Segmenting and Selecting Cross-sale Prospects using Dynamic Pricing

Fredrik Thuring[1], Jens Perch Nielsen[1], Montserrat Guillén[2] and Catalina Bolancé[2]

[1]*Faculty of Actuarial Science and Insurance, Cass Business School, 106 Bunhill Row, London, EC1Y 8TZ, U.K.*

[2]*Riskcenter-IREA, Dept. Econometrics, University of Barcelona, Diagonal 690, 08034 Barcelona, Spain*

Keywords: Cross-Sales, Call Center, Dynamic Pricing, Price Elasticity of Demand, Closed form Expressions, Multivariate Bühlmann-Straub Credibility, Financial Services, Insurance Industry.

Abstract: In this paper we consider segmentation of a company's customer data base with respect to the future expected profit, emerging from not yet sold products. We consider a situation where the company is interested in contacting a subset of the customers in the data base to offer additional products. The price at which the products are offered may vary and by utilizing an estimate of the price elasticity of demand, we are able to find closed a form expression of the optimal price as well as the corresponding expected profit, for each customer prospect. We implement the methodology and test it using real data from a major Scandinavian insurance company. We underline that a closed form expression of the optimal price is seldom found in the dynamic pricing literature, suggesting that our pricing formula would be of interest to a broad audience of economists, econometricians, pricing managers and actuaries.

1 INTRODUCTION

This paper addresses the challenge of setting an optimal price and segmenting a customer data base with respect to the profit that is expected to emerge from a cross-sale attempt. Of particular interest is the financial services industry where companies have significant data bases and a traditional long relationship with each customer, once they purchase their products. Normally, the cross sale challenge is associated with using the company's specific knowledge of e.g. the probability of a cross sale, the cost of a cross sale attempt, the average discounted future profit and the uncertainty of the profit of the entire cross sale attempt for that individual. In this paper we propose an extension to this approach by introducing an estimate of the price elasticity of demand and deriving an expression for the optimal price at which the additional product should be sold, in order to maximise profit to the company.

Profit emerging from contacting subsets of customers has been considered by e.g. Bult and Wansbeek (1995), Venkatesan and Kumar (2004), Gönül and Hofstede (2006) and Kaishev et al. (2012). In Bult and Wansbeek (1995), the optimal selection is based on the customer response (sale or no sale) to a direct marketing offer of books, periodicals and music to households by a retailer in the Netherlands. Given

sale, it is assumed that the marginal, i.e. per customer, return (profit) is deterministic. Customer selection based on customer life time value is considered in Venkatesan and Kumar (2004) where the model utilizes classical regression techniques. Gönül and Hofstede (2006) takes the approach further by introducing optimisation objectives such as profit maximisation, customer retention and utility maximisation. They apply their methodology to the problem of setting optimal sales catalogue mailing strategies. In Kaishev et al. (2012) a model is presented for the customer specific profit given a cross-sale approach. The profit to the company is generated by the stochastic income, at point of sale, minus the cost of contacting a specific customer minus the stochastic cost generated by the customer as an effect of his/her actions. The model is primarily directed towards applications in the insurance and banking sector where customers can be associated with either claims or loan defaults, affecting the company's profit. None of the studies consider changing the price at which products are offered in order to boost revenue and profit. In this paper we argue that by introducing elasticity of demand and dynamic pricing into customer segmentation studies, as described above, the customer selections would improve further and increase the profit to the company. The literature on traditional cross-sale models is vast and we refer the interested reader to papers by Ka-

makura et al. (1991), Knott et al. (2002), Kamakura et al. (2003), Kamakura et al. (2004), Li et al. (2005), Kamakura (2007), and Li et al. (2010).

We consider the financial service industry primarily because of the large customer data bases normally available as well as the specific structure of the offered products. Financial services offered by banks and insurance companies, such as mortgage contracts and other types of loans, household, car and motorcycle insurance policies, and other types of personal lines insurance products, differ in several ways from other conventional retail products and services which other companies offer. There is a policy duration specified at the date of sale of a financial product and also the cost associated with a specific customer is stochastic and becomes known to the organization at some random time after the sales date. For example, the cost generated by an insurance policy is mainly determined by the claim amount which depends on the occurrence and severity of the related insured event. In a mortgage setting, a holder of a mortgage contract may default on his/her loan repayment at some random moment within the duration of the contract, which may lead to a loss for the lending bank or its insurance company, of unknown (random) size.

