DEPARTMENT OF MANAGEMENT

AFDELING FOR VIRKSOMHEDSLEDELSE

Working Paper 2006-4

Expected Present Value of a Customer

Niels Stender

UNIVERSITY OF AARHUS • DENMARK

ISSN 1398-6228

Expected Present Value of a Customer

Niels Stender
University of Aarhus
niels@innohead.com

May 10, 2006

Abstract

A quantitative framework for evaluating expected present values of future customer transactions is presented and applied to a telecommunications company's customer database. The framework is applicable to continous service providers in general.

Contents

1 Introduction

Is is described in Grant & Schlesinger (1995) how competetive pressure and information technology have changed the strategic focus of companies from cost reduction and mass marketing, towards segmented markets and eventually management of every single customer relation.

Managerial aspects of this development has been treated by numerous authors in the Customer Relationship Management literature. For instance, Reichheld & Teal (1996) and Reichheld (1996) argues that more loyal customers strongly affect profitability and recommend aligning organizational resources as to obtain stronger degrees of loyalty.

Marketing research into loyalty and its determinants has mainly been conducted through a qualitative framework based on questionnaire data, but a need for quantitative research has risen as companies are increasingly seeing customer relationships as assets in line with other more tangible assets. What is becoming known as the Customer Lifetime Value literaturee is addressing this need, by quantitatively modeling and measuring constructs such as loyalty and profitability based on observed behavior. See Jain & Singh (2002) for a review of this young topic. Examples of contributions are Schmittlein & Peterson (1994), who present a valuation-model of the customer base, accounting for stochastics in customer retention, repurchase intensity and purchase levels. Blattberg & Deighton (1996), Dwyer (1997), Berger & Nasr (1998), Reinartz & Kumar (2000) and Jacobs et al. (2001) are other examples of authors utilizing lifetime value concepts either as a primary study or as a tool for testing other hypothesis.

The possibility of adressing the issue quantitatively has eased, as computing power and data availability such as customer records and scanner data are routinely maintained in more industries. For the services industries, continous service providers such as telecommunications and financial services have had access to customer specific information for years due the nature of their business. Therefore this type of company will serve as a demonstration of the quantitative models presented.

The models presented are useful for contractual settings, i.e. settings in which customers subsribe to a service and is required to take explicit action to end the relation. In non-contractual settings, as Reinartz & Kumar (2000) points out, a different methodological approach is be called for that explicitly accounts for the unobservability of an agent ceasing to consider himself a customer.

Any model of customer profitability should ultimately be linked to actionable parameters for management, such as in the framework put forward by Rust et al. (1995). Such a link is only touched upon indirectly due to data availability, but the modeling approach could be extended to account for other types of data. As an example, Bolton (1998) is seen to include questionnaire data.

Decreased cost of customers over time, due to for instance learning, and value of word-of-mouth[1] should also be accounted for, but unavailability of data and proper methodology prevents further pursuit of these issues at present.

The structure of the paper is as follows. In section (2) a flexible statistical model of customer retention is studied, followed by a short discussion of a behavioral model implied. Then in section (3) a measure of the present value of customer spending, adjusted for the risk of customers exiting, is introduced.

[1] As considered theoretically by Jacobs et al. (2001)

2 Customer Duration

2.1 Statistical Model

A statistical framework for forming inference of the lifetime of a customer is put forward here. The framework is known in the statistical literaturee as a duration model and is a special case of event history analysis. These models are widely used in medicine, engineering, labour market economics, insurance and is spreading to other areas as data on durations become available. In the marketing literaturee duration models have been used to model for instance inter-purchase times, see Vilcassim & Jain (1991) and Gönul & Srinivasan (1993); and brand-switching, see Wedel et al. (1995). Li (1995) argued for the applicability of duration models to customer duration times in a phone company setting, while Bolton (1998) studied life-times in relation to satisfaction in a model with no unobservable heterogeneity.

2.1.1 Basics

Let the stochastic variable ⋆ describe the duration of a customer, that is, the time measured from the customer enters a …nancial relationship of interest till the end of the relationship. Then the hazard function for ⋆ is de…ned as

$$\star(\star) = \lim_{\star\star \,!\, 0} \frac{\Pr(\star \cdot \star \cdot \star + \star\star j \star \,\, \star)}{\star\star} \star$$

The hazard can be interpreted as the instantaneous rate of customers leaving per unit time, at time ⋆. The probability of surviving up and at least to time ⋆ is captured by the survivor function $\star^{1}(\star)$. If ⋆ had a proper distribution function, this could be calculated as $\star^{1}(\star) = 1 \; \text{i} \; \mathrm{R}_0^{\star} \star(\star) \star\star\star$ However, the distribution function may not be proper, since some customers could be thought to have in…nite duration. The hazard function is de…ned even in this instance, which is one reason why it is often used as the point of departure in duration modeling. Given a hazard, one can derive the survivor function as

$$\star^{1}(\star) = \star^{\text{i} \; \text{l}_0^{\star} \star(\star)\star\star} \star$$

Let us now specify a hazard function allowing for heterogeneity among customers. The purpose of this modeling is to explain di¤erences in duration attributable to some observed variables ⋆. The hazard of exiting the state of being a customer is assumed to consist of three components: A baseline hazard common to all, observed heterogeneity and unobservered heterogeneity. Marketing researchers are increasingly unwilling to assume that the set of variables available for modeling, captures all essential information governing agent model behavior. Hence, to explicitly incorporate this aspect in a given model is increasingly popular.

For reaching this goal, the mixed proportional hazard model is often chosen among the duration models. A typical speci…cation and one we shall utilize, is seen in eq. (1).

$$\star(\star j \star\star\star; \star) = \star\star^{\star\star} \star_0(\star; \star_0)\star \quad \star\star \; 0 \star\star = \mathrm{f}\star\star\star_0 \mathrm{g} \tag{1}$$

The baseline hazard $\star_0$ is a positive function common to all agents, being a function of time alone with parameter vector $\star_0\star$ Agents are allowed to display heterogeneity through a proportional factor $\star\star^{\star\star}\star$ where ⋆ is a vector of observable characteristics for each agent and ⋆ is a vector of unknown parameters. ⋆ is a scalar unobserved heterogeneity term, de…ned to have the distribution $\star\,(\star;\star)$ with unknown parameters $\star$. Independence between $\star$ and $\star$ is assumed.

The key bene…t received for the seemingly restrictive proportionality assumption, lies in the identi…ability of the model as demonstrated by Elbers & Ridder (1982), even for $\star$ from the unobservered heterogeneity distribution. This assumption of proportionality may not be realistic, but there are ways of testing it and generalizing it. A simple generalization would be to partition some sample by a criteria believed to generate heterogeneity over $\star_0$, letting each partition $\star$ be governed by a seperate baseline hazard $\star_{0\star}$. A more general and complicated approach would be to follow Horowitz (1999), but this approach puts severe demands on data and may converge slowly for samples of realistic size.

2.1.2 A Semi-Parametric Speci…cation

Since $\star_0$ is unknown and we have no real guidiance on its form, we shall opt for a ‡exible speci…cation. One can partition the timeline into …nite intervals at the points $f\star_\star g_{\star=0}^{\star}$ and assume constant hazard in these as in eq. (2). At the cost of adding more parameters to the model, a completely ‡exible speci…cation can be achieved.

$$\star_0(\star;\star_0) = \begin{cases} \star_0 \text{ for } \star_0 \cdot \star\star\ \star_1 \\ \star_1 \text{ for } \star_1 \cdot \star\star\ \star_2 \\ \star\star\star \end{cases} \qquad \star_0 = f\star_0\star\star\star\star\star_\star g\star\ \star_\star \ \text{¸}\ 0 \tag{2}$$

In the applied case later on, data is grouped in monthly intervals. For grouped data, (Lancaster 1990, p. 180) shows that the maximum likelihood estimator given any baseline hazard for the underlying continous process, yields a model identical to eq. (2).

Adding everything together, we get a hazard function with a heterogeneity term in eq. (3) and a derived survivor function in eq. (4).

$$\star\,(\star j\star\star\star;\star) = \star\star^{\star^0\star} \sum_{\star=1} 1\,(\star_{\star\text{¡}1} \cdot \star\star\ \star_\star)\,\star_\star \tag{3}$$

$$\begin{aligned} \star^1\,(\star j\star\star\star;\star) &= \exp\left(\text{¡}\,\star\star^{\star\star}\int^{\star}\star\,(\star j\star\star\star\star\star)\,\star\star\right) \\ &= \exp\left(\text{¡}\,\star\star^{\star\star}\left[\sum_{\star=0}^{\star 0}(\star_\star\ \text{¡}\ \star_{\star\text{¡}1})\,\star_\star + (\star\ \text{¡}\ \star_{\star\,\text{¡}\,1})\,\star_\star\right]\right) \end{aligned} \tag{4}$$

where

$$\star = \sup f\star j\star_\star \cdot \star g$$

Since $\star$ is unobserved by de…nition, it is necessary to infer parameters from marginal distributions, practically integrating $\star$ out. Given the distribution $\star$ of the unobserved heterogeneity , the marginal survivor function is derived as

$$\star^1\,(\star j\star;\star) = \int_{<^+} \star^1\,(\star j\star\star\star;\star)\,\star\ (\star\star)\,\star$$

Working on grouped data, …rstly de…ne $\star^1\,(\star j\star\star\star;\star) = \star^1\,(\star_\star j\star\star\star;\star)\ \star\ \star 2\ f0\star\star\star\star\ \text{¡}\ 1g$ and $\star^1\,(\star j\star;\star) = \star^1\,(\star_\star j\star\star\star;\star)$ (this makes it convenient to work with intervals instead of $\star$)$\star$

A point mass distribution for ⋆ such as

$$⋆ (⋆;⋆) = \begin{cases} ⋆ & \text{for } ⋆ = ⋆_1 \\ 1 ¡ ⋆ & \text{for } ⋆ = ⋆_2 \\ 0 & \text{otherwise} \end{cases} \quad ⋆⋆ = f⋆⋆⋆_1⋆⋆_2 g ⋆⋆ 2\ [0;1]⋆$$

is often used in the duration model literature. In a MC-experiment, Heckman & Singer (1984) shows that if the true distribution ⋆ is approximated by a discrete point mass distribution $⋆^0$, convergence of $⋆^0$ up to ⋆ may be slow, but the estimate of ⋆ display reasonably robust convergence properties. Also, allowing the number of mass points in the discrete distribution to be estimated along with the other parameters in the model, tends to make estimates more robust against some forms of misspeci…cation.

