Predicting Customer Wallet without Survey Data

Nicolas Glady[a]

Christophe Croux[b] [1]

[a] Nicolas.Glady@econ.kuleuven.be: Faculty of Business and Economics, K.U.Leuven

[b] Christophe.Croux@econ.kuleuven.be: Faculty of Business and Economics, K.U.Leuven

Abstract

A single company provides only a part of the total volume of products or services required by a customer. From the company perspective, this total business volume conducted by a customer, the customer's Size-of-Wallet, is generally unobservable. The percentage of this business done with the company, the customer's Share-of-Wallet, is unobservable as well. This paper focuses on the prediction of these values and on the derived concept of Potential-of-Wallet, which is the difference between the Size-of-Wallet and the actual business volume the customer does with the focal company. In the existing literature, the models predicting the customer wallet need survey data to estimate the model parameters. We propose an approach to predicting customer wallet without using survey data. In the empirical application, we show that a company can generate substantial gains by targeting customers with a large Potential-of-Wallet.

Keywords: Customer Relationship Management, Prediction, Retail Banking, Share-of-Wallet.

Introduction

A single company provides only a part of the total volume of products or services required by a customer. From the company perspective, this total business volume conducted by a customer, the customer's *Size-of-Wallet* (SioW), is generally unobservable. The percentage of this business done with the company, the customer's *Share-of-Wallet* (ShoW), is unobservable as well.

In a competitive market, it is crucially important for companies to fully exploit their existing customer base. Different approaches to identifying customers according to their loyalty, profitability or other customer metrics have been proposed. Recent studies (Zeithaml 2000, Gupta and Zeithaml 2006) have discussed the identification of the key drivers of service quality, customer retention and profits, showing that retention positively affects profits. In order to improve the profitability of the customer base, the ShoW can be used in two ways. First, it allows companies to detect the customers with the highest loyalty, who should be prevented from defecting. Second, the ShoW can be used to identify customers with a high growth potential. For instance, a customer with a ShoW of 25%, if motivated to increase business with the focal company, should become more profitable. In this case, the Size-of-Wallet does not change, but the Share-of-Wallet does. We call the difference between the Size-of-Wallet and the actual business volume a customer does with the company the *Potential-of-Wallet* (PoW) of this customer. To increase an existing customer's profitability, the PoW should be a key driver of customer relationship management.

Many recent studies have proposed models for estimating the ShoW or customer defection (e.g. Magi 2003, Buckinx and Van den Poel 2005, Perkins-Munn et al. 2005, Buckinx et al. 2007). In the retail banking business, which is the context of our empirical application, Keiningham et al. (2003), Garland (2004), Baumann et al. (2005) and Du et al. (2007) have provided interesting insights into the topic. All these studies rely on collected survey data.

A Customer's Size-of-Wallet and Share-of-Wallet are generally unobservable and, to our knowledge, no method has so far been proposed in the literature to predict

these quantities without ``additional'' information. This ``additional'' information consists of survey data collected from customers who give an estimate of their own ShoW. Conducting a survey, however, is very expensive, so the research objective of this paper is to propose an approach to predicting customers' Size-of-Wallet and Share-of-Wallet (and therefore the Potential-of-Wallet) without using survey data. Our method only requires information that a company can easily obtain: some RFM (recency, frequency and monetary) variables and some socio-demographics variables, such as a customer's age and address.

In the sections ``Data Description'' and ``Results'', we apply our model to the savings deposited by customers at a major Belgian retail banker. This retail banker provided all the information required for estimating and assessing the model. The model predicts the volume of savings deposited by a customer (defined as the amount on the savings account), and the proportion of these savings deposited outside the focal company. We show that a company can generate substantial gains by targeting customers with a large Potential-of-Wallet.

In addition, we investigate the association between the SioW, the ShoW and other variables generally used in marketing studies. This reveals that the Share-of-Wallet is non-linearly related to the observed activity and the Size-of-Wallet. In our application, the customers with a ShoW close to 75% are, on average, those with the largest balance in their savings account. We also observe that customers with a ShoW close to 25% or 75% are, on average, those with the highest reported Size-of-Wallet.

The next section of this paper defines the key concepts of this study, the Size-of-Wallet, the Share-of-Wallet and the Potential-of-Wallet. In the ``Methodology'' section, the *Generalized Binomial Model* (GBM) is proposed as a tool for predicting the Size-of-Wallet and the Share-of-Wallet without using survey data. The ``Data Description'' section presents the empirical application, i.e. an analysis of customer activity in the savings accounts of a Belgian retail banker. This section also describes the survey data required to assess model performance. We stress, however, that these survey data are not used for model estimation. In the ``Results'' section, we discuss the implementation of the model and its performance to predict the Size, Share and Potential-of-Wallet in our empirical application. Finally, the last section summarizes our conclusions.

Size, Share and Potential of the Customer Wallet

Each consumer does a certain volume of business within a certain product category. This volume, defined as his/her Size-of-Wallet, is different for every customer, and depends on the customer's needs and preferences. Unlike other studies (e.g Du et al. 2007), we do not make a distinction between a customer's Size-of-Wallet and his/her ``category requirement'' for a specific type of product or service, since our empirical application deals with only one product, namely the money held in a savings account. Generalization to several products is nevertheless straightforward, as the Share-of-Wallet can be defined as the (weighted) average of a customer's category requirements.

A customer's *Share-of-Wallet* (ShoW) is defined as the proportion of business this customer does with the focal company. As such, it is a measure of both behavioral loyalty and potential profitability. It is a measure of loyalty in the sense that the customers with a higher ShoW are the company's most loyal customers. For example, in our application, the savings accounts of a Belgian retail banker, a customer with a ShoW of 100 % deposits all his/her savings at the focal bank and is therefore perfectly loyal. The ShoW is also a measure of potential profitability, as the customers with an average ShoW could increase their activity and thus become more profitable from the focal company's perspective. For instance, someone could save 500 euro a month at two different banks. If this customer was motivated to increase his/her ShoW with the focal company, his/her profitability to the company would increase accordingly. The ShoW is therefore a positive measure of loyalty, but a negative measure of the growth potential of a customer's profitability. This leads us to define the *Potential-of-Wallet* (PoW) of a customer as the difference between the customer's Size-of-Wallet and the actual business volume he/she conducts with the focal company.

For many years, the literature has insisted on replacing companies' traditional ``product-centric'' approach by a ``customer-centric'' one, placing the customer at the center of marketing activities. In a recent paper, Shah et al. (2006) identified the financial metrics deterring a firm from becoming customer-centric. The authors argue that, instead of measuring the total market share by product or subbranch, loyalty studies should be

driven by an analysis of each customer's Share-of-Wallet. Consequently, (as in Wirtz et al. 2006, for loyalty programs) the impact of an action on customer loyalty should be assessed using the Share-of-Wallet.

In what follows, business volumes are measured in numbers of ``transactions''. The Size-of-Wallet, then, is the total number of transactions a customer makes. The Share-of-Wallet is the probability that a transaction is made with the focal company, and a customer's Potential-of-Wallet is the number of transactions this customer has with the company's competitors. A transaction is a ``typical'' flow of money depending on the business application. In our empirical application, we define a transaction as a 100-euro deposit into a savings account.

We model customer activity in terms of the number of transactions, but one might also be interested in modeling the customer profitability instead. This problem can be approached by first predicting the number of transactions, for example using the GBM model discussed in the next section. Second, multiplying the predicted number of transactions by a specified profit margin, which may be different for every individual customer, yields an estimate of total customer profitability.

