Submitted to *Expert Systems with Applications*
manuscript 1

Selecting prospects for cross-selling financial products using multivariate credibility

Fredrik Thuring, Jens Perch Nielsen, Montserrat Guillén, Catalina Bolancé *
Cass Business School, City University, London, United Kingdom, Fredrik.Thuring.1@city.ac.uk (corresponding author)
Cass Business School, City University, London, United Kingdom, jens.nielsen.1@city.ac.uk
Department of econometrics, RISC-IREA, University of Barcelona, Spain, mguillen@ub.edu
Department of econometrics, RISC-IREA, University of Barcelona, Spain, bolance@ub.edu

Insurance policies or credit instruments are financial products that involve a long-term relationship between the customer and the company. For many companies a possible way to expand its business is to sell more products to preferred customers in its portfolio. Data on the customers' past behaviour is stored in the company's data base and these data can be used to assess whether or not more products should be offered to a specific customer. In particular, data on past claiming history, for insurance products, or past information on defaulting, for banking products, can be useful for determining how the client is expected to behave in other financial products. This study implements a method for using historical information of each individual customer, and the portfolio as a whole, to select a target group of customer to whom it would be interesting to offer more products. This research can help to improve marketing to existing customers and to earn higher profits for the company.

Key words: Cross-sale selections; Financial services industry; Multivariate credibility.

1. Introduction

Cross-selling means approaching the present customers of a company and encouraging them to increase their engagement with the company by purchasing one or many additional products. It is one of the main tools for managers to strengthen the customer relationship (Kamakura et al., 1991). In the financial sector, customers have a long-term relationship with their service provider and data on their characteristics, transactions, demographics and behaviour is stored in the company's data base, see Seng and Chen (2010) and Liao et al. (2011). This information can be used to select preferred customers and cross-sell them products they do not yet possess.

* The authors acknowledge support by the Spanish Ministry of Science ECO2010-21787-C01-03.

We present a method that describes how to model past behaviour in one or many financial products in order to estimate a customer specific risk profile for a certain product not yet owned by him or her, see e.g. Bae and Kim (2010) for other examples of modelling customer behaviour. Thereafter, the risk profile estimate is used to select which customers, from the company's portfolio, to approach and attempt to make a cross-sale. Knowledge about past customer behaviour in one or many financial products is known to explain the performance in other related products, of the same customer (see e.g. Englund et al., 2009). However, there has been no attempt to implement this as a cross-selling marketing instrument. Our objective is to show a case study of this method and to explain how such a system can be implemented in practice. The general procedure is described in Figure 1 where we see, from data analysis to customer selection, how a financial company can select a target group of customers in order to cross-sell them a certain product. Initially, data

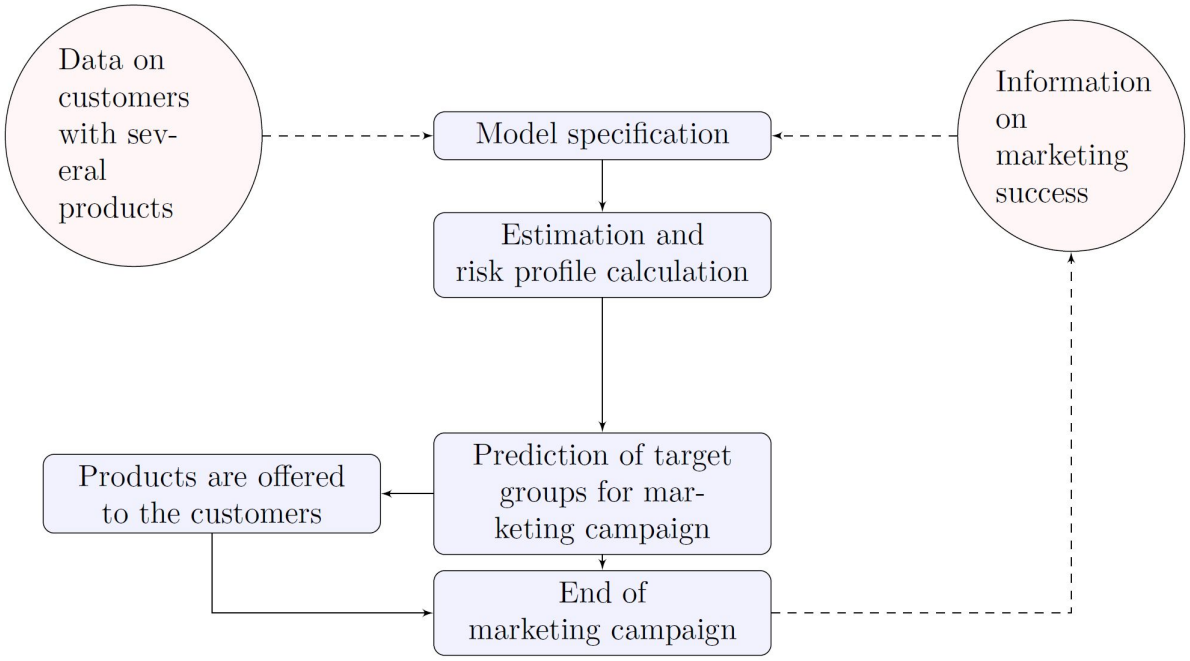

Figure 1 Workflow for cross-selling in the financial sector

on customers with several products are analysed and a model is specified. The model predicts an individual score for each customer with respect to a financial product he/she does not own. In our paper the score is called a risk profile because it predicts the customer behaviour, for a particular

not owned product, given the individual information about the behaviour in other owned and related products. The company can then select a target group for a marketing campaign based on the predicted risk profiles, offer a specific product to this group and thereafter the success of the cross-sale campaign can be analysed to refine the model, see also Malthouse (2010).