Imposing a simple two-point mass distribution we get

$$⋆^1 (⋆j⋆;⋆) = ⋆⋆^1 (⋆j⋆_1⋆⋆;⋆) + (1 ¡ ⋆)⋆^1 (⋆j⋆_2⋆⋆;⋆)⋆$$

A similar structure obviously arises when ⋆ has additional mass points. The current model is identi…ed up to a constant, which is why a normalization constraint is imposed. Often this is taken to be ⋆ (⋆) = 1⋆ but to facilitate the approach of Baker & Melino (2000) the …rst level in the baseline hazard of the piecewise-constant hazard is normalized to 1, so $⋆_0 = 1$ is imposed.

2.1.3 Censoring

One important advantage of this type of modeling, is the inherent ability to treat case censoring. When observing customer durations, not all customers will be observed to exit, since our window of observation will be limited. De…ne

$$⋆_⋆ = \begin{cases} 1 & \text{if customer } ⋆ \text{ is observed to exit at some point} \\ 0 & \text{if customer } ⋆ \text{ is censored at some point} \end{cases} ⋆$$

Given a set of parameters, we can state the likelihood of observing a particular set of events. There are two cases: 1) When observing a single customer exiting between $⋆_⋆$ and $⋆_{⋆+1}$, not observing the point of exit. The corresponding likelihood would be $⋆^1 (⋆j⋆;⋆) ¡ ⋆^1 (⋆ + 1j⋆;⋆)⋆$ This is the probability of surviving up to at least time $⋆_⋆$ less the probability of of surviving up to at least time $⋆_{⋆+1}⋆$ 2) If the customer is known to survive up to time $⋆_⋆$, but censored from here on, the corresponding likelihood would be simply $⋆^1 (⋆j⋆;⋆)$.

Observing ⋆ customers, with customer ⋆ observed up to or exiting in interval $⋆_⋆$ the likelihood is

$$⋆ (⋆) = \left[⋆^1 (⋆_⋆ ¡ 1j⋆;⋆) ¡ ⋆^1 (⋆_⋆j⋆;⋆)\right]^{⋆_⋆} \left[⋆^1 (⋆_⋆j⋆;⋆)\right]^{1 ¡ ⋆_⋆} ⋆ \quad (5)$$

Given a …xed number of mass points in ⋆⋆eq. (5) is fully speci…ed and identi…ed.

For the model put forward here to be valid, a special sampling scheme must be used. In the duration literaturee, two schemes are mainly referred a) Stock-sampling, which amounts to sampling from an existing population at a given time and b) Flow-sampling, amounting to sampling among the arrival of new individuals in the population. Flow-sampling is appropriate here. When accounting for unobservered heterogeneity in a duration model, matters are complicated to the intractable when working in stock-samples, while failing to account for either a stock-sample property present in data or unobservered heterogeneity, will bias any investigation severely.

2.2 Behavioral Model

The statistical model put forward above can be seen as the reduced form of a simple structural model. Let us think of a two-state model of a customer and his relations to the service provider. At each point in time he can either continue to consume the current service or switch to another provider.

Assume a customer with individual characteristics vector $\star\star$ receives net bene…ts $\star(\star)$ continously from consuming a service. O¤ers from competitiors arrive as in a Poisson process with intensity parameter $\star_\star(\star)$. Each o¤er yields potential alternative bene…ts $\star$ and is drawn from from a population described by a distribution $\star\ (\star;\star\star\star)$ conditional on $\star$ and $\star\star$ The customer accepts the new o¤er if $\star\star\ \star(\star)$ and rejects otherwise.

For a small period of time $\star\star$ the approximate risk of the customer receiving a new o¤er would be $\star_\star(\star)\star\star$ The probability of the customer accepting the o¤er, would be $1\ ¡\ \star\ (\star(\star);\star\star\star) = \bar{\star}^{1}(\star\star\star)\star$ so the approximate risk of the customer exiting would be $\star_\star(\star)\bar{\star}^{1}(\star\star\star)\star\star$ The corresponding hazard $\star\ (\star;\star)$ is therefore $\star\ (\star;\star) = \star_\star(\star)\bar{\star}^{1}(\star\star\star)\star$

A multiplicative hazard model is derived if the choice of $\bar{\star}^{1}$ is restriced, such that $\bar{\star}^{1}(\star\star\star) = \star\ (\star)\star\ (\star)\star$

To arrive at the model above, we could assume $\star = \left[\ \star_0\ \ \star_1\ \right]^0$ and set $\star_\star(\star) = \star^{\star_0\star_0}$ and $\bar{\star}^{1}(\star\star\star) = \star^{¡\star_1\star_1}\star_0(\star)\star\star_0$ being de…ned in eq. (2). This amounts to the Poisson process di¤ering in intensity through a subset of the individual characteristics and to the alternative o¤ers $\star$ being exponentially distributed with parameter $\star = \log\star_0(\star) + \star_1\star_1\star$

We would arrive at a similar reduced form model if we exchanged $\star_0$ and $\star_1$ for $\star\star$ but the parameters $\star_0$ and $\star_1$ would not be identi…ed.

The distributional assumptions are quite restrictive. A host of other distributions could be utilized or even left unspeci…ed, resulting in a non- or semi-paramteric statistical model. However, when leaving the multiplicative hazard model, it becomes unwieldly to include terms accounting for unobservered heterogeneity.

This model bears a loose resemblance to models of job-search. In these models it is usually assumed that the customer knows $\star$ and acts optimally according to this information. Here in contrast, the customer seems to act naively by utilizing a simple decision rule. Firstly it would be easy to interpret $\star(\star)$ as bene…ts including switching costs. Secondly, modeling a more sophisticated decision process for an optimizing agent, would usually put more demands on data than what is expected to be currently available in most company databases.

2.2.1 Perspective to Measures of Intention

As an alternative to monitoring repurchase behavior directly, constructs such as customer satisfaction or repurchase intention[2] is often used as an approximating measure due to the ease of collecting this information and for the ease of interpretation. Mittal & Kamakura (2001) studies the link between the intentional and realized variable in the automobile industry and demonstrates how reported intentions and subsequent realized behavior is moderated by customer characteristics. This …nding is supported by other authors.

The behavioral model put forward here provides a perspective through which the repurchase intention methodology can be viewed, when dealing with contractual settings. In a continous services market with a number of near substitutes, any current customer of some service would not yet have received a signi…cantly better o¤er from competitors, because then he would already have switched. Asking the customer if he is satis…ed with

[2] The opposite concept would be "continuation behavior" and "continuation intentions", which seems more appropriate in a contractual setting

the current service should yield a positive response, since to his current knowledge, no signi…cantly better service exists. This would not be unlike asking an investor if he is satis…ed with his current portfolio of …nancial assets. If he believed another portfolio would give a better risk-yield ratio, deducting transaction cost, he would already be in the process of executing transactions towards such a portfolio.

Measuring individual repurchase intentions with the object of predicting customer defections, should not display much predictive power over the longer term. It would primarily tag consumers already having perceived a better o¤er, currently moving through a window of time created by technical and administrative barriers.

This conforms with the authors experience of instability in segments generated on the basis of simplistic customer satisfaction measures. The phone company from which the data in this article originates from, made a study of customer satisfaction along two dimensions in a random sample of customers. In questionnaire data they measured customers perception of a) the degree of satisfaction/attractiveness of the companys current product o¤erings and b) the attractiveness of competitors o¤ers on the market. A loyalty index was formed, such that for instance customers …nding a competitors o¤er attractive while …nding the current companys o¤er unattractive, were categorized as high risk customers. When measuring these dimensions a year later, it turned out that a very large portion of customers had moved to another level of loyalty. This could be interpreted in the light of the views put forward earlier in this section, although other explanations are possible.

The point of view applies only to continous service providers. In non-contractual settings, the de…nition of a customer is obfuscated by the inability of the company to formally measure the current status of the customer in the accounting system or elsewhere. Measuring repurchase intentions in this setting then, in a way amounts to measuring the extent of the customer base.

2.3 Application

Now let us demonstrate an application towards modeling customer potential value. Given a customer base from a telco, accounting information on usage levels and a commercial demographic area-keyed database, we shall draw inference on the expected value to be derived from various customer segments.

2.3.1 Data

A sample of 4712 telephone customers is investigated. The group is ‡ow-sampled at random, from the subpopulation of customers in the telcos' database arriving on or after 1st of January 1999 until November 2001, an observation window of 35 months. It contains observations on date of the beginning of the relationship, monthly usage of phone denominated in minutes and economic worth. When a customer ceases to uphold his subscription or buys all his minutes from another operator, he is recorded as churned. This in turn, limits the scope of the application to modeling the variable element of the cash‡ow from a given customer. However, extending the model to include the …xed monthly fee of the subscription is simple.

In the appendix some descriptors relevant for duration modeling is presented using the Proc LIFETEST in SAS. It is seen how standard errors on the survivor probability estimates gets larger, as the number of observations thins out. Looking at the hazard it is seen to decrease steeply and then increase again towards the end. The last part of the

Group	Variable	Note
Aggregated Area-Specific Variables		
Family Structure	FSINK	Singles, no children
	FDINK	Couples, no children
	FSIK	Singles w/children
	FDIK	Couples w/children
Housing	HDT	Detached Houses
	HT	Terraced Houses
	HMF	Houses with multiple families
	HO	Other types of buildings
Employers	EAGRI	Agriculture
	EINCO	Industry and Construction
	ESERVICE	Services
	EPUBLIC	Public
	ENA	Unknown
Job Status	JSE	Self-Employed
	JHPL	High-pay Labour
	JMPL	Medium-pay Labour
	JLPL	Low-pay Labour
	JU	Unemployed
	JNP	Not participating in workforce
	JNA	Unknown Job Status
Age	AGE0_17	People aged from 0 to 17 years
	AGE18_32	
	AGE33_45	
	AGE46_60	
	AGE61_99	
Other	INC250PLUS	People earning more than $30.000/year
Individual Specific Variables		
	SEX	A guess derived from firstname
	VERTICAL	#subscribers on same coordinate
	DENS20	#subscribers in a 20sq.m area
Calendar Time		
	START	Month of Arrival, counting from January 1999

Figure 1: An overview of explaining variables

hazard is estimated using a very small number of churners, so the increasing hazard in the end may be a spurious result.

