Forecasting sales in industrial services: modeling business potential with installed base information

(Please cite the further developed version of this article that has been published: Stormi, K., Laine, T., Suomala, P., & Elomaa, T. (2018). Forecasting sales in industrial services: Modeling business potential with installed base information. Journal of Service Management, 29(2), 277-300.)

Forecasting sales in industrial services: modeling business potential with installed base information

Kati Stormi(1), Teemu Laine(1), Tapio Elomaa(2), Petri Suomala(1)

(1) Cost Management Center, Tampere University of Technology, Tampere, Finland
(2) Laboratory of Mathematics, Tampere University of Technology, Tampere, Finland

Abstract

Purpose – The purpose of the study is to examine how installed base information could help servitizing original equipment manufacturers (OEMs) forecast and support their industrial services sales, and thus increase OEMs' understanding regarding the dynamics of their customers lifetime values (CLVs).

Design/methodology/approach – This work constitutes a constructive research aiming to arrive at a practically relevant, yet scientific model. It involves a case study that employs statistical methods to analyze real-life quantitative data about sales and the global installed base.

Findings – The study introduces a forecasting model for industrial service sales, which considers the characteristics of the installed base and predicts the number of active customers and their yearly volume. The forecasting model performs well compared to other approaches (Croston's method) suitable for similar data. However, reliable results require comprehensive, up-to-date information about the installed base.

Research limitations/implications – The study contributes to the servitization literature by introducing a new method for utilizing installed base information and, thus, a novel approach for improving business profitability.

Practical implications – OEMs can use the forecasting model to predict the demand for—and measure the performance of—their industrial services. To-the-point predictions can help OEMs organize field services and service production effectively and identify potential customers, thus managing their CLV accordingly. At the same time, the findings imply new requirements for managing the installed base information among the OEMs, to understand and realize the industrial service business potential. However, the results have their limitations concerning the design and use of the statistical model in comparison with alternative approaches.

Originality/value – The study presents a unique method for employing installed base information to manage the CLV and supplement the servitization literature.

1 Introduction

Industrial services are those required by manufactured products mainly in the business-to-business (B2B) context (Oliva and Kallenberg, 2003). They ensure effective, uninterrupted, and durable use of industrial equipment. Installation, training, repair, maintenance, and spare parts are examples of these services. Original equipment manufacturers (OEMs) seek to servitize, that is, to expand their business to industrial services, in order to differentiate themselves from competitors and compensate for declining sales volumes and profit margins (Wise and Baumgartner, 2000). In this context, the *installed base* (or the fleet) of equipment in use by customers is a central concept. The installed base includes all the pieces of equipment sold by the OEM to its customers that will be served by the OEM and other service providers during the equipment's lifecycle (Dekker *et al.*, 2013; Oliva and Kallenberg, 2003).

The increase in industrial service sales does not guarantee higher profitability. The effect of services on a firm's performance (e.g., profitability) is controversial (Kwak and Kim, 2016). Gebauer and Fleisch (2007) suggest that the share of total revenue attributable to services has a generally positive effect on *return on sales* (ROS). Suarez *et al.* (2013) found that the relationship between service integration and ROS is convex[1] in the software industry. However, there is also evidence of a concave[2] relationship between the share of services and operating profit (Kwak and Kim, 2016). The profitability problems of services have even led to deservitization; some OEMs have reduced the share of services or abandoned services and returned to square one as an equipment manufacturers (Finne *et al.*, 2013; Valtakoski, 2016).

Planning and controlling industrial service operations impacts the OEMs' performance. In fact, several studies have aimed to explain the reasons for the occassionally negative impact of increased service integration on firms' performance, although additional research about causes and consequences is still needed (Gebauer *et al.*, 2012). One identified reason is the increase in the transaction costs in certain circumstances. Typically, services require greater customization and interaction with customers, which, in turn, generates increased costs (Ulaga and Reinartz, 2011). Forecasting and stabilizing service demand are generally extremely difficult for OEMs (Tenucci and Cinquini, 2016). Industrial service sales vary from year to year, as well as within the year. Establishing a field service organization implies committing to fixed costs that do not adapt to fluctuating demand. During peak demand, OEMs must consider subcontracting, whereas during off-peak periods, some of the field technicians are unemployed. Broadly speaking, planning for field technicians' schedules can be challenging. Under these circumstances, a key resource for designing and delivering successful industrial services is field service organization and, particularly, the capability to manage it efficiently and mitigate execution risk (Ulaga and Reinartz, 2011). Therefore, there is a need for a greater understanding of the drivers of industrial service demand in the installed base as well as the lifetime value of the customers as owners of the installed base.

In response, the study examines how OEMs can find opportunities to increase service sales and predict service demand in advance, which can serve to enhance service operations management and extend the customer lifetime value (CLV), respectively. The paper argues that in doing so, OEMs require a better understanding of one of their key resources: the customers' installed base (Ulaga and Reinartz, 2011). Utilizing information about the installed base is a timely topic in

[1] The relationship between ROS and the share of services is U-shaped.
[2] The relationship between ROS and share of services takes an inverted U-shape.

business planning (Dekker *et al.*, 2013), and it holds extensive prospects in identifying and realizing the business potential embedded in servitization initiatives. *In this vein, the purpose of the study is to examine how installed base information could help servitizing OEMs forecast and support their industrial services sales, and thus increase OEMs' understanding regarding the dynamics of their customers lifetime values (CLV).* That said, increasing sales and predicting service demand in this environment based on installed base information can be beneficial for operational and economic purposes in a broad sense; the installed base provides a natural base for service businesses and long-term customer relationships.

This study presents a forecasting model that predicts demand for industrial services. The forecasting model predicts demand at the customer level and concentrates on services directly related to the installed base, such as repair, maintenance, and remanufacturing (Mathieu, 2001). Because the forecasting model predicts customer-specific demand, it could serve other purposes as well. For example, the forecasting model can identify different customer performances with respect to their lifetime values and determine unmet customer potential in contrast to their installed bases or lost industrial service customers with remarkable installed bases. Opportunities emerge from customers ordering less than one might expect based on their installed base. Thus, OEMs need to thoroughly understand the installed base and the related business potential. This could help them realize the business potential promoted in the literature (Baines *et al.*, 2009; Wise and Baumgartner, 2000).

Theoretically, the forecasting model constructed in this work is closely related to the CLV literature, because it provides information about the antecedents and dynamics in the activity of and the annual sales to the different customers. CLV is generally defined as the present value of all future profits obtained over the customer relationship (Gupta *et al.*, 2006; Holm *et al.*, 2012). One objective of CLV models is to make marketing actions more accountable and help select the most effective marketing strategies (Dwyer, 1997; Rust *et al.*, 2004; Venkatesan and Kumar, 2004). This indicates that CLV models typically use marketing actions as an explanatory variable and customer's transaction volume and transaction frequency as the dependent variable. This study uses the size of the customer's installed base as an explanatory variable for the customer's yearly industrial service volume as well as three alternative models to identify active customers, that is, customers that will make at least one service transaction.

The installed base at the customer level could be a driver of both the frequency of purchases and the size or the margin of each purchase in the industrial services context, namely, the central determinants of the CLV results. However, these aspects have not been studied in the existing CLV literature, despite several B2B contexts in the CLV literature (Kumar *et al.*, 2008; Kumar and Shah, 2009; Venkatesan and Kumar, 2004). This finding further specifies the research gap addressed in this paper; in addition to contributing to the servitization literature, this study contributes to the CLV literature by introducing a new forecasting model. The forecasting model predicts customer-specific industrial service sales as a factor to be included in CLV models in the future. Applying the model over the entire customer lifetime produces an estimate of the CLV.

The study constitutes a constructive research (Kasanen *et al.*, 1993), as it seeks to construct a forecasting model in response to real-life managerial challenges (forecasting industrial service sales and business potential) with theoretical implications. Thus, the results of this paper are evaluated in the discussion section in light of their implications for the literature and in practice, as well as in contrast to possible alternative models (Croston's (1972) method was chosen as a benchmark). Empirically speaking, the paper employs a case study in an OEM, and extensive data and quantitative methods were employed to build the forecasting model. The case company is a large OEM that distributes its products worldwide. The data includes real-life intermittent customer orders and fleet information for the period 2012–2015, as well as interviews and interactions with key personnel involved in demand forecasting. Project members arranged interviews and meetings in the fall of 2015 and the spring of 2016. In total, the data covers more than 100,000 invoices. The case company has more than 1,000 customers, and the size of the installed base encompasses more than a thousand pieces of equipment.

The remainder of the paper is organized as follows: Section 2 examines the servitization and CLV literature by providing the research background and contribution domains of the paper. Section 3 introduces the forecasting model for industrial services. Section 4 describes the research method and data used. Section 5 summarizes the results of the forecasting model. Section 6 discusses and evaluates the findings in light of the existing literature and practice and, finally, provides the conclusions and implications of the study.

2 Literature review

2.1 Examining the business potential of industrial services

Industrial services, in general, are services that ensure effective, uninterrupted, and durable use of manufacturing equipment. OEMs service equipment in use by the customer, that is, the fleet or the installed base (Oliva and Kallenberg, 2003). This paper uses the terms “fleet” and “installed base” interchangeably. Depending on the context, the terms can either refer to the entity of all pieces of equipment owned by one customer or to all pieces of equipment used by the OEM’s customers. The study focuses on equipment (and the installed base) that the customer uses in industrial production, such as facilities or production machinery. In addition, the focus is on typical examples of industrial services, such as spare parts and periodic maintenance.

Industrial services represent a potentially remarkable business potential for the OEMs (Vandermerwe and Rada, 1988), as OEMs engaged in servitization can offer their services to customers that use their machinery in their own operations. In developed countries, OEMs have largely increased their service offerings (Neely, 2008) and focused on developing their customer relationships with a more long-term orientation (Penttinen and Palmer, 2007). This servitization is due to financial, strategic, and marketing tactics (Baines *et al.*, 2009). From a financial point of view, services often provide a stable and increased revenue stream. Strategic drivers are concerned with gaining competitive advantage, and from the marketing perspective, services increase product sales (Baines *et al.*, 2009).

