

PREDICTIVE PERFORMANCE OF CUSTOMER LIFETIME VALUE MODELS IN E-COMMERCE AND THE USE OF NON-FINANCIAL DATA

Pavel Jasek, Lenka Vrana, Lucie Sperkova, Zdenek Smutny,
Marek Kobulsky*

Abstract

The article contributes to the knowledge of customer lifetime value (CLV) models, where extensive empirical analyses on large datasets from online stores are missing. Based on this knowledge, practitioners can decide about the deployment of a particular model in their business and academics can design or enhance CLV models. The article presents predictive performance of selected CLV models: the extended Pareto/NBD model, the Markov chain model, the vector autoregressive model and the status quo model. Six large datasets of medium and large-sized online stores in the Czech Republic and Slovakia are used for a comparison of the predictive performance of the models. Online stores have annual revenues in the order of tens of millions of euros and more than one million customers. The comparison of CLV models is based on selected evaluation metrics. The results of some of the models which use additional non-financial data on customer behaviour – the Markov chain model and the vector autoregressive model – do not justify the effort which is needed to collect such data. The advantages and disadvantages of the selected CLV models are discussed in the context of their deployment.

Keywords: CLV models, forecasting, online marketing management, e-commerce, methodology, online shopping, online marketing.

JEL Classification: C53, C55, M21, M31

1. Introduction

Models of the customer lifetime value (CLV) primarily allow companies to better allocate resources and formulate strategies in marketing (Ferrentino *et al.*, 2016). Companies use CLV mainly in the area of customer segmentation to build long-term relationships with customers and effectively manage investments in marketing activities (Weng and Huang, 2018). However, use of CLV can help answer a number of very different questions such

* Pavel Jasek, Faculty of Informatics and Statistics, University of Economics, Prague, Czech Republic (pavel.jasek@vse.cz);

Lenka Vrana, Faculty of Informatics and Statistics, University of Economics, Prague, Czech Republic (lenka.vrana@vse.cz);

Lucie Sperkova, Faculty of Informatics and Statistics, University of Economics, Prague, Czech Republic (lucie.sperkova@vse.cz);

Zdenek Smutny, Faculty of Informatics and Statistics, University of Economics, Prague, Czech Republic (zdenek.smutny@vse.cz);

Marek Kobulsky, Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University in Prague, Prague, Czech Republic (MerrekSK@gmail.com).

This article was prepared thanks to the grant of the University of Economics, Prague, under Grant agreement number F4/18/2014.

as decisions related to retaining and acquiring customers, investments in product image, or what is the company's long-term value (Lin *et al.*, 2017; Haenlein *et al.*, 2006).

Several models for different conditions, environments, and relationships have been proposed in the last decades for marketing purposes. The development of information and communication technology (ICT) and its acceptance in society have given rise to e-commerce. In Europe and the US, it is a rapidly growing retail market (Centre for Retail Research, 2017; Lone, 2017; Postnord, 2016). Online shopping, with a high level of competition, is a significant part of e-commerce that offers retail sales directly to end customers. E-commerce companies operating online stores have high data availability thanks to customers' interaction in their stores, *e.g.*, viewing, buying and discussing products (Chiang and Yang, 2018; Rathi and Betala, 2019). This data can be very useful for implementing a CLV model that uses available historical data to estimate CLV, which can help a company spend financial resources on marketing activities as efficiently as possible. However, e-commerce companies face issues of selecting the appropriate CLV model and data on interactions mainly in the case of enriching composite model types.

CLV models have been traditionally divided into six groups according to the modelling approach (Estrella-Ramón *et al.*, 2013; Gupta *et al.*, 2006). In this study, three of these model groups are represented by the following models: a probability model with RFM factors (extended pareto/NBD model), an econometric model (Markov chain model) and a persistence model (vector autoregressive model). These models are suitable for estimating customer value under online shopping conditions at the aggregated customer base level – see also Section 2.1. Apart from direct benefits for business practitioners' decision-making process on CLV modelling selection, this study also has significant value from a methodological perspective.

There are a number of review articles covering theoretical aspects of CLV modelling approaches, *e.g.* Kumar (2018), Estrella-Ramón *et al.* (2013), Chang *et al.* (2012), Fader and Hardie (2009), Damm and Monroy (2011). However, there are still not enough empirical analyses (Donkers *et al.*, 2007; Batislam *et al.*, 2007; Platzer & Reutterer, 2016; Chamberlain *et al.*, 2017) and very few of them are focused on online shopping. Current literature on online shopping mainly uses one dataset only. Therefore, by comparing multiple datasets and modelling approaches, this paper has a strong unique proposition to enhance the current knowledge of predictive performance of CLV models. This study presents comparable results of the deployment of selected CLV models using six different datasets. New insights help improve the knowledge base on predictive performance of individual models. Finally, this knowledge can be used by academics to design new models or enhance older ones (Gupta *et al.*, 2006).

The aim of the article is to make a comparison of the predictive ability and quality of selected CLV models based on statistical metrics. This article also helps companies to make decisions concerning the selection and application of a CLV model with a focus on the use of non-financial data. The article proposes two research questions based on managerial issues:

Which of the compared models for calculating CLV have good predictive performance for CLV in the purchase environment with customer relationships typical of online shopping as a part of e-commerce?

Good predictive performance is observed when stable quality prediction results can be achieved among all datasets used based on evaluation metrics and outperform the other models compared in this study.

Can non-financial data on customer behaviour improve the predictive performance of the selected CLV models in comparison with CLV models that use only financial data from online shopping as a part of e-commerce?

From a managerial point of view, there are two reasons for the second question. The first is higher financial costs of obtaining additional data (e.g., data from external analytical services such as Google Analytics or social media) beyond the basic transaction data that are commonly available to online retailers. The second reason is a limited ability of these companies to process large amounts of internal and external data into the form usable by the selected model.

This research methodically builds on our previous research presented in Jasek *et al.*, 2018. The remainder of the article proceeds as follows. Section 2 presents the methodology, selected CLV models, statistical metrics used and a description of datasets. The results and answers to the research questions are introduced in Section 3 and then discussed in Section 4.

2. Methodology and Data Collection

The research is conducted in several phases:

1. The desired research objectives are set in the initial phase, where the research question is formulated. The initial phase also justifies the necessity and appropriateness of the proposed research – see Section 1.
2. The second phase contains the process of selecting various suitable CLV models to be used by e-commerce businesses engaged in online shopping. At this stage, the implementation of the models is also performed according to the description in Section 2.1.
3. In the third phase, requirements and criteria for the data are defined, based on the models that were selected. They allow the researcher to determine the structure and timeframe of the data needed to perform the research – see Section 2.2.
4. The fourth phase describes collection of data with the required structure from various e-commerce companies. Furthermore, datasets meeting the proposed conditions are then pre-processed according to the needs of the individual CLV models. The data pre-processing is described in Section 2.2. Section 2.3 describes the datasets from the various e-commerce subjects.
5. The fifth phase of the research evaluates the selected CLV models with respect to statistical metrics listed in Section 2.4. In this section, the comparison process performed is comprehensively described, including the definition of a training and testing period.

6. In the final phase of the research, the research question is solved first in Section 3.1, then the results obtained are subjected to a wider discussion, also considering relevant managerial decision-making implications, in Section 4.