The area-keyed demographics database contains aggregated information on areas of on average 560 individuals with a standard deviation of 231 individuals. The variables measured includes income, age, type of family, housing, economic activity and jobs. All variables are speci...ed as numbers of individuals in the area falling within a given category. These are transfomed to proportions and included in the most general model.

In ...g. (1) an overview of variables is shown. Not all listed variables are included in the most general model however, since more groups of variables sum to one$\star$ From every group, the variables with the smallest correlation with an arti...cial[3] duration is left out.

2.3.2 Estimation

Optimizing w.r.t. $\star$ is not trivial, since the likelihood surface has several local optima and is singular at the boundaries of the parameter space of $\star$. This is a known problem not unrelated to the problem of estimating mixing parameters in mixture distributions among others.

Here a solver is used utilizing the Broyden-Fletcher-Goldfarb-Shanno algorithm of optimization, starting out with a model containing the full set of variables, but no unobserved heteregeneity. Estimates from this basic model then serves as starting values

[3] A variable $\star^{¤}$ was constructed, such that

$$\star^{¤} = \begin{cases} \star & \text{if uncensored} \\ 2\star & \text{if censored} \end{cases} \star$$

This implies the unrealistic assumption, that censoring when present, always happens in the middle of an interval.

for more complicated models, that is, models with more mass points in the unobserved heterogeneity distribution.

⋆ being unobserved and its distribution ⋆ approximated by a discrete mass point distribution, leaves the question of the number of mass points to include. Baker & Melino (2000) suggests, based on Heckman & Singer (1984), a speci…c-to-general approach in which the number of mass points is gradually increased by an iterative testing procedure that produces starting values for the continuing optimization. When an additional mass point doesn't increase the log-likelihood su¢ciently, the iteration is stopped and the number of mass points is …xed.

At this point, elimination of insigni…cant variables is undertaken, using a general-to-speci…c method on a 10% signi…cance level. The most general model includes squares of variables listed in …g. (1) in order to account for more complicated types of dependencies i.e. maybe extreme values of the proportion of self-employed generates higher risk of exit. To soften the issue of multicollinearity, any pair of variables with a correlation above 0.98 had its square removed at the outset. A high degree of correlation as such is also seen in data, as documented in the appendix in section 5.1, which could invalidate variance estimates.

The baseline hazard is speci…ed as a stepwise function containing 25 intervals of length 1 month, except the …rst and last two intervals, as seen i eq. (6). The …rst interval is given a length of 3 months, since administrative technicalities such as three months notice may be administered pragmatically in practice, regret and errors may induce a type of variability that is not informative to model in the current context. The last two intervals are also given a length of 3 months each, since the number of uncensored cases start to dwindle and therefore gives estimates of too short intervals a high variance.

$$\star_0(\star) = \begin{cases} \star_0 & \text{for } 0 \cdot \star\star\ 3 \\ \star_1 & \text{for } 3 \cdot \star\star\ 4 \\ \vdots & \vdots \\ \star_{21} & \text{for } 23 \cdot \star\star\ 24 \\ \star_{22} & \text{for } 24 \cdot \star\star\ 27 \\ \star_{23} & \text{for } 27 \cdot \star\star\ 30 \end{cases} \tag{6}$$

Durations longer than 30 months are simply treated as censored.

2.3.3 Model Performance

Given the full customer database, we shall be interested in measuring how well the model predicts who exits. Lift is a measure often used for classi…cation problems and it will be used here.

A baseline guess on the probability that a given customer exits, would simply be the proportion of exits. It amounts to #churners/#customers= 887⋆4712 = 0⋆1882⋆ Given a sample of 100 customers from the database, we would expect around 19 to be churners. Given our model, the survivor probability of surviving from birth to the right-censoring barrier is calculated. The probabilities are sorted in increasing order. Let ⋆ (⋆) ⋆ ⋆ integer, be the number of churners among the …rst ⋆ customers in the sorted set. Then ⋆⋆⋆ ⋆ (⋆) = $\frac{\star(\star)}{0\star 1882\star}$ i.e. how much better the model predicts the number of churners compared to the baseline guess.

The lift measure is graphed in …g. (6) for 1000 observations. Notice that given ⋆ customers, it will always be the case that ⋆ (⋆) = 1⋆ The lift of the duration model is well above 1 for the …rst …fth of the observations, which is comforting.

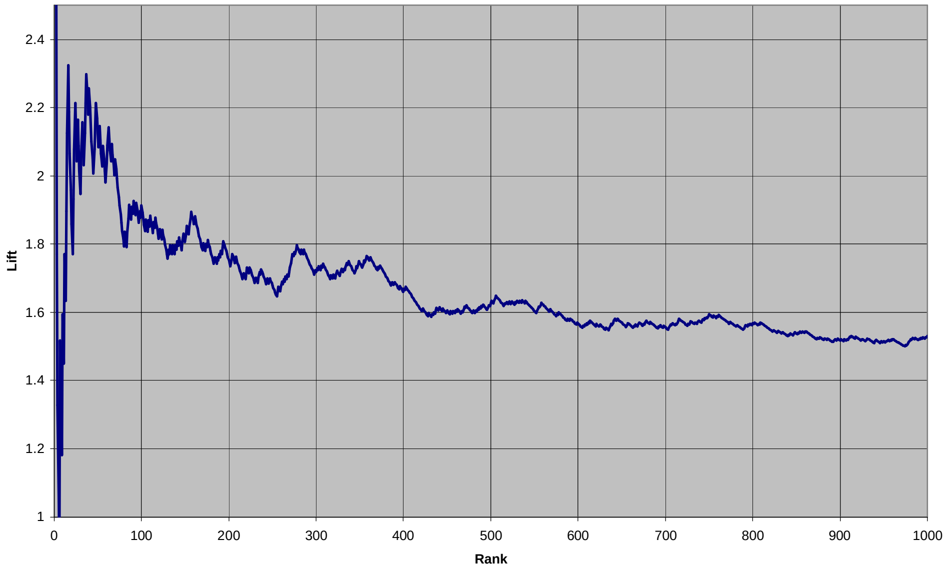


Figure 2: The lift of the duration model

	Coefficient	Std.Error	t-value	t-prob
CSTARTD100	41.2862	3.86	10.70	0.000
CSTARTD100SQ	-109.9410	11.79	9.33	0.000
FDINK	-10.9232	2.03	5.39	0.000
FSIK	-11.6604	3.56	3.27	0.001
HMF	-2.7909	0.95	2.94	0.003
EINCO	-22.2349	5.28	4.21	0.000
EPUBLIC	24.2383	3.81	6.37	0.000
JHPL	-21.8063	5.13	4.25	0.000
JU	80.0718	21.09	3.80	0.000
FSIKSQ	32.3020	10.69	3.02	0.003
FDIKSQ	-14.4429	4.33	3.33	0.001
HMFSQ	2.3890	0.86	2.78	0.005
EPUBLICSQ	-40.8960	6.12	6.68	0.000
JSESQ	-86.6225	19.90	4.35	0.000
JHPLSQ	40.2622	16.25	2.48	0.013
JUSQ	-725.3990	234.00	3.10	0.002
INC250SQ	9.6630	1.86	5.19	0.000
EINCOSQ	60.4031	13.98	4.32	0.000

Figure 3: Beta Parameter Estimates

I have been told by analysts at the phone company, that the industry standard lift is around 3 or 3.5 when using individual specific information such as other products consumed, age and so on with datamining algorithms such as classifier trees and neural networks[4]. The lift obtained here of around 1.5 was obtained by using nothing but area-specific demographics, which supports the model.

Other types of tests based on more sound statistical principles should be utilized here, but it is not trivial to perform for instance residual analysis when unobserved heterogeneity is accounted for along with censoring.

2.3.4 Results

In table (3) estimates of β is reported.

In table (5) estimates of the baseline hazard gamma parameters is reported and a graph is presented in fig. (6). The baseline hazard is seen to increase by customer duration,

[4] See Hand et al. (2001) for an introduction to data mining as a discipline.

	Coefficient	Std.Error
Gamma1	0.2926	0.1910
Gamma2	0.4396	0.2228
Gamma3	0.7903	0.2341
Gamma4	1.1334	0.2422
Gamma5	1.3245	0.2565
Gamma6	1.5056	0.2730
Gamma7	2.0064	0.2735
Gamma8	1.6932	0.3259
Gamma9	1.9849	0.3316
Gamma10	2.1682	0.3468
Gamma11	2.8560	0.3381
Gamma12	2.2892	0.4123
Gamma13	3.0991	0.3786
Gamma14	2.8205	0.4365
Gamma15	2.6460	0.4897
Gamma16	3.3199	0.4448
Gamma17	3.4803	0.4646
Gamma18	3.8808	0.4747
Gamma19	4.2541	0.4820
Gamma20	4.4225	0.5064
Gamma21	3.8165	0.6200
Gamma22	5.0660	0.5059
Gamma23	5.9130	0.5789

Figure 4: Gamma Parameter Estimates

which for a given type of customer heterogeneity says that customers are more prone to churn, the longer they have been customers. On an aggregated level this will not be case however, since high-risk customers will churn faster than others and therefor leave back the low-risk customers.

In …g. (7) and (8) point estimates of the standardized heterogeneity distribution is reported. The levels are scaled, such that $\star(\star) = 1\star$

Interpreting the coe¢cients is a bit more messy in a duration model as compared to OLS. If a coe¢cient $\star_\star$ is positive, it naturally increases the hazard if $\star_\star$ is increased. If a variable enters with a square term, the direction of the e¤ect would be $(\star_{\star 1} + 2\star_{\star 2}\star_{\star\star})\star$ since the functional form in which the variable enters is $\star_{\star 1}\star + \star_{\star 2}^2\star^2\star$ To get an idea of the direction of e¤ects of variables with a square term, the 2nd degree polynomium $\star_{\star 1}\star + \star_{\star 2}\star^2$ is plotted in section (5.2) along with a plot of their distribution.