Methodology

With regard to the retail banking business, Baumann et al. (2005) and Du et al. (2007) have proposed models to estimate the Share-of-Wallet. These studies first acquired survey data for a part of the population of interest, and then applied the estimated model to the rest of the customer base. In this paper, we are interested in predicting a customer's Size-of-Wallet and Share-of-Wallet without survey data, using only easily obtainable information on customers and their purchasing behavior at the focal company. The population of interest comprises the financial institution's present customers, but also individuals who were clients in the past. We first propose a model expressing the number of transactions as a function of the SioW and the ShoW. This model is referred to as the *Generalized Binomial Model* (GBM). Next, we establish a typology of the variables which are used in the empirical application.

The Generalized Binomial Model

The *Generalized Binomial Model* (GBM) models the number of transactions a customer makes. Firstly, the total number of transactions a customer i makes, denoted by n_i, is assumed to be drawn from a Poisson distribution with individual parameter λ_i,

$$f_1(n_i \mid \lambda_i) = \frac{e^{-\lambda_i} \lambda_i^{n_i}}{n_i!}. \tag{1}$$

Note that n_i, the total number of transactions the customer i makes, i.e. this customer's Size-of-Wallet, is unobserved. The individual parameter λ_i is modeled using observable predictor variables,

$$\lambda_i = \exp(\boldsymbol{\beta}'\mathbf{v_i}), \tag{2}$$

where $\mathbf{v_i} = (v_{i,1},\ldots,v_{i,L})$ are the values of the L regressors accounting for the Size-of-Wallet of customer i. Once β is estimated, the estimated value of λ_i is given by $\hat{\lambda}_i = \exp(\hat{\boldsymbol{\beta}}'\mathbf{v_i})$. Model (2) resembles a Poisson regression model, but with an unobserved dependent variable.

Secondly, the number of transactions made by customer i at the focal company,

denoted by x_i, is assumed to follow a binomial distribution,

$$f_2(x_i \mid \pi_i, n_i) = \binom{n_i}{x} \pi_i^{x_i} (1-\pi_i)^{n_i-x_i}. \tag{3}$$

The values x_i, in contrast to n_i, are observed by the company. In (3), π_i stands for the probability that customer i will choose the focal company when making a transaction, i.e. this customer's Share-of-Wallet. In our empirical application, this would imply that a customer, when making a 100-euro deposit, would have a certain probability to make this deposit either at the focal company (success) or outside the focal company (failure).

The Share-of-Wallet, π_i, is modeled by the logistic regression model

$$\ln(\frac{\pi_i}{1-\pi_i}) = \boldsymbol{\alpha}' \mathbf{w_i}, \tag{4}$$

where $\mathbf{w_i} = (w_{i,1}, \ldots, w_{i,J})$ are J regressors accounting for the Share-of-Wallet information. Once the parameter α is estimated, the estimated ShoW follows as

$$\hat{\pi}_i = \frac{1}{1+e^{-\hat{\alpha}'\mathbf{w_i}}}. \tag{5}$$

Recall that only the x_i are observed, and that the Size-of-Wallet and Share-of-Wallet are unobserved. However, these values can be predicted after estimating the model parameters α and β.

The parameters α and β are estimated by *Maximum Likelihood* (ML), where the likelihood of an individual observation equals

$$f(x_i \mid \alpha, \beta) = \int f_2(x_i \mid n, \alpha) f_1(n \mid \beta) dn,$$

with f_1 and f_2 given in (1) and (3). Hence,

$$f(x_i \mid \alpha, \beta) = \sum_{n=x_i}^{\infty} \binom{n}{x_i} \pi_i^{x_i} (1-\pi_i)^{n-x_i} \frac{e^{-\lambda_i} \lambda_i^n}{n!},$$

where π_i is a function of α and λ_i is a function of β, see equations (4) and (2). By substituting $m_i = n - x_i$, we have

$$f(x_i \mid \alpha, \beta) = \frac{(\lambda_i \pi_i)^{x_i} e^{-\lambda_i}}{x_i!} \sum_{m_i=0}^{\infty} \frac{[\lambda_i(1-\pi_i)]^{m_i}}{m_i!},$$

or

$$f(x_i \mid \alpha, \beta) = \frac{(\lambda_i \pi_i)^{x_i} e^{-\lambda_i \pi_i}}{x_i!},$$

yielding the probability density function of a Poisson distribution with parameter $\lambda_i \pi_i$. The estimates $\hat{\alpha}$ and $\hat{\beta}$ are then obtained by minimizing

$$-\log L(x_1, \ldots, x_N \mid \alpha, \beta) \propto \sum_{i=1}^{N} [\lambda_i \pi_i - x_i \log(\lambda_i \pi_i)], \qquad (6)$$

where N is the total number of observations (customers). The estimated value of λ_i and π_i follows from (2) and (5).

It remains to predict the total number of transactions, n_i, being the Size-of-Wallet. For each individual customer i, we know that $n_i \geq x_i$ because the total number of transactions a customer makes is always greater than or equal to the number of transactions this customer makes with the focal company. Hence, we set

$$\hat{n}_i = E[n_i \mid n_i \geq x_i]$$
$$= \frac{1}{P(n_i \geq x_i)} \sum_{n=x_i}^{\infty} n f_1(n \mid \lambda_i),$$

which can be rewritten as

$$\hat{n}_i = \lambda_i \frac{1 - F(x_i - 2, \lambda_i)}{1 - F(x_i - 1, \lambda_i)}, \qquad (7)$$

where $F(x, \lambda)$ is the cumulative distribution function of a Poisson distribution with parameter λ at the value x. The predicted value of the Size-of-Wallet is consequently obtained by (7), by replacing λ_i by its estimate. The Potential-of-Wallet is then predicted as $\hat{n}_i - x_i$. Notice that the prediction of n_i only requires an estimate of λ_i and not of π_i.

The assumptions made by the GBM model are quite stringent. First, it is assumed that the Size-of-Wallet follows a Poisson distribution. Second, we do not allow for consumer heterogeneity in equations (2) and (4). A way to introduce more heterogeneity in the model is to decompose the Poisson distribution parameter into the product of a deterministic coefficient (λ_i) and a random customer effect. A common choice for the distribution of the customer effect is a gamma distribution with an expected value of 1 (see Green 2003, pp 744-745). The distribution of the Size-of-Wallet is then a Negative Binomial Distribution, which is more flexible than the standard Poisson distribution. It can be verified that, under these circumstances, equation (6) corresponds to the log-likelihood of a Negative Binomial regression model. When this approach was implemented in our specific empirical application, it was not found to improve prediction

results.

It is clear that estimating the unobserved Size-of-Wallet and Share-of-Wallet in the absence of survey data is not feasible without imposing rather strong model assumptions. We therefore advocate performing a goodness-of-fit test after estimating the model, as will be shown in the ``Model Validation'' subsection. If such a goodness-of-fit test indicates severe model misspecification, the use of a more flexible model should be considered. A major advantage of the GBM model, however, is the simplicity of its likelihood, which facilitates the computation of the ML-estimates.

It should also be stressed that the GBM model requires the business volume to be measured as a certain number of transactions, and not as a continuous quantity. In most empirical applications, this requires a preliminary step, i.e. discretization, whereby business volume is measured as a finite number of basic transactions. The empirical application detailed in the ``Data Description'' section gives an example of such a discretization.