To be able to introduce elasticity of demand and dynamic pricing into a cross-sale model, we consider the model from Kaishev et al. (2012). This model involves three random quantities, a binary random variable, modelling the event of cross-selling, a random variable modelling the price of the offered product and another random variable, modelling the cost associated with a specific customer for the cross-sale product. We generalise the model by introducing a relation between the probability of a successful cross-sale and a price change variable $c_{ik} > -1$ for customer i and product k. Also the offered price Π_{ik} is related to c_{ik} and a tariff price Π_{0ik} as $\Pi_{ik} = (1 + c_{ik})\Pi_{0ik}$. With these introductions to the model of Kaishev et al. (2012) we are able to derive a closed form expression of the customer specific optimal price change c_{ik}^* which maximises expected profit from each customer.

2 A MODEL FOR CROSS-SALE PROFIT UNDER THE INFLUENCE OF PRICE ELASTICITY

We consider a company, offering K different products, and a set of $i = 1, \ldots, I$ customers for which the stochastic profit H_{ik}, for a product $k \in \{1, \ldots, K\}$, is

of interest. We assume that all I customers are in possession of at least one product $k' \in \{1, \ldots, k-1, k+1, \ldots, K\}$, but not product k, wherefore the profit, with respect to product k, is of stochastic nature. The company may contact the customers to cross-sell product k, by which the customer response (purchase/no purchase) is modeled by a Bernoulli random variable A_{ik}, with success probability p_{ik}. The company has influence over the offered price Π_{ik}, to a specific customer i, by adjusting the tariff price Π_{0ik} with a price change parameter $c_{ik} > -1$ as $\Pi_{ik} = (1 + c_{ik})\Pi_{0ik}$. The tariff price Π_{0ik} is a stochastic variable since it may be influenced by customer decisions such as add-ons and other customisation not known to the company prior to the cross-sale contact. We assume that the success probability p_{ik} is dependent of the price change parameter c_{ik}, a cross-sale probability p_{0ik} and a price elasticity $e_k > 0$ as $p_{ik} = \frac{p_{0ik}}{(1+c_{ik})^{e_k}}$. It should be noted that p_{ik} decreases as the price change parameter c_{ik} increases and that $p_{ik} = p_{0ik}$ for $c_{ik} = 0$ (i.e. no price change). The cross-sale product k is associated with a stochastic cost S_{ik} as well as a deterministic cost ω_{ik} for performing the cross-sale contact to the customer. We assume that a customer's response to a cross-sale contact A_{ik} is not related to the tariff price Π_{0ik} or the stochastic cost S_{ik}, i.e. A_{ik} is independent of Π_{0ik} and S_{ik}. Inspired by the model of Kaishev et al. (2012), we propose an extended model taking the price elasticity e_k into account as the following.

$$H_{ik} = \mathbf{1}_{\{A_{ik}\}}(\Pi_{ik} - S_{ik}) - \omega_{ik} \quad (1)$$

$$\mathbf{1}_{\{A_{ik}\}} = \begin{cases} 1, p_{ik} \\ 0, 1 - p_{ik} \end{cases} \qquad p_{ik} = \frac{p_{0ik}}{(1+c_{ik})^{e_k}}$$

$$\Pi_{ik} = (1 + c_{ik})\Pi_{0ik}$$

The model of Kaishev et al. (2012) is a special case of the proposed model with $c_{ik} = 0$. We denote by $\mu_{ik} = \mathbb{E}[H_{ik}]$ the expectation of the stochastic variable H_{ik} for which we have the expression

$$\mu_{ik} = \mathbb{E}[H_{ik}] =$$

$$\frac{p_{0ik}}{(1+c_{ik})^{e_k}}((1 + c_{ik})\mathbb{E}(\Pi_{0ik}) - \mathbb{E}(S_{ik})) - \omega_{ik} \quad (2)$$

where we have used that A_{ik} is independent of Π_{0ik} and S_{ik}.

3 OPTIMISATION OF THE EXPECTED PROFIT

We are interested in finding the value of the price change parameter c_{ik} for which the expected profit (2)

is maximised. We present the following proposition.