Using the computed estimates, we can now form statistical statements about future exit-behavior using the survivor-function and given a set of characteristics $\star$. Let $\star_{\star\star}$ be the probability of customer $\star$ exiting in period $\star$, as de…ned in eq. (7). A variant of these probabilities will be used as an element in a discounting device in the next section, so it is warranted to study their properties and relation to model parameters.

$$\star_{\star\star} = \star^{1}(\star ¡ 1j\star_\star) ¡ \star^{1}(\star j\star_\star)\star \quad \star\star \; 0 \tag{7}$$

Let $\star_{\star\star}$ be the $\star$'th element of the observed characteristics for customer $\star\star$ The sensitivity of an exit probability $\star_{\star\star}$ with respect to a small change in a model variable $\star_{\star\star}$, would be calculated from the expression in eq. (8).

$$\frac{\star\star_{\star\star}}{\star\star_{\star\star}} = \star_\star \sum_{\star=1} \star_\star \left[\star^{1}(\star ¡ 1j\star_\star\star\star_\star) \log \star^{1}(\star ¡ 1j\star_\star\star\star_\star) ¡ \star^{1}(\star j\star_\star\star\star_\star) \log \star^{1}(\star j\star_\star\star\star_\star) \right] \tag{8}$$

For variables entering with a square, the equvivalent expression would be expression would be $\frac{\star\star_{\star\star}}{\star\star_{\star\star}} = (\star_{\star 1} + 2\star_{\star 2}\star_{\star\star})[¢]\star$ To get a feel of the magnitude of these e¤ects, the sensitivities of customers having been retained for 12 with respect to the sample average of $\star$ is reported in …g. (9). The elasticity is also calculated at $\bar{\star}$ in order ease a comparison of magnitudes.

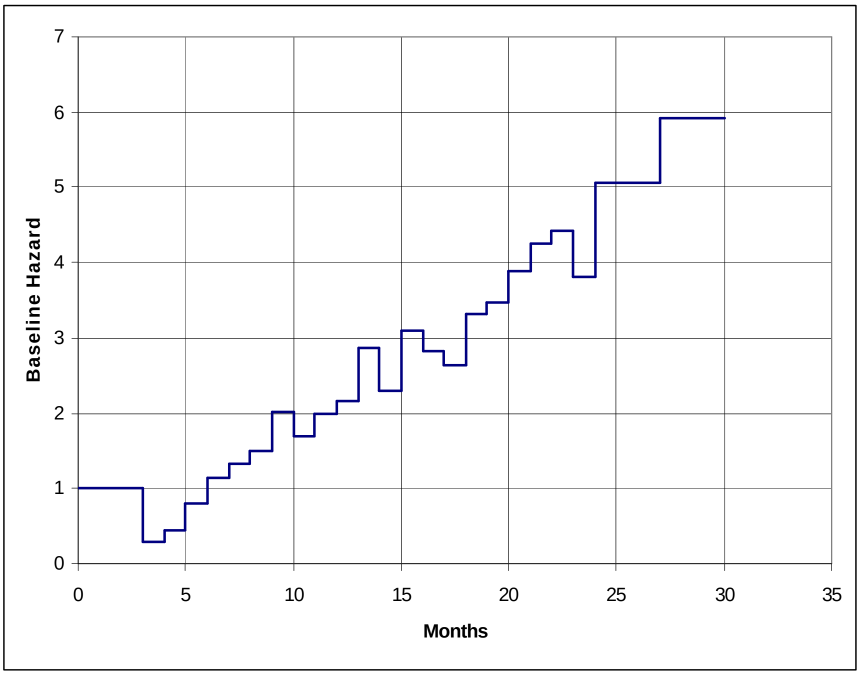


Figure 5: A plot of the baseline hazard

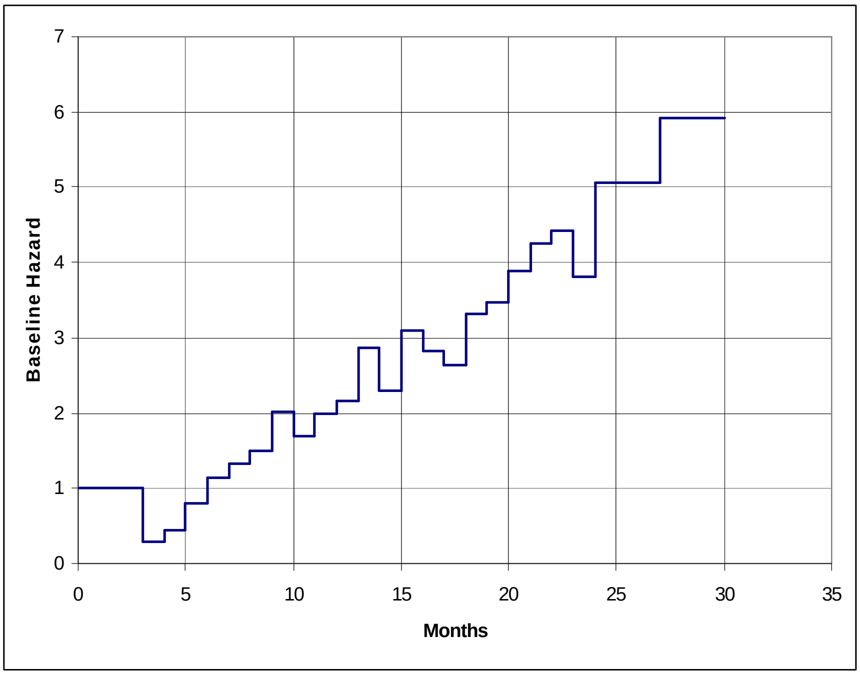


Figure 6: A plot of the baseline hazard

v	p
0.0115	0.8480
1.2537	0.0560
9.5912	0.0959

Figure 7: Mass Point Estimates for Heterogeneity Distribution

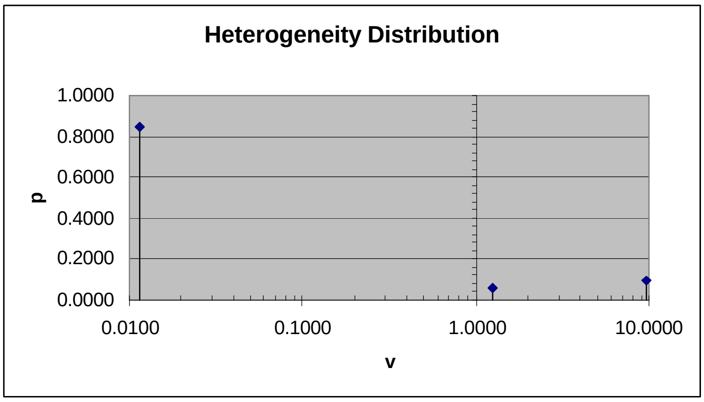


Figure 8: Log Mass Point Graph

	Derivative	Elasticity
CSTART	0.0077	0.8372
FDINK	-0.0164	-2.7224
FDIK	-0.0068	-0.7064
FSIK	-0.0132	-0.3901
HMF	-0.0006	-0.2009
EINCO	-0.0089	-0.7961
EPUBLIC	0.0128	1.6362
JHPL	-0.0235	-1.2038
JU	0.0401	0.9811
JSE	-0.0129	-0.4207
INC250	0.0120	3.2649

Figure 9: Sensitivity Measures on $\star_{¢12}$ w.r.t. changes in $\star$

Due to the squared variables entering the model it is dangerous to comment on the e¤ects of changes except for a small ball around $\bar{\star}\star$ It is seen that more recent customers on average is more risky customers, while moving towards "double-income no-kids" areas decreases risk dramatically. "Double-incomes with kids" areas does the same thing on a smaller scale, as "single-income with kids"-areas on an even smaller scale. The interpretaion here is a bit tricky though, since moving towards areas with SIKs will on average decrease the proportion of DINK's, because these are correlated with $\star$ = ¡ 0$\star$37$\star$

Moving away from the average towards areas with "multiple family houses" i.e. buildings such as apartments, seems to decrease risk. A closer look at the graph in section (5.2) reveals that if we move su¢ciently away from the mean, which is near 0$\star$5, risk is poised to increase. Using the mean of HMF may be inappropriate or misleading, since most probability mass are situated at the extremes, as seen in section (5.2.1). This variable can be seen as a pointer as to whatever we are situated in or outside the city. In places with no apartements at all there is higher risk, as in places with nothing but apartements. A mixed landscape, i.e. HMF near 0.5, minimizes risk.

Moving in the direction of areas with industry and construction businesses seems to decrease risk, while areas with a larger proportion of public employment is far more risky. Areas with more high-pay labour and self-employment seems less risky, while areas while larger degrees of unemployment is more risky. Risk increases with income, but again, there is a great deal of correlation between high-pay labour and income.

A suggestion for a better way to look at estimates, would be to segment the areas along the model variables, characterize each resulting pro…le and look at its risk.

Note that the object of this modelling e¤ort is not necessarily to provide results with a meaningful interpretation, but simply to …nd stable relations between observed variables and risk. Though of course it is easier to have faith in the stability of a relation if it is interpretable.

3 Present Value

3.1 Consumption Level

In a continous-service setting, the level of individual product consumption each period can be thought of as a continous stochastic process. The value of the consumption, however, is often reported in discrete intervals due to accounting practice and technicalities in other management information systems.

Let the consumption in period $\star$ for consumer $\star$ be $\star_{\star\star}\star$ For each customer we have a vector of observations $\star_{\star\star} = \{\star_{\star 1}\star\ \star_{\star 2}\star\ \star\star\star\star\ \star_{\star\star}\}'$ available. We could choose to model each series individually using an ARMA or ARIMA specification, but usually the number of observations would be prohibitively small. Even if a few series of sufficient size were available and an estimated time-series model would fit each series with different sets of parameters, this information would be hard to apply to the remaining shorter series. It is unclear how shorter series should be forecasted, given knowledge of parameters in longer series. Alvarez et al. (2002) provide an interesting approach to simulating income processes.

Another issue is the stability in the environment generating the series. Any series with a large number of observations would be generated over a long period of time and therefore be suspectible to change to a larger degree; this could be effects from competitor reactions and evolution in products and product features.

The goal in this section is to forecast consumption a number of periods ahead, using a small number of observations. It is additionally assumed that we have no exogenous variables available correlated with consumption, although this could be the case in many scenarios.