Despite its business potential, the impact of industrial services on a company’s performance is controversial, and the performance measures used to examine the impact often vary. Some studies use revenue growth as a performance outcome (Kohtamäki *et al.*, 2013; Neely, 2008), while many others use profitability, especially ROS (namely, the operating profit margin) (Gebauer, 2008; Gebauer *et al.*, 2010b; Gebauer and Fleisch, 2007; Kwak and Kim, 2016; Neely, 2008). Some studies use Tobin’s *q* (Fang *et al.*, 2008) or different combinations of factors (Eggert *et al.*, 2014). Gebauer and Fleisch (2007) suggest that the share of total revenue attributable to services (share of services or service integration) has a generally positive effect on the general operating margin. However, Neely (2008) shows that although manufacturing firms that have undergone servitization are larger than traditional manufacturing firms in terms of sales revenues, at the aggregate level, these manufacturing firms also generate lower profits as a percentage of sales. Several studies have shown that the relationship between services and a firm's profitability is not self-evident (Gebauer *et al.*, 2012). External and internal moderating factors exist, such as firm size, organizational structure, and a competitive environment.

Further, the level of service integration (service revenue as a percentage of total revenue) seems to affect a firm's profitability. Suarez *et al.* (2013) found that the relationship between service integration and return on sales is convex in the software industry, but there is also conflicting evidence from the concave relationship (Gebauer, 2008; Kwak and Kim, 2016). The relationship is concave if, at the first phase, the level of service integration has a positive effect on ROS. However, the marginal impact on ROS decreases. In other words, an equal increase in service integration causes a diminishing increase in ROS. After a certain threshold, service integration has a negative impact on ROS; the increase in service integration causes ROS to decrease at a faster rate. Kwak and Kim (2016) found evidence of a concave relationship in their research, which covered more than 70 percent of the total production of the Korean machinery and equipment sector. The convex relationship (Suarez *et al.*, 2013), on the other hand, means that a diminishing negative effect turns into an increasing positive effect after a certain threshold.

Early studies suggest that OEMs' transition from products to services is a continuum, a process that follows logical and sequential phases, starting as a pure equipment manufacturer and ending as a solution provider (Mathieu, 2001; Oliva and Kallenberg, 2003). However, recent studies have shown that different strategies co-exist, and only a few companies comply with the traditional product-service continuum (Kowalkowski *et al.*, 2011; Kowalkowski *et al.*, 2015). There is also support on the opposite development continuum, starting from solution services and ending with product services (Cusumano *et al.*, 2015). In addition, some OEMs have decided to deservitize; that is, they have abandoned industrial services (Gebauer and Kowalkowski, 2012; Valtakoski, 2016).

Relevant literature has shown that several external and internal factors affect service performance: service orientation in corporate culture and organization structure (Gebauer *et al.*, 2010a; Gebauer *et al.*, 2010b), the percentage of loyal customers (Eggert *et al.*, 2014), the company's available service-enabling resources (Raddats *et al.*, 2015), and the service provision offered (Kwak and Kim, 2016), to mention a few. Ulaga and Reinartz (2011) identify four critical resources and five capabilities for successful hybrid offerings. According to them, product use and process data about the installed base and field service organization are two key resources; further, service-related data processing capability and hybrid offering deployment capability are among the key success factors. In addition, Tenucci and Cinquini (2016) point out that forecasting the demand for industrial services is difficult.

According to an analysis of the literature, no previous studies have addressed the problem of forecasting uncertain demand in the context of industrial services. Further, research on how to take advantage of installed base information remains rare (Dekker *et al.*, 2013). In this context, the installed base provides a natural and potential information source, especially if the purpose is to predict the demand for product-related services, such as maintenance and repair. Demand for these services is highly dependent on the size of the installed base. To fill this gap, the study offers a forecasting model that predicts customer-specific demand for industrial services. The model could eventually enhance OEMs' service operations management and partly addresses the problem of diminishing profitability after service integration has reached a certain level (Kwak and Kim, 2016). This means that collecting, maintaining, and processing installed base information could have a moderating effect on the negative impact of service integration on profitability. For example, the concave relationship can turn a diminishing operating profit into stable or even increasing profit (Figure 1).

[Insert Figure 1 here]

Other relationships assumed between the percentage of services and the ROS could also be affected by an enhanced understanding of industrial service demand and customer behavior, respectively.

Overall, an extended understanding of the antecedents and dynamics of industrial service demand and the industrial service sales realized would lead to the improved management of industrial service operations. Besides, such an extended understanding is directly connected to the potential annual revenues of different customers (with different kinds of fleets), and indirectly connected to the economic effectiveness of industrial service operations related to those customers. Therefore, modeling the economic value of a customer relationship is strongly connected to this study, and as it stems from the CLV literature, this stream of literature is discussed next.

2.2 CLV literature

Different customers vary greatly in terms of the industrial service business potential (e.g., related to the size of the installed base) and the actual revenues achieved from those customers (e.g., related to loyalty and other characteristics of the customer relationships). Overall, in the servitization context, the dynamics of the industrial service business potential for different customers is not sufficiently understood (Lindholm *et al.*, 2017). Moreover, CLV applications are lacking in this practical need.

CLV is defined as the net present value of the customer relationship (Gupta *et al.*, 2006). By definition, the CLV should contain all revenues and costs during the entire lifetime of the customer. Specifically, the costs should cover customer acquisition and retention, the costs of goods sold, as well as marketing costs, such as advertising, loyalty programs, one-to-one marketing, and distribution. However, most models are not comprehensive and instead concentrate on certain viewpoints, such as customer acquisition (Gupta *et al.*, 2006).

Future revenues and costs, (i.e., CLV) only form one part of the customer's value for a firm (Jaakkola and Alexander, 2014; Kumar *et al.*, 2010a). The customer engagement value (CEV) comprises three more value sources: customer referral value, customer influencer value, and customer knowledge value. These values are created when a customer affects other customers' buying behavior or when a customer contributes to product development (Kumar *et al.*, 2010a). The value created through customer referrals and customer knowledge development might be significant in the industrial services context. However, this study is limited to CLV considerations, since data for the total CEV is not reliable and readily available.

CLV includes future revenues and costs. Thus, calculating CLV requires a model that reliably predicts a customer's future transactions and the volume of each transaction. The modeling is based on an understanding of customer behavior. Jackson (1985) divides industrial customer behavior into two categories: lost-for-good and always-a-share. Lost-for-good behavior means that customers, once lost, are typically gone forever, and a returning customer is treated as a new one.

Always-a-share (migration model) (Berger and Nasr, 1998) means that customer commitment to one provider is low, and customers change providers constantly.

Some customer relationships are based on contracts, in what is known as a contractual setting. A typical example of a contractual setting is a journal subscription. A customer pays a subscription price and gets access to the journal for a given period, such as one year. After the subscription period, a customer decides whether to discontinue or renew the contract and the customer relationship. From the company's point of view, a customer who does not renew the contract is lost-for-good. In a contractual setting, CLV models are typically based on the retention rate, that is, the probability that a customer will renew the subscription. Simple models assume a constant retention rate across customers (Berger and Nasr, 1998). Models that are more sophisticated assume that the retention rate of an individual customer is constant over time. However, the retention rate can vary across customers (Fader and Hardie, 2007; Fader and Hardie, 2010).

Non-contractual settings can be discrete or continuous in time (Fader and Hardie, 2009). A non-contractual context is considered discrete if the transaction can occur only at certain predetermined times (hereafter, NCD), such as participation in an annual conference (Fader and Hardie, 2009). However, most non-contractual settings are continuous in time (hereafter, NCC), which means that a transaction can occur at any time, such as in the case of retail shopping.

Typically, CLV models in non-contractual settings consist of four main components: (i) the transaction frequency, (ii) contribution margin, (iii) marketing costs, and (iv) customer lifetime (Kumar *et al.*, 2006). Calculating the CLV is simple if one can reliably predict a customer's transaction rate, the contribution margin of each transaction, and the lifetime of the customer relationship. In NCD settings, the beta-geometric/beta-Bernoulli (BG/BB) model is an appropriate predictor of transaction frequency and customer lifetime (Fader *et al.*, 2010).

The study divides the transaction frequency and lifetime models in the NCC context into three categories: Pareto/NBD models, hierarchical Bayes models based on the generalized gamma distribution (HB/GG), and joint regression models for purchase incidence, contribution margin, and marketing costs. Pareto/NBD models separately predict an individual customer's lifetime and transaction frequency. While active, individual customer i makes purchases according to a Poisson process with purchasing rate λ_i (i.e., the customer executes on average λ_i transactions per time period) (Schmittlein *et al.*, 1987). However, the data is often insufficient to model the λ_i of each individual customer separately. The problem of insufficient data is solved by assuming that the purchasing rate λ for individual customers is distributed according to the gamma distribution across customers. In other words, the cross-sectional heterogeneity in transaction frequency rate λ is gamma distributed across customers (Fader and Hardie, 2009). This is the NBD part of the Pareto/NBD model.

Each customer remains active for a lifetime that has an exponentially distributed duration, with the death rate μ_i and the heterogeneity in the death rate μ is the distributed gamma across customers (Schmittlein *et al.*, 1987). This is the Pareto part of the Pareto/NBD model.

Pareto/NBD models assume that λ_i and μ_i remain constant over time. Allenby *et al.* (1999) developed a hierarchical Bayes model of customer interpurchase times (transaction frequency) based on the generalized gamma distribution. The HB/GG model allows cross-sectional and temporal heterogeneity, which is introduced through the component mixture model being dependent on lagged covariates (Allenby *et al.*, 1999). Lagged covariates can be, for example, the cross-buying and frequency of different marketing contacts, such as email or telephone (Venkatesan and Kumar, 2004). Such models are particularly useful when a company wishes to ascertain the impact of its marketing actions on the CLV.

