Chapter 8
Finite Mixture Partial Least Squares Analysis: Methodology and Numerical Examples

Christian M. Ringle, Sven Wende, and Alexander Will

Abstract In wide range of applications for empirical data analysis, the assumption that data is collected from a single homogeneous population is often unrealistic. In particular, the identification of different groups of consumers and their appropriate consideration in partial least squares (PLS) path modeling constitutes a critical issue in marketing. In this work, we introduce a finite mixture PLS software implementation which separates data on the basis of the estimates' heterogeneity in the inner path model. Numerical examples using experimental as well as empirical data allow the verification of the methodology's effectiveness and usefulness. The approach permits a reliable identification of distinctive customer segments along with characteristic estimates for relationships between latent variables. Researchers and practitioners can employ this method as a model evaluation technique and thereby assure that results on the aggregate data level are not affected by unobserved heterogeneity in the inner path model estimates. Otherwise, the analysis provides further indications on how to treat that problem by forming groups of data in order to perform a multi-group path analysis.

8.1 Introduction

Structural equation modeling (SEM) and path modeling with latent variables (LVP) are applied in marketing research to measure complex cause-effect relationships (Fornell and Larcker 1981; Steenkamp and Baumgartner 2000). Covariance structure

C.M. Ringle
University of Hamburg Institute for Industrial Management and Organizations
Von-Melle-Park 5, 20146 Hamburg, Germany
e-mail: c.ringle@smartpls.de
and
University of Technology Sydney Centre for Management and Organisation Studies
PO Box 123, Broadway NSW 2007, Australia

S. Wende and A. Will
University of Hamburg Institute for Industrial Management and Organizations
Von-Melle-Park 5, 20146 Hamburg, Germany
e-mail: s.wende@smartpls.de, mail@alexwill.de

V. Esposito Vinzi et al. (eds.), *Handbook of Partial Least Squares*, Springer Handbooks of Computational Statistics, DOI 10.1007/978-3-540-32827-8_9,

analysis (CSA) (Jöreskog 1978) and partial least squares analysis (PLS) (Lohmöller 1989) constitute the two corresponding, yet different (Schneeweiß 1991), statistical techniques for estimating such models. An important research issue in SEM and LVP is the measurement of customer satisfaction (Fornell et al. 1996; Hackl and Westlund 2000), which is closely related to the requirement of identifying distinctive customer segments (ter Hofstede et al. 1999; Wu and Desarbo 2005).

In SEM, segmentation can be achieved based on the heterogeneity of scores for latent variables in the structural model (DeSarbo et al. 2006). Jedidi et al. (1997) pioneer this field of research and propose a procedure that blends finite mixture models and the expectation-maximization (EM) algorithm (McLachlan and Krishnan 2004; Wedel and Kamakura 2000). However, this technique extends CSA but is inappropriate for PLS path modeling. For this reason, Hahn et al. (2002) propose the finite mixture partial least squares (FIMIX-PLS) approach that joins a finite mixture procedure with an EM algorithm specifically regarding the ordinary least squares (OLS)-based predictions of PLS. Sarstedt (2008) reviews existing segmentation techniques for PLS path modeling and concludes that FIMIX-PLS can currently be viewed as the most comprehensive and commonly used approach to capture heterogeneity in PLS path modeling.

Building on the guiding articles by Jedidi et al. (1997) and Hahn et al. (2002), this paper presents FIMIX-PLS as it is implemented for the first time in a statistical software application (SmartPLS; Ringle et al. 2005). Thereby, this methodology for segmenting data based on PLS path modeling results is made broadly applicable for research in marketing, management and other social sciences disciplines. This kind of analysis is typically performed in two stages. In the first step, FIMIX-PLS (see Chap. 8.2) is applied for different numbers of classes using standard PLS path modelling estimates. If distinctive groups of observations in the overall set of data cause heterogeneity in the inner PLS path model estimates, FIMIX-PLS results permit detection of this heterogeneity and provide implications how to treat it by segmentation. In the second step (see ex post analysis in Chap. 8.3), an explanatory variable must be uncovered that entails both, similar clustering of data, as indicated by evaluated FIMIX-PLS outcomes, and interpretability of the formed groups of observations. Then, correspondingly separated sets of data are used as new input for segment-specific LVP computations with PLS facilitating multigroup analysis (Chin and Dibbern 2010, Chap. 8.7). Both analytical steps frame a comprehensive application of the FIMIX-PLS approach and are carried out by numerical examples with experimental (see Chap. 8.4) and empirical data (see Chap. 8.5) in this paper. The numerical examples reveal some important methodological implications that have not been addressed, yet.