When forecasting, a typical criteria used for evaluating the quality of the forecast is mean square error (m.s.e.)[5]. Forecasting $\star_{\star\star+\star}$ by $\star^{*}_{\star\star+\star}$, the m.s.e. is equal to $\star\left[\left(\star^{*}_{\star\star+\star} - \star_{\star\star+\star}\right)^2\right]\star$ An optimal forecast in the sense that it minimizes m.s.e. given the observations $\star_{\star\star}$ is $\star^{*}_{\star\star+\star} = \star\,(\star_{\star\star+\star}|\star_{\star\star})$. Postulating $\star^{*}_{\star\star+\star}$ is a linear function of $\star_{\star\star}$ ensures ease of estimation, though there may exist better non-linear estimators.

The method used here, resembles that of a panel data set, but with timevarying coefficients. Given observations at time $\{1\star\star\star\star\star_1\}$, consumption at points $\{\star_1+1\star\star\star\star\star_1+\star\}$ is modelled for consumer $1\star\star\star$ as in the system in eq. (9).

[5] Hamilton (1994) is a solid reference on time-series forecasting.

$$
\begin{aligned}
\star^{¤}_{1\star\star1+1} &= \star^{0}_{1}\begin{bmatrix}\star_{11}\\ \star_{12}\\ \vdots\\ \star_{1\star1}\end{bmatrix} = \star^{0}_{1}\star_{1\star1} \qquad (9)\\
\star^{¤}_{1\star\star1+2} &= \star^{0}_{2}\star_{1\star1}\\
&\vdots\\
\star^{¤}_{1\star\star1+\star} &= \star^{0}_{\star}\star_{1\star1}\\
\star^{¤}_{2\star\star1+\star} &= \star^{0}_{1}\star_{2\star1}\\
&\vdots\\
\star^{¤}_{\star\,\star\star1+\star} &= \star^{0}_{\star}\star_{\star\,\star1}
\end{aligned}
$$

$\star_\star$ is a vector of coe¢cients. Notice that we are imposing a restriction across agents in $\star_\star$ and include no constant, such that the only way in which forecasts can vary is through their lagged values. Coe¢cients f$\star_1\star\star\star\star\star\star_\star$g is then estimated by OLS on series containing observations f1$\star\star\star\star\star\star_1\star\star\star\star\star\star_1 + \star$g$\star$ If the level of consumption is independent of churning behavior, there would be no bias from excluding shorter series. Such a proposition should be tested of course.

When will this approach be exactly correct? If there is no trend in the series and $\star_{\star\star}$ is generated as a linear function of lagged values having the same coe¢cients for all agents the approach would work. If this is not these, which is likely, the method will at least have minimized the forecasting error over the sample and no special importance will be tied to the weights f$\star_\star$g$\star$The assumption of a low number of observations in each series to be forecasted prevents more advanced speci…cations from being utilized.

3.2 Expected Value of Consumption

The goal in this section is to arrive at an expected present value of each consumer. Worries of structural changes in the environment makes it a questionable practice to forecast or form expectations very far out in the future, so instead a …xed period is used. To form expectations of the one month-ahead or one year-ahead present value of a customer strikes a balance between allowing for comparisons between customers and to utilize a given structure in observered data instead of relying too heavily on assumptions of consumer behavior.

Remeber $\star$ is the duration of a consumer. In the section above, we implicitly calculated an estimate of $\star\,(\star_{\star+\star}|\star_\star\star\star\ \ \text{¸}\ \star)\star$ that is, the expected value of consumption in period $\star+\star$ given that a consumer survives until that period. The present value in period $\star$ of consumption in period $\star+\star$ would be calculated as

$$
\begin{aligned}
\star\star\,(\star\star\star+\star) &= (1+\star)^{\,\mathrm{i}\,\star}\star\,(\star_{\star+\star}|\star_\star\star\star\ \ \text{¸}\ \star+\star)\,\mathrm{Pr}\,(\star\ \ \text{¸}\ \star+\star) \qquad (10)\\
&= (1+\star)^{\,\mathrm{i}\,\star}\star^{¤}_{\star+\star}\Big[\star^{1}(\star)\ \mathrm{i}\ \star^{1}(\star)\Big],
\end{aligned}
$$

$\star$ being a suitably normalized discount rate and $\star^{1}$ being a survivor function as in section (2). Here any dependence on exogenous variables $\star$ is surpressed.

A few limitations should be mentioned here. First of all, the sensitivity of PV with respect to variables controlled by management would be interesting to study, since it could serve as an indicator of returns to investment in assets that typically are beyond measurement. This would be in line with questions addressed by Rust et al. (1995). However, in practice such management control variables are often not available in sufficient detail and length to be of use in a formal model. Should any be available, such as special product features, price or quantity discounts; they would be more suspectible to competitor reactions and potentially invalidating inference as compared to demographics or similar variables. See Heil & Helsen (n.d.) for a study of price wars.

Secondly, to compare investments in customers as an asset with return on investments in regular financial assets, one would need a model of volatility associated with the expected present value and even more so as mentioned in the paragraph above, a valid idea of profit sensitivity to management control variables.

Thirdly, an extra level of general discounting could be introduced to account for risks relating to systemwide events, such as the arrival of superior technologies or products and changes in the economical environment. A simple approximating solution would be to opt for a higher level of discounting $\star$.

3.3 Application

The models put forward in the previous sections are applied to time-series of customer consumption, using the same databases as described in section (2.3.1). It is choosen to estimate a one-year ahead PV using six months of consumption observations, i.e. $\star_1 = 6$ and $\star = 12$ in the notation of section (3.1).

Investigating individual series longer than 30 months ($\star$ = 954) often indicated the presence of a unit-root, though the power of unit-root tests versus near-integrated series are very low. Also, when forecasting for a short period of time, accounting for a unit-root or estimating using a near-unit root specification can be seen as a matter of approximation. Many series can be described by ARMA(1,1) representations or ARIMA(0,1,1), but convergence of estimation routines are often a problem along with the general validity of basing inference on a small number of observations.

A regression of consumption of a random month on available explanatory variables as in fig. (1, p. 12), yields a $\star^2$ not far from 0.03 although a few individual variables are significant. This indicates that the approach of forecasting consumption by lagged values is a fair one.

For all consumers having a series of observations with at least 18 points of observations ($\star$ = 1791), the first 6 observations are used to forecast the coming 12$\star$(7 $\cdot$ $\star$ $\cdot$ 18)$\star$The OLS results are seen in fig. (10). As to be expected, the root mean squares are seen to display a upward slope as forecasts are taken farther into the future. Also, the first period consumption is seldom significant and enters the forecast with a very low weight. For some customers this is probably due to a "first-month effect" in which the customer starts using the phone for the first time and/or starts consuming later in the month.

For the calculation of expected PV (EPV), all customers arriving before 6 months of the right censoring window and having at least 6 observations without being censored, is included ($\star$ = 2172). Using eq. (10), inserting the survivor function from section (2) and utilizing an annual discount rate of 8%, a one year-ahead PV is calculated as $\sum_{\star=1}^{12} \star\star(\star_1\star\star_1 + \star)$ for each customer. Additionally an actual one-year ahead PV (APV) or realized present value is calculated, also using the same discount rate.

The EPV is plotted against APV in section (5.3, p. 26). The Pearson correlation is 0$\star$6614 and the root mean square error is 1326$\star$78$\star$ Forecasting the same data, using the most recent observation as the predictor, yields a correlation of 0$\star$5999 and a root mean square error of 1795$\star$94 which is 35% higher than the more advanced model. The

t	alpha_t,1	alpha_t,2	alpha_t,3	alpha_t,4	alpha_t,5	alpha_t,6	RMSE
7	*0.0150*	*0.0164*	0.0682	0.0537	0.1045	0.6182	112.5156
8	*0.0125*	0.0641	0.0788	0.1536	0.0603	0.4796	125.4644
9	*-0.0113*	0.1435	0.0823	0.0835	0.1784	0.3452	156.1863
10	*-0.0022*	0.1477	0.0729	0.1177	0.1481	0.3435	147.1055
11	*-0.0364*	0.1323	0.1265	0.1940	0.1078	0.3070	152.8533
12	*-0.0203*	0.1064	0.1369	0.1186	0.1214	0.3226	149.0183
13	*-0.0173*	*0.0645*	0.1938	0.1694	*0.0558*	0.3564	181.3494
14	*-0.0235*	0.1736	*0.0734*	0.0961	0.1697	0.3248	185.9879
15	-0.1104	0.3359	*0.0280*	*0.0528*	0.1407	0.3286	203.2479
16	-0.0994	0.3353	*-0.0005*	0.1134	*0.0604*	0.3703	189.8678
17	*0.0217*	*0.0319*	0.2789	*0.0272*	0.0938	0.3193	154.0699
18	*0.0043*	0.1224	0.1511	0.0959	*0.0762*	0.3292	175.1303

Figure 10: ⋆ coe¢cients for predicting ⋆⋆⋆Italics indicate insigni…cance on a 5% level.

gain from using an advanced model must be said to be relatively modest, which could be blamed on the low level of customers exiting in the database studied. A larger degree of costumer exits relative to the forecasting period and better explaining variables yields larger di¤erences by the authors experience, especially when variables speci…c to the individual are introduced.

4 Conclusion

An approach to modeling individual customer expected present value has been presented. The main idea was to separate the analysis of consumption levels from the duration of customer relationships. The ideas were applied to a database of phone customers and predicted features of the data reasonably well.

More evidence in favor of such an approach is needed for this particular model, and for other types of marketing models presented in the quantitative literaturee as such, but the evidence available points to the expediency in marketing managers embracing quantitative tools more fully in their decision making process.

4.1 Limitations and Extensions

Time-varying variables in the duration model, more traditional testing of the duration model speci…cation and more advanced models of consumption levels dealing with mixture distributions for instance, would be obvious candidate extensions to the existing framework. It would also be interesting to see a similar model applied to other types of continous providers, such as the insurance-, banking- and cellular provider sectors. Building a stronger link to variables controlled by management could be another valuable direction to go, since the goal in …rst place was to improve decision making. Competitive reactions was also an issue mentioned several times as being potentially invalidating for a model of the current type.

References

Alvarez, J., Browning, M. & Ejrnæs, M. (2002), Modelling income processes with lots of heterogeneity, paper d2-3, in '10th International Conference on Panel Data, July 5-6, 2002', International Conferences on Panel Data, Berlin.