Typically, HB/GG models do not account for the entire customer lifetime. Instead, they calculate the CLV over a predetermined future period, such as three years (Kumar *et al.*, 2006). Theoretically, CLV models should estimate customer value over the customer's entire lifetime. However, in most companies, the business environment remains stable for around three years. Thereafter, changes in competitors and technology alter the business climate substantially (Kumar *et al.*, 2008).

In the third category, joint regression models together examine (i) the level of marketing contacts directed toward the customer during a specific time period, such as one month; (ii) the probability that the customer buys in each period; and (iii) the contribution margin of the transaction, provided that the customer buys (Kumar *et al.*, 2008). All three components of the CLV are correlated and jointly modeled through regression models. Similar to the HB/GG models, joint regression models are particularly useful when a company wants to find the link between firm (marketing) actions and the CLV (Kumar *et al.*, 2008).

Many CLV models simply assume that the contribution margin of each transaction remains stable or increases over time (Berger and Nasr, 1998). Models that are more sophisticated assume that the value of each transaction varies randomly around the customer's unobserved mean transaction value (gamma/gamma models) (Fader *et al.*, 2005). The contribution margin can also depend on recency, that is, the time of last purchase (Pfeifer and Carraway, 2000). Many studies use regression models to estimate the contribution margin (Kumar *et al.*, 2006; Kumar *et al.*, 2008; Kumar and Shah, 2009; Kumar *et al.*, 2010b; Kumar *et al.*, 2010a).

Table 1 shows the customer behavior, empirical environment, and the frequency and contribution margin models in some CLV studies. Table 1 reveals that all example studies but one fall in the NCC context. CLV studies in the business-to-consumer (B2C) marketing context are more common than in the B2B context. Additionally, CLV studies in the B2B manufacturing context are almost nonexistent. A rare example is Brashear Alejandro *et al.* (2013), who present a CLV model in a make-to-order steel manufacturer.

[Insert Table 1 here]

3 The forecasting model for industrial services

The study introduces a forecasting model for product-related industrial services to be used by OEMs. In this context, the installed base is a prerequisite for the demand and sales of industrial services, and significant fluctuations in customers' annual volume and purchase frequency are typical. The number, type, intended use, and age of the fleet determine the overall demand for services. The installed base of some customers covers hundreds of pieces of equipment, whereas many customers have only one piece of equipment. In addition, an OEM's share of the total demand varies from customer to customer even if their fleet sizes are the same. Some customers maintain their fleets annually with proactive service activities, while others respond only to cases of failure. Some customers maintain their fleets by themselves or use competitors. Some customers order once every five years, while others make dozens of orders each year. Further, the number of customers is small, typically a few hundred.

OEMs usually divide their product-related industrial services into two main categories: spare parts and labor services. Labor services are those that have only labor requirements, such as periodic maintenance and installation. Some OEMs separate wear parts from spare parts. Wear parts are spare parts that wear evenly when the equipment is used. This study models only industrial labor services. However, similar models can estimate the spare and wear parts as well. The forecasting model has its roots in the CLV literature. However, the purpose of the model is not to calculate the customer's value over the total lifetime of the relationship. Instead, the forecasting model predicts two things: whether the customer will be active in the following year and the active customer's yearly volume. Hence, the forecasting model comprises two sub-models: the customer activity model and yearly volume model. The forecasting model does not account for the impact of different marketing actions on service demand. Rather, the forecasting model uses the size of the installed base as the driver of demand. Naturally, the discounting of cash flows is unnecessary because the analysis covers only the ensuing year. The following sections present the yearly volume and transaction frequency models in detail.

3.1 The yearly volume model

This research uses a panel data regression analysis to model yearly volume (Venkatesan and Kumar, 2004). The dependent variable is the 2015 volume. The independent variables are the number of different equipment types and lagged volume. Different equipment types vary by size and purpose and, therefore, require a different number of industrial services. The case company members divided their equipment into 14 categories; the categories were determined by the device size and intended use. Lagged volume refers to the sum of sales for the years 2012–2014. The model uses the sum of sales as the lagged volume since customers do not necessarily purchase services yearly. For example, one customer purchased services in 2012 and 2015 but not in 2013 and 2014. In this case, the sum of sales for the years 2012–2014 only applies to the volume in 2012. Typically, CLV models predict the contribution margin instead of the yearly volume. However, this study is interested in the yearly volume and uses it as a dependent variable in the regression model (Equation (1)).

$$V_{it} = \beta_0 + \beta_1 V_{lagged} + \beta_2 QT1_i + \dots + \beta_{14} QT14_i + e_{i,t} \tag{1}$$

where

V_{it}	=	Volume of customer i in year t
V_{lagged}	=	Sum of sales for the years 2012–2014
$QT1_i - QT14_i$	=	The number of type 1–14 devices that customer i owns
β_0	=	Intercept
$\beta_1 - \beta_{14}$	=	Unknown parameters
e_{it}	=	Error term
i	=	Customer index
t	=	Time index

3.2 The customer activity models

This study uses three different models to identify customers as active or inactive in a given year: the binary logistic regression (BLR), zero-inflated negative binomial regression (ZINBR), and Pareto/NBD models. Typically, CLV models predict customer transaction frequency. However, in this study, customer activity is the primary object of interest, even though the ZINBR and Pareto/NBD models also predict transaction frequency.

BLR is a regression model where the dependent variable can have only two values, “0” or “1.” In the customer activity model, “0” indicates that a customer is inactive in a given year; that is, the customer will not purchase any industrial services. Likewise, “1” implies that a customer will make at least one service order in that period. The independent variables in the BLR model are the lagged volume (sum of sales for the years 2012–2014); the total size of a customer’s installed base; and a categorical variable, which takes value “1” if the customer has purchased services in each of the years over 2012–2014 and “0” otherwise (Equation (2)).

$$log\left(\frac{p_t}{1-p_t}\right) = \beta_0 + \beta_1 V_{lagged} + \beta_2 IB\_SUM + \beta_3 STEADY\_ORDER + e_{it} \tag{2}$$

where,

p_t	=	Probability that a customer will purchase services in year t
V_{lagged}	=	Sum of sales for the years 2012–2014 (€1000)
IB_SUM	=	The total size of the installed base
$STEADY\_ORDER$	=	“1” if customer has orders for each of the years 2012–2014, “0” otherwise
$\beta_1 - \beta_3$	=	Unknown parameters
e_{it}	=	Error term
i	=	Customer index
t	=	Time index

The ZINBR model (Scott Long, 1997) applies to count variables with excessive zeros. It actually comprises two separate models: the zero-inflation model (binomial with logit link) for excessive zeros and a count model (negative binomial with log link) for modeling a count of interest, for example, the number of orders that an active customer will make. The explanatory variables in the ZINBR model are the same as those in the BNR model (Equation (2)).

Pareto/NBD models predict customer lifetime (Pareto) and transaction frequency (NBD) separately. This model is suitable for continuous non-contractual settings, namely, settings where transactions can occur at any time. OEMs often sign service contracts with customers. However, in the industrial services context, the content of a contract varies, and despite the contract, customers often execute orders as well. Thus, the conclusion of the service contract does not make the customer relationship a truly contractual setting.

4 Method and data

The study is a constructive research (Kasanen *et al.*, 1993) with theoretical implications intended for the servitization and CLV literature. It presents a case study that uses real-life quantitative empirical data and methods to create a model that forecasts sales for product-related industrial services.

The case company—called Alpha, to ensure confidentiality—is a large Nordic OEM. While Alpha manufactures equipment for various industry categories and distributes products worldwide, this analysis covers only one product category, which includes more than 50 types of equipment. Alpha provides labor services (maintenance and repair) and provides spare and wear parts. The percentage of total revenue attributable to services varies from year to year but is typically more than 50 percent. The challenge that Alpha faces is fluctuating demand for spare parts and labor services. The field service organization is basically a fixed cost that does not quickly adapt to changes in service demand. Demand peaks increase subcontracting, and subcontracting on a tight schedule is often costly. During off-peak periods, in turn, the field force organization is partly unemployed.

The statistical methods used in this study are based on two databases that Alpha collects and maintains: a comprehensive database of the installed base and the database that contains all service invoices. The installed base (device) database contains the delivery date, device type, location, and customer information per piece of equipment sold. The installed base database, which holds thousands of records, covers devices sold since 1975. However, the installed base database is incomplete because it does not contain equipment sold to all customers.

In the study, invoice data covers the service invoice history for the period 2012–2015. Alpha records the date, total volume, service category (labor, spare or wear part), target device type, and customer information for each invoice. In total, the invoice data comprises hundreds of thousands of invoices. This implies tens of thousands of invoices in each industrial service category: labor services, spare, and wear parts. This study assumes that one invoice is sent per order. Thus, the number of orders (known as “transactions” in the CLV literature) is equals the number of invoices. The invoice data includes invoices sent to all customers (totaling thousands). However, the installed base data is not comprehensive and only covers the equipment of some customers. In addition, not all customers were active over 2012–2015, resulting in a sample size of approximately 700 customers.

For confidentiality reasons, the given numbers are not precise. However, Table 2 contains a rough estimate of the size of the database.

[Insert Table 2 here]

Initially, the data, which the researchers received in January 2016, was presented in a spreadsheet. The project members also organized two face-to-face meetings with two Alpha employees during the spring of 2016 to ensure the correct interpretation and quality of the data. Alpha employees are responsible for collecting, maintaining, and analyzing business-critical data. As required, the researchers requested refinements and additions to the data via phone calls or email. The product category covered in this study comprises more than 50 types of equipment. Sales volumes for some device types are small, including only a few pieces. Thus, for this research, product types were classified based on how much they required labor services, on average. Alpha employees conducted the grouping, which resulted in 14 categories.