Typology of the Variables

The model presented above predicts the Size-of-Wallet and the Share-of-Wallet on the basis of the explanatory variables $\mathbf{v}_i$ and $\mathbf{w}_i$. These explanatory variables need to be selected with care, and according to the particular business application. To avoid overfitting, we prefer to use a rather limited set of explanatory variables. In marketing studies, two kinds of variables are generally used: transactional and socio-demographic variables. Transactional variables are related to a customer's business activity from the focal company's perspective, e.g. the number of past purchases or the number of these transactions. The RFM (recency, frequency and monetary) framework fits in here. In contrast, the socio-demographic variables are customer characteristics, unrelated to their relationship to the company, e.g. their age, address, etc. Due to the unavailability of socio-demographic variables, many studies only consider transactional data. Nevertheless, Mittal and Kamakura (2001) show that consumers with different characteristics have different response patterns, and that it is preferable to consider both transactional and socio-demographic explanatory variables.

Another distinction that can be made is between wealth-related and wealth-unrelated variables. For instance, the lifetime of the customer's relationship with the company or his/her recency are unrelated to this customer's wealth. In contrast, the savings made, the number of products purchased at the company, or the customer's job are wealth-related.

Table 1 displays the explanatory variables used in the empirical application according to the two dimensions previously described. The value of a customer's ``average savings'' is defined as the average savings balance of this customer during the last three months, i.e. over the July-September 2006 period.[1] This period precedes the October-December period used for defining the $x_i$, the number of transactions made by customer $i$. The ``average savings'' variable is obviously wealth-related and it is also a transactional variable, as it is a function of the relationship between the customer and the company. A customer's address is not transactional, but is related to wealth: the value of a house, hence its owner's wealth, depends on its location. Customer's recency, expressed as the number of time units between this customer's first and last transaction, is a transactional and wealth-independent variable. Finally, a customer's age is a socio-demographic variable, which may be considered either as wealth-related, or as wealth-unrelated. This information is usually provided by a customer when opening a savings account.

In our empirical application, average savings, recency, and age are continuous variables. The address, though, is coded by a sequence of dummy variables for 10 different categories which are described in Table 2, with ``missing'' as a reference category. The variable ``address'' was provided by the financial company and gives the area type of the customer's place of residence. The labels of the categorical variable ``address'' refer to typical inhabitants of this area type.

Data Description

The company that provided the data for our empirical application is a major retail banker, active in many countries, and selling several categories of financial products. The group of customers selected for our study is a sample of the company's Belgian customers. The number of transactions made by a customer, x_i, is defined as the balance in this customer's savings account, averaged over the October-December 2006 period. This number is expressed in units of 100 euro rounded to the closest integer. For example, a deposit of 1030 euro in the savings account is considered as 10 basic transactions of 100 euro. A savings account is a special bank account in Belgium, with an interest rate higher than that on a checking account, but it is less liquid, and cannot be negative. If an individual is not a customer of the company, and therefore has no savings account, the number of observed transactions is 0.

In this paper, we propose to apply the GBM model to predict customer wallet in the absence of survey data. However, without survey data, it is impossible to assess the model's predictive performance, since the customers' Size-of-Wallet and Share-of-Wallet are unobserved. Fortunately, the focal company conducted a survey, requesting customers to report the proportion of their savings held at the company, i.e. their Share-of-Wallet. This survey is used to assess the quality of the model's predictions. We emphasize again that these survey data are kept out-of-sample during the estimation of the model parameters.

The customer base was stratified on the basis of variables such as age, gender, the customer's longevity with the company and the level of business activity. The survey, sent by email to this stratified sample of 20,000 customers, was carried out from October 2006 to December 2006. Respondents self-reported their Share-of-Wallet (expressed as a percentage: 0%, 10%, …, 100%), which may have resulted in some respondent accuracy bias, but the direction of such bias is unknown. From the whole sample of 20000 customers, 3125 individuals responded to the survey. The survey data provide the self-reported Share-of-Wallet, π_i. The Size-of-Wallet of customer i is then simply $n_i = x_i/\pi_i$, for the customers with a reported Share-of-Wallet greater than 0. The

Potential-of-Wallet follows as $n_i - x_i$.

Figure 1 displays the histogram of the ShoW reported by respondents. Two peaks can be observed, the highest one at 100% of ShoW and a smaller one at 0% of ShoW.[2] Figure 2 reports the boxplots of savings distribution at the time of the survey as a function of the reported ShoW. We see that, in general, savings increase with the Share-of-Wallet. However, customers with a ShoW around 75% have, on average, higher savings than the customers with a ShoW of 100%.

Figure 3 displays the boxplots of the distribution of the reported Size-of-Wallet as a function of the reported Share-of-Wallet. Customers with the highest Size-of-Wallet are those with a reported Share-of-Wallet around 25% or 75%, which suggests that the most wealthy customers have more than one financial services provider, whereas a customer with smaller savings is more likely to have only one savings account. Consequently, the ``full customers'' (at 0% or 100% of ShoW) are not the ones with the highest SioW. In contrast, ``partial customers'', with a ShoW around 25% or 75%, have the highest SioW. This observation was also made by Du et al. (2007), who found, first, that customers' Size-of-Wallet and Share-of-Wallet were sometimes negatively correlated, suggesting that customers with small shares within the focal firm tend to transact a large volume outside it. Second, they observed that customers with higher incomes tend to balance Share-of-Wallet across firms.

From a customer relationship management perspective, these observations imply that the most interesting customers, those with the highest Potential-of-Wallet, are more likely to have a ShoW close to 25%. Since the ``100% customers'' are already making all their transactions with the focal company, and the ``75% customers'' are likely to go on conducting a part of their business outside the company; the ``25% customers'' are the customers with the largest business volume to be acquired. In what follows, we emphasize the prediction of the Potential-of-Wallet ($n_i - x_i$) instead of the prediction of the Size-of-Wallet. From an acquisition perspective, the PoW should drive the customer relationship strategy, whereas the SioW is actually less useful as a workable concept. For a company, it is not a consumer's wealth which is interesting, but it is the proportion of this wealth that could generate a profit.

In order to justify the choice of the explanatory variables, Table 3 reports the

correlations between the explanatory variables and the variables to predict. Note that these correlations cannot be computed without survey data. First, we see that the Size-of-Wallet is positively correlated with a customer's past average savings and age. Second, the Share-of-Wallet is positively correlated with a customer's past average savings and recency, but negatively with age. Finally, living in a wealthier area is associated with a larger Size-of-Wallet.

Results

In this section, we describe how the parameters of the model are estimated. Next, we propose an approach to assess the accuracy of the model predictions. Finally, we discuss the results, and compare the performance achieved by the Generalized Binomial Model for predicting the Potential-of-Wallet, the Share-of-Wallet, and the Size-of-Wallet, with that of two benchmark models.

First, we apply the GBM model, using only transactional variables, and call it the Transactional-Generalized Binomial Model (Transactional-GBM). Using a RFM-like approach, the Size-of-Wallet (via $\mathbf{v}_i$) is only a function of the average savings, and the Share-of-Wallet (via $\mathbf{w}_i$) is only a function of customer recency. Next, in order to assess the extent to which socio-demographic variables could improve model performance, we combine all the variables presented in Table 1, and refer to this as the Full-Generalized Binomial Model (Full-GBM). Since a customer's Share-of-Wallet is a function of the interaction between this customer and the focal company, the predictor w_i explaining π_i contains only the two transactional variables. In contrast, a customer's Size-of-Wallet is also a function of this customer's socio-demographic characteristics. Hence, v_i, which explains λ_i, contains all four variables of Table 1.