In an insurance company the method presented in Section 3 can be used to detect customers likely to report few insurance claims, with respect to a not yet owned insurance coverage, and cross-sell them that specific coverage at possibly a discounted premium level. Insurance companies normally have models for the expected (yearly) claim frequency, given certain characteristics of the customer and the insured object, which have been estimated based on collateral data on historic claims reported by past and present customers of the company, see Denuit et al. (2007) for details on claims frequency models. When predicting the claim frequency of a specific customer, such models do not usually take into consideration the individual claims experience of that customer, but predict the claims frequency based on a risk categorization which is a function of the characteristics with respect to the customer and the object. Since there can be customers with more or less risk adverse (individual) behaviour, there are cases for which the claim frequency model over-estimates or under-estimates the claim occurrence. If, for a certain customer, the claim frequency model over-estimates the claim occurrence the customer is reporting "fewer claims than expected", on the other hand if the claim frequency model under-estimates the claim occurrence the customer reports "more claims than expected". By knowing about the individual behaviour (more or less claims than expected) in one or many of a customer's existing coverages, a similar behaviour can be expected for another coverage, not yet owned by the customer. For instance, someone who has a motor insurance policy coverage and who claimed less than expected is probably also going to be claiming less than expected in other coverages such as house insurance. This phenomenon can be explained by the attitude that individuals have towards risk (see Slovic et al., 2004 and Harrison et al., 2007). People that are very much risk adverse drive carefully and also maintain their houses and belongings in good conditions. As a result, there is a correlation between the number of claims that they report to their insurance company in two different insurance coverages. On the other hand,

some individuals have a completely different attitude towards risk, they are more aggressive when driving and are therefore expected to be careless about their properties too. So, when cross-selling house coverage to individuals who already have motor insurance with the company, it would be wise to take into consideration the observed number of car claims (for the specific customer) in comparison to the expected number of car claims. Note that the reverse is also true, the number of past house insurance claims can help to predict future car insurance claims.

A similar argument can be made for the banking sector. Customers that have not defaulted in the past on their loans and/or have a flawless credit card payment history, are the ones also expected to be profitable for other credit instruments. As for insurance companies, banks and other credit institutions have models and assessments for the likelihood of a customer not being able to repay credit card loans or mortgages and the concept of "fewer incident than expected" and "more incidents than expected" is applicable here as well. In the proceeding, we will refer to the all events leading to a customer induced loss for a financial company (insurance claims, loan defaults, non-repayment of credit card loans, etc.) as incidents.

The rest of the paper is organized as follows. In Section 2 we present the background of cross-selling and marketing of financial products. We show that selecting customers, based on behaviour in other related products, is an issue that has not been explicitly discussed in existing works. Section 2 also provides a short overview of credibility theory, which we use to estimate the individual risk profile. In Section 3 we briefly show how the risk profile can be obtained, in the cross-selling case, and Section 4 presents a real case study on customers from the database of a Swedish insurance company. The results illustrate how the methods can be used in practice, they show that implementation is straightforward and can lead to substantial profit improvement compared to a strategy, for cross-selling, where customers are selected randomly. Finally, Section 5 concludes.

2. Background

We first review recent cross-sale studies and thereafter the concept of credibility theory, which is the technique used for evaluating cross-sell prospects in this paper.

2.1. Cross sale models

Understanding and using cross-selling techniques is crucially important for a company because as the customers acquire more products from the same provider, the switching cost, associated with leaving for a competitor, increases (Kamakura et al., 2003). Therefore, cross-selling is considered a strong driver for lowering the customer churn, increasing the number of loyal customers and obtaining higher customer lifetime value (Akura and Srinivasan, 2005). In addition to this, considering product features allows significant contributions for managers striving for valuable and strong relationship with their current customer base (Larivière and Van den Poel, 2004). Another important, but not as obvious, benefit from cross-selling is that companies can learn more about the customers' preferences and buying behaviour (Kamakura et al., 2003) and cumulate various types of data to their data warehouse e.g. demographic information (Ahn et al., 2011). Such information can be used as explanatory variables to predict certain behaviours of the customers such as customer retention and profitability outcomes (Larivière and Van den Pol, 2005).

Other studies focus on modeling the probability of a successful cross-sale attempt. In an early study by Kamakura et al. (1991) probabilistic predictions are made on whether or not a customer would purchase a particular product/service based on their ownership of other products/services. In Knott et al. (2002), different models are applied to predict which product a customer is expected to buy next and the approach is further developed in Li et al. (2005), where also the appropriate time to approach a specific customer is studied.

Even though many studies have been made on cross-selling as a method for increasing a company's revenue, only few discuss potential heterogeneity in the profitability of the cross-sale prospects. As pointed out in Larivière and Van den Pol (2005), financial products are not the typical grocery products such as milk, coffee or cookies, but products that are bought and owned for a specific period in time. In addition to this, financial products are associated with uncertain costs which are determined at some (uncertain) time after the product is sold. Therefore it is not guaranteed that a successful cross-sale attempt, to a specific customer, will generate profit to the company. Instead if the cross-sold product generates claims (for an insurance company) or a loan

default (for a lending bank) the financial product actually generates a loss to the company, in most cases far greater than the income at the point of sale (insurance premium or interest payment). Englund et al. (2008) suggest that their multivariate credibility estimator could be used for evaluating cross-sale prospects by taking into account only information from the other insurance products of these specific prospects. The resulting estimate of the risk profile can be used to identify the expected profitable customers (having less than expected number of claims or loan defaults) and hence increase the company's total profit from cross-selling.

2.2. Credibility theory

In actuarial science, credibility theory is a technique widely used to price different insurance coverage such as health, life and property insurance (Frees, 2003). In general, the idea is to weight data, associated with an individual policyholder, with data from a collective of policyholders using a credibility weight α,

$$\text{individual estimate} = \alpha \times \text{individual data} \; + \; (1-\alpha) \times \text{collective data} \; .$$

A historical review of credibility theory starts with the papers by Mowbray (1914) and Whitney (1918) in which the credibility weight is determined ad hoc, focusing on practical applications, and not yet founded on concrete mathematical grounds. In Bühlmann (1967) (and in the more general Bühlmann and Straub, 1970, where the Bühlmann-Straub credibility model is presented) this was changed by viewing the determination of α as an optimisation problem where only the first and second order moments of the data is needed for the optimal estimator (Norberg, 2004). The generalisation of the credibility estimator to higher dimensions was introduced in Jewel (1973) and later in a multivariate hierarchical framework by Venter (1985). In Jewel (1989) the specific problem of multivariate predictions of first and second order are investigated, while a comprehensive reference to (multivariate) credibility in general is Bühlmann and Gisler (2005). A specific interpretation of the Bühlmann-Straub credibility model is found in Englund et al. (2008) and Englund et al. (2009) where the dimensions, in the multidimensional credibility model, are interpreted as different insurance coverages, between which the claim occurrence can be more or less correlated.