On average, customers' installed bases and order volumes are relatively small. However, there are considerable differences in the sizes of the installed bases, order volumes and transaction frequencies between the customers. For example, in 2012-2014, the largest order volumes per customer were over 1000 times larger than the smallest order volumes. In addition, the order volume and number of transactions are strongly correlated (.912). Consequently, the order size does not vary significantly. The correlation between the installed base and volume (.528) and the number of transactions (.550) is positive and statistically significant, albeit lower than the correlation between volume and the number of transactions. Table 3 shows the proportion of active customers in different installed base size categories. For example, 38 percent of customers with extra small installed base were active, whereas 78 percent of customers with extra large were active. According to Table 3, the increase in the installed base also increases the activity.

[Insert Table 3 here]

To extract better insights from the data, Figure 2 depicts the sales volume (from 2012 to 2014) of sample customers against the total size of the customer's installed base (the numbers are hidden for confidentiality). Figure 2 does not distinguish between different equipment types, and the size of the installed base covers all types of equipment. The figure shows two important points. First, the dotted blue line (Linear (Normal)) roughly describes the relationship between the installed base and volume. The line estimates how much demand one piece of equipment approximately produces per year. Second, quite many customers order considerably less than expected, based on their installed bases. These customers are 'underachievers' from Alpha's service business perspective (as conveyed in Figure 2). In addition, some individual customers order more than expected ('overachievers'). However, according to Figure 2, there is an average positive relationship between the size of the installed base and customer volumes (and number of transactions). In addition, the relationship seems to be linear. It is noteworthy, that even though underachievers (overachievers) order less (more) than expected based on their installed bases, their annual order volumes remain typically quite stable. Remarkably, neither underachievers nor overachievers are outliers in the yearly volume model (Equation 1), as the model uses also lagged volume as an explanatory variable in addition to the size of the installed base. However, underachievers inevitably reduce the correlation between the installed base and the yearly volume.

[Insert Figure 2 here]

There can be many reasons for the lower than expected sales. Clearly, a huge installed base requires maintenance services. However, based on the data, Alpha does not serve the installed base of orange customers. Orange customers might use competitors or they can service the installed base themselves. In addition, reporting practices and limited and incorrect data might cause underachievers. Alpha has its operations in several countries. Not all units necessarily allocate costs in the same way, and they might use different customer identification systems. For example, subsidiaries can allocate an installed base to a different customer identification than the industrial services.

5 Results of the forecasting model

5.1 The yearly volume model

The yearly volume model is a linear regression model (Equation (2)). The sample comprised customers who made at least one transaction over 2012–2014 (718 customers).

Table 4 shows the results for the yearly volume model. The sum of sales over 2012–2014 (V_{lagged}) and product types 1, 3, 4, 5, and 7 are significant at the 1 percent level. Some product types seem to have had a negative impact on yearly volume. However, the installed base of different equipment types varies in size, from thousands of devices to a few dozen. Thus, the coefficients for the smallest equipment types are not necessarily precise; they require interpretation and possibly a combination of some device types. For clarity, this analysis uses the equipment categorization made by the case company. In total, the yearly volume model provides an adjusted R^2 of .80. This means that the model explains 80 percent of the variation in the customer's service demand.

[Insert Table 4 here]

5.2 The customer activity models

The study uses three alternative models to predict customer activity: the BLR, ZINBR, and Pareto/NBD models. Results for the BLR model are below (Table 5). The table shows that all other independent variables except installed base size (IB_SUM) are statistically significant. According to the results, customers who ordered during each of the years over 2012–2014 ($STEADY\_ORDER$ = "1") have a higher probability, by 8.3 times, of making at least one order in 2015.

[Insert Table 5 here]

Similarly, Table 5 shows results for the ZINBR model. The table separates the count and zero-inflation models. The total volume over 2012–2014 (V_{lagged}) is statistically significant in both models. However, the total size of the installed base (IB_SUM**)** and $STEADY\_ORDER$ are statistically significant only in the count model. Hence, it seems that the installed base size does not affect the customer's annual activity but does affect transaction frequency. There are at least two possible explanations. First, some customers purchase industrial services more regularly, regardless of the size of the installed base. Second, the total size of the installed base correlates with the lagged volume, and the impact of the installed base on customer activity is contained in the lagged volume.

The Pareto/NBD model parameters are based on transaction data for the calibration period, starting at the beginning of 2012 and ending on December 31, 2014. The year 2015 is the holdout period to evaluate the goodness of fit of the transaction frequency model. The goodness of fit evaluation follows Wadsworth (2012). Evaluating the goodness of fit depends on how well the model holds over the calibration period and, more importantly, the holdout period. Figure 3 shows the number of customers against the frequency of repeat transactions during the calibration period. Figure 3 plots the repeat transactions, because the Pareto/NBD model is concerned with those, not the total number of transactions during the calibration period. Zero repeat transactions mean that a customer executes, in total, one transaction during the calibration period: the initial transaction and zero repeat transactions.

[Insert Figure 3 here]

Figure 3 shows that most customers (more than 200) made zero repeat transactions during the calibration period. However, about 140 customers executed seven or more transactions during the calibration period. The figure also shows that the model overestimates the number of customers executing zero repeat transactions or 7+ repeat transactions.

Figure 3 evaluates the goodness of fit of the Pareto/NBD model during the calibration period. However, it is more interesting how well the model stands in the holdout period (2015). Table 6 divides customers into eight bins based on how many repeat transactions were made during the calibration period (from 0 to 7+). The table shows the actual and expected frequency and number of customers in each bin. For example, 237 customers executed zero repeat transactions during the holdout period. The customers' expected frequency was 0.3, and they actually made, on average, 0.3 transactions. The table indicates that the Pareto/NBD model predicts transactions for customers that rarely make orders quite accurately (bins 0, 1 and 2, highlighted in the table). This finding is significant for two reasons. First, the study is interested in customer activity, not in how many transactions a customer will make. Thus, the model must be able to identify customers that will be active: those that will make at least one service order. Second, the total number of customers in bins 0, 1, 2 is 442, which amounts to 61 percent of all customers. It is, therefore, essential that the model identify infrequent customers with sufficient precision.

[Insert Table 6 here]

Table 7 compares three different customer activity models: BLR, ZINBR, and Pareto/NBD. In total, 359 customers were non-active (observed activity = 0); 352 customers made at least one order (observed activity = 1) in 2015. Among the non-active customers, the specificity of the BLR model was 0.897 percent, whereas the model sensitivity was only 0.543 percent, resulting in a 72.2 percent percentage accuracy overall. Similarly, the overall percentage accuracy was 70.89 percent and 74.4 percent for the ZINBR and Pareto/NBD models, respectively.

[Insert Table 7 here]

As Table 7 shows, the overall percentage accuracy for all three models is almost the same: 72.2 percent, 70.89 percent, and 74.4 percent for the BLR, ZINBR, and Pareto/NBD models, respectively. However, the BLR model clearly overestimates the number of non-active customers; model specificity is 0.897, and sensitivity is only 0.543. The ZINBR and Pareto/NBD models, instead, estimate active and non-active customers much more evenly; specificity and sensitivity for both models is around 70 percent. For the BLR, ZINBR, and Pareto/NBD models, the F_1 scores are 0.659, 0.713, and 0.736, respectively. This result also supports the superiority of the ZINBR and Pareto/NBD models over the BLR model.

6 Discussion and conclusions

6.1 Synthesis, evaluation, and implications of the findings

This section discusses (1) the rationale of the study, (2) the results of the yearly volume and customer activity models, (3) the overall credibility of the forecasting model, and (4) the implications of the findings.

First, the study examines how installed base information could help servitizing OEMs forecast and support their industrial service sales and, thus, augment their understanding of the dynamics of CLV. The paper argues that the industrial service business potential is embedded in the installed base (Lindholm *et al.*, 2017; Gebauer *et al.*, 2012; Vandermerwe and Rada, 1988); hence, information on the installed base enables OEMs to forecast and support industrial service sales. This viewpoint has thus far been under researched in the literature (Lindholm *et al.*, 2017). The paper also responds to the real-life profitability challenge faced by OEMs: increased service provision typically leads to increased interactions with customers and fluctuating service demand, which, despite increased revenues, may lead to decreasing profitability (Neely, 2008; Lindholm *et al.*, 2017).

Second, as central results concerning the constructed models, the yearly volume model predicted customer-specific sales quite well, although it used only the customers' lagged volume and the size of the installed base as explanatory variables. The yearly volume model coefficient of determination (R^2) was 80 percent: the model explains 80 percent of the variation in customers' yearly volume. However, lagged volume has more predictive power than the size of the installed base. This is partly explained by customers who purchase a small number of services as compared to the size of their installed base (underachievers in Figure 2): their demand remains almost the same from year to year but does not align with the size of their installed base. Accordingly, for these customers, lagged volume is a good predictor whereas the installed base is not. However, this does not mean that the installed base does not require maintenance services or that the nature of the installed base would not explain their demand for industrial services; instead, they just do not use (only) the OEM to serve their installed base.

Altogether, there seems to be a positive linear relationship between the size of the installed base and service volume. Furthermore, segmenting customers based on their performance from Alpha's service business perspective provides further insights. More particularly, creating separate models for 'underachievers' and 'other customers' (Figure 2) could strengthen the correlation between the sizes of the installed bases and the yearly service volumes among those 'other customers'. This study, however, included both segments in the same model, since one aim of the study was to actually pinpoint low performing customers, i.e. underachievers, to be more actively and effectively managed by the OEM. Customer segmentation for the models could be based on a company's knowledge about customers and their maintenance practices. Alternatively, a company could use purchase history as a proxy for the customer performance, to continuously compare customers' behavior against the installed base in different segments.

In this study, only a limited number of customers were active yearly. Therefore, the customer activity model predicts customers that will be active. Among the three customer activity models evaluated here, the BLR model predicts only customer activity; the ZINBR and Pareto/NBD models predict transaction frequency as well. The overall percentage accuracy for each customer activity model was more than 70 percent. However, the size of the installed base predicts transaction frequency but does not seem to predict customer activity. Again, underachievers (Figure 2) are a prominent explanation: if the size of the installed base explains activity, then customers with a large installed base should have many transactions on a yearly basis. In addition, different maintenance practices among customers might be an explanation for this. In the industrial services context, industrial companies are customers as well. Some customers schedule proactive maintenance to service breaks that they organize, for example, annually or even less frequently.

