10.26398/IJAS.0032-016

HIGHER-ORDER PLS-PATH MODELING: A USEFUL TOOL FOR THE REDUCTION OF THE LENGTH OF THE CUSTOMER SATISFACTION QUESTIONNAIRE

Rosanna Cataldo[1], Roberto Galasso, Maria Gabriella Grassia, Marina Marino

Department of Social Science, University of Naples Federico II, Italy

Abstract *Customer satisfaction is a key issue for every company wishing to increase customer loyalty and thereby create a better business performance. As such, considerable research and revenue have been invested in developing an accurate survey process to assess consumer satisfaction. The questionnaire represents a vital part of the survey process, but its length, often, affects cooperation rates in surveys. This work focuses on using Structural Equation Modeling (SEM), specifically Partial Least Squares-Path Modeling (PLS-PM) estimated through the PLS-PM Regression Approach, as an alternative method for the analysis and study of Customer Satisfaction. In particular, the aim is to reduce the length of the questionnaire, by employing a classic questionnaire with items related to the general satisfaction of the user, a classic PLS-PM model (using the Manifest Variables (MVs) of the higher-order satisfaction block) and a higher-order PLS-PM model (not considering the MVs of the higher-order satisfaction block). The objective is to demonstrate that, by eliminating the MVs related to satisfaction and, consequently, by reducing the length of the questionnaire, a higher-order PLS-PM model produces results similar to those obtained using a classic model, both in terms of the validation of the model and the interpretation of the results.*

Keywords: *Higher-Order PLS-Path Modeling; Customer Satisfaction; Questionnaire*

1. INTRODUCTION

Without question, customer satisfaction is one of the most widely studied themes in marketing. Customer satisfaction and retention are key issues for organizations in today's competitive market place. As such, considerable research and revenue have been invested in developing accurate ways of assessing consumer satisfaction at both the macro (national) and micro (organizational) level, facilitating comparisons in performance both within and between industries. The concept of customer satisfaction occupies a central position in marketing thought and practice. Satisfaction is a major outcome of marketing activity and serves to link

[1] Corresponding author: Rosanna Cataldo Department of Social Science, University of Naples Federico II, Vico Monte della Pietá 1, 80138, Naples, Italy. E-mail: rosanna.cataldo2@unina.it

processes culminating in purchase and consumption with post-purchase phenomena such as attitude change, repeat purchase and brand loyalty. This topic has recently become a particularly salient concern for companies interested in assessing the determinants of consumer satisfaction and dissatisfaction, the centrality of the concept being reflected by its inclusion in the marketing concept that profits are generated through the satisfaction of consumer needs and wants. For this reason, for many years, organizations, both public and private, have carried out customer satisfaction surveys as an instrument for interpreting their present situation and determining their future direction.
The phases of a Customer Satisfaction survey, in analytical terms, can be considered as:

- Survey Planning;
- Data Collection;
- Processing and Interpretation of Data; and
- Presentation and Use of the Results.

During the planning of the survey, one of the most delicate stages is the formulation of the questionnaire. The questionnaire represents just one part of the survey process, but it is a vital part of that process. It is clear to anyone undertaking data collection through a questionnaire survey that the questionnaire is a crucial element in its success. The questionnaire requires an effective planning, as it is the scheme of a highly structured interview, whose purpose is to gather information on the variables, qualitative and quantitative, the subject of the survey. However, numerous studies demonstrated that going into too much detail it requires long questionnaires, and it can result in various problems, notably the consumer's abandonment of the questionnaire. Numerous studies have shown significant effects of questionnaire length on response rates in surveys. Very long questionnaires can lead to a decrease in response rates, interruptions and in completion and contents with many missing values (Dillman et al., 1993); (Stanton et al., 2002). Respondents who are not very involved or interested in the topic investigated tend to abandon the questionnaire, which can possibly lead to a sampling bias (Moore et al., 2002). MacElroy (2000) also mentions the questionnaire length among the main four factors influencing the abandonment rate. Many approaches have been introduced in the literature to face this problem finding solutions from a strictly psychometric point of view. Burisch (1997), for example, proposed a shortened versions of existing instruments through selecting items with the best psychometric properties; Smits and Vorst (2007) write that "this approach does no longer provide data on all of the original variables and this may be undesirable when a researcher wishes to completely cover the domain associated with a construct." They propose a structurally incomplete design as a good alternative to the procedure suggested by Burisch (1997) with the advantages that the prior knowledge of the psychometric properties of the items is not needed, and the data will provide information on all of the original variables. Other studies have even proposed that a single item is preferable to reduce the response burden (Bize and Plotnikoff, 2009); (Rolstad et al., 2011). Wieland et al. (2017) in their work propose a framework that provides a novel overview allowing researchers to more fully understand the structure and methods available and make better-informed scale-purification decisions.