Baker, M. & Melino, A. (2000), 'Duration dependence and nonparametric heterogeneity: A monte carlo study', Journal of Econometrics 96(2), 357–393.

Berger, P. D. & Nasr, N. I. (1998), 'Customer lifetime value: Marketing models and applications', Journal of Interactive Marketing 12(1), 17–30.

Blattberg, R. & Deighton, J. (1996), 'Manage marketing by consumer equity test', Harvard Business Review pp. 136–144.

Bolton, R. N. (1998), 'A dynamic model of the duration of the customer's relationship with a continous service provider: The role of satisfaction', Marketing Science 17(1), 45–65.

Dwyer, F. R. (1997), 'Customer lifetime valuation to support marketing decision making', Journal of Direct Marketing 3(4), 8–15.

Elbers, C. & Ridder, G. (1982), 'True and spurious duration dependence: The identi…ability of the proportional hazard model', The Review of Economic Studies 49, 403–409.

Gönul, F. & Srinivasan, K. (1993), 'Consumer purchase behaviour in a frequently bought product category: Estimation issues and managerial insights from a hazard function model with heterogeneity', Journal of the American Statistical Association 88, 1219–1227.

Grant, A. W. H. & Schlesinger, L. A. (1995), 'Realize your customers' full pro…t potential', Harvard Business Review pp. 59–72.

Hamilton, J. D. (1994), Time Series Analysis, Princeton University Press.

Hand, D., Mannila, H. & Smyth, P. (2001), Principles of Data Mining, MIT.

Heckman, J. & Singer, B. (1984), 'A method for minimizing the impact of distributional assumptions in econometric models for duration data', Econometrica 52, 271–320.

Heil, O. P. & Helsen, K. (n.d.), 'Toward an understanding of price wars: Their nature and how they erupt', International Journal of Research in Marketing 18, 83–98.

Horowitz, J. L. (1999), 'Semiparametric estimation of a proportional hazard model with unobserved heterogeneity', Econometrica 67(5), 1001–1028.

Jacobs, F. A., Johnston, W. & Kotchetova, N. (2001), 'Customer pro…tability - prospective vs. retrospective approaches in a business-to-business setting', Industrial Marketing Management 30, 353–363.

Jain, D. & Singh, S. S. (2002), 'Customer lifetime value research in marketing: A review and future directions', Journal of Interactive Marketing 16(2), 34–46.

Lancaster, T. (1990), The Econometric Analysis of Transition Data, Cambridge University Press.

Li, S. (1995), 'Survival analysis', Marketing Research 7(4), 16–24.

Mittal, V. & Kamakura, W. A. (2001), 'Satisfaction, repurchase intent, and repurchase behavior: Investigating the moderating e¤ect of customer characteristics', Journal of Marketing Research 38, 131–142.

Reichheld, F. F. (1996), 'Learning from customer defections', Harvard Business Review pp. 56–69.

Reichheld, F. & Teal, T. (1996), The Loyalty E¤ect, Harvard Business School Press, Boston.

Reinartz, W. J. & Kumar, V. (2000), 'On the pro…tability of long-life customers in a noncontractual setting: An empirical investigation and implications for marketing', Journal of Marketing 64, 17–35.

Rust, R. T., Zahorik, A. J. & Keiningham, T. L. (1995), 'Return on quality (ROQ): Making serive quality …nancially accountable', Journal of Marketing pp. 58–70.

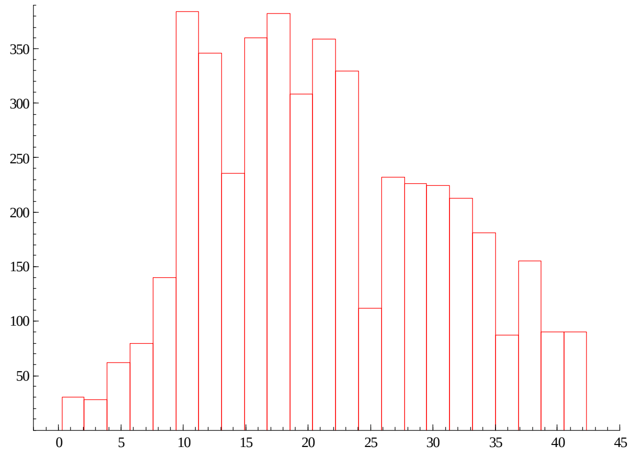
Schmittlein, D. C. & Peterson, R. A. (1994), 'Customer base analysis: An industrial purchase process application', Marketing Science 13(1), 41–67.

Vilcassim, N. J. & Jain, D. C. (1991), 'Modeling purchase-timing and brand-switching behavior incorporating explanatory variables and unobserved heterogeneity', Journal of Marketing Research 28, 29–41.

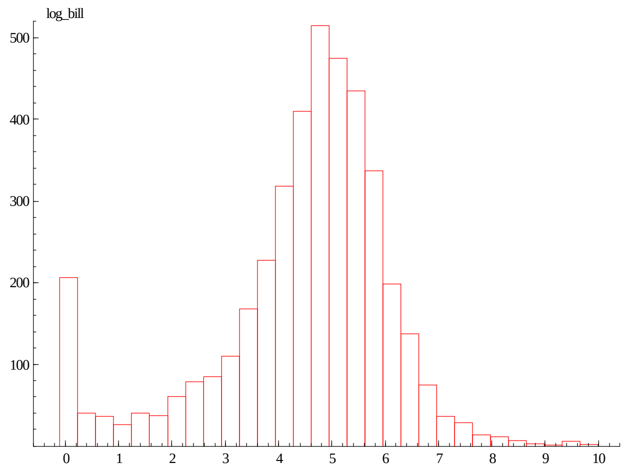
Wedel, M., Kamakura, W. A., Desarbo, W. S. & Hofstede, F. T. (1995), 'Implications for asymmetry, nonproportionality, and heterogeneity in brand switching from piece-wise exponential mixture hazard models', Journal of Marketing Research 32, 457–462.

5 Appendix

5.1 Descriptive Statistics of Data



A histogram of the number of monthly consumption records on each consumer. This does not correspond to lifetimes due to left-censoring of records.

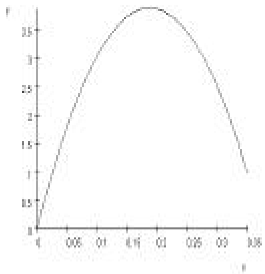


A histogram of log(⋆⋆⋆⋆⋆⋆ ⋆⋆⋆⋆⋆ + 1) across customers on January 2002.

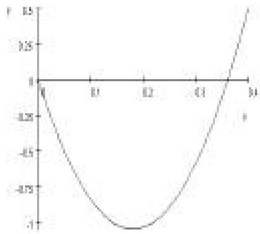
Descriptive statistics is currently placed in a seperate document at the end of this section.

5.2 Sensitivity

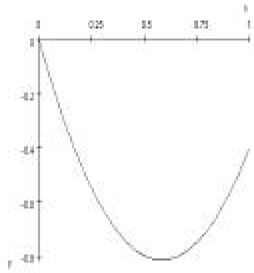
For all variables entering with a square, the function $\star_{\star 1}\star + \star_{\star 2}\star^2$ is plotted here. $\star_{\star 1}\star\star_{\star 2}$ being the estimated coe¢cient for the variable and its square.