Nonetheless, the data does not fully support this assumption: the correlation between the volume and number of transactions is high (.912). It is noteworthy that the study findings should not only be assessed based on the strength of the relationship between the concepts in the models. For planning and controlling purposes, it is important to reveal the overall relationships between the different concepts, namely, installed base and industrial service volumes, and to identify and explain unexpected customer behavior regarding these overall relationships (e.g., how to identify customers with unused industrial service business potential).

Third, in terms of the overall credibility of the forecasting model, Figure 4 illustrates customers and their actual (light blue) and predicted (dark blue) volumes in 2015. The predicted volume is a combination of the yearly volume and customer activity models. The customer activity model distinguishes the active customers from the inactive ones. The yearly volume model predicts the volume of active customers. Customers are shown in ascending order by their actual volumes. The vertical axis shows each customer's volume on a logarithmic scale. Up to 312 customers did not make orders; accordingly, the model predicted that they were inactive (for clarity, they are not shown in the figure). Customers 313 through 378 were not active but the model predicted that they would have made orders. Typically, these customers have a considerable installed base. However, they do not order services as expected. For example, customer 378 ("Huge installed base" in Figure 4) has the second largest installed base. Regardless, the customer did not make any orders in 2015. Most customers depicted by orange circuits in Figure 2, namely, potential customers, belong to this category. Customers 379 through 711 were all active. However, the model predicted that some of them are inactive, for example, customer 699 ("Sudden increase" in Figure 4). The reason for the prediction error is an instant, significant change in the customer's volume. In addition, the installed base of the customer is small. Hence, neither the installed base size nor lagged volume indicated the actual volume, causing the model to predict that the customer is inactive.

[Insert Figure 4 here]

Comparing the forecasting model with other models, such as time series models, gives an overview of the model's performance. Before evaluating the model against other approaches, it is noteworthy that customers typically show intermittent demand in the industrial services context. Intermittent demand is infrequent and irregular, and years with zero demand are typical. Traditional time series forecast models, such as exponential smoothing and moving average, do not fit well in forecasting intermittent demand. Instead, Croston's method (Croston, 1972) and its variants (Syntetos and Boylan, 2001) are widely used to forecast intermittent demand. Both metrics used for the comparison, root mean squared error (RMSE) (For confidentiality, the actual figures are not shown) and symmetric mean absolute percentage error (SMAPE), 38 % for the forecasting model and 46 % for the Croston's method, indicate the superiority of the forecasting model. However, the prediction history in Croston's method was relatively short, at merely three years. A longer demand history would probably improve the prediction accuracy of Croston's method with a similar dataset.

Fourth, in terms of overall implications, the findings imply that OEMs could clearly benefit from utilizing advanced models to forecast industrial service demand and reveal the dynamics in customers' buying behavior. More specifically, as demonstrated in the study, the installed base information may provide additional details on sales forecasts, which can help proactive resource planning. Besides, the models, when designed properly, may specify the industrial service demand connected to a certain type of fleet or a certain use condition, which would then enable effective marketing campaigns and other activities regarding the industrial service operations serving the installed base.

Advanced forecasting models could significantly support sales and marketing in specifying business potential and pinpointing categories of customer relationships with varying performances in contrast to their potential lifetime values. The sales and marketing functions could particularly use the forecasting model to assess the realized industrial service sales against the potential of the installed base; discover potential customers that do not purchase industrial services despite their installed base; and above all, manage customer relationships toward a larger lifetime value. More specifically, the volume plotted against the size of the installed base gives an overview of the relationship between industrial service volume and the size of the installed base (as presented in Figure 2). In the case presented in this paper, there was a positive linear relationship between the size of the installed base and service volume. In addition, the forecasting model revealed customers with very small volumes relative to the size of their installed base, indicating the potential to increase their industrial sales.

As a practical and more detailed implication, OEMs can use the results of the study in a detailed manner. As noted, Figure 2 and the forecasting model (Figure 4) reveal underachievers that an OEM should correct. There might be a natural explanation for the underachievers; for example, the installed base database may not include all the pieces of machinery (i.e., may not be comprehensive) or may not connect them to the correct industrial service sales (i.e., may not be up-to-date). If there are no mistakes regarding the data, the OEM should take appropriate actions to increase their market share among customers having remarkable potential.

After correcting the identified underachievers, OEMs can use the sales forecasting model to identify active customers (customer activity model) and the relationship between the installed base and the sales volume (yearly volume model). Forecasts would help OEMs enhance their service operations management. Further, the model may reveal the danger of losing customers that have been inactive even though they should have been active according to the model. Finally, underachievers represent the possibility for increased sales. At minimum, OEMs should find out why certain customers' volumes are below expectations.

Creating such models, however, requires a comprehensive, detailed, and up-to-date database on the installed base. Besides, users of the database need to know how to deal with limitations that exists in any database in practice. For example, disproportionately large service volumes can be a sign of product problems or an indication that customers simply do not know how to use the product correctly. When the databases do become available and with sufficient understanding about their contents after capturing customer behavior, advanced forecasting models could hold implications for various operations related to industrial services, including research and development, technical support, training, and sales and marketing.

6.2 Contribution of the findings to the servitization and CLV literature

The existing service literature broadly discusses the impact of servitization on OEMs' performance and the causes of the potential negative (or positive) impact on profitability (Gebauer *et al.*, 2005; Lindholm *et al.*, 2017). However, there are few studies that provide solutions for processing installed base information in order to realize service business potential; Dekker *et al.* (2013) is an exception. The present study provides a means for filling this gap by combining the installed base data and financial data in the same forecasting models and examining the potential implications of these models in the servitization context.

First, this study contributes to the servitization literature in terms of a comprehensive, up-to-date installed base database embedded in the forecasting models that could facilitate effective, proactive planning and controlling industrial service operations. Recently, the occassional negative effect of increased service integration on profitability has been acknowledged in the servitization literature (Kwak and Kim, 2016). However, the means of recognizing such effects and responding to them have not yet been addressed in this context. The impact of service integration (as a percentage of the total revenue attributable to services) on a firm's profitability is controversial and depends on several factors. In terms of shape, the relationship can generally be negative (Neely, 2008), concave (Kwak and Kim, 2016), convex (Suarez *et al.*, 2013), or generally positive (Gebauer and Fleisch, 2007). One identified reason for the negative impact is the increase in transaction costs, caused by the growing interaction with customers (Kwak and Kim, 2016; Nordin *et al.*, 2011). Further, an increase in services causes an increase in the size of the field service organization. This, in turn, means that OEMs must commit to fixed costs that do not adapt to fluctuating service demand. Forecasting models, like the one constructed in this paper, could reveal the dynamics of the industrial service demand for multiple planning purposes. In principle, approaches for serving the installed base in a profitable manner could be refined after each major change in the installed base, thus connecting the management of the industrial service business more tightly with the management of equipment sales and enabling more comprehensive management control (Lindholm *et al.*, 2017).

Second, this study contributes to the CLV literature, especially in application to servitization. It does this by presenting a new approach to employing business data in predicting key elements, thereby affecting the value of customer relationships, namely, the customer activity and the annual sales volumes arising from customers having differently sized installed bases. The yearly sales forecast model predicts customers' yearly service volumes based on the size of their installed base. Thus, the size of the installed base can be used to predict customers' yearly volumes and, finally, to calculate the CLV. Advanced CLV models (Kumar *et al.*, 2008; Venkatesan and Kumar, 2004) use, for example, the level of marketing actions as an explanatory variable for customer sales. However, the novelty of this research is to use the installed base as an explanatory variable. Such an approach could be integrated in future CLV models. By definition, the CLV should include all customers' future purchases, including both equipment and service sales. Hence, the CLV is the present value of the customer's equipment and industrial service purchases. Sales of new equipment (or any other change in the installed base) may affect business potential and the lifetime value of customers in question. Acknowledging and managing such dynamics in future CLV models would make them more suitable and effective for OEMs under servitization.

6.3 Conclusion, limitations, and future research

In sum, the study contributes to the servitization literature by introducing a new method of utilizing installed base information and, thus, a new approach for improving business profitability. As an implication, utilizing installed base information to a full extent would enable a stronger understanding of CLV. Nonetheless, the findings imply new requirements for managing installed base information among OEMs, in order to understand and realize the business potential of industrial services.

The real-life business context notably represents practical limitations for the present study and its implications. Naturally, there were several challenges related to the design and implementation of the model and, thus, in capturing its desired benefits. The financial data (for 2012–2015) comprised hundreds of thousands of invoices. Editing and using such a vast dataset is time-consuming. In practice, this means that companies do not necessarily have enough know-how or resources to implement the model regularly. However, the findings suggest that when forecasting models become broadly implemented and understood by the OEM(s), the benefits of collecting, maintaining, and processing installed base data will likely overcome the costs of designing and implementing these new practices.

The research methodology used is not without its limitations. This study constitutes constructive research given its objective to respond to a real-life management challenge by constructing a model that OEMs can employ to better understand and realize the business potential of the installed base for industrial services. Regarding the (weak) market test of the construct (Kasanen et al., 1993), at the moment, no major decisions have been made with its help. However, the model has been reviewed and thoroughly discussed with (and between) business analysts in the case company, and they see remarkable potential in it. The model addresses an existing managerial challenge and can extract new information about the dynamics of industrial service business, which conveys a valid construct in this context. It is noteworthy that neither this nor perhaps any other study could clearly pass the strong market test, that is, connecting better performance to the use of the construct (Kasanen et al., 1993), although the model presented in this paper could actually help OEMs perform better in their industrial service businesses.