Proposition. The price change c_{ik}^* which maximises μ_{ik} in (2) is the following

$$c_{ik}^* = \frac{\mathbb{E}(S_{ik})\,e_k}{\mathbb{E}(\Pi_{0ik})\,(e_k-1)} - 1. \qquad (3)$$

Proof. We can rewrite (2) as follows

$$\mu_{ik} = p_{0ik}\mathbb{E}(\Pi_{0ik})\,(1+c_{ik})^{(1-e_k)} -$$

$$p_{0ik}\mathbb{E}(S_{ik})\,(1+c_{ik})^{-e_k} - \omega_{ik}.$$

Differentiating with respect to c_{ik} gives

$$\frac{d\mu_{ik}}{dc_{ik}} = p_{0ik}\mathbb{E}(\Pi_{0ik})\,(1-e_k)\,(1+c_{ik})^{-e_k} -$$

$$p_{0ik}\mathbb{E}(S_{ik})\,(-e_k)\,(1+c_{ik})^{-e_k-1} =$$

$$= p_{0ik}\left(\frac{\mathbb{E}(\Pi_{0ik})(1-e_k)}{(1+c_{ik})^{e_k}} + \frac{\mathbb{E}(S_{ik})e_k}{(1+c_{ik})^{e_k+1}}\right)$$

and by equating to zero (noting that $p_{0ik} > 0$) we get

$$\frac{\mathbb{E}(\Pi_{0ik})(1-e_k)}{(1+c_{ik})^{e_k}} + \frac{\mathbb{E}(S_{ik})e_k}{(1+c_{ik})^{e_k+1}} = 0.$$

The above can be rewritten as

$$(1+c_{ik})^{-e_k}\left(\mathbb{E}(\Pi_{0ik})\,(1-e_k)\,(1+c_{ik}) + \mathbb{E}(S_{ik})\,e_k\right) = 0.$$

Noting that $c_{ik} > -1$ we get

$$\mathbb{E}(\Pi_{0ik})\,(1-e_k)\,(1+c_{ik}) + \mathbb{E}(S_{ik})\,e_k = 0$$

which is solved by

$$c_{ik} = \frac{\mathbb{E}(S_{ik})e_k}{\mathbb{E}(\Pi_{0ik})(e_k-1)} - 1.$$

$\square$

From (3) it should be noted that $e_k > 1$ for the optimal price change c_{ik}^* to be feasible. In economic terms, this means that the demand for the good need to be elastic or relatively elastic. The maximum expected profit $\mu_{ik}(c_{ik}^*)$ for customer i is received by insertion of (3) into (2) and after some algebraic manipulations it yields

$$\mu_{ik}(c_{ik}^*) = p_{0ik}\left(\frac{\mathbb{E}(\Pi_{0ik})}{e_k}\right)^{e_k}\left(\frac{e_k-1}{\mathbb{E}(S_{ik})}\right)^{e_k-1} - \omega_{ik}. \qquad (4)$$

With the expression for the optimal price change (3) and the corresponding profit (4) the company is able to assess whether or not a certain customer should be contacted and offered the product k, at the price $\Pi_{ik} = (1+c_{ik}^*)\Pi_{0ik}$. The possibility of $\mu_{ik}(c_{ik}^*) < 0$ should be noted, i.e. even at the optimal price change some customers are expected to contribute negatively to the profit of the company and should therefore not be contacted.

4 VARIABLE SPECIFICATION AND PARAMETER ESTIMATION

For specification of the variables Π_{0ik}, S_{ik} and ω_{0ik} in the model (2) we follow the approach of Kaishev et al. (2012) and consider the insurance business where problems regarding which customers to cross-sell to, at which price, are common. For the expected value of the stochastic variable describing the tariff price Π_{0ik} we use the mean value from a collateral data set of the company as $\mathbb{E}(\Pi_{0ik}) = \bar{\pi}_{0k}$.