2.1 Selection of models for comparison and their description

In this study, probability, econometric and persistence models were selected, as they are suitable for online shopping conditions in e-commerce business. For the comparison, three previously published models were selected: the extended Pareto/NBD model (Fader *et al.*, 2005), the Markov chain model (Haenlein *et al.*, 2007; Paauwe *et al.*, 2007), and the vector autoregressive model of persistence (Villanueva *et al.*, 2008). The choice of the specific individual models is always debatable, whether it is the number of the models or their nature. In this research, the authors tried to use models as different as possible to exploit the intriguing aspect of CLV evaluation, which is that so many different models exist in the first place. Hence, the most interesting differing representatives for calculating CLV were chosen. It is also notable that the CLV models used have not yet been compared together on the same dataset(s). The status quo model was added to represent the naïve approach to calculating CLV, and to make the other individual approaches easier to compare.

Nevertheless, the chosen models satisfy the following conditions of suitability, which are related to the purchase environment and the customer relationships typical of e-commerce (Jasek *et al.*, 2018):

- **Non-contractual relation:** Customers are not contractually bound, and it is only up to them whether and when they make a purchase from the given retailer.
- **Non-membership:** Customers do not have to be members of a club. Many retailers have their loyalty schemes, but with regard to selecting a model, there should be a universal approach to customers, lifting this prerequisite.
- **Always-a-share:** A customer who stopped shopping can return at any time.
- **Variable-spending environment:** The retailer offers a broad portfolio of products at varying prices (the opposite of a specialised shop focusing on a single core product).
- **Continuous:** The customer can make a purchase anytime, repeatedly and several times a day.

Status quo model

Given that the focus of this article is to compare the performance of different types of models, it seems appropriate not only to compare models among themselves but also to designate a baseline status quo model. This model should be the simplest to calculate. With such a baseline defined, it is then possible to compare the true benefits of the more sophisticated methods tested in this article.

Donkers *et al.* (2007) describe a status quo model used in a contractual environment as a model based not on shopping behaviour, but on the retention and total profit per customer. Their status quo model is defined as

$$Profit_{i,t+1} = Profit_{i,t} \quad (1)$$

where $Profit_{i,t}$ is the profit from customer i in time t , and therefore assumes consistent profit over time.

Given that a suitable alternative for non-contractual relations might not be found, the model described above was modified. The status quo model in this paper has two prerequisites:

- 1) A customer who has not made a purchase for more than a year is considered inactive.
- 2) Active customers are assumed to make a purchase every following week that has the same value as their average weekly purchase in the last year of the period (52 weeks).

These assumptions require an addition to the status quo model – if the customer was active in the last year of the period, they are not expected to leave during the period that is being forecast. Thus, the model can be written as

$$Profit_{i,p+j} = \frac{\sum_{t=p-52}^{p-1} Profit_{i,t}}{52} \quad j = 0, 1, 2, 3, \dots, h-p, \quad (2)$$

where p is the threshold of the prediction and h is its horizon. The period of one year is chosen only as a rule of thumb to keep the described model as simple as possible. In practice, analysts often define customers as leaving just after their inactivity exceeds one year (Chamberlain *et al.*, 2017), which is why it was chosen here as well.

This model does not aim to predict the most accurate CLV, but thanks to its results the true benefits of other methods can be better assessed. It is possible that the performance of another model would be comparable (or even lower) than the results of the status quo model, even though the other models use advanced statistical methods or require wider data sources.

Extended Pareto/NBD model

The extended Pareto/NBD model (EP/NBD) introduced in Schmittlein and Peterson (1994) is an extension of the original Pareto/NBD model (Schmittlein *et al.*, 1987) using revenue data of individual orders. By stating that recency and frequency are independent of monetary value, it becomes necessary to solve two submodels: one for the expected number of transactions, the other for the expected average order value. Multiplying the results from the individual submodels yields the CLV. Therefore, the EP/NBD model consists of a Pareto/NBD submodel and a gamma-gamma spending submodel. This article uses the gamma-gamma spending submodel described in Fader *et al.* (2005), further clarified in Fader and Hardie (2013). Since Fader *et al.* (2005) provide a thorough description of the approach with all the appropriate derivations, it is deemed unnecessary to describe it here any further.

Markov chain model with decision tree learning

The application of the Markov chain model (MC) for CLV calculation was described mainly by Pfeifer and Carraway (2000) and further developed by others – *e.g.*, Haenlein *et al.* (2007) and Paauwe *et al.* (2007). The MC model oftentimes uses two approaches

to define states. The first approach operates with RFM variables as suggested by Pfeifer and Carraway (2000), *i.e.*, recency (time elapsed since the last purchase), frequency (total number of purchases) and monetary value (total generated income) of a customer. Different values of recency are then used for creating individual states. The second approach takes into account other profitability drivers, such as age, demographics/lifestyle, type and intensity of product ownership and engagement level as predictor variables. It uses classification and regression tree (CART) analysis for segmentation into individual subgroups (states) (Haenlein *et al.*, 2007). In the MC model used in this comparative study, the following profitability drivers were considered to define states: region, time since last purchase (recency), number of purchases (frequency), marketing traffic sources, average day of the order, month of the order, order rank, and delivery price. Monetary value is used as the dependent variable in the decision tree to separate the possible customer states.

The CLV of a customer is defined as the discounted sum of state-dependent contribution margins, weighted by their corresponding transition probabilities. Given the fact that this concerns a first-order MC model, the limitation of this approach is the dependence of transition probabilities only for behaviour during the latest period (although the behaviour could also have been influenced by earlier periods). Another limitation is the assumption that the transition matrix remains stable and constant in time. This solution is suitable mainly for medium-term forecasts (Haenlein *et al.*, 2007).

Vector autoregressive model

The vector autoregressive model (VAR) is a method used most often for multivariate time series analysis, similar to multivariate multiple regression. The model is well described in specialised literature (see Tsay, 2013). The application of the VAR model to customer equity predictions is described in Villanueva *et al.* (2008). They researched the impacts of customer acquisition on a company's performance. Their model focused on capturing the dynamic relationships between three time series: the number of customers acquired by marketing actions, the number of customers acquired by word of mouth and the firm's performance. This model was modified here and instead of marketing actions and word of mouth it uses the division according to Burcher (2012) into earned media (EM) – *e.g.*, social networking services, paid media (PM) – *e.g.*, pay-per-click campaigns, and owned media (OM) – *e.g.*, organic search. The reason for this modification was a clearer and easier segmentation of customers according to traffic sources.

The model is designed as a classical VAR(p) model. It captures the dynamic relationships between four stationary time series: the number of customers acquired by EM, PM, OM and the firm's performance (*VALUE*). We use differencing (once or more times) to stationarize the series if necessary. The model is defined as

$$\begin{pmatrix} EM_t \\ PM_t \\ OM_t \\ VALUE_t \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{pmatrix} + \sum_{l=1}^p \begin{pmatrix} a_{11,l} & a_{12,l} & a_{13,l} & a_{14,l} \\ a_{21,l} & a_{22,l} & a_{23,l} & a_{24,l} \\ a_{31,l} & a_{32,l} & a_{33,l} & a_{34,l} \\ a_{41,l} & a_{42,l} & a_{43,l} & a_{44,l} \end{pmatrix} \begin{pmatrix} EM_{t-l} \\ PM_{t-l} \\ OM_{t-l} \\ VALUE_{t-l} \end{pmatrix} + \begin{pmatrix} e_{1,t} \\ e_{2,t} \\ e_{3,t} \\ e_{4,t} \end{pmatrix} \quad (3)$$

where t stands for time, the vector $(c_1 \ c_2 \ c_3 \ c_4)'$ contains the constant terms and the vector $(e_{1,t} \ e_{2,t} \ e_{3,t} \ e_{4,t})'$ contains the error terms with Gaussian white noise properties. The VAR(1) model in this form can describe the following relationships:

- *direct effects* of acquisition on the firm's performance (coefficients $\alpha_{41,t}$, $\alpha_{42,t}$ and $\alpha_{43,t}$),
- *cross-effects* between three types of customer acquisition (coefficients $\alpha_{12,t}$, $\alpha_{13,t}$, $\alpha_{21,t}$, $\alpha_{23,t}$, $\alpha_{31,t}$ and $\alpha_{32,t}$),
- *feedback effects*, which state how the firm's performance affects acquisition in subsequent time periods (coefficients $\alpha_{14,t}$, $\alpha_{24,t}$ and $\alpha_{34,t}$),
- *reinforcement effects*, where the value of series in time t affects its future values, e.g., customers acquired by word of mouth would spread positive information about the firm, which would lead to more acquisitions (coefficients $\alpha_{11,t}$, $\alpha_{22,t}$, $\alpha_{33,t}$ and $\alpha_{44,t}$).