Besides, the design and use of the statistical model has its limitations. For example, the Pareto/NBD model has an over-forecasting bias, that is, the model overestimates the number of transactions (Table 7). However, the accuracy of the model is adequate for the purposes of this study, since it aims to identify active customers. The actual number of transactions is not significant. When reflecting on the results of the paper and their limitations, one needs to remember that the main purpose of the article is not to provide the best available statistical model to predict demand for industrial services. Instead, the research provides OEMs with a new way to augment their understanding of customers. A practical approach also poses challenges to the model. It must be easy to implement; any adaptation of the model should not significantly increase the amount of data collected.

Thus, in addition to the limitations of the model itself, the source data also has some limitations that restrict the practical implications of the results. Installed base databases may not always be up to date, in practice. Customers may, for example, sell or disable equipment without notifying the OEM. Such limitations may cause inaccuracies in the use of the model outlined in this paper.

This study is one example of how OEMs can process installed base data. The suitability of a similar forecasting model to predict spare part consumption is one interesting research direction. Besides, future studies can expand the presented model to a CLV model or even a CEV model (Kumar *et al.*, 2010a) for industrial customers. Indeed, a comprehensive CLV model would also include customers' equipment purchases. Besides, the CLV model should cover the customer's entire lifetime and should incorporate marketing activities, which have not been considered here.

Finally, almost every model leaves room for improvement in terms of actually capturing business practices. For example, a gamma generalized linear model (gamma GLM) might be an option to simultaneously model customer volume and activity. In addition, this study does not take the fleet's age into account. Typically, an aging installed base requires more servicing than a new one. Moreover, the intended use of the equipment varies, leading to remarkable variations in industrial service demand. Some customers use their devices in problematic conditions (e.g., wet or cold), and in other cases, they simply do not use their devices appropriately. This study delves into only one way to utilize installed base data with several implications. However, the potential of systematically collected and maintained installed base data remains primarily unrealized.

References

Allenby, G.M., Leone, R.P. and Jen, L. (1999), "A dynamic model of purchase timing with application to direct marketing", *Journal of the American Statistical Association,* Vol. 94 No. 446, pp. 365-374.

Baines, T.S., Lightfoot, H.W., Benedettini, O. and Kay, J.M. (2009), "The servitization of manufacturing: A review of literature and reflection on future challenges", *Journal of Manufacturing Technology Management,* Vol. 20 No. 5, pp. 547-567.

Berger, P.D. and Nasr, N.I. (1998), "Customer lifetime value: Marketing models and applications", *Journal of Interactive Marketing,* Vol. 12 No. 1, pp. 17-30.

Brashear Alejandro, T., Tukel, O.I. and Dixit, A. (2013), "Application of customer lifetime value model in make-to-order manufacturing", *Journal of Business & Industrial Marketing,* Vol. 28 No. 6, pp. 468-474.

Croston, J.D. (1972), "Forecasting and stock control for intermittent demands", *Operational research quarterly,* Vol. 23 No. 3, pp. 289-303.

Cusumano, M.A., Kahl, S.J. and Suarez, F.F. (2015), "Services, industry evolution, and the competitive strategies of product firms", *Strategic Management Journal,* Vol. 36 No. 4, pp. 559-575.

Dekker, R., Pinçe, Ç, Zuidwijk, R. and Jalil, M.N. (2013), "On the use of installed base information for spare parts logistics: A review of ideas and industry practice", *International Journal of Production Economics,* Vol. 143 No. 2, pp. 536-545.

Dwyer, F.R. (1997), "Customer lifetime valuation to support marketing decision making", *Journal of Interactive Marketing,* Vol. 11 No. 4, pp. 6-13.

Eggert, A., Hogreve, J., Ulaga, W. and Muenkhoff, E. (2014), "Revenue and profit implications of industrial service strategies", *Journal of Service Research,* Vol. 17 No. 1, pp. 23-39.

Fader, P.S. and Hardie, B.G. (2007), "How to project customer retention", *Journal of Interactive Marketing,* Vol. 21 No. 1, pp. 76-90.

Fader, P.S. and Hardie, B.G. (2009), "Probability models for customer-base analysis", *Journal of interactive marketing,* Vol. 23 No. 1, pp. 61-69.

Fader, P.S. and Hardie, B.G. (2010), "Customer-base valuation in a contractual setting: The perils of ignoring heterogeneity", *Marketing Science,* Vol. 29 No. 1, pp. 85-93.

Fader, P.S., Hardie, B.G. and Lee, K.L. (2005), "RFM and CLV: Using iso-value curves for customer base analysis", *Journal of Marketing Research,* Vol. 42 No. 4, pp. 415-430.

Fader, P.S., Hardie, B.G. and Shang, J. (2010), "Customer-base analysis in a discrete-time noncontractual setting", *Marketing Science,* Vol. 29 No. 6, pp. 1086-1108.

Fang, E., Palmatier, R.W. and Steenkamp, J.E.M. (2008), "Effect of Service Transition Strategies on Firm Value", *Journal of Marketing,* Vol. 72 No. 5, pp. 1-14.

Finne, M., Brax, S. and Holmström, J. (2013), "Reversed servitization paths: a case analysis of two manufacturers", *Service Business,* Vol. 7 No. 4, pp. 513-537.

Gebauer, H., Edvardsson, B., Gustafsson, A. and Witell, L. (2010a), "Match or mismatch: strategy-structure configurations in the service business of manufacturing companies", *Journal of Service Research,* Vol. 13 No. 2, pp. 198-215.

Gebauer, H., Fleisch, E. and Friedli, T. (2005), "Overcoming the service paradox in manufacturing companies", *European management journal,* Vol. 23 No. 1, pp. 14-26.

Gebauer, H., Ren, G., Valtakoski, A. and Reynoso, J. (2012), "Service-driven manufacturing: provision, evolution and financial impact of services in industrial firms", *Journal of Service Management,* Vol. 23 No. 1, pp. 120-136.

Gebauer, H. (2008), "Identifying service strategies in product manufacturing companies by exploring environment–strategy configurations", *Industrial Marketing Management,* Vol. 37 No. 3, pp. 278-291.

Gebauer, H., Edvardsson, B. and Bjurko, M. (2010b), "The impact of service orientation in corporate culture on business performance in manufacturing companies", *Journal of Service Management,* Vol. 21 No. 2, pp. 237-259.

Gebauer, H. and Fleisch, E. (2007), "An investigation of the relationship between behavioral processes, motivation, investments in the service business and service revenue", *Industrial Marketing Management,* Vol. 36 No. 3, pp. 337-348.

Gebauer, H. and Kowalkowski, C. (2012), "Customer-focused and service-focused orientation in organizational structures", *Journal of Business & Industrial Marketing,* Vol. 27 No. 7, pp. 527-537.

Glady, N., Baesens, B. and Croux, C. (2009), "A modified Pareto/NBD approach for predicting customer lifetime value", *Expert Systems with Applications,* Vol. 36 No. 2, pp. 2062-2071.

Gupta, S., Hanssens, D., Hardie, B., Kahn, W., Kumar, V., Lin, N., Ravishanker, N. and Sriram, S. (2006), "Modeling Customer Lifetime Value", *Journal of Service Research,* Vol. 9 No. 2, pp. 139-155.

Holm, M., Kumar, V. and Rohde, C. (2012), "Measuring customer profitability in complex environments: an interdisciplinary contingency framework", *Journal of the Academy of Marketing Science,* Vol. 40 No. 3, pp. 387-401.

Jaakkola, E. and Alexander, M. (2014), "The role of customer engagement behavior in value co-creation: a service system perspective", *Journal of Service Research,* Vol. 17 No. 3, pp. 247-261.

Jackson, B.B. (1985), *Winning and keeping industrial customers: The dynamics of customer relationships,* Lexington books, Lexington (Mass.).

Kasanen, E., Lukka, K. and Siitonen, A. (1993), "The constructive approach in management accounting research", *Journal of management accounting research,* Vol. 5, pp. 243.

Kohtamäki, M., Partanen, J., Parida, V. and Wincent, J. (2013), "Non-linear relationship between industrial service offering and sales growth: The moderating role of network capabilities", *Industrial Marketing Management,* Vol. 42 No. 8, pp. 1374-1385.

Kowalkowski, C., Daniel Kindström and Witell, L. (2011), "Internalisation or externalisation?: Examining organisational arrangements for industrial services", *Managing Service Quality,* Vol. 21 No. 4, pp. 373-391.

Kowalkowski, C., Windahl, C., Kindström, D. and Gebauer, H. (2015), "What service transition? Rethinking established assumptions about manufacturers' service-led growth strategies", *Industrial Marketing Management,* Vol. 45 No. 2, pp. 59-69.

Kumar, V., Aksoy, L., Donkers, B., Venkatesan, R., Wiesel, T. and Tillmanns, S. (2010a), "Undervalued or overvalued customers: capturing total customer engagement value", *Journal of Service Research,* Vol. 13 No. 3, pp. 297-310.

Kumar, V., Petersen, J.A. and Leone, R.P. (2010b), "Driving profitability by encouraging customer referrals: who, when, and how", *Journal of Marketing,* Vol. 74 No. 5, pp. 1-17.

Kumar, V. and Shah, D. (2009), "Expanding the role of marketing: from customer equity to market capitalization", *Journal of Marketing,* Vol. 73 No. 6, pp. 119-136.

Kumar, V., Shah, D. and Venkatesan, R. (2006), "Managing retailer profitability—one customer at a time!", *Journal of Retailing*, Vol. 82 No. 4, pp. 277-294.

Kumar, V., Venkatesan, R., Bohling, T. and Beckmann, D. (2008), "Practice prize report-The power of CLV: Managing customer lifetime value at IBM", *Marketing Science,* Vol. 27 No. 4, pp. 585-599.

Kwak, K. and Kim, W. (2016), "Effect of service integration strategy on industrial firm performance", *Journal of Service Management,* Vol. 27 No. 3, pp. 391-430.

Lindholm, A., Laine, T.J. and Suomala, P. (2017), "The potential of management accounting and control in global operations: Profitability-driven service business development", *Journal of Service Theory and Practice,* Vol. 27 No. 2, pp. 496-514.