As segmentation is a key element for marketers to form and improve their targeted marketing strategies, these analyses allow us to demonstrate the potentials of FIMIX-PLS for identifying homogeneous clusters of consumers with regard to the benefits they seek or in their response to marketing programs. This research is important to expand the methodological toolbox for analyzing LVP with PLS. Like the confirmatory tetrad analysis to empirically test whether a measurement model is reflective or formative (Gudergan et al. 2008), researchers and practitioners should employ FIMIX-PLS as a standard procedure to evaluate their PLS path modeling results. They thereby assure that outcomes on the aggregate data level are not affected by unobserved heterogeneity in the inner path model estimates. Otherwise, the analysis provides further indications on how to treat that problem by forming groups of data in order to perform a multi-group path analysis. Significantly distinctive group-specific path model estimations impart further differentiated interpretations of PLS modeling results and may foster the origination of more effective (marketing) strategies (Rigdon et al. 2010; Ringle et al. 2010a; Sarstedt et al. 2009).

8.2 Methodology

The first methodological step is to estimate path models by applying the basic PLS algorithm for LVP (Lohmöller 1989). Then, FIMIX-PLS is employed as formally described and discussed by its developers (Hahn et al. 2002) using the estimated scores of latent variables and their modified presentation of relationships in the inner model (see Table 8.7 in the appendix for a description of all of the symbols used in the equations presented in this paper):

$$B\eta_i + \Gamma\xi_i = \zeta_i \tag{8.1}$$

Segment-specific heterogeneity of path models is concentrated in the estimated relationships between latent variables. FIMIX-PLS captures this heterogeneity. The distributional function for each segment is defined as follows, assuming that η_i is distributed as a finite mixture of conditional multivariate normal densities $f_{i|k}(\cdot)$:

$$\eta_i \sim \sum_{k=1}^{K} \rho_k f_{i|k}(\eta_i|\xi_i, B_k, \Gamma_k, \Psi_k) \tag{8.2}$$

Substituting $f_{i|k}(\eta_i|\xi_i, B_k, \Gamma_k, \Psi_k)$ results in the following equation:

$$\eta_i \sim \sum_{k=1}^{K} \rho_k \left[\frac{|B_k|}{\sqrt[M]{2\pi}\sqrt{|\Psi_k|}} e^{-\frac{1}{2}(B_k\eta_i+\Gamma_k\xi_i)'\Psi_k^{-1}(B_k\eta_i+\Gamma_k\xi_i))} \right] \tag{8.3}$$

It is sufficient to assume multivariate normal distribution of η_i. Equations (8.4) and (8.5) represent an EM-formulation of the likelihood function and the log-likelihood (lnL) as the corresponding objective function for maximization:

$$L = \prod_i \prod_k [\rho_k f(\eta_i|\xi_i, B_k, \Gamma_k, \Psi_k)]^{z_{ik}} \tag{8.4}$$

$$lnL = \sum_i \sum_k z_{ik} ln(f(\eta_i|\xi_i, B_k, \Gamma_k, \Psi_k)) + \sum_i \sum_k z_{ik} ln(\rho_k) \tag{8.5}$$

The EM algorithm is used to maximize the likelihood in this model in order to ensure convergence. The "expectation" of (8.5) is calculated in the E-step, where z_{ik} is 1 if subject i belongs to class k (or 0 otherwise). The relative segment size ρ_k, the parameters ξ_i, B_k, Γ_k and Ψ_k of the conditional probability function are given, and provisional estimates (expected values) for z_{ik} are computed as follows according to Bayes' theorem:

$$E(z_{ik}) = P_{ik} = \frac{\rho_k f_{i|k}(\eta_i|\xi_i, B_k, \Gamma_k, \Psi_k)}{\sum_{k=1}^{K} \rho_k f_{i|k}(\eta_i|\xi_i, B_k, \Gamma_k, \Psi_k)} \tag{8.6}$$

Equation (8.5) is maximized in the M-step. Initially, new mixing proportions ρ_k are calculated by the average of adjusted expected values P_{ik} that result from the previous E-step:

$$\rho_k = \frac{\sum_{i=1}^{I} P_{ik}}{I} \tag{8.7}$$

Thereafter, optimal parameters for B_k, Γ_k, and Ψ_k are determined by independent OLS regression (one for each relationship between latent variables in the inner model). ML estimators of coefficients and variances are assumed to be identical to OLS predictions. The following equations are applied to obtain the regression parameters for endogenous latent variables:

$$Y_{mi} = \eta_{mi} \tag{8.8}$$

$$X_{mi} = (E_{mi}, N_{mi})' \tag{8.9}$$

$$E_{mi} = \begin{cases} \{\xi_1, \ldots, \xi_{A_m}\}, A_m \geq 1, a_m = 1, \ldots, A_m \wedge \xi_{a_m} \text{ is regressor of } m \\ \emptyset \text{ else} \end{cases} \tag{8.10}$$

$$N_{mi} = \begin{cases} \{\eta_1, \ldots, \eta_{B_m}\}, B_m \geq 1, b_m = 1, \ldots, B_m \wedge \eta_{b_m} \text{ is regressor of } m \\ \emptyset \text{ else} \end{cases} \tag{8.11}$$