In the marketing field, Brosnan et al. (2018) propose using constrained principal component analysis to shorten the survey length in a data-driven way by identifying optimal items with maximum information. The method allows assessing item elimination potential, and explicitly identifies which items provide maximum information for a specified number of items. Business consultant Reichheld (2003) proposed a single question as a sufficient measure and indeed the best, of customer satisfaction. Customers are asked *'How likely is it that you would recommend [brand or company X] to a friend or colleague?'* and can respond by choosing a number between 0 to 10, with 0 labeled 'not at all likely', 5 labeled 'neutral', and 10 labeled 'extremely likely'. The responses are then aggregated and transformed into a single summary statistic, the *Net-Promoter Score (NPS)*. A company's NPS is the difference between the proportion of customers placing themselves at points 9 or 10 (called 'promoters') and the proportion of customers placing themselves between 0 and 6 (called 'detractors'). Respondents on scale points 7 and 8 are called 'neutrals'. The overall score is simply calculated as the percentage of promoters minus the percentage of detractors (excluding the neutrals). Scores range from -100 (everyone is a detractor) to +100 (everyone is a promoter). In industry, a positive score is well regarded, and scores over 50 are thought to highlight good performance (Hamilton et al., 2014). According to Reichheld (2003) and his collaborators (Reichheld and Covey, 2006) the NPS question is all a company needs to ask in their customer satisfaction surveys. Their conviction that likelihood to recommend is the best measurement for businesses to understand the state of their customer relations is quite strong: "an individual's propensity for recommending a company to friends or colleagues may be the most direct gauge of customer loyalty and ultimately, financial success" (Schneider et al., 2008). However, a great deal of subsequent research on questionnaire design suggests that the measurement used for the NPS might not be optimal (Krosnick, 2018). In fact, when this type of indicator is low, it becomes impossible to understand what are the aspects that make the customer dissatisfied. In such cases, customer satisfaction analysis, while being very laborious and expensive, remains one of the few valid tools to fully understand the customers' preferences, respond to their needs and appreciate which aspects to evaluate. As a result it becomes necessary to address the problems concerning Customer Satisfaction in a different way. How then, in these cases, can we shorten the questionnaire? Partial least squares path modeling (PLS-PM) approach can be used to answer to this question because, as Tenenhaus et al. (2005) write, "the PLS approach has been applied extensively in customer satisfaction studies". The main advantage of this approach is that no scale information is required, and no scale changes are made. Bearing in mind the studies on hierarchical models, we can apply the model of Customer Satisfaction to one of these models, considering precisely satisfaction as a Latent Variable (LV), a concept that is not directly observed but are rather inferred (through a mathematical model) from other variables that are observed (directly measured) called Manifest Variables (MVs). Not wanting to eliminate the direct questions relating to the tangible aspects of the product / service / company, the elimination of generic questions concerning customer satisfaction has been proposed. This allows us to shorten the questionnaire by eliminating between 3 and 6 questions and consequently also to reduce the time of completion. The objective of this work is exactly that of demonstrating that, by eliminating the MVs related to satisfaction and, consequently, by reducing the length of the questionnaire, a higher-order PLS-PM produces results similar to those obtained using a classic model, both in relation to the validation of the model and the interpretation of the results.

2. THE THEORETICAL FRAMEWORK

2.1. PRINCIPLES OF PLS-PM

In this work, Structural Equation Modeling (SEM) (Kaplan, 2008), specifically PLS-PM (Tenenhaus et al., 2005) has been applied. In literature, there are two types of SEM: covariance-based SEM (CB-SEM) and Partial least squares SEM (PLS-SEM, also called PLS path modeling (PLS-PM)) (Hair Jr et al., 2016). For many years, CB-SEM was the dominant method for analyzing complex interrelationships between observed and latent variables. In recent years, the number of published articles using PLS-PM increased significantly relative to CB-SEM (Hair et al., 2017), and it is now widely applied in many social science disciplines

(Hair et al., 2019). PLS-PM consists of an iterative algorithm that computes the estimation of the LVs, measured by a set of MVs, and the relationships between them, by means of an interdependent system of equations based on multiple and simple regression. The idea is to determine the scores of the LVs through a process, that, iteratively, computes first an outer and then an inner estimation. As Lauro et al. (2018) write in their work, "the PLS-PM algorithm is named *soft modeling* in contrast to CB-SEM defined *hard modeling*. PLS-PM approach is aimed at model fitting, and is prediction oriented, i.e. the aim is to obtain the best prediction of the LVs; (2) the parameters of each block are estimated separately, as in Path Analysis, and by simple/multiple regression; (3) PLS-PM has the best estimation of the measurement model, because it optimizes the prediction of the LVs and the relationship between the MVs and LVs; (4) there is no problem concerning the identification of the model; (5) the estimates become consistent when the sample size gets larger; (6) it can estimate the model also in presence of multicollinearity and missing data; and (7) it is possible to estimate the model even when the number of observations is smaller than the number of the MVs".

There are many referenced review papers, in the literature, on the PLS approach to SEM, for example Chin (1998), Hair Jr et al. (2016), Hair et al. (2019) and Tenenhaus et al. (2005). Recently, Lauro et al. (2018) have presented some current developments in PLS-PM for the treatment of non-metric data, hierarchical data, longitudinal data and multi-block data.

2.2. HIGHER-ORDER PLS-PM

PLS-PM is a suitable tool for the investigation of models with a high level of abstraction (Lohmöller, 2013), in cases where the building of a system of Composite Indicators (CIs) depends on different levels of construction. Higher-order constructs (also known as hierarchical component models in the context of PLS-SEM (Lohmöller, 2013); (Becker et al., 2012)) provide a framework for researchers to model a construct on a more abstract dimension (referred to as higher-order component (HOC)) and its more concrete subdimensions (referred to as lower-order components (LOCs)). As such, they extend standard construct conceptualizations, which typically rely on a single layer of abstraction (Sarstedt et al., 2019) (Figure 1).