For the variable S_{ik} we follow standard actuarial convention and identify it as the aggregate claims cost $S_{ik} = \sum_{t=1}^{N_{ik}} X_{ikt}$, where N_{ik} describes the stochastic number of insurance claims (for customer i) and X_{ik} describes the corresponding monetary size of each of these claims. Additionally, we let each customer i be associated with a latent risk variable Θ_{ik} modelling unobserved characteristics related to the risk of that customer, for which θ_{ik} is its realisation. The expectation of N_{ik}, conditioned on the latent random risk variable Θ_{ik}, is $\mathbb{E}[N_{ik} \mid \Theta_{ik} = \theta_{ik}] = \bar{\lambda}_k\theta_{ik}$ and X_{ik} has expectation $\mathbb{E}[X_{ik}] = \bar{x}_k$, $\bar{\lambda}_k$ is sometimes called the a priori expected number of claims. Both $\bar{\lambda}_k$ and $\bar{x}_k$ denote the mean value for which collateral data from the company is needed in the estimation. By assuming independence between N_{ik} and X_{ik} the expectation of S_{ik} (conditioned on Θ_{ik}) becomes

$$\mathbb{E}[S_{ik} \mid \Theta_{ik} = \theta_{ik}] = \mathbb{E}[N_{ik} \mid \Theta_{ik} = \theta_{ik}]\,\mathbb{E}[X_{ik}] = \bar{\lambda}_k\theta_{ik}\bar{x}_k. \quad (5)$$

We use multivariate credibility theory to estimate θ_{ik} and specifically the estimator presented in Thuring (2012). This estimator produces an estimate of θ_{ik}, for the cross-sale product k for which no customer specific data is available, using available information with respect to another product k'. We assume that customer specific data is available for the a priori expected number of claims $\lambda_{ik'}$ and observed number of claims $n_{ik'}$, with $n_{ik'}$ being a realisation of $N_{ik'}$, for all I customers.

$$\theta_{ik} = \theta_{0k} + \frac{\lambda_{ik'}\tau_{kk'}^2}{\lambda_{ik'}\tau_{k'k'}^2 + \sigma_{k'}^2}\left(\frac{n_{ik'}}{\lambda_{ik'}} - \theta_{0k'}\right) \qquad (6)$$

In order to be able to evaluate (6), estimates of θ_{0k}, $\tau_{kk'}^2$, $\tau_{k'k'}^2$, $\sigma_{k'}^2$ and $\theta_{0k'}$ need to be obtained from collateral data of the company, see e.g. Bühlmann & Gisler (2005) p. 185-186. For further details, see also Englund et al. (2009) and Thuring et al. (2012).

For the cost of a cross-sale contact ω_{ik} we consider the specific case of each cross-sale contact being equally costly for the company and therefore use

as estimator $\omega_{ik} = \omega_k$, where ω_k is estimated with appropriate data of the company.

The cross-sale probability p_{0ik} is estimated using a regression model $p_{0ik} = f_{p,k}\left(\mathcal{Y}_{p,ik}\right)$, where $f_{p,k}$ is an appropriate regression function, estimated based on collateral data from the insurance company and $\mathcal{Y}_{p,ik}$ is a set of customer specific antecedents of the contacted customer, see related studies of Knott et al. (2002) and Li et al. (2005) where suggestions for $f_{p,k}$ are given.

For the price elasticity e_k we consider the definition of price elasticity as the logarithm of the percentage change in demand over the logarithm of the percentage change in price. Furthermore we assume that we have collateral data available where the success of a large number of cross-sale contacts has been registered and that a random number of the contacts have been associated with a price change c_k. The price change have been introduced to measure its effect on sales success of the cross-sale contacts. To be able to estimate the price elasticity we denote by $\bar{p}_0$ the estimated probability of cross-sale success, for contacts associated with no price change. Furthermore, we denote by $\bar{p}_{c_k}$ the estimated probability of cross-sale success, for contacts associated with a price change of c_k. According to the definition of price elasticity we have

$$e_k = \frac{\ln\left(\frac{\bar{p}_0}{\bar{p}_{c_k}}\right)}{\ln\left(1 + c_k\right)}. \tag{7}$$

5 REAL DATA STUDY

We have available the data base of a major Scandinavian insurance company within non-life personal lines insurance (home, building, car, boat, etc.). To be able to estimate the necessary parameters for (3) and (4) we need data from different sources within company's data base. The average values of the tariff price $\bar{\pi}_{0k}$ is easily calculated from the data of all existing customer in possession of product k. Also the average a priori expected number of claims $\bar{\lambda}_k$ and the average claim severity $\bar{x}_k$ is available by considering all claims historically reported to the company in relation to the risk exposure (total number of years for all customers).