Most of the analysed time series show seasonal fluctuations; see Figure 2. However, the seasonality is also present in the EM, PM and OM series. Therefore, the seasonal effects in the VALUE series could be explained by the seasonal movements of the other series or by its lagged values. We also created a VAR model with exogenous seasonal dummy variables (VARX): the exogenous variables may affect the four analysed series, but the analysed series cannot affect the exogenous variables. For more information, see Bierens (2004). Our VARX is specified without the constant terms vector and with 12 seasonal dummy variables – each for one month. We cannot use weekly seasonal dummy variables, because each year can have a different number of weeks and holidays (such as Christmas) can occur in different weeks. However, we suppose that the analysed series are too short to enable training the seasonal factors and that the VAR model without the seasonal factors will provide better results.

To determine the new acquisitions made by each observed channel, the first year of training data serves only as a customer history. If a customer makes her first purchase after this year, she is considered a newly acquired customer. On the other hand, if she makes the first purchase during the first year, she is counted as an existing customer. This is necessary as we do not have the full history of each retail store or an indicator of newly registered clients. Unfortunately, this means that the training data are 52 weeks shorter for the VAR and VARX models.

As the VAR (and VARX) model is based on aggregated time series, its results are also aggregated. It is possible to describe the relationships between the input time series and predict the output value for each week, but the output value for each customer cannot be predicted. The impulse response function can analyse the general influence of a customer newly acquired by paid media and compare it with one acquired by earned media, but it cannot predict the value of a specific customer or evaluate which customer will be among the top 10%. In spite of that, the VAR model was included in this paper, as the authors considered models based on time series to be one of the principal modelling approaches. This empirical analysis acknowledges that the VAR model cannot be used for individual customer-level predictions.

2.2 Data collection and pre-processing

The models selected in Section 3.1 require specific data features and structure. To perform a successful comparison, a definition of required data was published in a publicly available call for data. Several medium and large-sized online stores in the Czech Republic and Slovakia were asked to participate in this research and provide their data. Six datasets met the conditions required for participation in the study, which included non-contractual business-to-consumer business settings, minimum volume of hundreds of transactions per month, and a significant degree of repeat purchases.

The data pre-processing included (i) descriptive analysis, (ii) data cleaning and (iii) selection of a feature subset from the datasets for individual models. On the basis of the descriptive analysis output and in cooperation with the given e-commerce company, the identified outliers were removed. All the datasets were cleaned to include only purchases from the same country as determined by the most frequent common country. Individual datasets were trimmed to whole weeks at the beginning and end because the week is a suitable basic unit for prediction that can be used by all models. All the datasets were aggregated at a weekly level with aggregation details described below. From these data, the datasets for individual models were then created for comparison.

Table 1 | Five randomly selected rows from the complete dataset, consisting of the minimal dataset and data added specifically for the MC and VAR models

| dataset (online store) | Common data used by all four models (minimal dataset) | | | | Data specifically added for MC and VAR models | | | | | | | Data only for MC model |
|------------------------|---|-------------|-------------|------------|---|--------------|------------------|-------------------|---------------|--------------------------|-------------------------|------------------------|
| | customer_id | week_number | monday_date | profit_EUR | channel_poe | channel_type | medium_source | avg_pur-chase_day | item_quantity | transaction_shipping_EUR | transaction_revenue_EUR | zip_firstchar |
| B | 282006 | 98 | 2014-11-17 | 7.69 | owned | email | email_newsletter | 3.0 | 1 | 1.19 | 23.90 | 1 |
| D | 65298 | 302 | 2014-09-15 | 11.84 | earned | organic | organic_google | 4.5 | 1 | 0.00 | 42.96 | 8 |
| F | 1182543 | 158 | 2013-01-07 | 7.37 | paid | email | email_newsletter | 4.0 | 4 | 0.00 | 33.52 | 3 |
| F | 883193 | 103 | 2011-12-19 | 3.48 | paid | cam-paign | cpc_google | 4.0 | 1 | 0.00 | 37.74 | 5 |
| F | 1349757 | 197 | 2013-10-07 | 6.96 | paid | feed | feed_heureka | 2.0 | 3 | 0.00 | 62.00 | 6 |

Source: Authors

The selected models were ordered by the number of features considered to be inputs: the status quo, EP/NBD, MC and VAR model, where all the models share a minimal dataset,

and the latter two use additional data from the complete processed dataset. Table 1 presents the dataset by showing randomly selected rows. Letters A-F are used to denote individual online store datasets. The records are based on each customer's weekly purchases, where the customer identifier *customer_id* is unique to each dataset, and the weeks (shown as *week_number*) correspond to the dataset's data range counting from 1. A specific week is also characterised by *monday_date*. *Profit_EUR* is the amount of gross profit attributed to these purchases. As several purchases could be made during one week, *avg_purchase_day* demonstrates the average of individual transaction weekdays (1 = Monday). Traffic sources are attributed to the most frequent or first source within the week – *channel_poe* stands for the paid, owned and earned media categorization according to Burcher (2012), *channel_type* and *medium_source* are detailed information about the traffic source. All the other purchase information is aggregated: *item_quantity* is the quantity of all products purchased, *transaction_shipping* is the amount spent on shipping, *transaction_revenue_EUR* is the total amount spent by a customer including VAT. Regional information about purchase delivery addresses was compressed into *zip_firstchar* as a first character of the postcode. There are only ten unique first postcode characters in the Czech Republic and Slovakia, so such a categorical variable provides sufficient details for the decision tree in the MC model.

2.3 Description of datasets

For the purposes of collecting data and carrying out a comparative analysis, selected online stores in the Czech Republic and Slovakia were addressed. The obtained data had to meet the criteria defined in Section 2.2. The required data concerning the total number of customers available can be seen in Table 2. In some cases, data for the entire time of the online shop's existence were available while in other cases they were not – compare Table 2 with detailed online store information below. The datasets used for this analysis are very recent – from between 2008 and 2016, ranging from 151 to 381 weeks of data among the online stores. The next part briefly summarizes the business verticals of the datasets. All these companies agreed to participate in this research on condition of anonymity and with a prohibition of disclosing the dataset to any third party.

Aggregate results of the descriptive analysis are given in Table 2, which provides summary information about the individual datasets of the online stores presented. Appendix A includes a time series plot of profit.

The company operating online store A focuses on the narrow area of games of all sorts, including board games. In addition to the online shop, they also have several brick-and-mortar stores in the Czech Republic. The revenues during the year are dominated by a strong Christmas season. In their marketing activities, they target mainly younger customers. The current (2016) total number of customers is nearly 15,000, and altogether they have made almost 20,000 orders.

The company running online store B focuses on sports equipment. In its field, this store is among the biggest in the Czech Republic. Apart from the online shop, they have

several brick-and-mortar stores as well as activities outside the Czech Republic. The sales are again dominated by the Christmas season and several other times of the year. The annual revenues are in the order of tens of millions of euros. They currently (2016) have over 90,000 customers and 136,000 orders in the online store alone.