Model Estimation

The model described in the ``Methodology'' section was implemented in Matlab 7.4. The likelihood was maximized by an interior-reflective Newton method (as described in Coleman and Li 1994, 1996).[3] This maximization method requires an expression for the likelihood gradient which we provide in the Appendix.

In order to test the significance of the explanatory variables, we apply a likelihood ratio test. Let L be the model's log-likelihood with all the explanatory variables, and L_j the model's log-likelihood with the variable j removed. We want to test whether the variable j contributes significantly to the model. The likelihood ratio test computes the test statistic $2(L-L_j)$, which is asymptotically chi-squared distributed, with one, two or

nine degrees of freedom.[4] Finally, for the categorical variable ``Address", the appropriate number of degrees of freedom is equal to nine. Table 4 reports the p-value of the significance test for each variable separately, and the average log-likelihood of the two models. As the results indicate, all variables contribute significantly to the models and cannot be excluded. The Full-GBM model is also significantly different from the Transactional model ($P < 0.0001$), showing that the socio-demographic variables are jointly significant in the GBM model.

While Table 4 indicates that all variables included are significant, it is also of interest to report the values of the estimated coefficients. For the Full-GBM model, ``average savings" variable is shown to have a positive effect on the SioW ($\hat{\beta} = 0.89$) and on the ShoW ($\hat{\alpha} = 0.07$), as one would expect. Indeed, customers with a larger amount of savings at the focal company are more likely to have both a larger Size-of-Wallet and a larger Share-of-Wallet. The ``recency" variable relates negatively to the SioW ($\hat{\beta} = -0.0015$), but positively to the ShoW ($\hat{\alpha} = 0.0042$). Furthermore, age has a positive effect on the SioW ($\hat{\alpha} = 0.002$). Finally, for the categorical variable ``address", the two most significative coefficients are the Upper Class and Immigrant dummies, with values of -0.11 and -0.06 for $\hat{\beta}$. This means that, all other variables (including past average savings) being fixed, customers living in Upper Class and Immigrant areas have a lower Size-of-Wallet at the retail bank of interest.

Model Validation

To assess the validity of the model assumptions, we implement a non-parametric bootstrap goodness-of-fit test.[5] We first compute the standardized residuals, defined as

$$r_i = \frac{x_i - \hat{\lambda}_i \hat{\pi}_i}{\sqrt{\hat{\lambda}_i \hat{\pi}_i}}. \tag{8}$$

Note that $\hat{\lambda}_i \hat{\pi}_i$ equals both the expected value and the variance of the Poisson distribution assumed by the GBM model for the number of transactions x_i. We then compute the summary statistic

$$X^2 = \sum_{i=1}^{N} r_i^2, \tag{9}$$

serving as a measure of the goodness-of-fit. A sample from the original data set with

replacements was drawn 1000 times. These new samples have N observations, as many as the original one, but an observation from the original data set can appear several times (bootstrap samples). We estimate the parameters of the GBM model for each bootstrap sample, and each time we compute a new set of residuals and the corresponding goodness-of-fit measure $X^{2,*}$. The percentage of times that a $X^{2,*}$ value exceeds the X^2 computed in (9) yields the p-value for testing the hypothesis that the model is correctly specified.

For our data set, we obtained p-values of 0.18 for the Transactional-GBM, and 0.25 for the full-GBM. Therefore, we can reasonably accept the validity of our model assumptions.

Baseline Models

To the best of our knowledge, no competing method is currently available for predicting the Size-of-Wallet or the Share-of-Wallet in the absence of survey data. We therefore propose two common sense methods as a benchmark.

The first baseline model assumes that all customers have a Share-of-Wallet equal to the focal company's market share. In our example, the market share is 10% (as provided to us by the company), and for every customer i, we set $\hat{\pi}_i = 10\%$ and $\hat{n}_i = 10x_i$. The best customers, according to their predicted Potential-of-Wallet, are therefore those with the largest observed savings. The disadvantage of such a baseline is that the Share-of-Wallet is constant and cannot be used to distinguish the customers.

The second baseline model we propose assumes that the expected Size-of-Wallet is the population mean. The Central Bank of Belgium provides the statistics required for computing this population mean (see BNB 2007). In the fourth quarter of 2006 (the time point when the prediction was made), the BNB reported 14451 euro in savings per inhabitant. This corresponds to 144.51 units of 100 euro. Since n_i is always greater than or equal to x_i, the second baseline states

$$\hat{n}_i = \begin{cases} x_i & \text{if } x_i \geq \bar{n}, \\ \bar{n} & \text{if } x_i < \bar{n}, \end{cases} \tag{10}$$

where $\bar{n}$ is the population mean of the n_i (equal to 144.51 transactions, as described

previously). The Share-of-Wallet is then

$$\hat{\pi}_i = \begin{cases} 1 & \text{if } x_i \geq \bar{n}, \\ x_i/\bar{n} & \text{if } x_i < \bar{n}, \end{cases} \tag{11}$$

While the advantage of this second baseline is that both $\hat{n}_i$ and $\hat{\pi}_i$ vary across customers, the disadvantage is that it requires knowing the population mean of the x_i, which may not be available in every empirical application.

Measures of Performance

In order to compare the four models (the full-GBM, the transactional-GBM, and the two baseline models presented in the previous section), we compared the predicted values, i.e. the values predicted by the model using the observable data, with the actual values from the survey data. The latter were kept out-of-sample during the estimation of the model.

Spearman's Correlation for the Potential-of-Wallet: the objective of this measure of performance is to assess the model's capacity to rank customers according to their Potential-of-Wallet, i.e. their business volume transactions outside the focal company. From an active customer relationship management perspective, this is the most useful value to predict. The first measure of performance is therefore Spearman's correlation between the predicted and reported PoW for each customer[6],

$$\rho_{PoW} = CORR(rank(n_i - x_i), rank(\hat{n}_i - x_i)). \tag{12}$$

Another advantage of Spearman's correlation is its robustness compared to the regular (Pearson's) correlation.

MAE of the Potential-of-Wallet: as a second measure of performance, we use the *Mean Absolute Error* (MAE) between the predicted PoW and the reported one. Note that this is also the MAE between the predicted and the reported SioW. We have

$$MAE_{PoW} = \frac{1}{N} \sum_{i=1}^{N} |n_i - \hat{n}_i|, \tag{13}$$

with N the number of observations in the sample.

Correlation for the Share-of-Wallet: since we are also interested in the quality of

the Share-of-Wallet prediction, the third measure of performance is the correlation between the predicted Share-of-Wallet and the reported one,

$$\rho_{ShoW} = Corr(\pi_i, \hat{\pi}_i). \tag{14}$$

MAE of the Share-of-Wallet: the fourth measure of performance is the MAE between the predicted Share-of-Wallet and the reported one,

$$MAE_{ShoW} = \frac{1}{N}\sum_{i=1}^{N}|\pi_i - \hat{\pi}_i|. \tag{15}$$

Spearman's Correlations of the Size-of-Wallet: the fifth measure of performance is Spearman's correlation between the predicted and the reported SioW,

$$\rho_{SioW} = CORR(rank(n_i), rank(\hat{n}_i)). \tag{16}$$

Gain per Customer: the last measure of performance gives an insight into the financial gains the company can generate thanks to the model's predictions. The model allows us to identify the top 1% customers with the highest potential. The customers with the highest potential can be identified in two ways: (i) those with the largest value of $(1 - Share - of - Wallet_i)$, (ii) those with the largest Potential-of-Wallet. Suppose that only this customer segment is targeted by a marketing campaign, and that it leads to an acquisition of 50% of their Potential-of-Wallet. If we assume a profit margin of 1%, this yields an average gain over the targeted customers of

$$Gain = \frac{1}{0.01N}\sum_{i \in \hat{T}}(n_i - x_i) \times 50\% \times 1\%, \tag{17}$$

with $\hat{T}$ the index set of all customers in the sample belonging to the top 1% customers with the highest potential as predicted by the model. This gain is expressed in euro per customer. This measure is very interesting from a managerial perspective, since it gives an insight into the profit the company could make by developing a plan of action according to the model predictions. In the next section, by way of comparison, we compute the gain generated by the model when customers are selected on the basis of their Potential-of-Wallet (Gain PoW), and the gain achieved when customers are

selected on the basis of their Share-of-Wallet (Gain ShoW).