In order to be able to evaluate (6), estimates of θ_{0k}, $\tau^2_{kk'}$, $\tau^2_{k'k'}$, $\sigma^2_{k'}$ and $\theta_{0k'}$ need to be obtained from a collateral data set consisting of present customers in possession of both product k and k'. The estimates are found in Table 1. These estimates are needed to estimate the customers specific risk profile $\hat{\theta}_{i2}$.

The constant cost of a cross-sale contact ω_k is estimated by analysing data on staffing cost and of-

Table 1: Estimates of the model parameters for estimating the customer specific risk profile $\hat{\theta}_{i2}$.

l	$\hat{\sigma}^2_l$	$\hat{\tau}^2_{l1}$	$\hat{\tau}^2_{l2}$	$\hat{\theta}_{0l}$
1	1.755	0.081	0.130	1.12
2	1.349	0.130	0.211	0.91

fice rent, collected at the call center of the company. The regression function $f_{p,k}\left(\mathcal{Y}_{p,ik}\right)$, for the cross-sale probability p_{0ik}, is based on data from past cross-sale attempts, details on variable selection and parameter estimates cannot be disclosed due to confidentiality agreements between the authors and the company. The estimate of the price elasticity e_k is received from a unique data set of insurance quotes subject to a random price change in order to measure its effect on the realisation of random variable A_{ik} (sale/ no sale).

Our validation data set, for implementation of the formulas (3) and (4), is a data set consisting of $I = 4463$ insurance customers who were targeted for a cross-sale campaign (in the following referred to as the campaign data set). These specific customers were approached, via telephone, and were, at the time of contact, in possession of a household insurance coverage (in the following referred to as product $k' = 1$). The aim of the telephone call was to cross-sell a car insurance coverage (in the following referred to as product $k = 2$). Not every customer accepted the cross-sale offer, of the 4463 contacted household policyholders, 177 purchased the car insurance coverage, i.e. $\sum_{i=1}^{I} \mathbf{l}_{\{A_{i2}\}} = 177$. Our aim with this data is to examine the expected profit from each of the customers, with respect to product $k = 2$, and how the profit is related to different values of the price change c_{i2}. Of particular interest is the optimal price change c^*_{i2} and its corresponding profit $\mu_{i2}\left(c^*_{i2}\right)$, according to (3) and (4), in comparison to $c_{i2} = 0$ which corresponds to the model of Kaishev et al. (2012). Table 2 presents summary statistics of all the necessary parameters from both the collateral data sets and the campaign data set.

We apply (3) and (4) to all the records of the campaign data set and present the results through histograms, see figure 1. In figure 1 we have also included a graph showing cumulative sums of the expected profit $\sum_{i=1}^{l} \hat{\mu}_{i2}$ (with no price change, $c_{ik} = 0$) and compared this to cumulative sums $\sum_{i=1}^{l} \hat{\mu}^*_{i2}$ (with optimal price change, c^*_{ik}), for $l = 1, \ldots, 4463$. The campaign data set is sorted by decreasing profit, $\hat{\mu}_{i2}$ and $\hat{\mu}^*_{i2}$ respectively, prior to cumulative summation. By this approach the curves show the total expected profit from an increasing subset of customers and the optimal number of customers to contact is where the curve has its maximum. For the expected profit as-

Table 2: Descriptive statistics of the campaign data set, note that $k' = 1$ represents household insurance coverage and that $k = 2$ represents car insurance coverage.

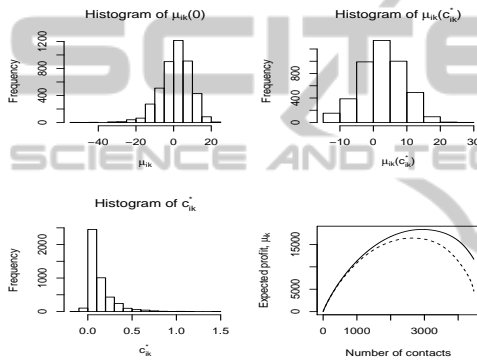|	Constant	Min	Max	Mean
$\hat{\lambda}_{i1}$	-	0.0083	3.92	0.64
$\hat{\theta}_{01}$	1.12	-	-	-
$\mathbf{l}_{\{A_{i2}\}}$	-	0	1	0.040
$\hat{p}_{i2}$	-	0.0040	0.13	0.069
$\hat{\theta}_{i2}$	-	0.71	2.05	0.95
$\hat{\theta}_{02}$	0.91	-	-	-
$\hat{\lambda}_2$	0.375	-	-	-
$\hat{m}_2$ (\$)	2,025	-	-	-
$\hat{\pi}_2$ (\$)	949	-	-	-
$\hat{\omega}_2$ (\$)	15	-	-	-
$\hat{e}_2$	3	-	-	-