Mathieu, V. (2001), "Service strategies within the manufacturing sector: benefits, costs and partnership", *International Journal of Service Industry Management,* Vol. 12 No. 5, pp. 451-475.

Neely, A. (2008), "Exploring the financial consequences of the servitization of manufacturing", *Operations Management Research,* Vol. 1 No. 2, pp. 103-118.

Nordin, F., Kindström, D., Kowalkowski, C. and Rehme, J. (2011), "The risks of providing services: Differential risk effects of the service-development strategies of customisation, bundling, and range", *Journal of Service Management,* Vol. 22 No. 3, pp. 390-408.

Oliva, R. and Kallenberg, R. (2003), "Managing the transition from products to services", *International journal of service industry management,* Vol. 14 No. 2, pp. 160-172.

Penttinen, E. and Palmer, J. (2007), "Improving firm positioning through enhanced offerings and buyer–seller relationships", *Industrial Marketing Management,* Vol. 36 No. 5, pp. 552-564.

Pfeifer, P.E. and Carraway, R.L. (2000), "Modeling customer relationships as Markov chains", *Journal of interactive marketing,* Vol. 14 No. 2, pp. 43-55.

Raddats, C., Burton, J. and Ashman, R. (2015), "Resource configurations for services success in manufacturing companies", *Journal of Service Management,* Vol. 26 No. 1, pp. 97-116.

Reinartz, W.J. and Kumar, V. (2000), "On the profitability of long-life customers in a noncontractual setting: An empirical investigation and implications for marketing", *Journal of Marketing,* Vol. 64 No. 4, pp. 17-35.

Rust, R.T., Kumar, V. and Venkatesan, R. (2011), "Will the frog change into a prince? Predicting future customer profitability", *International Journal of Research in Marketing,* Vol. 28 No. 4, pp. 281-294.

Rust, R.T., Lemon, K.N. and Zeithaml, V.A. (2004), "Return on marketing: Using customer equity to focus marketing strategy", *Journal of Marketing,* Vol. 68 No. 1, pp. 109-127.

Schmittlein, D.C., Morrison, D.G. and Colombo, R. (1987), "Counting Your Customers: Who-Are They and What Will They Do Next?", *Management science,* Vol. 33 No. 1, pp. 1-24.

Scott Long, J. (1997), *Regression models for categorical and limited dependent variables,* Sage Publications, Thousand Oaks, California.

Suarez, F.F., Cusumano, M.A. and Kahl, S.J. (2013), "Services and the business models of product firms: an empirical analysis of the software industry", *Management Science,* Vol. 59 No. 2, pp. 420-435.

Syntetos, A.A. and Boylan, J.E. (2001), "On the bias of intermittent demand estimates", *International Journal of Production Economics,* Vol. 71 No. 1, pp. 457-466.

Tenucci, A. and Cinquini, L. (2016), "The relevance of accounting Information in the servitization process", paper presented at Manufacturing Accounting Research Conference, June 15-17, Lisbon, Portugal.

Ulaga, W. and Reinartz, W.J. (2011), "Hybrid offerings: how manufacturing firms combine goods and services successfully", *Journal of Marketing,* Vol. 75 No. 6, pp. 5-23.

Valtakoski, A. (2016), "Explaining servitization failure and deservitization: A knowledge-based perspective", *Industrial Marketing Management,* Vol. 60, pp. 138-150.

Vandermerwe, S. and Rada, J. (1988), "Servitization of business: adding value by adding services", *European management journal,* Vol. 6 No. 4, pp. 314-324.

Venkatesan, R. and Kumar, V. (2004), "A Customer Lifetime Value Framework for Customer Selection and Resource Allocation Strategy", *Journal of Marketing,* Vol. 68 No. 4, pp. 106-125.

Wadsworth, E. (2012), "Buy'Til You Die-A Walkthrough", available at: ftp://gnu.cs.pu.edu.tw/network/CRAN/web/packages/BTYD/vignettes/BTYD-walkthrough.pdf (accessed 1 July 2016).

Wise, R. and Baumgartner, P. (2000), "Go downstream: The new profit imperative in manufacturing", *IEEE Engineering Management Review,* Vol. 28 No. 1, pp. 89-96.

Wübben, M. and Wangenheim, F.v. (2008), "Instant customer base analysis: Managerial heuristics often "get it right"", *Journal of Marketing,* Vol. 72 No. 3, pp. 82-93.

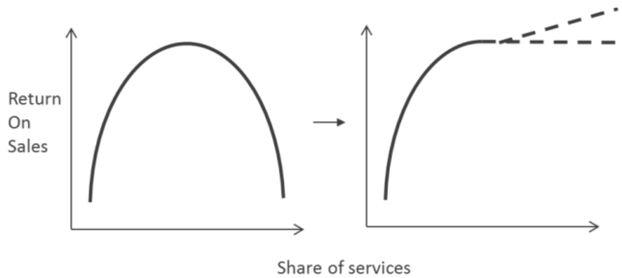

Figure 1 A possible scenario for how adequate demand forecasts can moderate the negative impact of service integration on firms' profitability