Model Performance

Table 5 reports the models' performance according to the MAE and correlations defined in the previous section. Both GBM models perform satisfactorily, with the Full-GBM apparently the best model overall. The difference in MAE for the Potential-of-Wallet (thus also for the Size-of-Wallet) between the two GBM models and the baseline models is considerable. The Full-GBM and the Transactional-GBM achieve comparable performances for the Size and Potential-of-Wallet related measures. But for the ShoW-related metrics, the Full-GBM outperforms the Transactional-GBM, as well as both baselines.

When looking at the last column of Table 5, we see that the two GBMs and the first baseline model achieve high correlations, above 90%. These values are explained by the fact that 73% of the customers have a ShoW of 100%. But this good performance is less informative, since we are more interested in the 27% of other customers, who have a (high) Potential-of-Wallet and are more interesting from an acquisition perspective.

The baselines actually perform quite well taking into account their simplicity. The first baseline has a constant ShoW, while the second baseline has a constant SioW for many customers. Nevertheless, the first baseline achieves a high value for Spearman's correlation for the Potential-of-Wallet, and the second baseline offers a good prediction of the Share-of-Wallet.

We should emphasize that predicting the Customer Wallet is a difficult task, and that highly accurate predictions are not to be expected. For example, the smallest value of the MAE for the ShoW is 39%, which is still a high value given that the share is itself a percentage. But, as can be seen from Figure 1, most observed values of the ShoW are close to the extremes of 0% and 100%, such that a prediction with an MAE of 39% remains informative. The values for $MAE_{PoW}=MAE_{SioW}$ are also rather large, given that the sample average of the n_i (which is unknown if no survey data are used) is of the

same order of magnitude, namely 76.5. But it seems that, for this empirical application, the reported measures of accuracy are the best that can currently be achieved. Indeed, to the best of our knowledge, no competing methods have so far been proposed in the literature for estimating Size-of-Wallet and Share-of-Wallet, and the GBM model performs better than the Baseline methods.

Table 6 presents the average Gain in Euro (see equation (17)) for the customer segment with the 1% highest potential selected according to their Potential-of-Wallet (Gain PoW) or Share-of-Wallet (Gain ShoW), as predicted by the different models. We see that the gain per customer is larger when the targeted customer group is selected on the basis of their Potential-of-Wallet than when selected on the basis of their Share-of-Wallet. Hence, the Potential-of-Wallet is the preferred metric to use for targeting customers from an acquisition perspective. When comparing the gain achieved by the different models, we see from Table 6 that the Full-GBM outperforms all the other models, and is consequently the most interesting from a managerial perspective.

It is reasonable to wonder to what extent results could be improved if the model parameters β and α were estimated using the Size-of-Wallet and Share-of-Wallet reported in the survey. If survey data are used, the Poisson model (2) and the logistic regression model (4) can be estimated directly. It turns out that the estimates of the β parameters, appearing in the Size-of-Wallet equation (2), were quite close and all had the the same signs as the ones obtained without using survey data. For the α parameter, the estimates were more different, although they kept the same signs as the estimates obtained without using survey data. In order to assess to what extent this difference in parameter estimates affects model performance, we computed the resulting performance measures. As might be expected, when survey data are used, the performance is slightly better. But the fact that the difference in overall performance is small gives additional validity to the newly proposed model.

In Figure 4, we present histograms of the estimated Share-of-Wallet for the two GBM models. We observe that the variance of the predicted ShoW is much smaller for the Transactional-GBM than for the Full-GBM. The better predictive power of the Full-

GBM may be due to the use of socio-demographic explanatory variables, whereas the other models do not use them. We conclude that, overall, the newly proposed GBM model performs better than the baselines methods we constructed.

Conclusions

In a competitive environment, the Size-of-Wallet and the Share-of-Wallet are key measures of customer relationship management. The Share-of-Wallet is a measure of loyalty and the derived concept of Potential-of-Wallet (the difference between the Size-of-Wallet and the business volume conducted at the focal company) is a measure of the growth potential of customer profitability. Problematically, these two quantities are generally unobservable.

In this paper, we have presented a model for predicting the Size-of-Wallet, the Share-of-Wallet and the Potential-of-Wallet, without using survey data. This was done by introducing the Generalized Binomial Model. This model outperforms the baseline models we considered and overall performs satisfyingly. Although there is still room for further improvement, the GBM already generates substantial financial gains, clearly outperforming the baselines in this matter, as we showed in the section ``Measures of Performance''. We also found that the GBM using both transactional and socio-demographic explanatory variables achieves better results than the model containing only transactional variables. This confirms previous studies showing that socio-demographic patterns are important drivers of purchasing behavior.

The survey data used for assessing the model's predictive performance also yielded interesting results. The customers with the highest Size-of-Wallet were those with a Share-of-Wallet close to 25% or 75%, not those with only one service provider. We also pinpointed that the customers with the highest Potential-of-Wallet were most likely those with a ShoW close to 25%, at least in our empirical application.

The model we proposed for predicting customer wallet is simple, and the computer code needed for estimating the model is publicly available.[7] A limitation of the GBM model is that it makes strong assumptions on the data-generating process. The approach is fully parametric, and it is important to validate the model after its estimation. A possible way to do this, using a bootstrapped goodness-of-fit test, is outlined in the ``Model Validation'' section. Estimating Share-of-Wallet and Size-of-Wallet in the absence of survey data does not seem feasible without making strong model

assumptions. To the best of our knowledge, no other paper deals with the problem of estimating Size-of-Wallet and Share-of-Wallet in the absence of survey data.

The proposed methodology was applied to only one empirical application within the banking context, limiting its generalizability. Hence, it is not inconceivable that that our findings may be specific to our data context. The problem of estimating the Size-of-Wallet and Share-of-Wallet is of major interest in other business sectors as well, for instance in the retail grocery sector. To test for the validity of the approach in other contexts, however, one should not only have another transactional customer database, but one should also have access to information on the value of the Share-of-Wallet, which would typically require survey data. Since survey data were only at our disposal for one particular application, we were unable to validate the predictions based on the GBM model for other products or businesses and leave this for future research.