Different approaches have been developed and proposed in the literature:

- the Repeated Indicator Approach (Lohmöller, 2013). It is the most popular approach, and it consists of taking all indicators of the LOCs and assigned them to the HOC (Sarstedt et al., 2019);

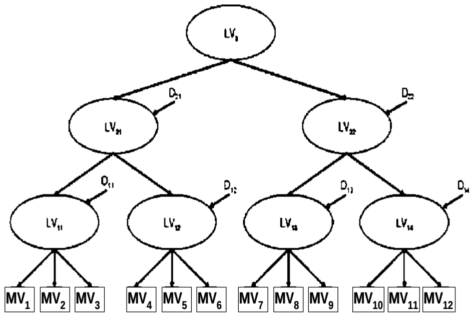

Fig: 1: Higher-Order Construct

- the Two-Step Approach (Ringle et al., 2012); (Wetzels et al., 2009). It estimates the construct scores of the first-order constructs in a first-stage model without the second-order construct present, and subsequently uses these first-stage construct scores as indicators for the higher-order latent variable in a separate second-stage analysis (Sarstedt et al., 2019);

- the Mixed Two-Step Approach (Cataldo et al., 2017). This approach begins with the implementation of the PLS-PM in the case of the Repeated Indicators Approach. In this way, the algorithm gives the scores of the LOCs. Then, the scores of the blocks are used as indicators of the HOC, and at this point the PLS-PM algorithm is performed again. This method has been implemented in order to use the component that is the best representative of its block and, at the same time, has the best predictive power on the second-order LV (Cataldo et al., 2017);

- the PLS Components Regression Approach (Cataldo et al., 2017). The PLS Component Regression Approach consists of three steps: (1) a HOC is formed of all the MVs of the LOCs; (2) PLS-Regression is applied in order to obtain h components for each block [2]; (3) once h components have been obtained, these are the MVs of the HOC and the PLS-PM algorithm is performed.

In this work, the Higher-Order PLS-PM model has been estimated with the PLS Component Regression Approach, using the “plspm” package in the R statistical software (Sanchez and Trinchera, 2012). Cataldo et al. (2017) have been demonstrated through a simulation study that this approach produces less biased and more stable parameter estimates than the approaches presented in the literature.

[2] All components that exceed a certain defined threshold are taken into consideration. See Cataldo et al. (2017)

3. AN APPLICATION CASE STUDY ON A MULTI-UTILITY COMPANY

The term “Multi-utility” refers to companies offering a wide range of services and/or products (Fortis and Liberati, 2001). In the business market, this type of service provision usually relates to energy, environmental services, waste treatment and/or telecom services. In the consumer market, it often concerns a combined provision in the field of energy and digital products and services (telephone, Internet and television). Providers like these are also referred to as multi-service providers. At the center of the multi-utility model there is the awareness that the relationship with the customer constitutes the most valuable asset and that the association of a large customer base and a brand strongly recognized in the market is a winning combination, just as there is the awareness that the strategic and efficiency advantages of this choice require a strong focus on the final phases of distribution, marketing and the sale of services.
In this scenario, it has been proposed to a medium-sized multi-utility company, operating throughout Italy, to undergo an analysis of the satisfaction of its customers. This is a pilot survey aimed at evaluating company “outcomes", attempting to understand the perceptions of the customers in terms of the quality of the services offered to them.

3.1. THE MEASUREMENT INSTRUMENT AND THE DATA

An ad-hoc survey was carried out in this pilot survey regarding the overall experience of the respondent with the company in order to monitor customer expectations and to plan the continuous improvement of the service. It is based on the ECSI (European Customer Satisfaction Index) [3] model (Fornell et al., 1996); (Bayol et al., 2000); (Tenenhaus et al., 2005) and questions regarding some of the seven constructs of the ECSI model are included (image, expectations, perceived product quality, perceived service quality, perceived value, customer satisfaction and customer loyalty) (Bayol et al., 2000). A structured questionnaire was used,

[3] The European Consumer Satisfaction Index (ECSI), developed by the European Organization for Quality (EOQ), the European Foundation for Quality Management (EFQM) and the European Academic Network for Customer-oriented Quality Analysis, and supported by the European Commission (DG III), was first introduced in 1999 across 12 European countries (Eklöf and Westlund, 2002). European experts, who developed the ECSI methodology (Committee et al., 1998), were inspired by the Swedish Customer Satisfaction Barometer (SCSB) (Fornell, 1992) introduced in 1989 and the American Customer Satisfaction Index (ACSI) (Fornell et al., 1996). The basic ECSI model is a structural equation model with unobservable latent variables. Each latent variables is operationalized by a set of measurement variables, observed by questions to customers, and the entire system is estimated using a partial least squares (PLS) method (Bayol et al., 2000); (Tenenhaus et al., 2005).

composed of scalar items, useful to investigate the dimensions listed above. As specified in the ECSI/ACSI methodologies, the responses were given on a 10 point scale: all the items were measured in relation to a Likert scale with scores ranging from 1 (strongly disagree) to 10 (strongly agree). In detail, the questionnaire is subdivided into five sections:

- *Consultant* (5 items): a section dedicated to understanding the customer satisfaction with the work of the Consultant

- *Product* (5 items): a section dedicated to understanding the customer satisfaction with the products/services provided

- *Customer Service* (5 items): a section dedicated to understanding the customer satisfaction with the customer service

- *Invoice* (4 items): a section dedicated to understanding the customer satisfaction with the invoice document and

- *Satisfaction* (6 items): a section dedicated to understanding how satisfied (or dissatisfied) the customers feel with the company.