3. Methodology

We use multivariate credibility theory to estimate a customer specific latent risk profile and thereafter evaluate if a specific additional product, of a specific customer, is expected to contribute positively to the profit of the company, if that product is cross-sold to the customer. The profit is measured as the customer specific deviation between the a priori expected number of incidents (insurance claims, loan defaults, etc) and the corresponding observed number. In the next paragraphs we present the methodology briefly and give reference to previous related work on the model and estimation technique.

3.1. Estimation of the risk profile

We use the standard multivariate Bühlmann-Straub credibility model, see e.g. Bühlmann and Gisler (2005, p. 178) and Englund et al. (2008). Individuals $i = 1, \ldots, I$ are customers to a financial company and have been so during time periods $j = 1, \ldots, J_i$. During these time periods, every customer has had $l = 1, \ldots, K$ different financial products. We alter between k, k' and l as index for financial products in general. For each customer i in time period j and product l, we have the a priori expected number of incidents $\lambda_{ijl} = e_{ijl} g_l(\mathcal{Y}_{ijl})$, which depends on the risk exposure $0 \leq e_{ijl} \leq 1$, a regression function g_l and of a set of explanatory variables $\mathcal{Y}_{ijl}$ characterising the customer and the insured object. This can be viewed as a categorisation of the customer and the insured object into one of a large (but finite) number of risk categories. The function g_l is common for all customers i and time periods j and can be estimated, using a generalised linear model, based on collateral data of the company. We assume that e_{ijl} can take values between $[0, 1]$, where $e_{ijl} = 0$ means that the l-th product is not active (not owned) for customer i in time period j and correspondingly, $e_{ijl} = 1$ means that the product l of customer i is active (owned) during the entire time period j. We assume a Poisson distribution for the random variable N_{ijl}, describing the actual number of incidents for customer i in time periods j and product l. The observation of N_{ijl} is n_{ijl}.

Consider another random variable Θ_{il} which represents hidden characteristics such as risk aversion, attitude, etc. that are not captured by the explanatory variables. Θ_{il} random variables are

often called the random effects. Let the pairs $(N_{1jl}, \Theta_{1l}), (N_{2jl}, \Theta_{2l}), \ldots, (N_{Ijl}, \Theta_{Il})$ be independent. We assume $E[N_{ijl}] = \lambda_{ijl}\theta_{0l}$ where $E[\Theta_{il}] = \theta_{0l}$ and $Cov[\Theta_{ik}, \Theta_{ik'}] = \tau^2_{kk'}$ for $k = 1, \ldots, K$ and $k' = 1, \ldots, K$. Further we assume that the conditional expectation is $E[N_{ijl} \mid \Theta_{il} = \theta_{il}] = \lambda_{ijl}\theta_{il}$. The risk profile θ_{il} describes the risk that is not captured by the model for the a priori expected number of claims, of customer i and product l, and, as mentioned above, is sometimes called random effect.

We define F_{ijl} as the deviation between the actual number of incidents N_{ijl} and the a priori expected number of incidents λ_{ijl},

$$F_{ijl} = \frac{N_{ijl}}{\lambda_{ijl}} \text{ and } F_{i\cdot l} = \frac{N_{i\cdot l}}{\lambda_{i\cdot l}} = \frac{\sum_{j=1}^{J_i} N_{ijl}}{\sum_{j=1}^{J_i} \lambda_{ijl}}.$$

Other definitions, of the deviation between the expected and observed risk, are possible see e.g. Guillén et al. (2011). We assume that $Cov[F_{ijk}, F_{ijk'} \mid \Theta_{ik}, \Theta_{ik'}] = 0$, for $k \neq k'$.

The homogeneous multivariate credibility estimator (1) is the best linear unbiased estimator of $\theta_i = [\theta_{i1}, \ldots \theta_{iK}]'$ (see Englund et al., 2009 and Bühlmann and Gisler, 2005, p. 181).

$$\theta_i = \theta_0 + \alpha_i (F_{i\cdot} - \theta_0) \tag{1}$$

with $\theta_0 = [\theta_{01}, \ldots \theta_{0K}]'$ and $F_{i\cdot} = [F_{i\cdot 1}, \ldots F_{i\cdot K}]'$. The credibility weight $\alpha_i = T\Lambda_i(T\Lambda_i + S)^{-1}$ where T is a K by K matrix with elements $\tau^2_{kk'}$, $k = 1, \ldots, K$ and $k' = 1, \ldots, K$. The matrices Λ_i and S are diagonal matrices with, respectively, $\lambda_{i\cdot l}$, $l = 1, \ldots, K$ and σ^2_l, $l = 1, \ldots, K$ in the diagonal. The parameter $\sigma^2_l = E[\sigma^2_l(\Theta_{il})]$, where $\sigma^2_l(\Theta_{il})$ is the variance within an individual customer i, for a product l (for further details see Bühlmann and Gisler, 2005, p. 81). We also refer to Bühlmann and Gisler (2005, pp. 185-186) for parameter estimation procedures of the matrices S and T and the vector θ_0.

Performing the matrix multiplication in (1) and considering element k of θ_i we get

$$\theta_{ik} = \theta_{0k} + \sum_{k'=1}^{K} \alpha_{ikk'} (F_{i\cdot k'} - \theta_{0k'})$$

where $\alpha_{ikk'}$ is element kk' of the matrix α_i. This can be rewritten as

$$\theta_{ik} = \theta_{0k} + \alpha_{ikk} \left(F_{i\cdot k} - \theta_{0k}\right) + \sum_{k' \neq k} \alpha_{ikk'} \left(F_{i\cdot k'} - \theta_{0k'}\right). \tag{2}$$

We now assume that if product k is not active (not owned) by customer i, the risk exposure $e_{ijk} = 0$ for all j and consequently $\lambda_{ijk} = \lambda_{i\cdot k} = 0$. It is possible to show that $\lambda_{i\cdot k} = 0$ implies that $\alpha_{ikk} = 0$ and (2) becomes

$$\theta_{ik} = \theta_{0k} + \sum_{k' \neq k} \alpha_{ikk'} \left(F_{i\cdot k'} - \theta_{0k'}\right), \tag{3}$$

where the $\alpha_{ikk'}$ is element kk' of α_i when taken into consideration that $\lambda_{i\cdot k} = 0$ in Λ_i.