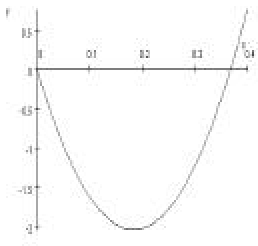
CSTART



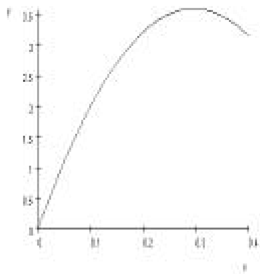
FSIK



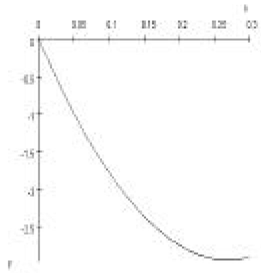
HMF



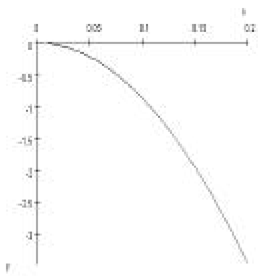
EINCO



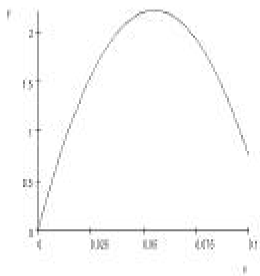
EPUBLIC



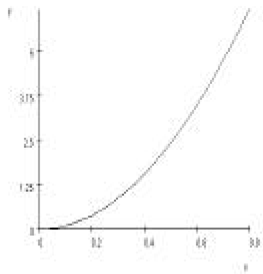
JHPL



JSE

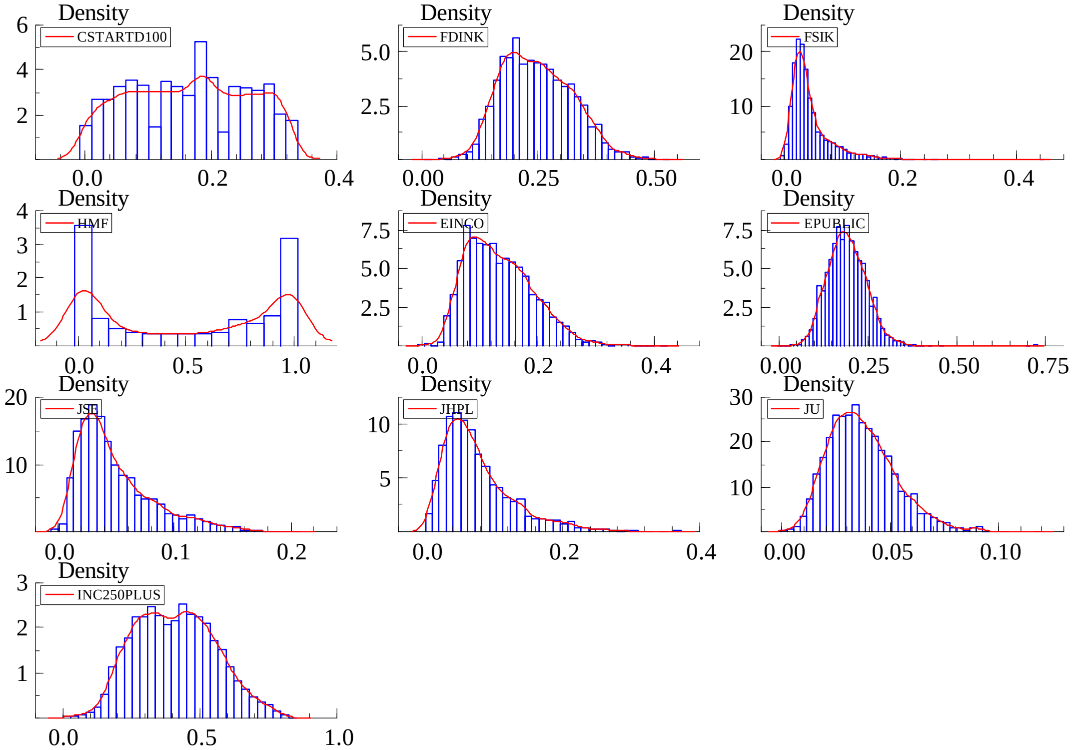


JU



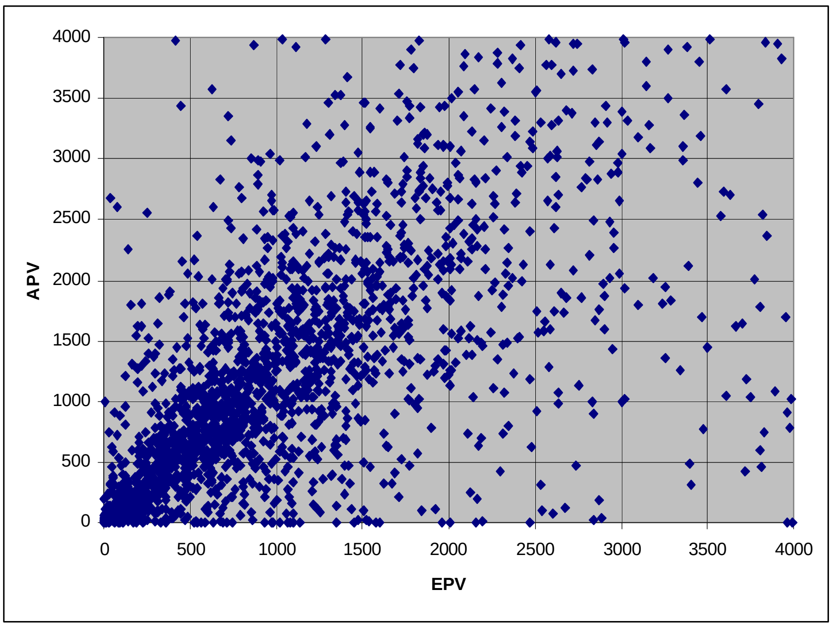
INC250

5.2.1 Histograms of variables

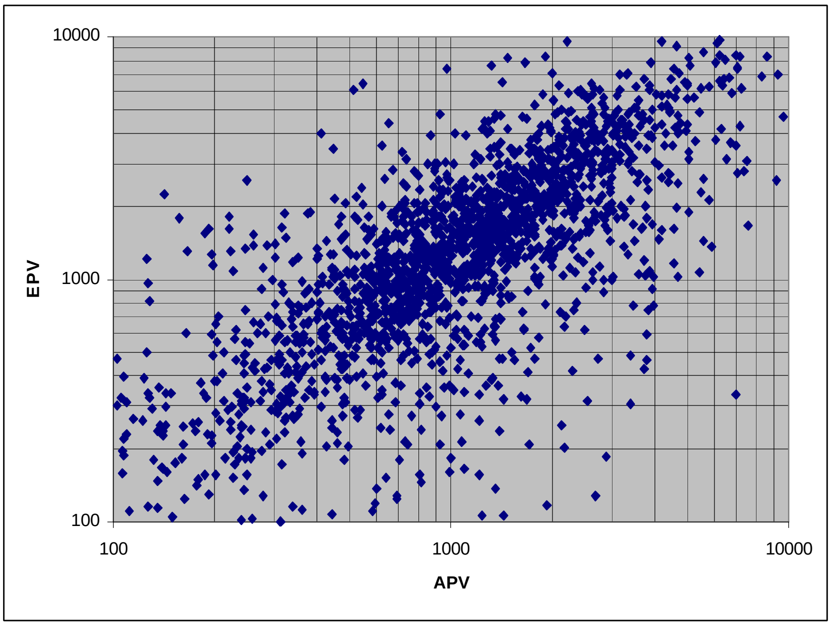


Histograms of variables entering the duration model along with kernel estimates of density

5.3 Present Value Graphs



Actual One Year Present Value plotted against Expected One Year Ahead Present Value



Same graph using log-scale

The LIFETEST Procedure

Life Table Survival Estimates

Interval [Lower,	Upper)	Number Failed	Number Censored	Effective Sample Size	Conditional Probability of Failure	Conditional Probability Standard Error	Survival	Failure
0	1	121	0	4712.0	0.0257	0.00230	1.0000	0
1	2	110	0	4591.0	0.0240	0.00226	0.9743	0.0257
2	3	74	0	4481.0	0.0165	0.00190	0.9510	0.0490
3	4	68	140	4337.0	0.0157	0.00189	0.9353	0.0647
4	5	44	159	4119.5	0.0107	0.00160	0.9206	0.0794
5	6	35	158	3917.0	0.00894	0.00150	0.9108	0.0892
6	7	37	106	3750.0	0.00987	0.00161	0.9026	0.0974
7	8	38	117	3601.5	0.0106	0.00170	0.8937	0.1063
8	9	33	119	3445.5	0.00958	0.00166	0.8843	0.1157
9	10	29	134	3286.0	0.00883	0.00163	0.8758	0.1242
10	11	35	108	3136.0	0.0112	0.00188	0.8681	0.1319
11	12	19	134	2980.0	0.00638	0.00146	0.8584	0.1416
12	13	20	118	2835.0	0.00705	0.00157	0.8529	0.1471
13	14	19	82	2715.0	0.00700	0.00160	0.8469	0.1531
14	15	28	125	2592.5	0.0108	0.00203	0.8410	0.1590
15	16	12	150	2427.0	0.00494	0.00142	0.8319	0.1681
16	17	21	132	2274.0	0.00923	0.00201	0.8278	0.1722
17	18	12	143	2115.5	0.00567	0.00163	0.8202	0.1798
18	19	8	110	1977.0	0.00405	0.00143	0.8155	0.1845
19	20	13	97	1865.5	0.00697	0.00193	0.8122	0.1878
20	21	12	97	1755.5	0.00684	0.00197	0.8065	0.1935
21	22	13	132	1629.0	0.00798	0.00220	0.8010	0.1990
22	23	14	109	1495.5	0.00936	0.00249	0.7946	0.2054
23	24	12	138	1358.0	0.00884	0.00254	0.7872	0.2128
24	25	5	102	1226.0	0.00408	0.00182	0.7802	0.2198
25	26	8	109	1115.5	0.00717	0.00253	0.7771	0.2229
26	27	8	105	1000.5	0.00800	0.00282	0.7715	0.2285
27	28	7	106	887.0	0.00789	0.00297	0.7653	0.2347
28	29	9	111	771.5	0.0117	0.00387	0.7593	0.2407
29	30	3	112	651.0	0.00461	0.00265	0.7504	0.2496
30	31	4	109	537.5	0.00744	0.00371	0.7470	0.2530
31	32	3	97	430.5	0.00697	0.00401	0.7414	0.2586
32	33	7	84	337.0	0.0208	0.00777	0.7362	0.2638
33	34	2	95	240.5	0.00832	0.00586	0.7209	0.2791
34	35	2	88	147.0	0.0136	0.00955	0.7149	0.2851
35	.	2	99	51.5	0.0388	0.0269	0.7052	0.2948

Figure 11: A standard life-table from the duration data produced by Proc LIFETEST in SAS

Evaluated at the Midpoint of the Interval

Interval [Lower,	Upper)	Median Standard Error	PDF	PDF Standard Error	Hazard	Hazard Standard Error
0	1	.	0.0257	0.00230	0.026013	0.002365
1	2	.	0.0233	0.00220	0.02425	0.002312
2	3	.	0.0157	0.00181	0.016652	0.001936
3	4	.	0.0147	0.00177	0.015803	0.001916
4	5	.	0.00983	0.00148	0.010738	0.001619
5	6	.	0.00814	0.00137	0.008976	0.001517
6	7	.	0.00891	0.00146	0.009916	0.00163
7	8	.	0.00943	0.00152	0.010607	0.001721
8	9	.	0.00847	0.00147	0.009624	0.001675
9	10	.	0.00773	0.00143	0.008864	0.001646
10	11	.	0.00969	0.00163	0.011223	0.001897
11	12	.	0.00547	0.00125	0.006396	0.001467
12	13	.	0.00602	0.00134	0.00708	0.001583
13	14	.	0.00593	0.00136	0.007023	0.001611
14	15	.	0.00908	0.00171	0.010859	0.002052
15	16	.	0.00411	0.00118	0.004957	0.001431
16	17	.	0.00764	0.00166	0.009278	0.002025
17	18	.	0.00465	0.00134	0.005689	0.001642
18	19	.	0.00330	0.00116	0.004055	0.001434
19	20	.	0.00566	0.00156	0.006993	0.001939
20	21	.	0.00551	0.00159	0.006859	0.00198
21	22	.	0.00639	0.00177	0.008012	0.002222
22	23	.	0.00744	0.00198	0.009405	0.002514
23	24	.	0.00696	0.00200	0.008876	0.002562
24	25	.	0.00318	0.00142	0.004087	0.001828
25	26	.	0.00557	0.00196	0.007197	0.002545
26	27	.	0.00617	0.00217	0.008028	0.002838
27	28	.	0.00604	0.00227	0.007923	0.002995
28	29	.	0.00886	0.00294	0.011734	0.003911
29	30	.	0.00346	0.00199	0.004619	0.002667
30	31	.	0.00556	0.00277	0.00747	0.003735
31	32	.	0.00517	0.00297	0.006993	0.004037
32	33	.	0.0153	0.00572	0.02099	0.007933
33	34	.	0.00600	0.00422	0.008351	0.005905
34	35	.	0.00973	0.00683	0.013699	0.009686
35	.	.	.	.	.	.