Table 2 | Summary information about individual online stores and available data

| Dataset (online store) | Number of transactions | Number of customers | Sum of profit (EUR) | Average transaction profit (EUR) | Data range (in weeks) |
|-----------------------------------|-----------------------------------|--------------------------------|--------------------------------|---|----------------------------------|
| A | 19,433 | 14,758 | 148,999 | 7.87 | 218 |
| B | 136,611 | 90,896 | 2,573,842 | 19.24 | 151 |
| C | 106,129 | 50,255 | 557,085 | 5.53 | 173 |
| D | 119,439 | 73,472 | 1,625,073 | 14.33 | 364 |
| E | 62,744 | 43,899 | 1,101,526 | 17.73 | 381 |
| F | 2,409,019 | 798,703 | 18,037,523 | 7.88 | 301 |

Source: Authors

The company operating online store C focuses on health products. Apart from their online shop, they also run a store in Slovakia. They have a strong year-over-year sales growth. At present (2016), they have over 50,000 customers, and nearly 110,000 orders made via the online store.

The company operating online store D focuses primarily on winter and adrenaline sports. They run a brick-and-mortar store complementary to the online shop. In their marketing activities, they target younger customers. They have a steady year-over-year sales growth, with annual sales in the higher millions of euros. They have over 73,000 customers at present and almost 120,000 orders.

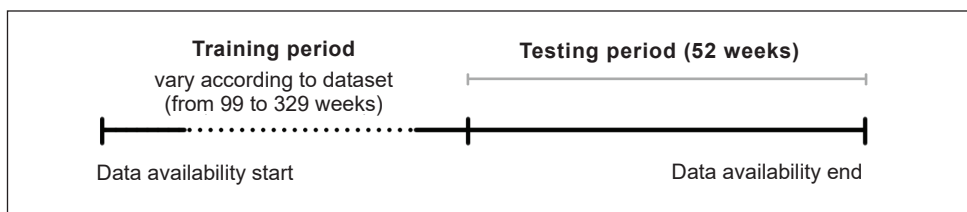
Online store E focuses on erotic and health accessories. It is one of the biggest erotic shops with a strong brand in the Czech Republic. Apart from the online shop, they also run brick-and-mortar stores. They have a well-functioning community and higher millions of euros in their annual revenues. They have 43,000 customers at present (2016), and nearly 63,000 orders made via the online store.

The last company, online store F, focuses on health and beauty products, especially cosmetics. They have several branches used mainly for goods delivery. The company has experienced strong year-over-year growth since 2007, with ongoing expansion also to other European countries. Due to the product portfolio, the company is impacted by a strong Christmas season. With almost 800,000 customers and 2.4 million orders in the last six years, this is the largest of our datasets.

2.4 Evaluation metrics

After the data pre-processing and exploration, the following procedure was established for the comparison of the individual models. First, the training and testing periods to carry out the prediction were defined. The testing period is 52 weeks long – see Figure 1. In order to make the most of the available data, the last 52 weeks from each dataset were separated, and the data were used for the evaluation of the models' predictive performance. The remaining data (the total number of weeks minus the final 52 weeks) serve as the training period, and its length differs for each data set.

Figure 1 | Data division for comparison.



Source: Authors

Data in the testing period were then cleansed of newly acquired customers, as the prediction was to be compared with reality only for the customers who were acquired before the testing period. The comparison was made from two perspectives – at the individual and the aggregated customer base level in accordance with Donkers *et al.* (2007).

The customer base level offers an overall perspective and compares the real situation with the prediction of CLV for all the customers together. The performance of the models for the whole period (forecast vs. actual metric) was evaluated and the weekly model performance was also assessed using the mean absolute percentage error (MAPE). Evaluation at the aggregated weekly level should assure that the MAPE has not faced observations with zero profit.

Another perspective compares the prediction at the individual customer level. The original intention was to construct the MAPE metric as well; however, this metric would use the actual value of profit in the denominator, which is often equal to zero (*e.g.*, for customers who did not make any purchase during the testing period). Therefore, the mean absolute error (MAE) was used instead.

Although MAE in the percentage of the average actual profit would be preferable, it seemed reasonable to use the same metric as Donkers *et al.* (2007) for the sake of comparison: MAE in the percentage of average CLV. The reason for this decision is the possibility of broader theorising about the results of various studies in case the researchers used the same statistical or evaluation metrics.

The success rate of selecting the most profitable customers was also evaluated. Using the individual models, 10% of the most profitable customers were selected in the 52-week

period and compared with the actual top 10%. To compare the performance of this classification, sensitivity (sometimes called true positive rate or recall) is computed.

To be able to compare the CLV models, the inherent limitation of the article is that we understand CLV as profit within 52 weeks and not as a lifetime. It is only by doing so that we can perform empirical comparisons of the selected CLV models, as was done for example by Chamberlain *et al.* (2017), Donkers *et al.* (2007) and Batislam *et al.* (2007).

3. Results

The aim of this article is to independently assess the predictive power of several CLV models using empirical analysis and evaluating different statistical metrics. The following performance metrics are presented for every model and dataset: forecast vs. actual, MAE (individual customer level), MAPE (weekly data), and sensitivity (identification of 10% of the most profitable customers). The results are obtained by comparing a 52-week period of calculated CLV against reality. A visual representation of the comparative analysis together with further details of the models' statistical estimates can be found in Appendix B.

Table 3 | Results for forecast vs. actual (in %)

| Forecast versus actual (in %, 52 weeks) | | | | | | |
|---|---------------|------------|--------|--------|--------|-------|
| Dataset (online store) | Actual profit | Status quo | EP/NBD | MC | VAR | VARX |
| A | 57,224 | 301.58 | 101.50 | 203.57 | 79.37 | 61.01 |
| B | 902,508 | 382.93 | 163.50 | 365.03 | 101.36 | – |
| C | 254,022 | 177.68 | 118.10 | 122.95 | 73.84 | 71.95 |
| D | 244,467 | 245.18 | 107.85 | 229.85 | 133.87 | 96.64 |
| E | 357,685 | 288.30 | 100.98 | 144.62 | 48.63 | 43.01 |
| F | 6,169,603 | 116.83 | 87.82 | 59.69 | 70.30 | 66.01 |
| Weighted mean (by actual profit) | | 161.78 | 98.64 | 106.26 | 74.97 | 58.61 |
| Relative standard deviation (%) | | 55.86 | 24.43 | 93.42 | 20.36 | 37.82 |

Source: Authors

In order to interpret the results, it is necessary to emphasize that in some cases, the quality of the results corresponds to the complexity of the problem: not only

do the models aim to estimate the probability of purchasing, buying frequency and average value, but most importantly the projected profit in time as well. It should be noted that the VAR model cannot make evaluation at the individual customer level and therefore cannot classify the 10% of the most profitable customers (*i.e.*, sensitivity) and be compared with actual profits (*i.e.*, MAE in %) – as was explained in detail in Section 2.1. Therefore, the respective tables show a cross in the case of the VAR model. It can be stated that undervaluation of the customer base offers more acceptable results than overestimation as possible applications have a lower risk of wrong investment decisions.