The questionnaire was administered to all customers who had provided an e-mail address at the time of registration of the contract (1300 customers were contacted). The data were collected during the spring of 2017. A total of 600 completed questionnaires were collected, almost equal to 50% of the sample. This is a satisfying result if we consider the e-mail administration method. Furthermore, the aim of this survey was not to build a representative sample to analyze the level of customer satisfaction, but to understand if reducing the length of the questionnaire the response rate increases.

Before making the Customer Satisfaction analysis, a pre-treatment of the data was performed. First of all, the questionnaires with more than 50% of data missing were discarded (34 questionnaires), for other questionnaires missing data the “nearest neighbor” was used, which consists in introducing a concept of similarity between the units, based on a distance function. Finally, 566 questionnaires were collected and used for the analysis. Next, the basic indicators (MVs) were trasformed according to a scale from 0 to 100, where 100 represents the best evaluation (Bayol et al., 2000).

3.2. CUSTOMER SATISFACTION ANALYSIS THROUGH THE PLS-PM MODEL

A classic PLS-PM model (using multiple variables of the higher-order *Satisfaction* block) (on the left-hand side in Figure 2) was compared with a higher-order PLS-PM model (not considering the variables of the higher-order block) (on the right-hand side in Figure 2).

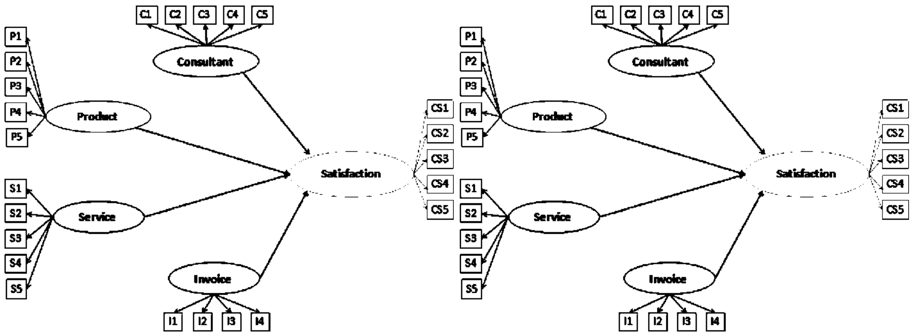

Fig. 2: The Classic Model and the Higher-order Model

Consultant, *Product*, *Customer Service* and *Invoice* are exogenous LVs with the role of input while *Satisfaction* is an endogenous LV with the role of output.
In the higher-order PLS-PM model, *Satisfaction* is conceived as a second-order latent construct affecting the first-order dimensions. This study focuses on a reflective-formative measurement model, a model resulting from the combination of reflective first-order and formative second-order constructs. In Table 1 the MVs are reported for each LV.

As the reflective blocks of MVs, it is necessary to check for unidimensionality in the meaning of factor analysis (Tenenhaus et al., 2005). For this purpose three tools are available to check the unidimensionality of a block: use of Principal Component Analysis (PCA) of the block MVs, Cronbach's Alpha and Dillon-Goldstein's Rho. A block is unidimensional if the second eigenvalue of the correlation matrix of the block MVs is smaller than 1, or at least very far from the first one. Cronbach's Alpha can be used to quantify the consistency of a block of positively correlated variables. A block is considered as internally consistent when the Cronbach's Alpha and/or the Dillon-Goldstein's Rho are larger than 0.7 (Tenenhaus et al., 2005). The statistics for checking the unidimensionality of each block are shown in Table 2, and they lead to an acceptation of the unidimensionality of all blocks for both models. This shows that the outer model is well specified and that the LVs are well measured by the MVs, their synthesis being good.

Table 3 reports the loadings, communality and redundancy indices for each

Tab. 1: LVs and MVs of the model

LVs		MVs (basic indicators)
Consultant	C1	Courtesy
	C2	Clarity
	C3	Competence
	C4	Reliability
	C5	Request Management Time
Product	P1	Satisfaction with the service offered
	P2	Satisfaction with the product
	P3	Satisfaction with the manager and individual contact
	P4	Satisfaction with the account
	P5	Satisfaction with the offer received
Customer Service	S1	Courtesy
	S2	Clarity
	S3	Competence
	S4	Reliability
	S5	Request Management Time
Invoice	I1	Satisfaction with the quality / price ratio
	I2	Ease in consulting the first page of the invoice
	I3	Ease in consulting the account page
	I4	Ease in consulting the details on the invoice
Satisfaction	CS1	Satisfaction with the Consultant
	CS2	Satisfaction with the Product
	CS3	Satisfaction with the Customer Service
	CS4	Satisfaction with the Invoice
	CS5	Company reliability
	CS6	Satisfaction with the Company

Tab. 2: The outer estimation of both the models

LVs	Mode	Cronbach Alpha	Dillon-Goldstein Rho	First eigenvalue	Second eigenvalue
Consultant	A	0.960	0.969	4.320	0.292
Product	A	0.962	0.970	4.340	0.223
Customer Service	A	0.958	0.968	4.290	0.353
Invoice	A	0.921	0.945	3.250	0.508
Satisfaction (classic model)	A	0.927	0.943	4.410	0.561
Satisfaction (higher-order model)	A	0.946	0.964	4.240	0.748

considered model. The communality and redundancy indices measure the goodness of the measurement model, measuring the percentage of variance explained by the LVs. Taking into account all the LVs , the values for these two indices are appreciably higher for all blocks, and, in particular, slightly higher if we estimate the component with the higher-order model. It is worth noting that for the classic model, the satisfaction block is determined by its MVs (SC1, SC2,...,SC6), while for the higher-order model, it is determined by the components resulting from the

PLS-Regression approach ($Y1_{u1}$, $Y2_{u1}$,....,$Y4_{u2}$) [4].