Equation (3) shows that even though a customer i does not have an active product k, it is possible to obtain his/her specific risk profile θ_{ik} (with respect to product k) by using data of $F_{i\cdot k'} = \frac{N_{i\cdot k'}}{\lambda_{i\cdot k'}}$ with respect to the other (owned) products $k' \in \{1, \ldots, k-1, k+1, \ldots, K\}$. From a company's perspective, customers with a low risk profile are preferred and therefore the estimate of θ_{ik} can be used to assess which customers to cross-sell product k to.

4. Empirical study

In this section we describe the data set collected to test the cross-sale selection methodology and our experiments with this data. We require a data set describing customers who own more than one financial product.

We conduct the experiment by neglecting the data with respect to one of the products and therefore imagine that this product is not owned by the customers. Instead the data for the other products is used to investigate if we are able to identify customers with fewer (or more) than expected number of incidents with respect to the discarded product.

4.1. Application data

The data sample is collected from the data base of a large Swedish insurance company writing business in both personal and commercial lines, however our sample consists solely of personal lines customers. The sample consist of a set of individuals who have been customers to the company

between 1999 and 2004 and who, during this time period, have owned all of the $K = 3$ main insurance coverages provided: motor, building and content insurance. The customers have not owned the coverages for equally long time so the policy duration spans between $J_i = 3$ and $J_i = 6$ years.

We have collected data from $I = 3395$ customers and for each customer i we estimate the a priori expected number of insurance claims $\hat{\lambda}_{ijl} = e_{ijl}\hat{g}_l(\mathcal{Y}_{ijl})$ (where $\hat{g}_l$ is estimated using a collateral dataset from the same company) and collect the number of claims n_{ijl} for each year $j = 1, \ldots, J_i$ and for each of the three coverages $l = 1$ (motor), $l = 2$ (building) and $l = 3$ (content). The a priori expected number of insurance claims $\hat{\lambda}_{ijl}$ has been assessed with the claim frequency model $\hat{g}_l$, in force at the time, using the characteristics of each customer and insured object. We present the mean and standard deviation of our data in Table 1, where it can be seen that the mean of the a priori expected number of claims $\hat{\lambda}_{ijl}$ is close to the mean of the observed number of claims n_{ijl} with the exception for product $l = 2$ (building coverage). Note that the standard deviation of the a priori expected number of claims is lower than the standard deviation of the observed number of claims, which is the result of the random effects and justifies credibility estimation.

Table 1. Descriptive statistics for Swedish insurer data from 1999 - 2004

		Mean	Std. dev.
Motor	Expected	0.084	0.053
	Observed	0.083	0.295
Building	Expected	0.064	0.033
	Observed	0.046	0.220
Content	Expected	0.051	0.028
	Observed	0.052	0.237

4.2. Experiment design and results

Our aim is to replicate the situation where the customers of a financial company have a set of products but lacking one of the products offered by the company. We assume that the company is interested in selecting customers with fewer than expected number of incidents. The company can achieve this by estimating the risk profile θ_{ik} for each customer i (with respect to the not owned product k) and select those with low risk profile. With our data set we imagine not knowing about

the data for one of the products k and thereafter estimate the risk profile θ_{ik} with data from the other products $1,\ldots,k-1,k+1,\ldots,K$. Thereafter we order the data set by increasing $\hat{\theta}_{ik}$ and partition it into a certain number M of subsets Φ_m (of size ϕ_m) with $m=1,\ldots,M$. The estimate of the risk profile $\hat{\theta}_{ik}$ is

$$\hat{\theta}_{ik}=\hat{\theta}_{0k}+\sum_{k'\neq k}\hat{\alpha}_{ikk'}\left(\hat{F}_{i\cdot k'}-\hat{\theta}_{0k'}\right),\ \text{where}\ \hat{F}_{i\cdot k'}=\frac{n_{i\cdot k'}}{\hat{\lambda}_{i\cdot k'}}=\frac{\sum_{j=1}^{J_i}n_{ijk'}}{\sum_{j=1}^{J_i}\hat{\lambda}_{ijk'}}. \tag{4}$$

The partitioning into subsets Φ_m is needed for presenting the results in an understandable way, we used different values of M and finally concluded that $M=5$ is an appropriate number of subsets. In this way, Φ_1 contains 20% of the customers associated with the lowest $\hat{\theta}_{ik}$, Φ_2 contains the next 20%, etc.. The number $\phi_m=679,$ for $m=1,\ldots 5$. Since the data sample is ordered by increasing $\hat{\theta}_{ik}$ before the partitioning into subsets Φ_m, we expect to capture customers with fewest incidents, compared to the a priori expected number, in subset Φ_1 and the customers with the most incidents, in comparison to the a priori expected number, in subset Φ_5. This can be validated by analysing the observed number of claims $n_{i\cdot k}$ in comparison to the a priori expected number $\lambda_{i\cdot k}$ for the customers in the different subsets Φ_m, with respect to the previously imagined not owned product k. For each subset Φ_m, we are interested in the deviation Δ of the observed number of claims in comparison to the a priori expected number expressed as a percentage as follows,

$$\Delta\left(\Phi_m\right)=100\left(\frac{\sum\limits_{i\in\Phi_m}n_{i\cdot k}}{\sum\limits_{i\in\Phi_m}\hat{\lambda}_{i\cdot k}}-1\right),\ m\in\{1,2,3,4,5\}. \tag{5}$$

Figure 2 describes our experiment with the data, for the situation where we are interested in identifying subsets Φ_m for product 2, using data from products 1 and 3. We use the notation $\hat{\theta}_{i,213}$ meaning that the risk profile θ_{i2} is estimated using data from products 1 and 3.

