estimated by standard statistical routines. Thus, in addition to be an adequate description of the dynamics, the SMM can be applied in practice easily using standard statistical packages.

### Acknowledgements

This work was supported by CTI, the Swiss innovation agency, under grant 10192.1.

### References

- Berger, P.D. and Nasr, N.I. (1998) 'Customer lifetime value: marketing models and applications', *Journal of Interactive Marketing*, Vol. 12, No. 1, pp.17–30.
- Berger, P.D., Eechambadi, N., Morris, G., Lehmann, D.R., Rizley, R. and Venkatesan, R. (2006) 'From customer lifetime value to shareholder value', *Journal of Service Research*, Vol. 9, pp.156–167.
- Blattberg, R.C. and Deighton, J. (1996) 'Manage marketing by the customer equity test', *Harvard Business Review*, Vol. 74, pp.136–145.
- Dwyer, F.R. (1989) 'Customer lifetime valuation to support marketing decision making', *Journal of Direct Marketing*, Vol. 3, No. 4, pp.8–15.
- Fader, P.S. and Hardie, B.G.S. (2009) 'Probability models for customer-base analysis', *Journal of Interactive Marketing*, Vol. 23, pp.61–69.
- Fielding, M. (2005) 'Get circulation going: DM redesign increases renewal rates for magazines', *Marketing News*, Vol. 1, pp.9–10.
- Gupta, S. and Lehmann, D.R. (2003) 'Customers as assets', *Journal of Interactive Marketing*, Vol. 17, No. 1, pp.9–24.
- Gupta, S., Hanssens, D., Hardie, B., Kahn, W., Kumar, V., Lin, N. and Sriram, N.R.S. (2006) 'Modeling customer lifetime value', *Journal of Service Research*, Vol. 9, No. 2, pp.139–155.
- Gupta, S., Lehmann, D.R. and Stuart, J.A. (2005) 'Valuing customers', *Journal of Marketing Research*, Vol. 41, pp.7–18.
- Heitz, C., Ruckstuhl, A. and Dettling, M. (2010) *Customer Lifetime Value Calculation with a Semi-Markov Model*, Research Report Zurich University of Applied Sciences.
- Jain, D. and Singh, S.S. (2002) 'Customer lifetime value research in marketing: a review and future directions', *Journal of Interactive Marketing*, Vol. 16, No. 2, pp.34–46.
- Kumar, V., Lemon, K.N. and Parasuraman, A. (2006) 'Managing customers for value: an overview and research agenda', *Journal of Service Research*, Vol. 9, No. 2, pp.87–94.
- Malthouse, E.C. and Blattberg, R.C. (2004) 'Can we predict customer lifetime value?', *Journal of Interactive Marketing*, Vol. 19, No. 1, pp.2–16.
- Pfeifer, P.E. and Carraway, R.L. (2000). 'Modeling customer relationships as Markov chains', *Journal of Interactive Marketing*, Vol. 14, No. 2, pp.43–55.

- Reichheld, F.F. and Teal, T. (2001). *The Loyalty Effect*, Harvard Business Press.
- Reinartz, W.J. and Kumar, V. (2003) 'The impact of customer relationship characteristics on profitable lifetime duration', *The Journal of Marketing*, Vol. 67, No. 1, pp.77–99.
- Ross, S.M. (1995) *Stochastic Processes*, Wiley.
- Rust, R.T., Lemon, K.N. and Zeithaml, V.A. (2004) 'Return on marketing: using customer equity to focus marketing strategy', *Journal of Marketing*, Vol. 68, pp.109–127.
- Rust, R.T., Zeithaml, V.A. and Lemon, K.N. (2000) *Driving Customer Equity: How Customer Lifetime Value is Reshaping Corporate Strategy*, The Free Press, New York.
- Villanueva, J. and Hanssens, D.M. (2007) 'Customer equity: measurement, management and research opportunities', *Foundation and Trends in Marketing*, Vol. 1, pp.1–95.

## Notes

- 1 The independent variables of the customers are given in categorical form, only a finite number of different parameter combinations exist. This leads to the fact that different customers are described by the same parameter values of the dynamical model. These customers may, however, still have different sojourn times.