Tab. 3: Loadings and Communality for both models

		Classical model				Higher-order model		
LVs	MVs	Loading	Comm	Red	MVs	Loading	Comm	Red
Consultant	C1	0.902	0.814	0	C1	0.903	0.816	0
	C2	0.956	0.915	0	C2	0.957	0.916	0
	C3	0.957	0.916	0	C3	0.958	0.917	0
	C4	0.948	0.899	0	C4	0.948	0.899	0
	C5	0.881	0.775	0	C5	0.877	0.769	0
Product	P1	0.909	0.827	0	P1	0.909	0.827	0
	P2	0.951	0.905	0	P2	0.952	0.906	0
	P3	0.929	0.863	0	P3	0.929	0.863	0
	P4	0.933	0.871	0	P4	0.935	0.874	0
	P5	0.933	0.871	0	P5	0.935	0.874	0
Customer Service	S1	0.849	0.721	0	S1	0.854	0.729	0
	S2	0.937	0.878	0	S2	0.939	0.881	0
	S3	0.960	0.921	0	S3	0.967	0.935	0
	S4	0.952	0.906	0	S4	0.958	0.918	0
	S5	0.929	0.863	0	S5	0.938	0.879	0
Invoice	I1	0.791	0.625	0	I1	0.806	0.650	0
	I2	0.929	0.863	0	I2	0.929	0.863	0
	I3	0.939	0.882	0	I3	0.938	0.880	0
	I4	0.938	0.879	0	I4	0.935	0.874	0
Satisfaction	SC1	0.831	0.690	0.642	Y1_u1	0.997	0.995	0.994
	SC2	0.862	0.744	0.691	Y2_u1	0.997	0.994	0.994
	SC3	0.746	0.556	0.517	Y3_u1	0.990	0.981	0,981
	SC4	0.832	0.692	0.643	Y3_u2	0.514	0.264	0.264
	SC5	0.946	0.895	0.831	Y4_u1	0.992	0.984	0.984
	SC6	0.915	0.837	0.778	Y4_u2	0.492	0.242	0.242

The results of the inner estimation and of the impact of the dimensions on *Satisfaction* are reported in Figure 3.
It is very interesting to note how the two approaches give very similar results also in the interactions between the exogenous and the endogenous LVs.
In the classic model, in order to measure *Satisfaction*, *Product* is the most important dimension, with an impact of 0.3564, followed by *Invoice* (0.2823), *Consultant* (0.2693) and *Customer Service* (0.2088). In the higher-order model, in order to measure *Satisfaction*, *Product* remains the most important dimension with an

[4] For the Customer Service block and for the Invoice block, two components are considered in the analysis respectively (for the Customer Service block $Y3_{u1}$, $Y3_{u2}$, for the Invoice block $Y4_{u1}$,$Y4_{u2}$), because both exceed the criteria imposed in the PLS-Component Regression Approach. For a detail see Cataldo (2016) and Cataldo et al. (2017).

impact of 0.3878, followed by *Invoice* (0.2622), *Consultant* (0.2383) and *Customer Service* (0.1896).

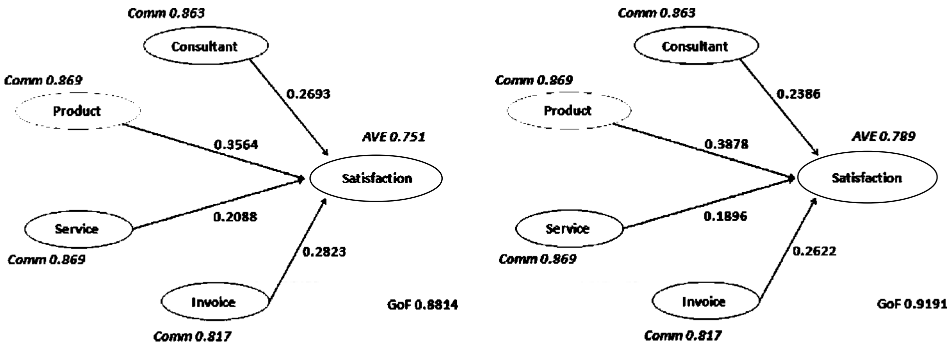

Fig. 3: The estimated classic model and the estimated higher-order model

From the Table 4 we can see that the company's customers are quite satisfied, the average values are all above sufficient levels. The customers are particularly satisfied with the *Consultant* while *Satisfaction* is the variable with the lowest scores.

Tab. 4: Average scores of the LVs

	Consultant	Product	Customer Service	Invoice	Satisfaction
Mean	7.166	6.806	6.743	6.650	6.587
St.Dev	2.974	2.923	2.526	2.610	2.917

Table 5 reports the main indices indicating the overall model quality: R^2 coefficient, the Communality index, the Average variance Extracted (AVE) and the Goodness of Fit (GoF) indices [5]. As we can see from the Table 5, in both cases all indices are very high, particularly the high R^2 indices show a good predictive power for both models (the endogenous LVs are well predicted by the explanatory LVs) and the overall model measured by GoF performs very well. However, the application case demonstrates how well the higher-order model fits over the classic model in terms of overall model quality and in terms of prediction.
In order to interpret the previous results, we proceeded with an Importance - Performance Matrix Analysis (IPMA), in which the critical aspects that have a negative impact on Satisfaction are highlighted. In particular, PLS-PM provides information on the relative importance of the dimensions in explaining other con-

[5] As concerning the goodness of fit, there is no overall fit index in PLS-PM. Nevertheless, the global criterion of GoF has been proposed by Tenenhaus et al. (2005).