Figure 1: Top left is a histogram over expected profits associated with no price change $\mu_{ik}(0)$, top right is a histogram over expected optimal profits $\mu_{ik}(c_{ik}^*)$, bottom left is a histogram of the optimal price changes c_{ik}^* and the bottom right graph shows cumulative sorted sums $\sum_{i=1}^{l} \hat{\mu}_{i2}$ (dotted) and cumulative sorted sums $\sum_{i=1}^{l} \hat{\mu}_{i2}^*$ (solid).

sociated with $c_{ik} = 0$ this is 2647, while applying the optimal price change to each customer c_{ik}^*, results in 2934 customers having a positive expected profit. The resulting total expected profit from these 2647 and 2934 customers is \$16,416 and \$18,338, an increase of almost 12%. The average value of price changes is calculated to 0.13, i.e. the optimal prices offered to the cross-sale customers is 13% above the tariff price.

6 CONCLUSIONS

In this paper, we have derived a closed form expression for the customer specific optimal price change and the corresponding optimal profit. The staring point is the model of Kaishev et al. (2012) which has been generalised towards a situation where the prob-

ability of success, of a cross-sale attempt, is dependent of the price at which the product is offered. The new approach is tested on data from a major Scandinavia insurance company where the target group of customers (associated with positive expected profit) increase as optimal price changes are applied. The resulting formula for optimal price change is not only of interest to researcher and practitioners of cross-selling but could be applied in a far broader context of pricing in general.

REFERENCES

Bult, J. R. and Wansbeek, T., 1995. Optimal selection for direct mail, *Marketing Science*, 14(4), 378-394.

Bühlmann, H. and Gisler, A., 2005. *A Course in Credibility Theory and its Applications* (Berlin, Germany: Springer Verlag).

Englund, M., Gustafsson, J., Nielsen, J. P. and Thuring, F., 2009. Multidimensional Credibility with Time Effects - An Application to Commercial Business Lines, *The Journal of Risk and Insurance*, 76(2), 443-453.

Gönül, F. F. and Hofstede, F.T., 2006. How to compute optimal catalog mailing decisions, *Marketing Science*, 25(1), 65-74.

Kaishev, V. K., Nielsen, J. P., Thuring, F. 2012. Optimal customer selection for cross-selling of financial services products, *Expert Systems with Applications*, forthcoming.

Kamakura, W.A., 2007. Cross-Selling: Offering the Right Product to the Right Customer at the Right Time in Lehzan Aksoy, Timothy Keiningham and David Bejou (Eds.) Profit Maximization through Customer Relationship Marketing , Haworth Press.

Kamakura, W. A., Kossar, B. and Wedel, M., 2004. "Identifying innovators for the cross-selling of new products," *Management Science*, 50, 1120-1133

Kamakura, W. A., Wedel, M., de Rosa, F. and Mazzon, J. A., 2003. Cross-selling through database marketing: a mixed data factor analyzer for data augmentation and prediction, *International Journal of Research in Marketing*, 20(1), 45-65.

Kamakura, W. A., Ramaswami, S. and Srivastava R., 1991. Applying latent trait analysis in the evaluation of prospects for cross-selling of financial services, *International Journal of Research in Marketing*, 8, 329-349.

Knott A., Hayes, A. and Neslin, S. A., 2002. Next-product-to-buy models for cross-selling applications, *Journal of Interactive Marketing*, 16(3), 59-75.

Li, S., Sun, B. and Montgomery, A. L. 2010. Cross-Selling the Right Product to the Right Customer at the Right Time, *Journal of Marketing Research*, 48(4), 683-700.

Li, S., Sun, B. and Wilcox, R. T., 2005. Cross-selling sequentially ordered products: an application to consumer banking services, *Journal of Marketing Research*, 42, 233-239.

Thuring, F., 2012. A credibility method for profitable cross-selling of insurance products, *Annals of Actuarial Science*, 6(1).

Venkatesan, R. and Kumar, V., 2004. A customer lifetime value framework for customer selection and resources allocation strategy, *Journal of Marketing*, 68, 106-125.