Table 4 | Results for MAPE (weekly level, in %)

| MAPE (weekly level, in %, 52 weeks) | | | | | | |
|--|---------------|------------|--------|--------|-------|-------|
| Dataset (online store) | Actual profit | Status quo | EP/NBD | MC | VAR | VARX |
| A | 57,224 | 269.73 | 66.48 | 151.74 | 24.32 | 51.53 |
| B | 902,508 | 346.08 | 40.44 | 325.23 | 23.14 | – |
| C | 254,022 | 92.94 | 49.14 | 38.04 | 26.66 | 28.88 |
| D | 244,467 | 217.40 | 34.48 | 198.87 | 73.25 | 26.48 |
| E | 357,685 | 226.14 | 26.41 | 67.88 | 48.35 | 55.71 |
| F | 6,169,603 | 43.78 | 18.88 | 33.36 | 27.67 | 41.34 |
| Weighted mean (by profit) | | 94.61 | 23.44 | 73.96 | 29.42 | 36.53 |
| Relative standard deviation (%) | | 108.82 | 39.83 | 128.08 | 30.80 | 17.50 |

Source: Authors

The results offer new insights into the performance of three fundamentally different sophisticated models and one naïve model. Six typical e-commerce businesses are compared. All the selected models share the same constraints considering the minimal dataset and data regarding recency and frequency of transactions. Moreover, the MC and VAR models were used to exploit their ability to process additional data as input parameters. An important common attribute of all the selected models is that regardless of the different data and methods used for CLV calculation, they all provide the same result: calculating customer lifetime value. This fact opens the possibility for a relevant comparative evaluation of CLV models and the results provided.

Table 5 | Results for MAE (customer-level, in %)

| MAE (customer level, in %, 52 weeks) | | | | | | |
|--------------------------------------|---------------|------------|--------|--------|-----|------|
| Dataset (online store) | Actual profit | Status quo | EP/NBD | MC | VAR | VARX |
| A | 57,224 | 97.20 | 105.24 | 138.29 | – | – |
| B | 902,508 | 101.42 | 118.03 | 103.58 | – | – |
| C | 254,022 | 90.25 | 102.55 | 138.94 | – | – |
| D | 244,467 | 108.72 | 146.70 | 130.54 | – | – |
| E | 357,685 | 105.93 | 164.67 | 148.81 | – | – |
| F | 6,169,603 | 91.86 | 111.62 | 197.56 | – | – |
| Weighted mean (by profit) | | 94.07 | 115.46 | 180.41 | – | – |
| Relative standard deviation (%) | | 5.13 | 10.84 | 18.24 | – | – |

Source: Authors

Table 6 | Results for sensitivity (customer level, in %)

| Sensitivity (customer level, in %, 52 weeks) | | | | | | |
|--|--------------------------|------------|--------|-------|-----|------|
| Dataset (online store) | Customers for validation | Status quo | EP/NBD | MC | VAR | VARX |
| A | 5,936 | 46.67 | 59.33 | 11.00 | – | – |
| B | 41,843 | 20.20 | 27.78 | 10.54 | – | – |
| C | 28,307 | 43.10 | 45.20 | 12.00 | – | – |
| D | 12,308 | 20.10 | 25.20 | 8.10 | – | – |
| E | 15,430 | 15.10 | 16.60 | 4.10 | – | – |
| F | 398,063 | 45.41 | 46.23 | 14.12 | – | – |
| Weighted mean (by customer base) | | 41.64 | 43.36 | 13.21 | – | – |
| Relative standard deviation (%) | | 25.07 | 20.36 | 18.61 | – | – |

Source: Authors

3.1 Answers to research questions

A conclusion can be drawn from all these results to answer the research questions from Section 1:

Which of the compared models for calculating CLV have good predictive performance for CLV in the purchase environment with customer relationships typical of online shopping as a part of e-commerce?

The results described in this part demonstrate that the EP/NBD model has outperformed the other selected models in the majority of the evaluation metrics and can thus be considered good and stable for online shopping within e-commerce business. Its predictive power was recognized in a 52-week period, and also when comparing individual predictions with overall customer base value prediction. However, the EP/NBD has not delivered the best performance in MAE (115% with standard deviation 11%), for which the status quo model had lower error results (a MAE of 94% with a standard deviation of 5%).

Can non-financial data on customer behaviour improve the predictive performance of the selected CLV models in comparison with CLV models that use only financial data from online shopping as a part of e-commerce?

The very dispersed nature of the MC model and the VAR model results related to balanced marketing mix features does not fully validate support for complex CLV models using other non-financial data on customer behaviour. On the other hand, the very good results of the status quo model experienced a drawback by overvaluation of the customer base, which was the consequence of high customer heterogeneity, low retention and low predictability of purchase frequency. These two insights can develop into a generalization that simple models like the status quo can be used only in well scrutinized business settings and complex models like the MC model require intense additional internal and external resources and data that need to be carefully selected with modelling goals in mind. The reason for proposing such intensive preparation is capturing relations between internal and external data in the specific business context. For complex models, a validation of the incremental impact of data enrichment on the predictive model is essential to achieve results more relevant to the set business goals. See also conclusions of Bradlow *et al.* (2017) and Donkers *et al.* (2007) for a comparison.

4. Discussion

Seasonality (*e.g.*, the Christmas period) constitutes a problem for all datasets containing predominantly seasonal purchase behaviour, as can be visually observed in Appendix A. That has a strong impact mainly for datasets B and D with the EP/NBD and VAR models, and this impact can be associated with the fact that the training period ended in late November, *i.e.*, in the midst of the Christmas season for online stores. Such seasonal effects were not discussed in the empirical studies by Donkers *et al.* (2007) and Batislam *et al.* (2007). The issue of seasonal effects in CLV modelling has begun to be tackled only recently, see Platzer and Reutterer (2016). We have suggested a VARX model with dummy variables to deal with seasonal effects. However, we suppose that the analysed datasets are not long enough to enable training the seasonal factors well. Following is a discussion of the results of each model and possible explanations of its performance.

The **status quo model** overestimates the customer base profit value in the majority of the datasets, as seen in Table 3. The best results were observed for the datasets C and F with an FA of 178% and 117%, respectively. The worst two results belong to the datasets B and A with an FA of 383% and 302%, respectively. Recency and frequency analysis of the training period reveal that both datasets C and F consist of a very high ratio of customers with low profitability and low purchase frequency with two completed transactions at most. The results of the rest of the datasets are not impacted by this customer cluster to such an extent. We also observed an undervaluation of customer-level profit prediction for segments of customers with interpurchase times higher than the threshold of 52 weeks chosen for the status quo model. This effect was mainly observed for the datasets A, B and E, which include approximately 30% of the customers with a second-transaction repurchase interval longer than 52 weeks. Customer segments with high purchase frequency result in an overestimation of the profit prediction by neglecting dropout signals. The status quo model also demonstrated strong results in the selection of the top 10% of the most profitable customers and had low individual customer level error rates, which could be seen as a good result for this naïve model construction.

These results indicate that the status quo model could be beneficial in situations with low or predictable dropout rates, with a stronger likelihood of higher purchase frequency and low variability in product margin – compare with Donkers *et al.* (2007). However, in comparison with other models used in this study, it must be concluded that the features of the status quo model are not suitable for datasets with long training time periods and large datasets with high customer heterogeneity including spontaneous purchase behaviour undergoing significant changes in time.

The **EP/NBD model** performed well in both individual predictions (the greatest sensitivity of 43%, good MAE of 115%) and overall customer base (FA of 99%, MAPE of 24%). The weaker results for the datasets D and E in terms of the MAE metric in the long period are attributed to lower purchase frequencies of the customers of the e-commerce retailers involved. This observation is also confirmed by the nature of the EP/NBD gamma-gamma submodel, which calculates the expected average order value. In this submodel, customers with a low number of total transactions (including single-order customers) are assigned the predicted lifetime average order value, which is constructed mostly by considering the majority of the customer base.