It is not uncommon that some customers of a financial company only have one of the many products offered by the company. The presented methodology works in this specific case as well by setting $e_{ijk}=0$ for the all products k which the customers does not own. I.e. for our data sample, we can also estimate the risk profile θ_{ik} by using information from only one of the two remaining

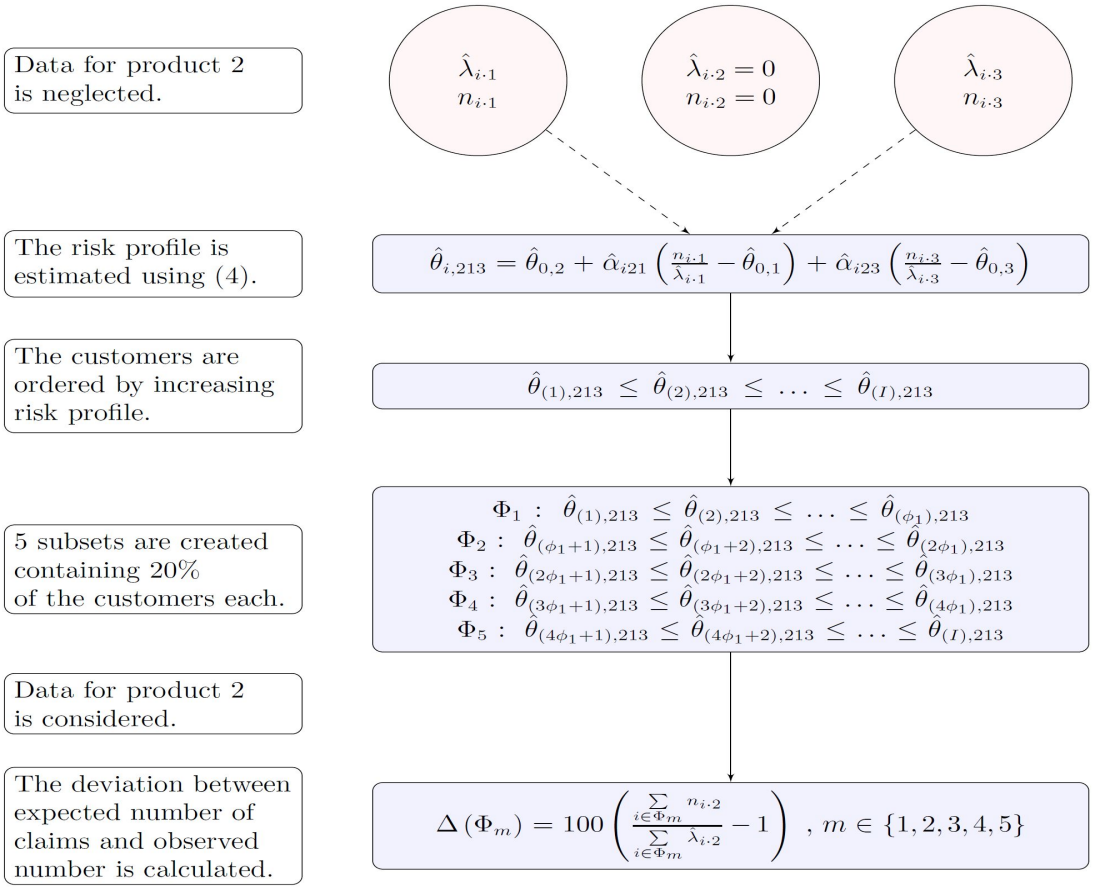

Figure 2 The design of the experiment for the particular case of creating subsets Φ_1 to Φ_5 based on the estimated risk profile for product 2 using information from products 1 and 3.

products in the data set. For instance, for the estimate of the risk profile of product $k=1$, θ_{i1}, we use the notation $\hat{\theta}_{i,12}$ if only data from product 2 is used in the estimation, and correspondingly for the other products.

The evaluation criteria (5), applied to investigate the deviation between of the observed number of claims $n_{i\cdot k}$ and the a priori estimated expected number $\hat{\lambda}_{i\cdot k}$, in the 5 subsets Φ_m, is presented in Figures 3 to 5. In Figure 3, we imagine that product $k=1$ (car coverage) is not owned by the customers and we use data from either product $k'=2$ (building coverage) or product $k'=3$ (content coverage) or data from both building and content coverage to estimate the risk profile θ_{i1}, with respect to product 1. Thereafter, for each of the three different estimators, we order the data set by increasing value of the risk profile estimate and partition the data into the subsets

Φ_m with $m = 1, \ldots, 5$ for calculation of $\Delta(\Phi_m)$, see equation (5).

As seen in Figure 3, the credibility estimator $\hat{\theta}_{i,12}$, which uses data from product 2, does only slightly differentiate the customers with respect to claiming ($n_{i\cdot 1}$) in comparison to the a priori expected claiming ($\lambda_{i\cdot 1}$) (left sub-figure of Figure 3). However, when ordering the data with respect to $\hat{\theta}_{i,13}$, which uses information from product 3, subset Φ_1 contains customers with on average 6% lower claims frequency than expected and subset Φ_5 contains customers with 22% more claims than expected, see center sub-figure of Figure 3. When using data from both product 2 and 3 the result is improved slightly and Φ_1 contains customers with 8% less claims than expected and Φ_5 contains customers with 26% more claims than expected.