Summary of the Number of Censored and Uncensored Values

Total	Failed	Censored	Percent Censored
4712	887	3825	81.18

Figure 12: A standard life-table from the duration data produced by Proc LIFETEST in SAS

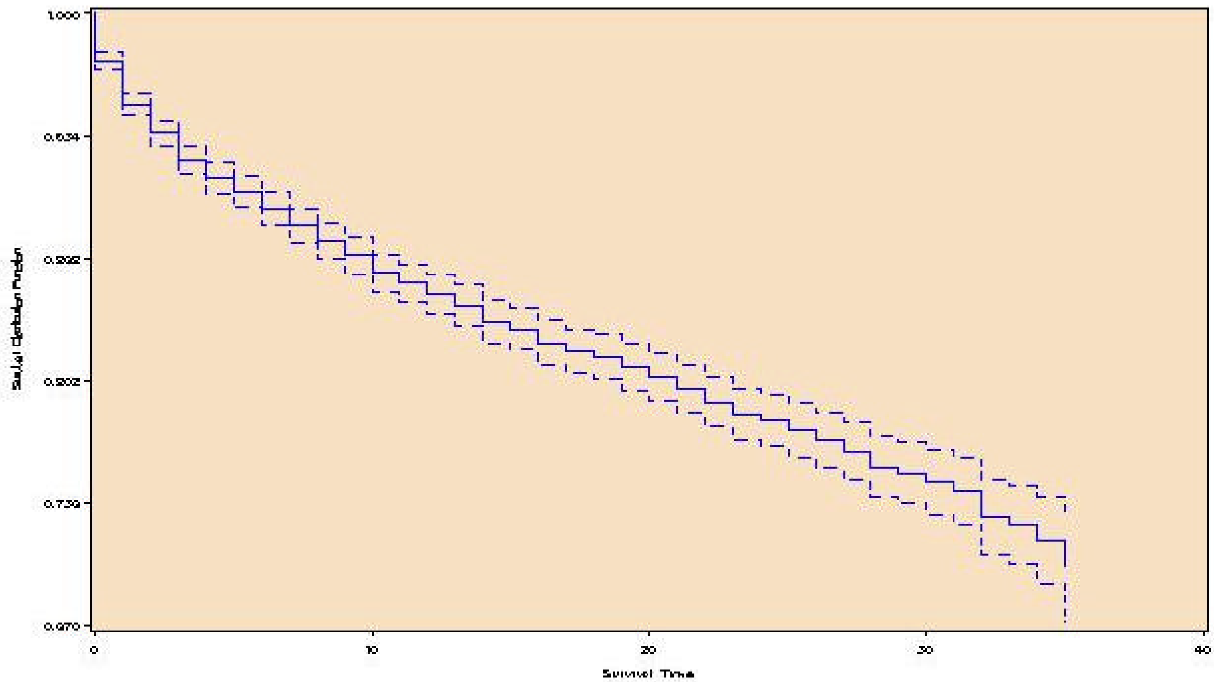


Figure 13: Empirical Survivor Function based on the Kaplan-Meier Estimator along with 95% con…dence bands

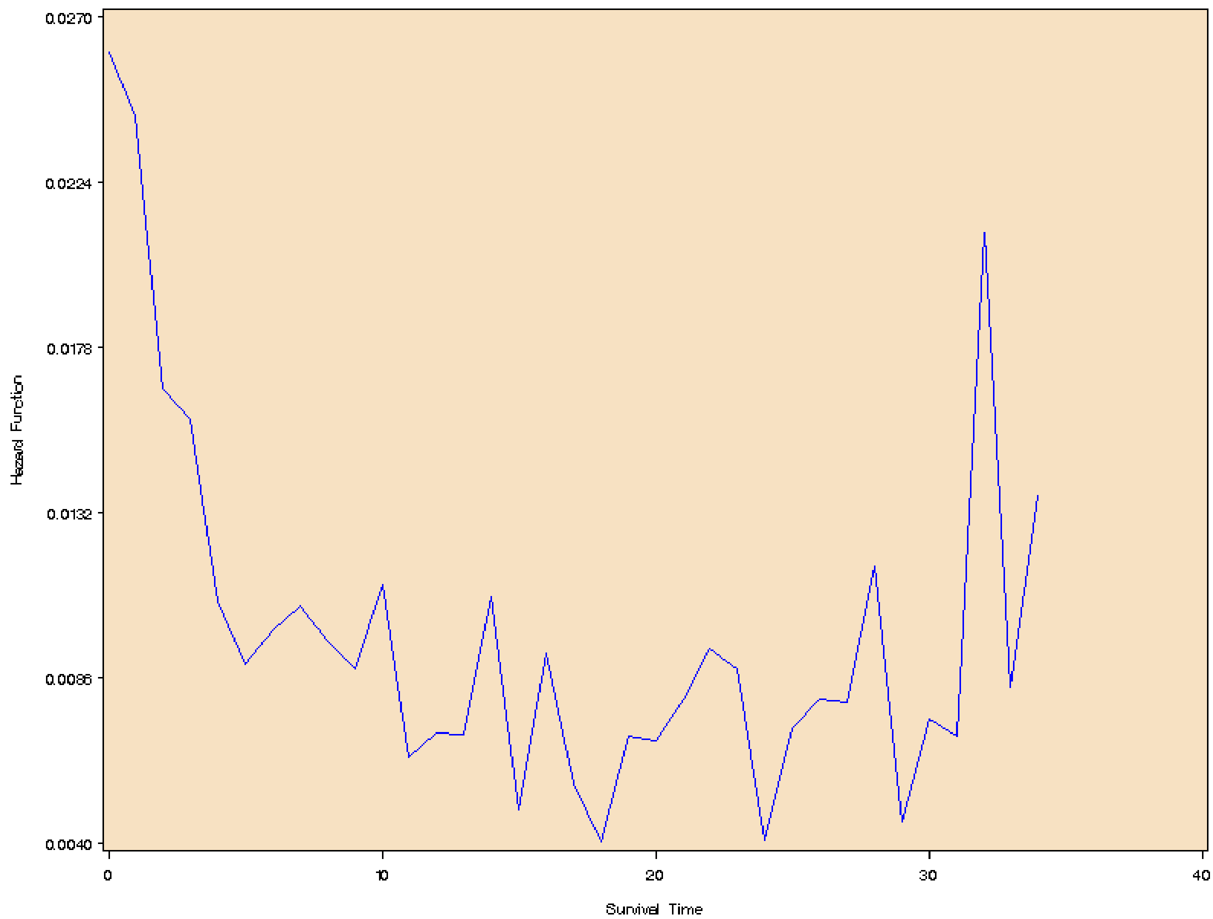


Figure 14: Empirical Hazard function

```
Descriptive statistics for 1 (1) to 4712 (1)
Means
           T          RC   CSTARTD100      AGE0_17     AGE18_32     AGE33_45
      16.127     0.81176      0.16464      0.18907      0.25357      0.18371
    AGE46_60    AGE61_99        FSINK        FDINK         FSIK         FDIK
     0.18319     0.19045      0.55002      0.24923     0.044361      0.15638
         HDT          HT          HMF           HO        EAGRI        EINCO
     0.35671     0.11004      0.50002     0.033227     0.018104      0.13538
    ESERVICE     EPUBLIC          ENA          JSE         JHPL         JMPL
     0.23332     0.19233      0.42086     0.049280     0.077227     0.081807
        JLPL          JU          JNP          JNA    DENSITY20     VERTICAL
     0.35167    0.036831      0.40141    0.0017801       8.1343      0.58531
         SEX
     0.49576
Standard deviations (using T-1)
           T          RC   CSTARTD100      AGE0_17     AGE18_32     AGE33_45
      10.028     0.39095     0.093248     0.077654      0.13726     0.038339
    AGE46_60    AGE61_99        FSINK        FDINK         FSIK         FDIK
    0.056952    0.094446      0.13774     0.074824     0.034704     0.086385
         HDT          HT          HMF           HO        EAGRI        EINCO
     0.36031     0.17554      0.40861     0.094268     0.030966     0.057789
    ESERVICE     EPUBLIC          ENA          JSE         JHPL         JMPL
    0.075760    0.056650      0.10693     0.033232     0.052961     0.037208
        JLPL          JU          JNP          JNA    DENSITY20     VERTICAL
    0.076452    0.015459      0.10458    0.0024349       13.021      0.49272
         SEX
     0.50004
```

Figure 15: Means and standard deviations of variabels entering the model

Working Papers published by the Department of Management

2005-1: Esben Koling Laustrup and Johannes Raaballe: Udbytteannonceringseffekten i Danmark.

2005-2: Carina Sponholtz: Separating the Stock Market's Reaction to Simultaneous Dividend and Earnings Announcements.

2005-3: Kyle Bagwell and Per Baltzer Overgaard: Look How Little I'm Advertising!

2005-4: Jerker Nilsson and Ole Øhlenschlæger Madsen: Cross-border Mergers Between Agricultural Co-operatives - a Governance Perspective.

2005-5: Bent Jesper Christensen and Morten Ørregaard Nielsen: The Effect of Long Memory in Volatility on Stock Market Fluctuations.

2005-6: Bent J. Christensen and Nicholas M. Kiefer: Investment in Advertising Campaigns and Search: Identification and Inference in Marketing and Dynamic Programming Models.

2005-7: Flemming Witt and Jørn Flohr Nielsen: The Difficult Empowerment in Danish Hospitals: Power to the Nurses!?

2006-1: Thomas Busch, Bent Jesper Christensen and Morten Ørregaard Nielsen: The Information Content of Treasury Bond Options Concerning Future Volatility and Price Jumps.

2006-2: Thomas Busch, Bent Jesper Christensen and Morten Ørregaard Nielsen: Forecasting Exchange Rate Volatility in the Presence of Jumps.

2006-3: Thomas Busch, Bent Jesper Christensen and Morten Ørregaard Nielsen: The Role of Implied Volatility in Forecasting Future Realized Volatility and Jumps in Foreign Exchange, Stock, and Bond Markets.

2006-4: Niels Stender: Expected Present Value of a Customer.