The more transactions the individual customer has completed, the more weight is attributed to his or her purchase behaviour and the remainder of the population is left with a minor role played in the calculation of the expected lifetime average order value for the given repeat customer. Hence, for datasets featuring a low number of repeat customers, the spending submodel prefers population estimates narrowing the distribution of the expected lifetime average order values. The resulting CLV predictions are then mostly driven by the EP/NBD recency and frequency submodel.

The results of the **MC model** appear to be fluctuating among all the datasets and statistical metrics. A high standard deviation was observed for the individual customer-level predictions (see Table 3) as well as the weekly trends (see Table 4). It is concluded from the analysis that the MC model estimate performance is highly dependent on the quality

of the initial subgroups at the beginning of the model execution. The estimation of the subgroups is performed using a decision tree, which should benefit from as many additional customer-level attributes as possible. However, the methodology described in Section 2.2 dictates only the features (profitability drivers) available for all datasets to be used. As a result, the quality of the subgroups is consistent with the weekly results and performance of the Markov chain submodel and further stresses the need for additional customer attributes in any future datasets. A recommendation is in place to identify and obtain further profitability drivers and maintain a thorough analysis of all the input dimensions for this particular model. An interesting exception to the aforementioned conclusion can be found in the dataset E, which initially identified clear subgroups in the CART submodel, but still achieved only 4% sensitivity in the detection of the most profitable customers (Table 6), which is very weak considering that a random selection would result roughly in the expected 10% sensitivity. Such an underperformance of the MC customer segmentation compared to a random selection had not been expected. The MC model performed the worst on all levels.

The **VAR model** performed consistently at good levels, undervaluing the customer base performance, as seen by the MAPE at the weekly level in Table 4. A strong overvaluation of the forecast vs. actual ratio in Table 3 can be observed only for the dataset D, an online store with heavy seasonal product purchases. Good performance was observed for the datasets B and D, which feature a balanced composition of paid, owned and other media in the marketing mix. The **VARX model** added seasonal dummy variables to the VAR model specification. As the training data were quite short, the model provided worse results than the original VAR model for most of the datasets. For the dataset B, this limitation in the date range resulted in no possibility to train the seasonal model. Better results were gained only for the dataset D, which is one of the longest datasets with regular seasonal effects (see Figure 2). More intensive work would be recommended on a company's marketing channel grouping to better reflect the marketing tactics in the channel classification used by the VAR models.

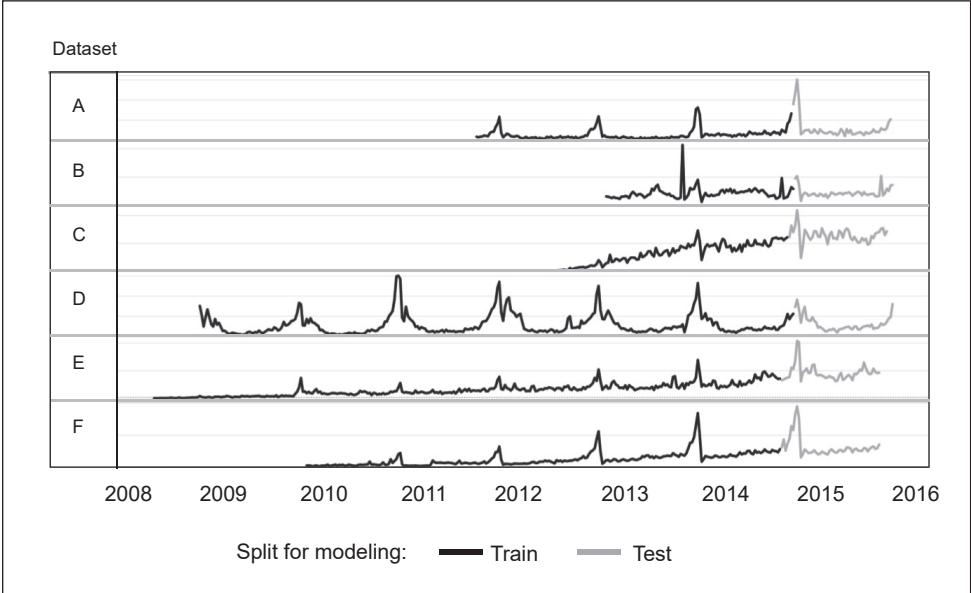
In this article, the benefits of the VAR model were limited as it was used for overall customer base predictions only. However, it provides more outputs, which could be beneficial for other types of business use (*e.g.*, analysis of performance and interactions of selected marketing channels).

The observed individual error rates can be considered very high, because even the best MAE of 90% for the dataset C and the status quo model still indicates a prediction error of almost the whole average yearly profit. Therefore, we suggest that finding ways to minimize such errors should be the main focus of a future research study.

There is an implicit restriction of this paper given by the fact that the datasets used for the empirical analysis tend to come from the Czech Republic and Slovakia. Thus, the results achieved in this paper should only be related to the context of Central and Eastern Europe, as the purchase behaviour is significantly impacted by cultural factors (Chiang and Yang, 2017; Kumar and Pansari, 2016). As a specific recommendation for a future research study, we suggest to focus on the cross-country differences in empirical results for international e-commerce.

Appendix A

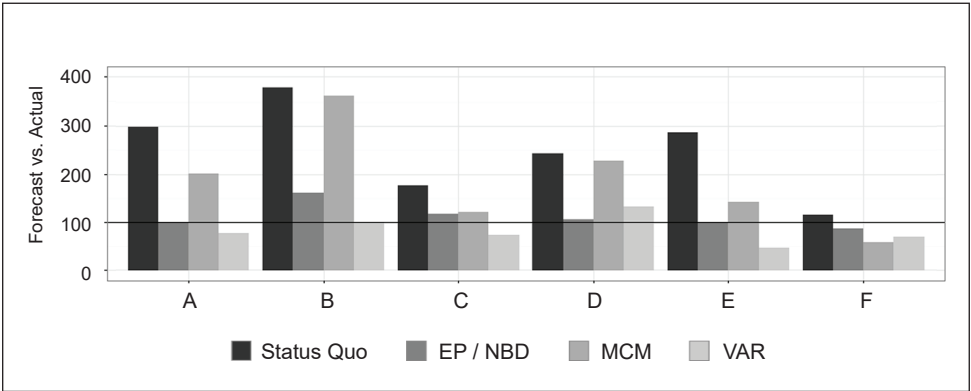
Figure 2 | Time series plot of profit for online stores A-F, with reference to training and 52-week testing period. Vertical axis has obfuscated absolute numbers due to data sensitivity.



Source: Authors

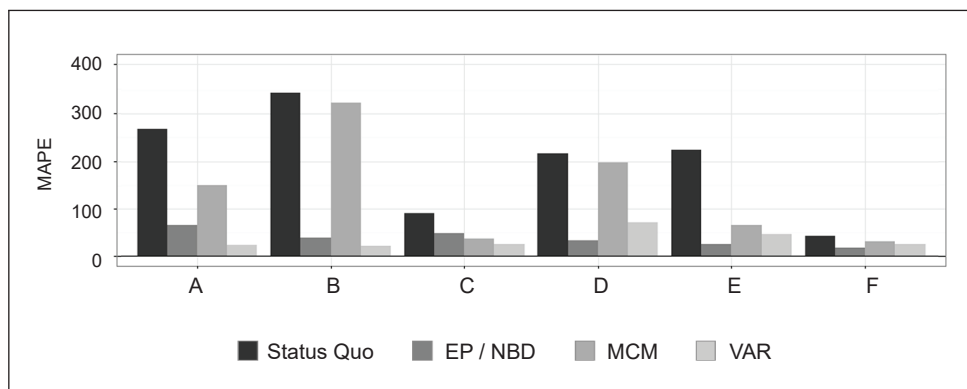
Appendix B

Figure 3 | Results for forecast vs. actual (in %) for online stores A-F.