Tab. 5: Overall model quality

	$R2$	Redundancies	AVE	GoF
Classic Model	0.9345	0.7021	0.7512	0.8814
Higher-order Model	0.9999	0.7893	0.7893	0.9191

structs in the structural model. Information on the importance of the constructs is relevant for drawing conclusions. For this reason, IPMA is a valuable decision making tool. The analysis is based on a scatter plot where each LV is positioned according to its mean and its pathway coefficient with respect to the target LV. The x-axis represents the total effect of the LVs on the target LV (i.e. their importance). The y-axis depicts the average construct scores of these LVs (i.e. their performance) Cataldo (2016). In Figure 4 the scatter plots for *Satisfaction* based on the two models are reported.

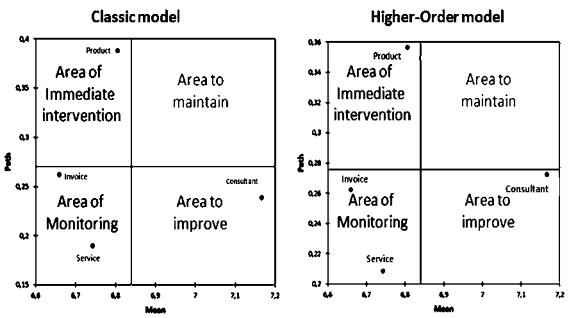

Fig. 4: The scatter plots for Satisfaction based on the two models

According to IPMA analysis, all the LVs are in the same location in both models. *Product* is in the critical area because it has a high impact on *Satisfaction* but a low mean value; *Invoice* and *Customer Service* are in the area of Monitoring in that they have a low value for the mean and the pathway coefficient; while *Consultant* is in the area to improve because it has a high mean value but a low pathway coefficient. A similar scatter plot can also be considered for the MVs for each block. In this kind of matrix, we have the possibility of analyzing the strengths, weaknesses, opportunities and threats of the constructs, that are considered in the model in order to estimate a latent concept. The graphs are valid for both approaches in that the first-order blocks are calculated in an identical way (Figure 5).

Analyzing the graphs in Figure 5 and focusing firstly on the area that in Figure 4 is in the Immediate Intervention quadrant, we can see how to improve the *Product*, namely that it is necessary to intervene on the *Satisfaction* with the *Account* (P4) and on the *Satisfaction* with the *Offer Received* (P5). Considering now the LVs

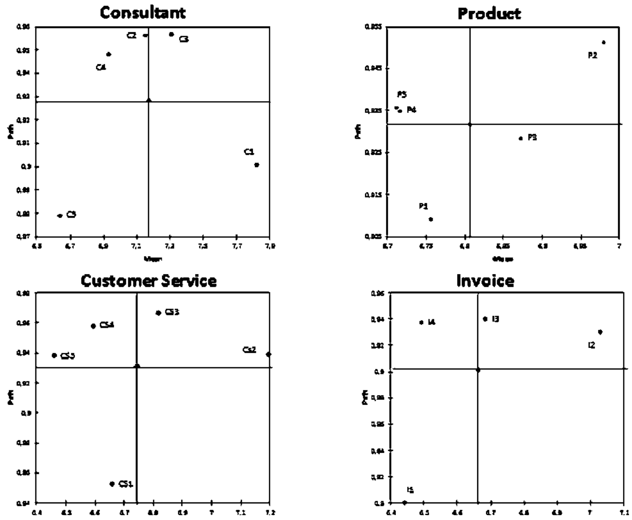

Fig. 5: The scatter plots of the basic indicators for each block

that are in the Monitoring area, we can see that in order to improve the Invoice it is necessary first to modify and make it easier to consult the *Details on the Invoice* (I4); while for the *Customer Service* a significant intervention must be made in the *Company Reliability* (CS5) and the *Satisfaction* with the *Invoice* (CS4). Finally, if you want to intervene also in the *Consultant* area, the aspects to be improved are the *Clarity of the Consultant* (C2) and, in this case, his/her *Reliability* (C4).

4. A SIMULATION STUDY

To understand if this type of result can be generalized, a simulated study has been presented. The aim of this simulation is to investigate, within the same simulation design, the performance of a classic PLS-PM model and the higher-order model estimated through the PLS Component Regression approach. The object of the simulation is to compare these performances using different sample sizes, in order to understand the effect of the sample dimension. The performances were evaluated by calculating the percentages of a well-specified pathway on the total. Starting from the IPMA it was decided to consider as well specified the pathways that are reported to be in the same quadrant of the graph in both approaches. Therefore, in mathematical terms, calculating the differences from the mean, they report the same sign. The Monte Carlo simulation was performed with the R language package. The data generation process was consistent with the procedure described by Paxton et al. (2001) for a Monte Carlo SEM study. As a first step, we defined the structure of the model and the parameters of the population. In the second step, we generated the second-order LV randomly and, given the parameters and the error terms, we estimated the first-order LVs. According to the outer parameters and error terms, in the last step, we generated the first and second-order MVs (Ciavolino and Nitti, 2013). The underlying population model used for the simulation consisted of one Second-Order LV (denoted by ξ^{II}) and four first-order LVs (denoted by ξ_1^I, ξ_2^I, ξ_3^I, and ξ_4^I), each of them formed of five MVs. The percentage of the well-specified pathway was calculated by applying classic and higher-order models to various data sets. The two approach performances were compared on the basis of the sample size (n = 100, 300, 500, 1000, 1500). The study design considered 500 replications for each condition.