APPENDIX

In this section, we develop the equations allowing us to compute the gradient of the log-likelihood function used during the Maximum Likelihood Estimation process. Equation (6) states that the log-likelihood is

$$\log L(\alpha,\beta) = \sum_{i=1}^{N}\left(\lambda_i\pi_i - x_i\log(\lambda_i\pi_i) + \log(x_i!)\right), \tag{18}$$

where λ_i and π_i are functions of β and α respectively. The derivative with respect to β is

$$\frac{\partial \log L(\alpha,\beta)}{\partial \beta} = \sum_{i=1}^{N}\left(\pi_i\frac{\partial\lambda_i}{\partial\beta} - x_i\frac{1}{\pi_i\lambda_i}\pi_i\frac{\partial\lambda_i}{\partial\beta}\right).$$

From equation (2), $\lambda_i = \exp(\beta'\mathbf{v_i})$, it follows that

$$\frac{\partial \log L(\alpha,\beta)}{\partial \beta} = \sum_{i=1}^{N}\left(\pi_i\lambda_i - x_i\right)\mathbf{v_i}. \tag{19}$$

The derivative of (18) with respect to α is

$$\frac{\partial \log L(\alpha,\beta)}{\partial \alpha} = \sum_{i=1}^{N}\left(\lambda_i\frac{\partial\pi_i}{\partial\alpha} - x_i\frac{1}{\lambda_i\pi_i}\lambda_i\frac{\partial\pi_i}{\partial\alpha}\right).$$

The logit function has the property

$$\frac{\partial\pi_i}{\partial\alpha} = \pi_i(1-\pi_i)\mathbf{w_i},$$

and with (5), we have

$$\frac{\partial \log L(\alpha,\beta)}{\partial \alpha} = \sum_{i=1}^{N}(\lambda_i\pi_i - x_i)(1-\pi_i)\mathbf{w_i}. \tag{20}$$

The two derivatives (19) and (20) provide the gradient used during the maximum likelihood estimation.

Finally, the Hessian Matrix is easy to compute and equals

$$\frac{\partial^2 \log L(\alpha,\beta)}{\partial\alpha\partial\beta} = \sum_{i=1}^{N}\pi_i(1-\pi_i)\lambda_i\mathbf{w_i}\mathbf{v_i}'. \tag{21}$$

Notes

[1] In the empirical application we use the log-transform of the average savings.

[2] The respondents have all been customers of the financial institution. At present, they may have no savings account at the focal company, but they purchased at least one product of this company in the past, and the bank consequently has information on these customers.

[3] The function *fminunc* from the Matlab Optimization Toolbox is used for the Maximum Likelihood Estimation.

[4] For instance, if the age variable is removed, the number of degrees of freedom is one, since age is only used for the estimation of the λ_i. If ``average savings" is removed, then the number of degrees of freedom is two, because this variable enters both in the equation for λ_i and π_i.

[5] The standard reference for bootstrap methods is Davison and Hinkley (2003).

[6] Customers with no savings at the focal company have no attributed Size-of-Wallet. Therefore, we discarded customers with a reported Share-of-Wallet smaller than 10%. Similarly for computing the MAE of the PoW and Spearman's correlation of the SioW.

[7] The Matlab code is available at http://www.econ.kuleuven.be/nicolas.glady/public.

References

Baumann, Chris, Suzan Burton, Greg Elliott (2005), Determinants of customer loyalty and share of wallet in retail banking. *Journal of Financial Services Marketing* 9(3) 231-248.

BNB (2007), Banque Nationale de Belgique - Nationale Bank van Belgie - Belgostat. Http://www.nbb.be/belgostat/startSDW.do.

Buckinx, Wouter, Dirk Van den Poel (2005), Customer base analysis: Partial defection of behaviorally-loyal clients in a non-contractual fmcg retail setting. *European Journal of Operational Research* 164(1) 252-268.

Buckinx, Wouter, Geert Verstraeten, Dirk Van den Poel (2007), Predicting customer loyalty using the internal transactional database. *Expert Systems with Applications* 32(1) 125-134.

Coleman, T.F., Y. Li (1994), On The Convergence of Reflective Newton Methods for Large-Scale Nonlinear Minimization Subject to Bounds. *Mathematical Programming* 67(2).

Coleman, T.F., Y. Li (1996), An Interior Trust Region Approach for Nonlinear Minimization Subject to Bounds. *SIAM Journal on Optimization* 6.

Davison, A.C., D.V. Hinkley (2003), *Bootstrap Methods and their Applications*. Cambridge: Cambridge University press.

Du, Rex Yuxing, Wagner A. Kamakura, Carl F. Mela (2007), Size and share of customer wallet. *Journal of Marketing* 71(2) 94-113.

Garland, Ron (2004), Share of wallet's role in customer profitability. *Journal of Financial Services Marketing* 8(3) 259-268.

Green, William H (2003), *Econometric Analysis*, 5th ed. New York: Prentice Hall.

Gupta, Sunil, Valarie A. Zeithaml (2006), Customer metrics and their impact on financial performance. *Marketing Science* 25(6) 718-739.

Keiningham, Timothy L., Tiffany Perkins-Munn, Heather Evans (2003), The impact of customer satisfaction on share-of-wallet in a business-to-business environment. *Journal of Service Research* 6(1) 37-50.

Magi, Anne W. 2003. Share of wallet in retailing: the effects of customer satisfaction, loyalty cards and shopper characteristics. *Journal of Retailing* 79(2) 97-106.

Mittal, Vikas, Wagner A. Kamakura (2001), Satisfaction, repurchase intent, and repurchase behavior: Investigating the moderating effect of customer characteristics. *Journal of Marketing Research* 38(2) 131-142.

Perkins-Munn, Tiffany, Lerzan Aksoy, Timothy L. Keiningham, Demitry Estrin (2005), Actual purchase as a proxy for share of wallet. *Journal of Service Research* 7(3) 245-256.

Shah, Denish, Roland T. Rust, A. Parasuraman, Richard Staelin, George S. Day (2006), The path to customer centricity. *Journal of Service Research* 9(2) 113-124.

Wirtz, Jochen, Anna S. Mattila, May Oo Lwin (2006), How effective are loyalty reward programs in driving share of wallet. *Journal of Service Research* 9(4) 327-334.

Zeithaml, Valarie A (2000), Service quality, profitability and the economic worth of customers: What we know and what we need to learn. *Journal of the Academy of Marketing Science* 28(1) 67-85.

Biography

Nicolas Glady is a doctoral candidate at the Faculty of Business and Economics at K.U. Leuven (Belgium). His research interests lie in data-based models to support marketing decisions. Current projects focus on the development of models for customer-base analysis, with a particular interest in Customer Lifetime Value, Share-of-Wallet and Churn Prediction.

Christophe Croux is Professor of Statistics and Econometrics at K.U. Leuven (Belgium). His research interests are robust statistics, multivariate data analysis, computational statistics, applied time series analysis, and marketing modeling. His work has been published in Biometrika, Journal of the American Statistical Association, Journal of Marketing Research, and The Review of Economics and Statistics, among others.

Table 1: Typology of the variables used in this study.

	Transactional	Socio-Demographic
Wealth Related	Average Savings	Address, Age
Wealth Unrelated	Recency	Age

Table 2: Description of the categorical variable accounting for the address.