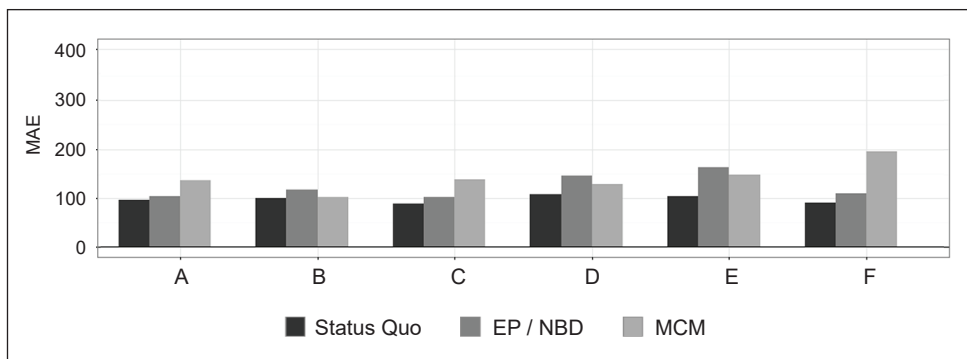
Source: Authors

Figure 4 | Results for mean absolute percentage error (MAPE, weekly-level, in %) for online stores A-F.



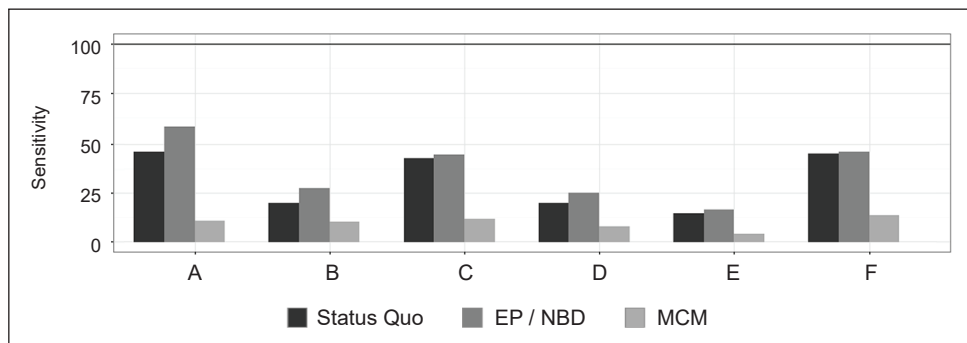
Source: Authors

Figure 5 | Results for mean absolute error (MAE, customer-level, in %) for online stores A-F. The comparison excludes the VAR and VARX models, which are not based on customer level.



Source: Authors

Figure 6 | Results for sensitivity (customer-level, in %) for online stores A-F. The comparison excludes the VAR and VARX models, which are not based on customer level.



Source: Authors

Table 7 | Estimated parameters for EP/NBD model

| Forecast versus actual (in %, 52 weeks) | | | | |
|---|--------|----------|--------|---------|
| Dataset (online store) | r | α | s | β |
| A | 1.7607 | 268.3112 | 0.1828 | 1.2031 |
| B | 1.4354 | 94.6908 | 0.1405 | 0.0954 |
| C | 0.7846 | 26.4615 | 0.1230 | 0.1405 |
| D | 1.3715 | 122.3069 | 0.3168 | 0.9132 |
| E | 1.7607 | 268.3112 | 0.1828 | 1.2031 |
| F | 0.5384 | 31.4624 | 0.0360 | 1.7935 |

Source: Authors

References

- Adomavicius, G., Tuzhilin, A. (2005). Toward the Next Generation of Recommender Systems: a Survey of the State-of-the-art and Possible Extensions. *IEEE Transactions on Knowledge and Data Engineering*, 17(6), 734–749, <http://doi.org/10.1109/TKDE.2005.99>
- Batıslam, E. M., Denizel, M., Filiztekin, A. (2007). Empirical Validation and Comparison of Models for Customer Base Analysis. *International Journal of Research in Marketing*, 24(3), 201–209, <http://doi.org/10.1016/j.ijresmar.2006.12.005>
- Bierens, H. J. (2004). *VAR Models with Exogenous Variables*. [Retrieved 2018-09-20] Available at: http://www.personal.psu.edu/hxb11/EasyRegTours/VAR_Tourfiles/VARX.PDF
- Bradlow, E. T., Gangwar, M., Kopalle, P., Voleti, S. (2017). The Role of Big Data and Predictive Analytics in Retailing. *Journal of Retailing*, 93(1), 79–95, <http://doi.org/10.1016/j.jretai.2016.12.004>
- Burcher, N. (2012). *Paid, Owned, Earned: Maximising Marketing Returns in a Socially Connected World*. Philadelphia: Kogan Page. ISBN 978-0749465629.
- Castéran, H., Meyer-Waarden, L., Reinartz, W. (2017). Modeling Customer Lifetime Value, Retention, and Churn, in Homburg, C., Klarmann, M., Vomberg, A., eds., *Handbook of Market Research*. Cham: Springer, pp. 1–33, http://doi.org/10.1007/978-3-319-05542-8_21-1
- Centre for Retail Research. (2017). *Online Retailing: Britain, Europe, US and Canada 2017*. [Retrieved 2018-09-20] Available at: <http://www.retailresearch.org/onlineretailing.php>
- Chamberlain, B. P., Cardoso, A., Bryan Liu, C. H., Pagliari, R., Deisenroth, M. P. (2017). Customer Lifetime Value Prediction Using Embeddings, in *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York: ACM, pp. 1753–1762, <http://doi.org/10.1145/3097983.3098123>
- Chang, W., Chang, C., Li, Q. (2012). Customer Lifetime Value: A Review. *Social Behavior and Personality*, 40(7), 1057–1064.
- Chiang, L. L. L., Yang, C. S. (2018). Does country-of-origin brand personality generate retail customer lifetime value? A Big Data analytics approach. *Technological Forecasting and Social Change*, 130, 177–187, <http://dx.doi.org/10.1016/j.techfore.2017.06.034>