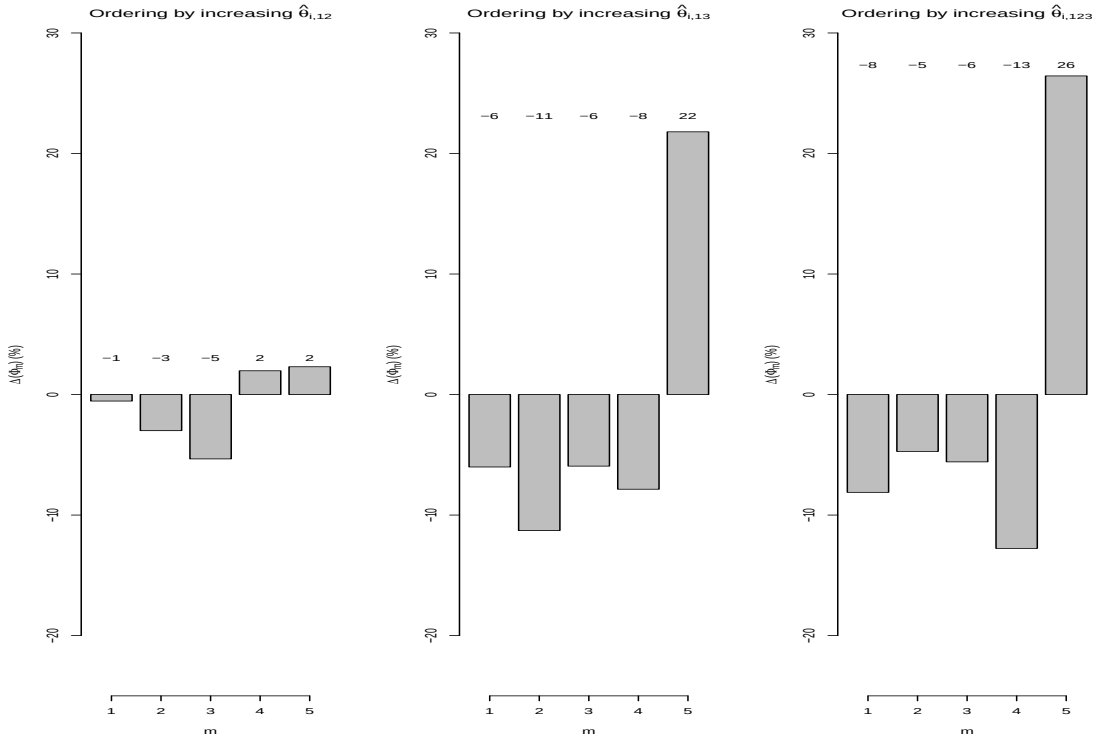

Figure 3 Average deviation between observed number of claims and a priori expected number with respect to product 1 (car coverage). The subsets Φ_m are created using only information from building coverage (left sub-figure), using only information from content coverage (center sub-figure) or using information from both building and content coverages (right sub-figure)

In Figure 4 we imagine that product 2 (building coverage) is not owned by the customers. We see

that almost all subsets Φ_m contain customers with fewer claims than expected because (according to Table 1) the average value of $\hat{\lambda}_{i\cdot 2}$ is far greater than the average value of $n_{i\cdot 2}$ since almost all customers have reported fewer claims than a priori expected. Still, the credibility estimators $\hat{\theta}_{i,23}$ (center sub-figure) and $\hat{\theta}_{i,213}$ (right sub-figure) is able to differentiate between subsets containing customers with less than expected claiming and more than expected claiming.

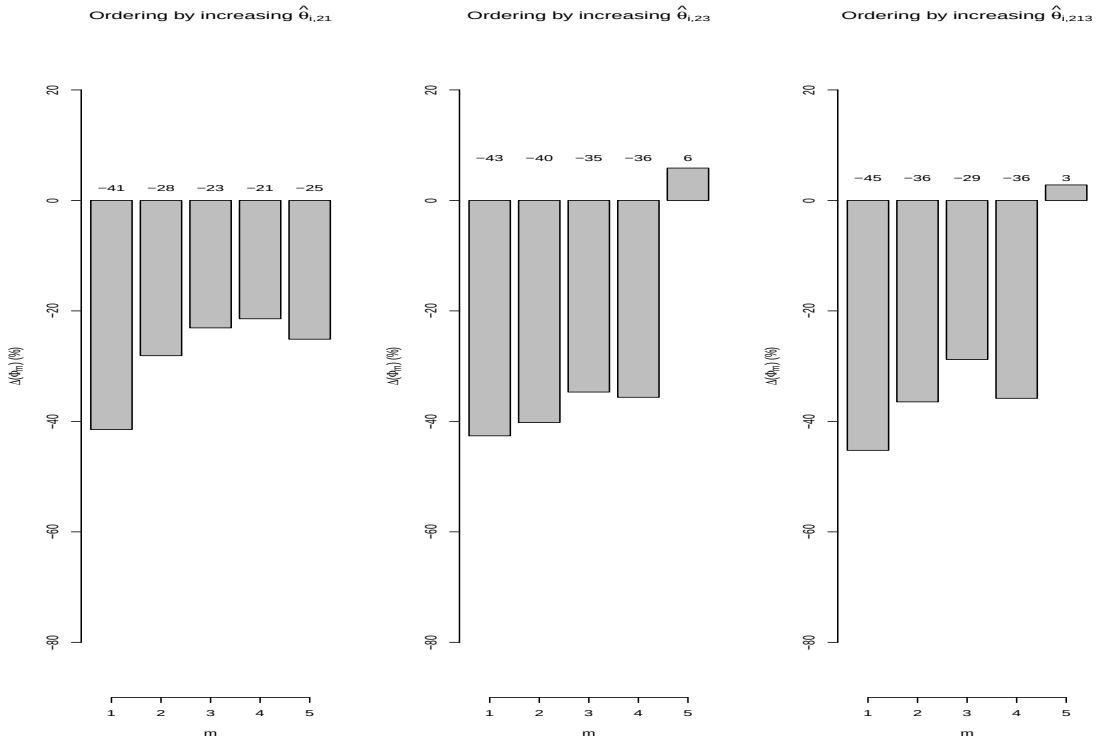

Figure 4 Average deviation between observed number of claims and a priori expected number with respect to product 2 (building coverage). The subsets Φ_m are created using only information from car coverage (left sub-figure), using only information from content coverage (center sub-figure) or using information from both car and content coverages (right sub-figure)

In Figure 5, we imagine that product 3 (content coverage) is not owned by the customers. We see that all credibility estimators ($\hat{\theta}_{i,31}$, $\hat{\theta}_{i,32}$, $\hat{\theta}_{i,312}$) are identifying the customers in subset Φ_5 as having much more claims than expected. Especially the estimator $\hat{\theta}_{i,312}$ (right sub-figure) is able to identify, in the subset Φ_5, customers who have on average 64% more claims than a priori expected while also identifying the customers in the subset Φ_1 with on average 10% less claims than a priori

expected.