Label	Area Type
1	Upper Class
2	High Standing
3	Semi-Rural
4	Urban
5	Middle Class
6	Factory Worker
7	Industrial Area
8	Rural
9	Immigrant
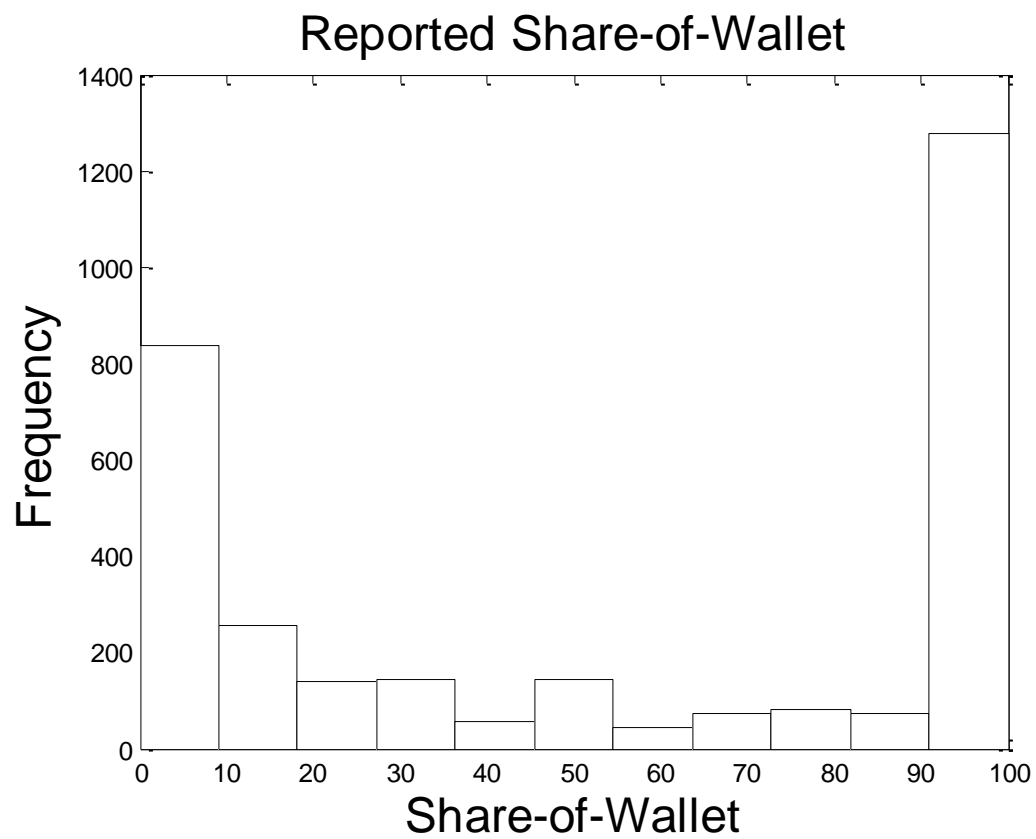
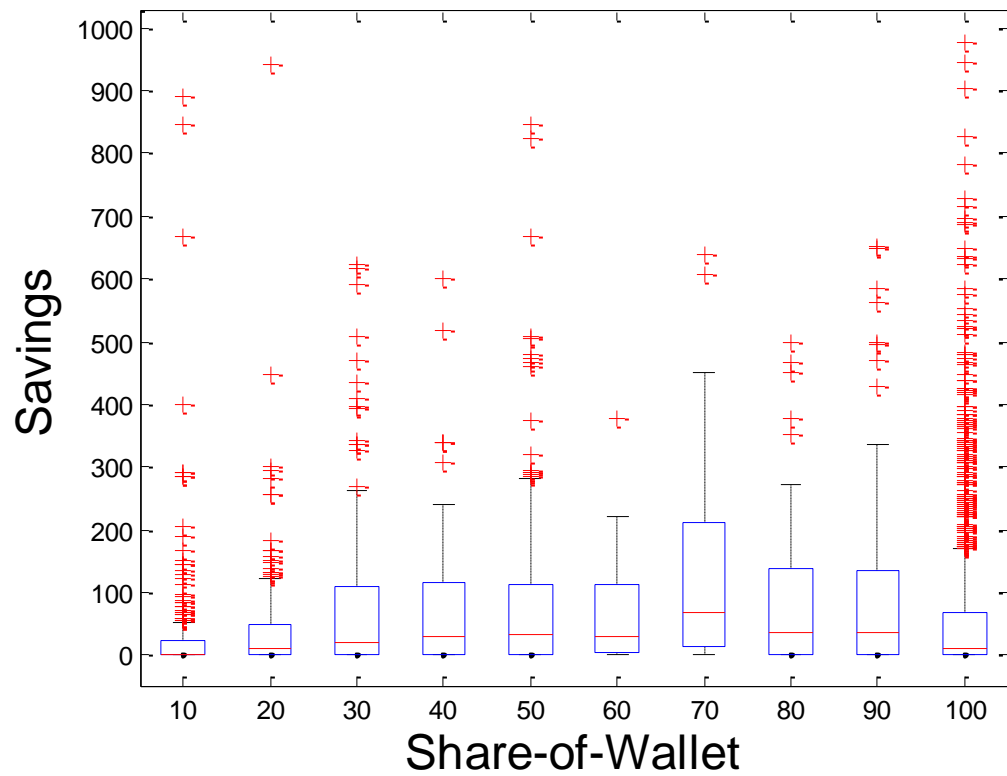Figure 1: Histogram of the reported Share-of-Wallet over a sample of 3125 customers.

Figure 2: Boxplot of the savings of 3125 customers, as a function of the Share-of-Wallet, divided in 10 categories, ranging from 10% to 100%.
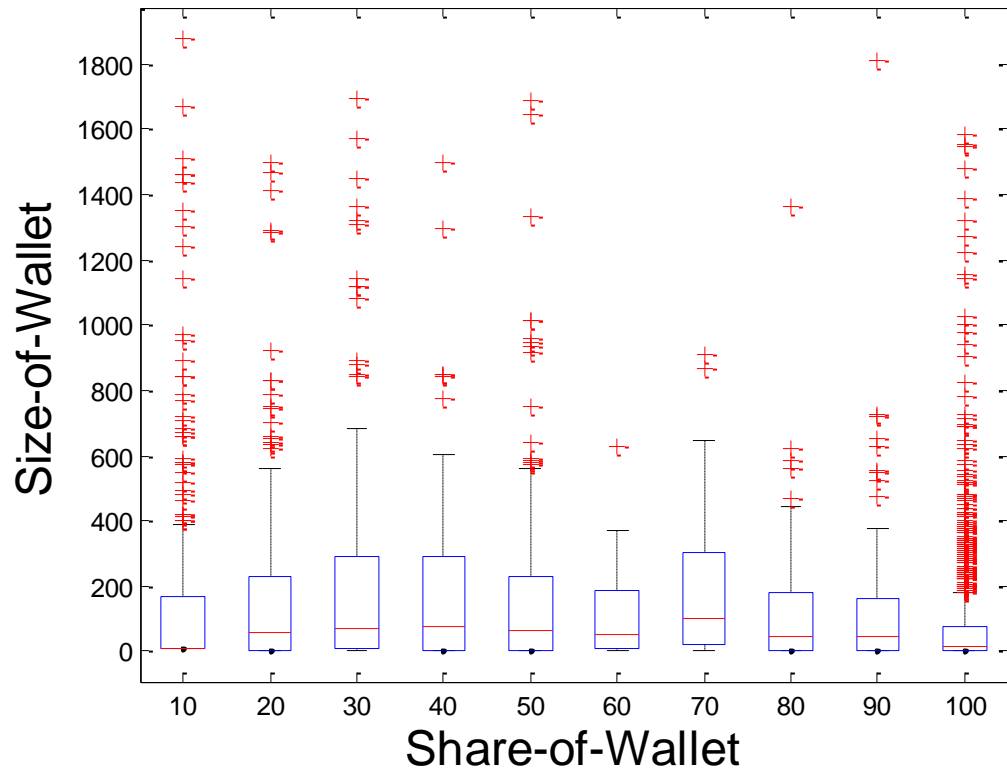
Figure 3: Boxplot of the Size-of-Wallet of 3125 customers as a function of the Share-of-Wallet, divided in 10 categories, ranging from 10% to 100%.