- Colombo, R., Jiang, W. (1999). A Stochastic RFM Model. *Journal of Interactive Marketing*, 13(3), 2–12, [http://doi.org/10.1002/\(SICI\)1520-6653\(199922\)13:3<2::AID-DIR1>3.0.CO;2-H](http://doi.org/10.1002/(SICI)1520-6653(199922)13:3<2::AID-DIR1>3.0.CO;2-H)
- Damm, R., Monroy, C. R. (2011). A Review of the Customer Lifetime Value as a Customer Profitability Measure in the Context of Customer Relationship Management. *Intangible Capital*, 7(2), 261–279, <https://doi.org/10.3926/ic.2011.v7n2.p261-279>
- Donkers, B., Verhoef, P.C., De Jong, M.G. (2007). Modeling CLV: A test of Competing Models in the Insurance Industry. *Quantitative Marketing and Economics*, 5(2), 163–190, <http://doi.org/10.1007/s11129-006-9016-y>
- Estrella-Ramón, A. M., Sánchez-Pérez, M., Swinnen, G., VanHoof, K. (2013). A Marketing View of the Customer Value: Customer Lifetime Value and Customer Equity. *South African Journal of Business Management*, 44(4), 47–64, <https://doi.org/10.4102/sajbm.v44i4.168>
- Fader, P. S., Hardie, B. G. S. (2009). Probability Models for Customer-Base Analysis. *Journal of Interactive Marketing*, 23(1), 61–69, <http://doi.org/10.1016/j.intmar.2008.11.003>
- Fader, P. S., Hardie, B. G. S. (2013). *The Gamma-Gamma Model of Monetary Value*. [Retrieved 2018-09-20] Available at: http://www.brucehardie.com/notes/025/gamma_gamma.pdf
- Fader, P. S., Hardie, B. G. S. (2001). Forecasting Repeat Sales at CDNOW: A Case Study. *Interfaces*, 31(3), S94–S107, <http://doi.org/10.1287/inte.31.3s.94.9683>
- Fader, P. S., Hardie, B. G. S., Lee, K. L. (2005). Counting your Customers the Easy Way: An Alternative to the Pareto/NBD Model. *Marketing Science*, 24(2), 275–284, <http://doi.org/10.1287/mksc.1040.0098>
- Ferrentino, R., Cuomo, M. T., Boniello, C. (2016). On the Customer Lifetime Value: a Mathematical Perspective. *Computational Management Science*, 13(4), 521–539, <http://doi.org/10.1007/s10287-016-0266-1>
- Gladya, N., Baesensa, B., Crouxa, C. (2009). A Modified Pareto/NBD Approach for Predicting Customer Lifetime Value. *Expert Systems with Applications*, 36(2), 2062–2071, <http://doi.org/10.1016/j.eswa.2007.12.049>
- Gupta, S., Hanssens, D., Hardie, B., Kahn, W., Kumar, V., Lin, N., Ravishanker, N., Sriram, S. (2006). Modeling Customer Lifetime Value. *Journal of Service Research*, 9(2), 139–155, <http://doi.org/10.1177/1094670506293810>
- Haenlein, M., Kaplan, A. M., Beeser, A. J. (2007). A Model to Determine Customer Lifetime Value in a Retail Banking Context. *European Management Journal*, 25(3), 221–234, <http://doi.org/10.1016/j.emj.2007.01.004>
- Haenlein, M., Kaplan, A. M., Schoder, D. (2006). Valuing the Real Option of Abandoning Unprofitable Customers When Calculating Customer Lifetime Value. *Journal of Marketing*, 70(3), 5–20, <http://doi.org/10.1509/jmkg.70.3.5>
- Jasek, P., Vrana, L., Sperkova, L., Smutny, Z., Kobulsky, M. (2018). Modeling and application of customer lifetime value in online retail. *Informatics*, 5(1), no. 2. <https://doi.org/10.3390/informatics5010002>
- Jerath, K., Fader, P. S., Hardie, B. G. S. (2011). New Perspectives on Customer “Death” Using a Generalization of the Pareto/NBD Model. *Marketing Science*, 30(5), 866–880, <http://doi.org/10.1287/mksc.1110.0654>
- Knox, G., van Oest, R. (2014). Customer Complaints and Recovery Effectiveness: A Customer Base Approach. *Journal of Marketing*, 78(5), 42–57, <http://doi.org/10.1509/jm.12.0317>
- Kumar, V., Pansari, A. (2016). National Culture, Economy, and Customer Lifetime Value: Assessing the Relative Impact of the Drivers of Customer Lifetime Value for a Global Retailer. *Journal of International Marketing*, 24(1), 1–21, <http://doi.org/10.1509/jim.15.0112>
- Kumar, V. (2018). A Theory of Customer Valuation: Concepts, Metrics, Strategy, and Implementation. *Journal of Marketing*, 82(1), 1–19, <http://doi.org/10.1509/jm.17.0208>

- Lin, H.-H., Li, H.-T., Wang, Y.-S., Tseng, T.H., Kao, Y.-L., Wu, M.-Y. (2017). Predicting Customer Lifetime Value for Hypermarket Private Label Products. *Journal of Business Economics and Management*, 18(4), 619–635, <http://doi.org/10.3846/16111699.2017.1308879>
- Lone, S. (2017). *European Ecommerce Report 2017*. [Retrieved 2018-09-20] Available at: <http://www.ecommercefoundation.org/download-free-reports>
- Paaauwe, P., Van der Putten, P., Van Wezel, M. (2007). DTMC: an Actionable E-customer Lifetime Value Model Based on Markov Chains and Decision Trees, in *Proceedings of the Ninth International Conference on Electronic Commerce*. New York: ACM, pp. 253–262, <http://doi.org/10.1145/1282100.1282147>
- Pfeifer, P. E., Carraway, R. L. (2000). Modeling Customer Relationships as Markov Chains. *Journal of Interactive Marketing*, 14(2), 43–55, [http://doi.org/10.1002/\(SICI\)1520-6653\(200021\)14:2<43::AID-DIR4>3.0.CO;2-H](http://doi.org/10.1002/(SICI)1520-6653(200021)14:2<43::AID-DIR4>3.0.CO;2-H)
- Platzer, M., Reutterer, T. (2016). Ticking Away the Moments: Timing Regularity Helps to Better Predict Customer Activity. *Marketing Science*, 35(5), 779–799, <http://doi.org/10.1287/mksc.2015.0963>
- Postnord (2016). *E-commerce in Europe 2016*. [Retrieved 2018-09-20] Available at: <http://www.postnord.com/globalassets/global/english/document/publications/2016/e-commerce-in-europe-2016.pdf>
- Rathi, N. A., Betala, A. S. (2019). How Marketing Decisions are Taken with the Help of Big Data, in Balas, V., Sharma, N., Chakrabarti, A., eds., *Data Management, Analytics and Innovation*. Singapore: Springer, pp. 101–112, http://doi.org/10.1007/978-981-13-1274-8_8
- Reinartz, W. J., Kumar, V. (2000). On the Profitability of Long-Life Customers in a Noncontractual Setting: An Empirical Investigation and Implications for Marketing. *Journal of Marketing*, 64(4), 17–35, <http://doi.org/10.1509/jmkg.64.4.17.18077>
- Schmittlein, D. C., Bemmaor, A. C., Morrison, D. G. (1985). Technical Note — Why Does the NBD Model Work? Robustness in Representing Product Purchases, Brand Purchases and Imperfectly Recorded Purchases. *Marketing Science*, 4(3), 255–266, <http://doi.org/10.1287/mksc.4.3.255>
- Schmittlein, D. C., Morrison, D. G., Colombo, R. (1987). Counting Your Customers – Who Are They and What Will They Do Next. *Management Science*, 33(1), 1–24, <http://doi.org/10.1287/mnsc.33.1.1>
- Schmittlein, D. C., Morrison, D. G. (1983). Prediction of Future Random Events with the Condensed Negative Binomial Distribution. *Journal of the American Statistical Association*, 78(382), 449–456, <http://doi.org/10.1080/01621459.1983.10477993>
- Schmittlein, D. C., Peterson, R. A. (1994). Customer Base Analysis: An Industrial Purchase Process Application. *Marketing Science*, 13(1), 41–67, <http://doi.org/10.1287/mksc.13.1.41>
- Singh, S. S., Jain, D. C. (2013). *Measuring Customer Lifetime Value: Models and Analysis*. INSEAD. Working Paper No. 2013/27/MKT, <http://doi.org/10.2139/ssrn.2214860>
- Tsay, R. S. (2013). *Multivariate Time Series Analysis: With R and Financial Applications*. New Jersey: John Wiley & Sons. ISBN 978-1-118-61775-5.
- Villanueva, J., Yoo, S., Hanssens, D. M. (2008). The Impact of Marketing-Induced Versus Word-of-Mouth Customer Acquisition on Customer Equity Growth. *Journal of Marketing Research*, 45(1), 48–59, <http://doi.org/10.1509/jmkr.45.1.48>
- Weng, C.-H., Huang, T. C.-K. (2018). Knowledge Acquisition of Association Rules from the Customer-lifetime-value Perspective. *Kybernetes*, 47(3), 441–457, <http://doi.org/10.1108/K-03-2016-0042>