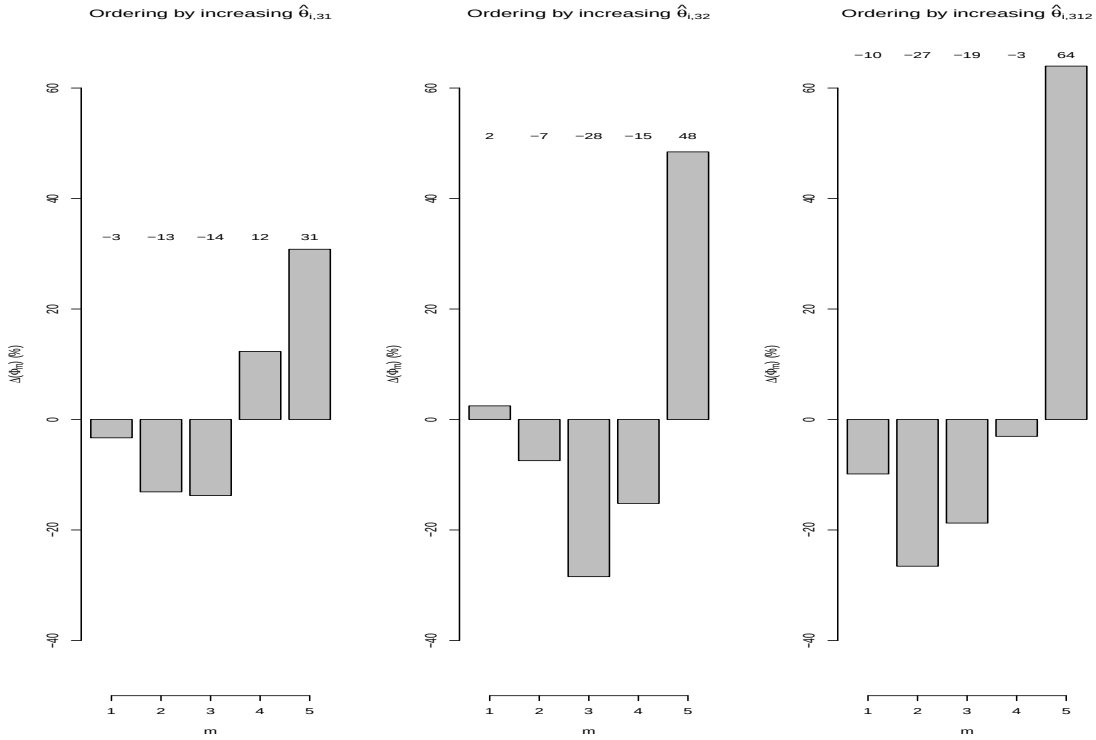

Figure 5 Average deviation between observed number of claims and a priori expected number with respect to product 3 (content coverage). The subsets Φ_m are created using only information from car coverage (left sub-figure), using only information from building coverage (center sub-figure) or using information from both car and building coverages (right sub-figure)

In Figures 3 to 5 it would be expected and preferred that the deviation of the observed number of claims in comparison to the a priori expected number, $\Delta(\Phi_m)$, would be lowest for $m = 1$. However, this is not the case for many of the estimators and especially for cross-selling product $k = 1$ (car) in Figure 3 the lowest $\Delta(\Phi_m)$ is recorded for $m = 3$, $m = 2$ and $m = 4$ for the credibility estimators $\hat{\theta}_{i,12}$, $\hat{\theta}_{i,13}$ and $\hat{\theta}_{i,123}$, respectively. A similar note can be made with regards to Figure 5. We draw the conclusion that for the collected data sample it is more efficient to identify a small group of customers to avoid to cross-sale to (Φ_5) than a small group of customers to target (Φ_1). Consequently, we find that by avoiding the 20% of the customers associated with the highest risk profile estimates $\hat{\theta}_{ik}$ (Φ_5) and targeting the remaining 80% the company would increase its

profit significantly. In Table 2 we compare $\Delta(\Phi_5)$ to $\Delta(\cup_{m=1}^{4}\Phi_m) = \Delta(\Phi_1 \cup \Phi_2 \cup \Phi_3 \cup \Phi_4)$ where $\Phi_m \cup \Phi_{m+1}$ denotes the union of Φ_m and Φ_{m+1}.

Table 2. Percentage deviation between observed and expected number of claims. Note that a positive value indicates that the subset of customers has reported more claims a priori expected.

Car			Building			Content		
Order	$\Delta(\cup_{m=1}^{4}\Phi_m)$	$\Delta(\Phi_5)$	Order	$\Delta(\cup_{m=1}^{4}\Phi_m)$	$\Delta(\Phi_5)$	Order	$\Delta(\cup_{m=1}^{4}\Phi_m)$	$\Delta(\Phi_5)$
Random	0%	-4%	Random	-29%	-30%	Random	4%	3%
$\hat{\theta}_{i,12}$	-2%	2%	$\hat{\theta}_{i,21}$	-30%	-25%	$\hat{\theta}_{i31}$	-4%	31%
$\hat{\theta}_{i,13}$	-8%	22%	$\hat{\theta}_{i,23}$	-40%	6%	$\hat{\theta}_{i32}$	-9%	48%
$\hat{\theta}_{i,123}$	-8%	26%	$\hat{\theta}_{i,213}$	-38%	3%	$\hat{\theta}_{i312}$	-13%	64%

Table 2 shows that by selecting the 80% ($\cup_{m=1}^{4}\Phi_m$) most favorable customers, with respect to the estimate of the risk profile θ_{ik}, the company is able to avoid customers associated with up to 64% more claims than a priori expected (content coverage, product 3). In the table we have also included results produced when the data sample has been randomly ordered and partitioned into 80% of the data and 20% of the data. The random order does not differentiate between subsets of customers with respect to percentage deviation between observed and expected number of claims. We see a similar pattern for product 1 (car) where the 80% most favorable customers are associated with 8% less claims than a priori expected while the remaining 20% are associated with 26% more claims than a priori expected. The performance of the credibility estimators in Product 2 (building) is difficult to interpret because almost all customers are associated with lower observed claim occurrence than a priori expected. However, even for this particular situation a subset Φ_5 can be identified consisting of customers with on average 6% more claims than a priori expected. Note that this is not received for the credibility estimator which uses all available information ($\hat{\theta}_{i,213}$) but the estimator which only uses data from the content product $k' = 3$, $\hat{\theta}_{i,23}$.