The starting point was the generation of the first-order LVs ξ_i^I (i=1...4) from a Gaussian distribution with $\mu = 0$ and $\sigma^2 = 1$ ($\xi_i^I \sim N(0,1)$). The data generated were rescaled in the interval [1, 100].

Because the structural model is supposed formative, the second-order construct ξ_j^{II} (the j^{th} higher-order construct) was computed as the product of ξ_i^I by the path coefficient vector β_{ij} with the addition of an error component ζ_j according to the equation (1):

$$\xi_j^{II} = \sum_{(i:\xi_i^I \to \xi_j^{II})} \beta_{ij}\xi_i^I + \zeta_j \tag{1}$$

Each vector of the error component ζ_j is drawn from a univariate normal distribution (Hulland et al., 2010) with a mean equal to zero and a standard deviation, var(ζ_j).

For the measurement model, which is reflective, the MVs are generated starting from the LVs, given the lambda coefficients, following the equation (2):

$$X_{n,k} = \xi_{k,1}^I * (\lambda^I)_{1,k}^{-1} + \delta_{n,k} \tag{2}$$

where k is the number of MVs related to each LV and the error term was distributed as a continuous uniform: $\delta \sim U(-1,1)$.

Following Ciavolino and Nitti (2013) simulation, the path coefficients vector (β) of the structural model is assumed to have the elements equal to 0.8, while all the elements of the vector loadings (λ^I and λ^{II}) for the measurement models, are set at 0.7.

We estimated each approach with a centroid inner weighting scheme.

Tab. 6: Percentages of well-specified pathways with 5 LVs and 5 MVs per block

	N=100	N=300	N=500	N=1000	N=1500
Mean	0.933	0.931	0.934	0.936	0.937
St.Error	0.012	0.011	0.011	0.011	0.010

As can we see in the Table 6, the well-specified pathways are almost always above 90%. The percentage also tends to increase as the sample size increases.Moreover, the standard error is always lower than 0.05, thus defining itself as a stable measure.

5. CONCLUSIONS

Starting from studies on hierarchical models, a further use of such models has been proposed. In some socio-economic areas, the data are often collected by means of a questionnaire. However, the administration of questionnaires is sometimes complex due to time and money issues, and it has therefore become essential to shorten the length of this survey tool, so as to reduce time, costs and abandonment rates. This work focused on using higher-order PLS-PM as an alternative method for the analysis and study of Customer Satisfaction. In particular, the aim of the work was to demonstrate that, by eliminating the MVs related to satisfaction and, consequently, by reducing the length of the questionnaire, a higher-order PLS-PM model produces results similar to those obtained using a classic PLS-PM model with the complete questionnaire, both in terms of the validation of the model and the interpretation of the results. In order to validate this finding, a simple simulation study has been carried out which has given satisfactory results, so confirming the generalization of this type of application. This allows us to use this type of model without altering the decisions made and at the same time shortening the length of the questionnaire, thus reducing survey time, costs and abandonment rates. Higher-order PLS-PM thus proves to be a useful tool for the reduction of the length of Customer Satisfaction questionnaires. Future works will test whether the higher-order Model is always the best choice, including in the path diagram the mediation and moderation effects that characterize the ECSI/ACSI models. Furthermore, the PLS-PM approach will be compared with other methods known in the literature that could be used to reduce the length of the questionnaire.

References

Bayol, M.P., de la Foye, A., Tellier, C. and Tenenhaus, M. (2000). Use of PLS path modeling to estimate the European consumer satisfaction index (ECSI) model. In *Statistica Applicata*, 12 (3): 361–375.

Becker, J.M., Klein, K. and Wetzels, M. (2012). Hierarchical latent variable models in pls-sem: guidelines for using reflective-formative type models. In *Long range planning*, 45 (5-6): 359–394.

Bize, R. and Plotnikoff, R.C. (2009). The relationship between a short measure of health status and physical activity in a workplace population. In *Psychology, health & medicine*, 14 (1): 53–61.

Brosnan, K., Grün, B. and Dolnicar, S. (2018). Identifying superfluous survey items. In *Journal of Retailing and Consumer Services*, 43: 39–45.

Burisch, M. (1997). Test length and validity revisited. In *European Journal of Personality*, 11 (4): 303–315.

Cataldo, R. (2016). *Developments in PLS-PM for the building of a System of Composite Indicators*. Ph.D. thesis, University of Naples Federico II.

Cataldo, R., Grassia, M.G., Lauro, N.C. and Marino, M. (2017). Developments in Higher-Order PLS-PM for the building of a system of Composite Indicators. In *Quality & Quantity*, 51 (2): 657–674.

Chin, W.W. (1998). The partial least squares approach to structural equation modeling. In *Modern methods for business research*, 295 (2): 295–336.

Ciavolino, E. and Nitti, M. (2013). Simulation study for PLS path modelling with high-order construct: A job satisfaction model evidence. In *Advanced dynamic modeling of economic and social systems*, 185–207. Springer.

Committee, E.T. et al. (1998). European customer satisfaction index: foundation and structure for harmonized national pilot projects. In *Report prepared for the ECSI Steering Committee*.