Table 3: Correlation between the explanatory variables and the Size and Share-of-Wallet.

Variables	Size-of-Wallet	Share-of-Wallet
Average Savings	0.27	0.35
Age	0.12	-0.12
Recency	0.03	0.21
Upper Class (Dummy)	0.0426	-0.0369
High Standing (Dummy)	0.0372	-0.0251
Semi-Rural (Dummy)	-0.0045	0.0091
Urban (Dummy)	-0.0190	0.0065
Middle Class (Dummy)	-0.0152	-0.0018
Factory Worker (Dummy)	-0.0287	-0.0025
Industrial Area (Dummy)	-0.0063	0.0597
Rural (Dummy)	-0.0274	-0.0172
Immigrant (Dummy)	-0.0239	0.0123

Table 4: Significance tests of the variables in the Full and Transactional-Generalized Binomial Model. P-values are reported, as well as the average log-likelihood of the two models.

Variables	Transactional-GBM	Full-GBM
Average Savings	< 0.001	< 0.001
Recency	< 0.001	< 0.001
Address	N.A.	< 0.001
Age	N.A.	< 0.001
Average Log-Likelihood	-2.97	-2.94

Table 5: The Correlation and the MAE between the predicted and reported values for the Potential, Share, and Size-of-Wallet.

Model	ρ_{PoW}	MAE_{PoW}	ρ_{ShoW}	MAE_{ShoW}	ρ_{SioW}
Full-GBM	27.29 %	94.69	29.23 %	0.39	91.46 %
Transactional-GBM	28.69 %	106.38	20.93 %	0.42	92.66 %
Baseline 1	29.15 %	589.22	0.00 %	0.49	93.56 %
Baseline 2	-8.42 %	155.51	26.48 %	0.41	67.87 %

Table 6: Average Gain in euro for the segment of customers with the 1% highest potential, selected according to their Potential-of-Wallet (Gain PoW) or Share-of-Wallet (Gain ShoW), as predicted by the different models.

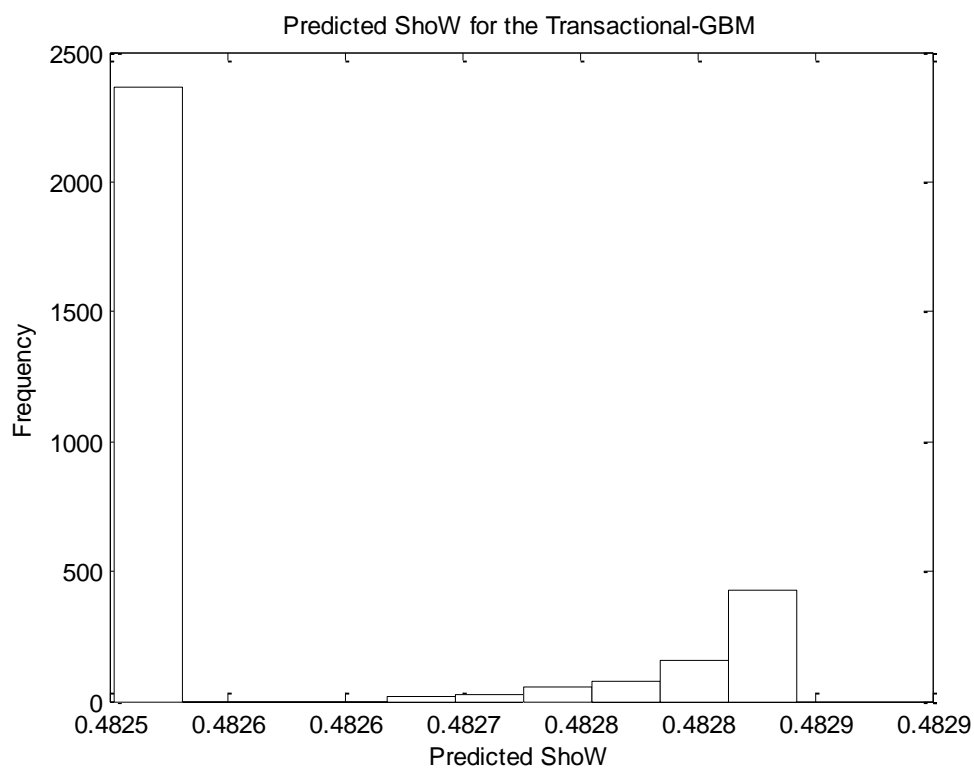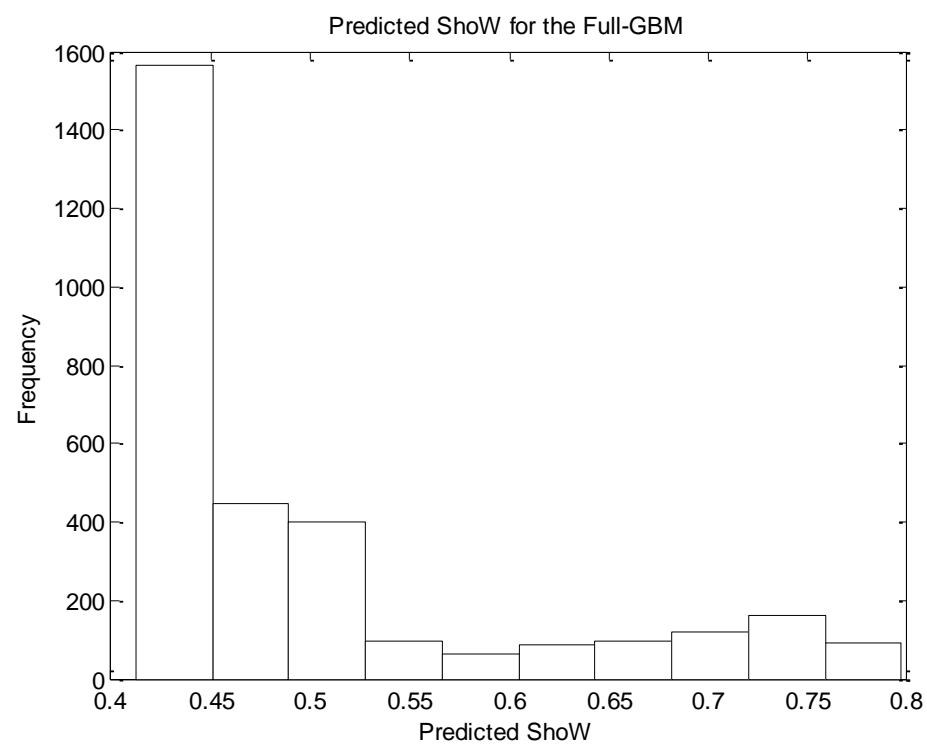| Model	Gain PoW	Gain ShoW
Full-GBM	327.16 €/Customer	109.03 €/Customer
Transactional-GBM	210.23 €/Customer	35.72 €/Customer
Baseline 1	190.14 €/Customer	38.24 €/Customer
Baseline 2	3.17 €/Customer	1.03 €/Customer

Figure 4: Histograms of the predicted Share-of-Wallet based on the Full-GBM (top figure) and the Transactional-GBM (bottom figure).