5. Discussion

This study investigates identification of customers to whom additional products should be offered, by estimating a customers specific risk profile with the use of behavioural data from other products of the specific customers. We use a standard multivariate credibility model applied to a portfolio of customers, of a financial company, owning several financial products from the company. The

model allows us to take into consideration the possible (positive) correlation in customer behaviour between different financial products and estimate the customer specific risk profiles, for a specific product not owned by the customer, without having observed any customer specific information with respect to that particular product. Instead, data on customer behaviour, with respect to the other (owned) products, is the only necessity for estimating the risk profile.

The methodology uses only two observables: the a priori expected number of incidents and the observed number of incidents. We assume that the financial company has a model for the expected number of incidents or is able to assess a value specific for each customer or category of customers. When estimating such models it is unusual to incorporate information about the number of incidents related to a specific customer. Instead the company finds patterns which can be used to categorise the customers, with respect to the expected occurrence of incidents, based on customers' characteristics. It is not uncommon that customers are associated with more or less number of incidents, than suggested by the categorisation, based on their attitude towards risk. In our methodology we use that the attitude towards risk seems to be similar across different financial products. I.e. if a customer is associated with more or less number of incidents, than a priori expected in some products, it is likely that this pattern will also emerge in other related products.

With the presented credibility estimators we are able to assign, to each customer, a specific estimate of his/her risk profile based on data which the company has available. We use the estimate, of each customer's risk profile, to identify subsets from the data containing customers associated with more or less incidents than a priori expected. In this way the company receives knowledge about which customer to target for cross-selling and which to avoid.

In our empirical study we analyse our methodology on real data from a large Swedish insurance company, consisting of personal lines customer with three different insurance coverages. We find that there are subsets of the data sample with large heterogeneity with respect to claiming in comparison to expected claiming. Furthermore, we find that these subsets are identifiable by using

an appropriate credibility estimator of the risk profiles. The appropriateness of a specific credibility estimator is dependent of the considered product, but in most cases an estimator which uses all available information is preferable. We find that it is easier to identify the 20% of the data containing customers to avoid than the 20% of the data containing customers to target. In fact, by targeting all customers but the worst 20%, the company could expect a subset of customer associated with less claims than a priori expected indifferent of which product is considered. The remaining 20% of the data sample consist of customers with up to 64% more claims than a priori expected.

References

Ahn, H., Ahn, J.J., Oh, K.J. and Kim, D.H. (2011). Facilitating cross-selling in a mobile telecom market to develop customer classification model based on hybrid data mining techniques. *Expert Systems with Applications*, 38(5), 5005-5012.

Akura, M.T. and Srinivasan, K. (2005). Research Note: Customer Intimacy and Cross-Selling Strategy. *Management Science*, 51(6), 1007-1012.

Bae, J., K. and Kim, J. (2010). Integration of heterogeneous models to predict consumer behavior. *Expert Systems with Applications*, 37(3), 1821-1826.

Bühlmann, H. (1967). Experience rating and credibility, *Astin Bulletin*, 4, 199-207.

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

Bühlmann, H. and Straub, E. (1970). Glaubwürdigkeit für Schadensätze, *Bulletin of Swiss Association of Actuaries*, 70, 111-133.

Denuit, M., Marechal, X., Pitrebois, S. and Walhin, J.F. (2007) *Actuarial modelling of claim counts: risk classification, credibility and bonus-malus systems* John Wiley and Sons, Ltd. New York.

Englund, M., Guillén, M., Gustafsson, J., Nielsen, L.H. and Nielsen, J.P. (2008). Multivariate latent risk: A credibility approach, *Astin Bulletin*, 38, 137-146.

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.

Frees, E.W. (2003). Multivariate credibility for aggregate loss models. *North American Actuarial Journal*, 7(1), 13-37.

Guillén, M., Pérez-Marn, A.M. and Alcañiz, M. (2011). A logistic regression approach to estimating customer profit loss due to lapses in insurance. *Insurance Markets and Companies: Analyses and Actuarial Computations*, 2 (2), Forthcomming.

Harrison, G.W., Lau, M.I. and Rutstrm, E.E. (2007). Estimating risk attitudes in Denmark: A field experiment. *Scandinavian Journal of Economics* 109(2), 341-368.

Jewel, W.S. (1973). Multidimensional credibility. *Actuarial Research Clearing House* 4.

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.

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, 45-65.

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 Wilcox, R.T. (2005). Cross-selling sequentially ordered products: an application to consumer banking services. *Journal of Marketing Research*, 42, 233-239.

Liao, S-H., Chen, Y-J. and Hsieh, H-H. (2011). Mining customer knowledge for direct selling and marketing. *Expert Systems with Applications*, 38(5), 6059-6069.

Larivière, B. and Van den Poel, D. (2004). Investigating the role of product features in preventing customer churn, by using survival analysis and choice modeling: The case of financial services. *Expert Systems with Applications*, 27(2), 277-285

Larivière, B. and Van den Poel, D. (2005). Predicting customer retention and profitability by

using random forest and regression techniques. *Expert Systems with Applications*, 29(2), 472-484.

Malthouse, E.C. (2010). Accounting for the long-term effects of a marketing contact. *Expert Systems with Applications*, 37(7), 4935-4940.

Mowbray, A.H. (1914). How extensive a payroll exposure is necessary to give a dependable pure premium? *Proceedings of the Casualty Actuarial Society*, 1, 24-30.

Norberg, R. (2004). Credibility theory. Encyclopedia of Actuarial Science. John Wiley & Sons. 398-406.

Seng, J-L., and Chen T.C. (2010). An analytical approach to select data mining for business decision. *Expert Systems with Applications*, 37(9), 8042-8057.

Slovic, P., Finucane, M., Peters, E., and MacGregor, D.G. (2004). Risk as Analysis and risk as feelings: Some thoughts about affect, reason, risk and rationality. *Risk Analysis*, 24(2), 311-322.

Whitney, A.W. (1918). The theory of experience rating *Proceedings of the Casualty Actuarial Society*, 4, 274-292.