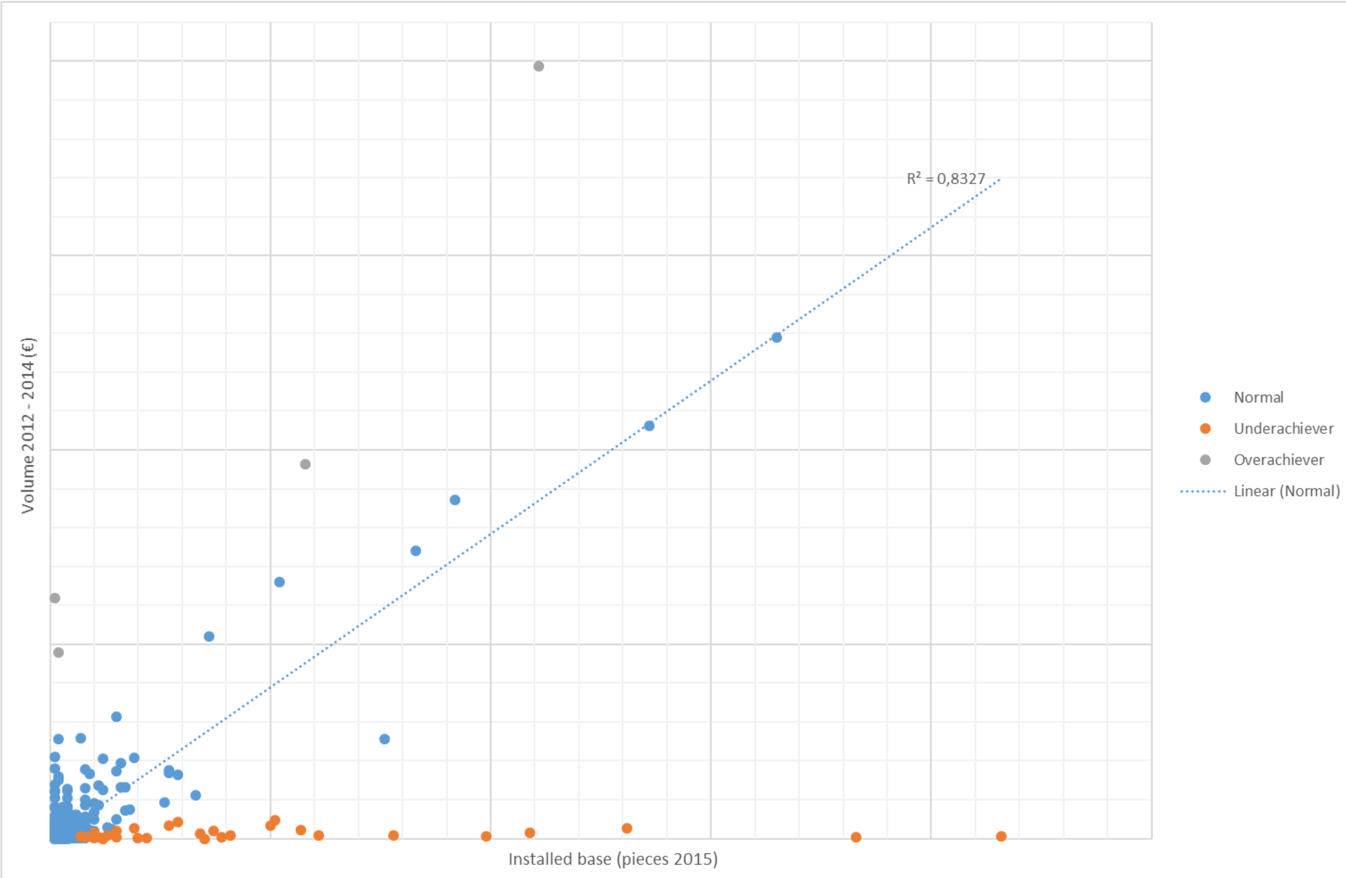

Figure 2 Customer volumes against installed base

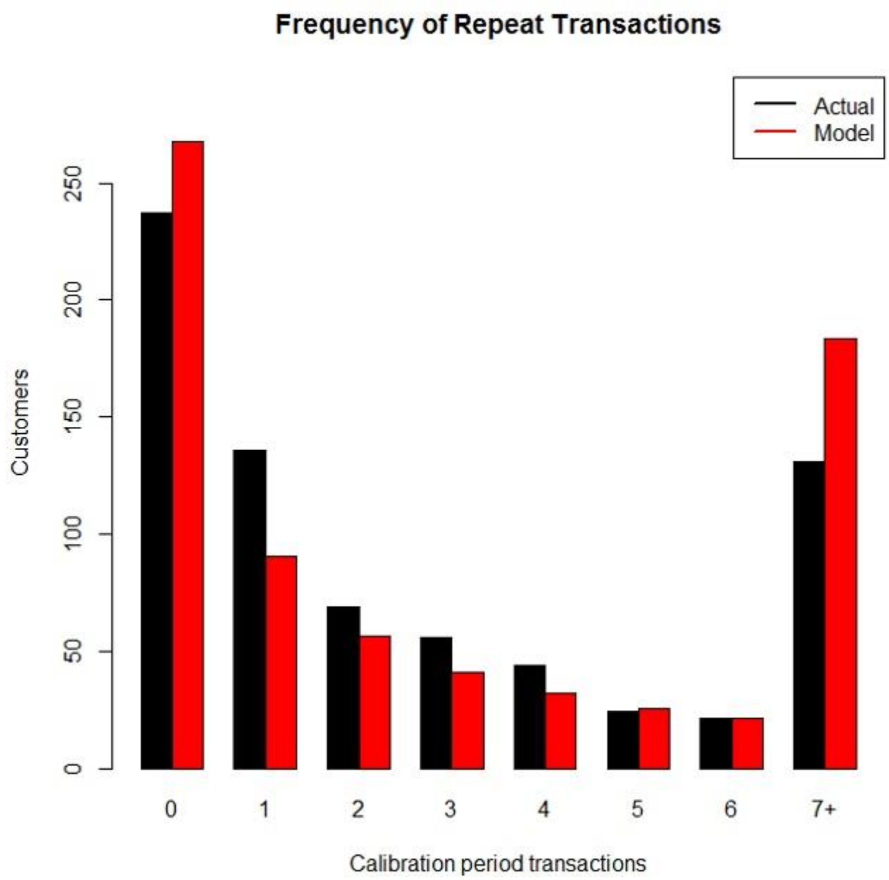

Figure 3 Number of customers against the frequency of transactions

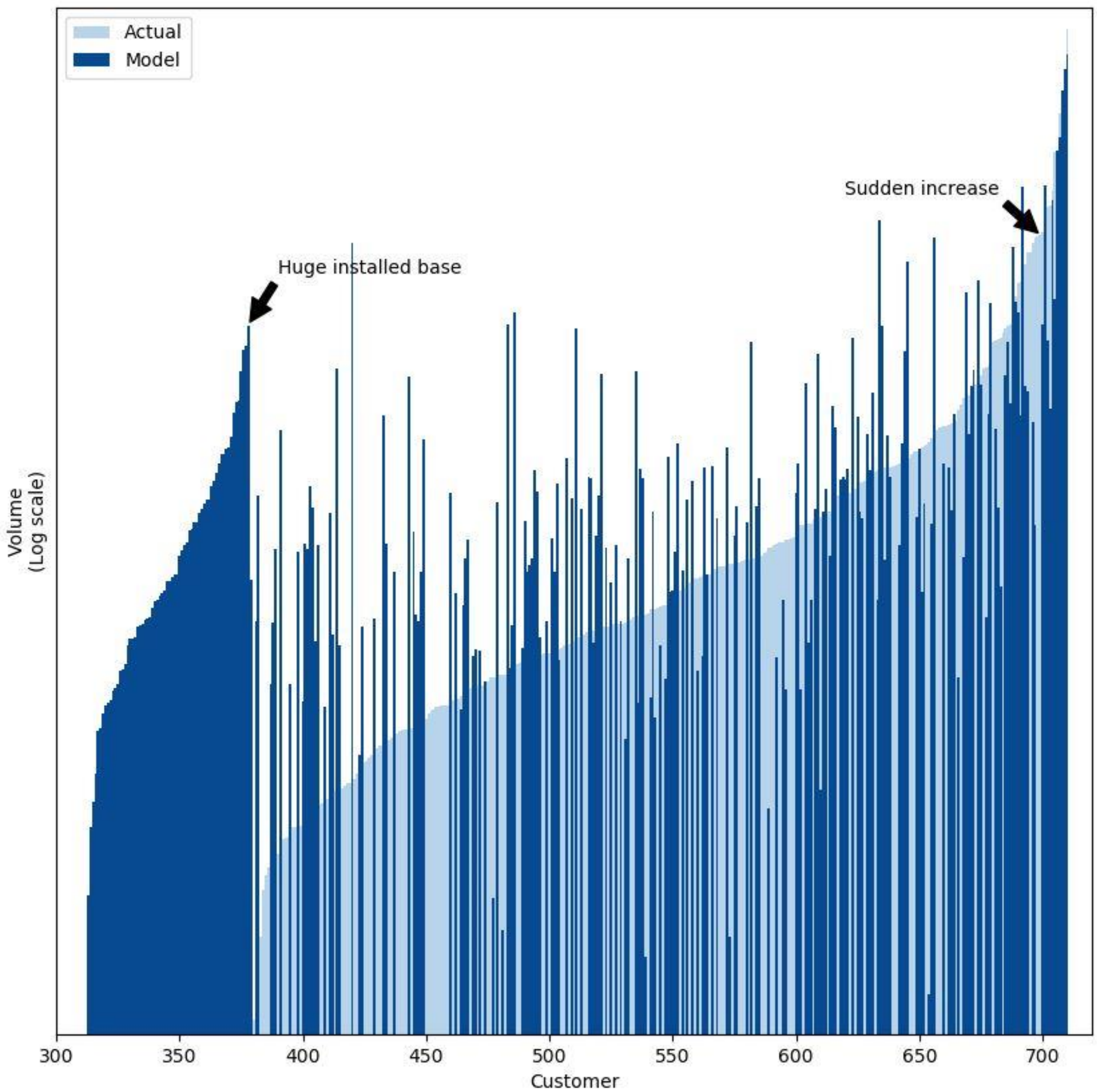

Figure 4 Actual and predicted customer volumes in 2015

Customer behavior	Industry		Frequency model	CM-model	Article
NCD	Annual donations		BG/BB	-	(Fader *et al.*, 2010)
NCC	Theoretical paper		Pareto/NBD	-	(Schmittlein *et al.*, 1987)
NCC	General retailer	B2C	Pareto/NBD	Average	(Reinartz and Kumar, 2000)
NCC	Apparel retailer, Airlines, Specialty retailer (CD)	B2C	Pareto/NBD (e.g.)		(Wübben and Wangenheim, 2008)
NCC	Investment services	B2C	Pareto/NBD (modified)	Gamma/Gamma	(Glady *et al.*, 2009)
NCC	Investment services	B2C	HH/GG	-	(Allenby *et al.*, 1999)
NCC	Computer hardware and software	B2B	HH/GG	Regression	(Venkatesan and Kumar, 2004)
NCC	Apparel retailer	B2C	HH/GG	Regression	(Kumar *et al.*, 2006)
NCC	Computer hardware	B2B	Joint regression	Joint regression	(Kumar *et al.*, 2008)
NCC	Computer hardware & software, Apparel retailer	B2B & B2C	Joint regression	Joint regression	(Kumar and Shah, 2009)
NCC	Financial services	B2C	Joint regression	Joint regression	(Kumar *et al.*, 2010b)
NCC	Computer hardware and software	B2B	Joint regression	Joint regression	(Rust *et al.*, 2011)
NCC	Iron & Steel	B2B	Retention	Regression	(Brashear Alejandro *et al.*, 2013)

Notes:
NCC non-contractual continuous in time (a transaction can occur at any time)
NCD non-contractual discrete in time (a transaction can occur only at predetermined times)
BG/BB beta-geometric/beta-Bernoulli

Table 1 Customer behavior and used models in some CLV studies

Data	Quantity
Invoices, total	> 100,000
Invoices, labor services	> 10,000
Invoices, spare parts	> 10,000
Invoices, wear parts	> 10,000
Equipment	> 1,000
Customers, total	> 1,000
Customer sample size	700

Table 2 Summary of data

Installed base size	**Active customers (%)**
Extra small	38
Small	46
Medium	53
Large	65
Extra large	78

Table 3 The proportion of active customers in different installed base size categories

Variable	*Coefficient*
Intercept	N.S.
$\boldsymbol{V_{lagged}}$	0.219***
$\boldsymbol{QT1}$	–9544***
$\boldsymbol{QT2}$	N.S.
$\boldsymbol{QT3}$	–2 631***
$\boldsymbol{QT4}$	1 335***
$\boldsymbol{QT5}$	3 983***
QT6	N.S.
$\boldsymbol{QT7}$	–14 522***
QT8	2 983**
QT9	N.S.
QT10	N.S.
QT11	1 193*
QT12	–839**
QT13	–1 185**
QT14	N.S.

Notes:
N.S. Not significant
*Significant at $\alpha = .10$
** Significant at $\alpha = .05$
*** Significant at $\alpha < .01$

Table 4 The results of the yearly volume model

Variable	*Coefficient (β)*	*Exp(β)*
BLR		
Intercept	–.844***	0. 43
V_{lagged}	0.017***	1. 17
IB_SUM	N.S.	1. 001
$STEADY\_ORDER$	2. 122***	8.
ZINBR		
Count model		
Intercept	0.295**	
V_{lagged}	0.003***	
IB_SUM	0.013*	
$STEADY\_ORDER$	0.874***	
Zero-inflation model		
Intercept	0.601*	
V_{lagged}	–0.331***	
IB_SUM	N.S.	
$STEADY\_ORDER$	N.S.	

Notes:
N.S. Not significant
*Significant at α = .10
** Significant at α = .05
*** Significant at α < .01

Table 5 Results for the BLR and ZINBR customer activity models

	0	1	2	3	4	5	6	7+
Actual frequency	0.3	0.6	0.9	1.3	1.8	1.7	1.6	8.9
Expected frequency	0.3	0.7	1	1.6	2	2.0	2.2	10.5
Number of customers	237	136	69	56	44	24	21	131

Table 6 Conditional expected transactions and number of customers

Observed		*Predicted*		
		Activity		Percentage Accuracy
		0	1	
		BLR		
Activity	0	322	37	89.7
	1	161	191	54.3
Overall Percentage				72.2
		ZINBR		
Activity	0	247	112	68.8
	1	95	257	73.0
Overall Percentage				70.9
		Pareto/NBD		
Activity	0	275	84	76.6
	1	98	254	72.2
Overall Percentage				74.4

Table 7 Comparison of the customer activity models