Dillman, D.A., Sinclair, M.D. and Clark, J.R. (1993). Effects of questionnaire length, respondent-friendly design, and a difficult question on response rates for occupant-addressed census mail surveys. In *Public opinion quarterly*, 57 (3): 289–304.

Eklöf, J.A. and Westlund, A.H. (2002). The pan-european customer satisfaction index programme - current work and the way ahead. In *Total Quality Management*, 13 (8): 1099–1106.

Fornell, C. (1992). A national customer satisfaction barometer: The swedish experience. In *Journal of marketing*, 56 (1): 6–21.

Fornell, C., Johnson, M.D., Anderson, E.W., Cha, J., and Bryant, B.E. (1996). The american customer satisfaction index: nature, purpose, and findings. In *Journal of marketing*, 60 (4): 7–18.

Fortis, M. and Liberati, E.B. (2001). *Le imprese multiutility: Aspetti generali e prospettive dei settori a rete*. Il mulino.

Hair, J.F., Hult, G.T.M., Ringle, C.M., Sarstedt, M. and Thiele, K.O. (2017). Mirror, mirror on the wall: a comparative evaluation of composite-based structural equation modeling methods. In *Journal of the Academy of Marketing Science*, 45 (5): 616–632.

Hair, J.F., Risher, J.J., Sarstedt, M. and Ringle, C.M. (2019). When to use and how to report the results of pls-sem. In *European Business Review*.

Hair Jr, J.F., Hult, G.T.M., Ringle, C. and Sarstedt, M. (2016). *A primer on partial least squares structural equation modeling (PLS-SEM)*. Sage publications.

Hamilton, D., Lane, J.V., Gaston, P., Patton, J., Macdonald, D., Simpson, A. and Howie, C. (2014). Assessing treatment outcomes using a single question: the net promoter score. In *The bone & joint journal*, 96 (5): 622–628.

Hulland, J., Ryan, M.J. and Rayner, R.K. (2010). Modeling customer satisfaction: A comparative performance evaluation of covariance structure analysis versus partial least squares. In *Handbook of partial least squares*, 307–325. Springer.

Kaplan, D. (2008). *Structural equation modeling: Foundations and extensions*, vol. 10. Sage Publications.

Krosnick, J.A. (2018). Questionnaire design. In *The Palgrave Handbook of Survey Research*, 439–455. Springer.

Lauro, N.C., Grassia, M.G. and Cataldo, R. (2018). Model based composite indicators: New developments in partial least squares-path modeling for the building of different types of composite indicators. In *Social Indicators Research*, 135 (2): 421–455.

Lohmöller, J.B. (2013). *Latent variable path modeling with partial least squares*. Springer Science & Business Media.

MacElroy, B. (2000). Variables influencing dropout rates in web-based surveys. In *Quirk's Marketing Research Review*, 71.

Moore, K.A., Halle, T.G., Vandivere, S. and Mariner, C.L. (2002). Scaling back survey scales: How short is too short? In *Sociological Methods & Research*, 30 (4): 530–567.

Paxton, P., Curran, P.J., Bollen, K.A., Kirby, J. and Chen, F. (2001). Monte Carlo experiments: Design and implementation. In *Structural Equation Modeling*, 8 (2): 287–312.

Reichheld, F.F. (2003). The one number you need to grow. In *Harvard business review*, 81 (12): 46–55.

Reichheld, F.F. and Covey, S.R. (2006). *The ultimate question: Driving good profits and true growth*. Harvard Business School Press Boston, MA.

Ringle, C.M., Sarstedt, M. and Straub, D.W. (2012). Editor's comments: A critical look at the use of pls-sem in" mis quarterly". In *MIS quarterly*, iii–xiv.

Rolstad, S., Adler, J. and Rydén, A. (2011). Response burden and questionnaire length: is shorter better? a review and meta-analysis. In *Value in Health*, 14 (8): 1101–1108. Sanchez, G. and Trinchera, L. (2012). plspm: Partial Least Squares data analysis methods. In *R package version 0.2-2, URL http://CRAN. R-project. org/package= plspm*.

Sarstedt, M., Hair Jr, J.F., Cheah, J.H., Becker, J.M. and Ringle, C.M. (2019). How to specify, estimate, and validate higher-order constructs in pls-sem. In *Australasian Marketing Journal (AMJ)*, 27 (3): 197–211.

Schneider, D., Berent, M., Thomas, R. and Krosnick, J. (2008). Measuring customer satisfaction and loyalty: Improving the ÔNet-Promoter' score. In *Poster presented at the Annual Meeting of the American Association for Public Opinion Research, New Orleans, Louisiana*.

Smits, N. and Vorst, H.C. (2007). Reducing the length of questionnaires through structurally incomplete designs: An illustration. In *Learning and Individual Differences*, 17 (1): 25–34.

Stanton, J.M., Sinar, E.F., Balzer, W.K. and Smith, P.C. (2002). Issues and strategies for reducing the length of self-report scales. In *Personnel Psychology*, 55 (1): 167–194.

Tenenhaus, M., Vinzi, V.E., Chatelin, Y.M. and Lauro, C. (2005). Pls path modeling. In *Computational Statistics & Data Analysis*, 48 (1): 159–205.

Wetzels, M., Odekerken-Schröder, G. and Van Oppen, C. (2009). Using pls path modeling for assessing hierarchical construct models: Guidelines and empirical illustration. In *MIS quarterly*, 177–195.

Wieland, A., Durach, C.F., Kembro, J. and Treiblmaier, H. (2017). Statistical and judgmental criteria for scale purification. In *Supply Chain Management: An International Journal*.