The closed form OLS analytic formulation for τ_{mk} and ω_{mk} is given as follows:

$$\tau_{mk} = \left((\gamma_{a_m mk}), (\beta_{b_m mk})\right)' = \left[\sum_i P_{ik}(X'_{mi} X_{mi})\right]^{-1} \left[\sum_i P_{ik}(X'_{mi} Y_{mi})\right] \tag{8.12}$$

$$\omega_{mk} = \text{ cell } (m \times m) \text{ of } \Psi_k = \frac{\sum_i P_{ik}(Y_{mi} - X_{mi}\tau_{mk})(Y_{mi} - X_{mi}\tau_{mk})'}{I\rho_k} \tag{8.13}$$

The M-step computes new mixing proportions. Independent OLS regressions are used in the next E-step iteration to improve the outcomes for P_{ik}. Based on an a priori specified convergence criterion, the EM-algorithm stops whenever the lnL hardly improves (see Fig. 8.1). This is more a measure of lack of progress than a measure of convergence, and there is evidence that the algorithm is often stopped too early (Wedel and Kamakura 2000).

When applying FIMIX-PLS, the EM-algorithm monotonically increases lnL and converges towards an optimum. Experience shows that FIMIX-PLS frequently

---

// initial E-step
set random starting values for P_{ik}; set $last_{lnL} = V$; set $0 < S < 1$

repeat do
begin
    // the M-step starts here
    $\rho_k = \frac{\sum_{i=1}^{I} P_{ik}}{I} \forall k$
    determine $B_k, \Gamma_k, \Psi_k, \forall k$
    calculate $current_{lnL}$
    $\Delta = current_{lnL} - last_{lnL}$

    // the E-step starts here
    if $\Delta \geq S$ **then**
    begin
        $P_{ik} = \frac{\rho_k f_{i|k}(\eta_i|\xi_i, B_k, \Gamma_k, \Psi_k)}{\sum_{k=1}^{K} \rho_k f_{i|k}(\eta_i|\xi_i, B_k, \Gamma_k, \Psi_k)} \forall i, k$
        $last_{lnL} = current_{lnL}$
    end
end
until $\Delta < S$

Fig. 8.1 The FIMIX-PLS algorithm

stops in local optimum solutions, caused by multimodality of the likelihood, so that the algorithm becomes sensitive to starting values. Moreover, the problem of convergence in local optima seems to increase in relevance whenever component densities are not well separated or the number of parameters estimated is large and the information embedded in each observation is limited (Wedel and Kamakura 2000). This results in relatively weak updates of membership probabilities in the E-step. Some examples of simple strategies for escaping local optima include initializing the EM-algorithm from a wide range of (random) values or using sequential clustering procedures, such as K-means, to obtain an appropriate initial partition of data. If alternative starting values of the algorithm result in different local optima, then the solution with the maximum value of likelihood is recommended as best solution. An issue for future research is to address concerns whether this kind of an unsystematically selected solution reaches the global optimum.

Another crucial aspect is that FIMIX-PLS only applies mixtures to the regressions in the inner model while this is not possible for the outer model. The algorithm's static use of latent variable scores does not entail dynamically forming new groups of data and computing group-specific outer and inner PLS path model estimates in every iteration compared to a prediction oriented segmentation algorithm presented by Squillacciotti (2010), Chap. 9. Eventhough, computational experiment for various data constellations show that FIMIX-PLS performs better or equally well compared with those alternative PLS segmentation approaches such as PLS-GAS (Ringle et al. 2010b), PLS-TPM and REBUS-PLS (Esposito Vinzi et al. 2007). In FIMIX-PLS, one regression equation for each segment captures the predictor-outcome relationships at the same time that the uncovered segments are captured in the inner model and, thus, reliably accounts for heterogeneity in the relationships of latent variables as demonstrated by two numerical in the Chaps. 8.4 and 8.5. Although, FIMIX-PLS results ought not instantaneously be analyzed and interpreted. In a second analytical step, the ex-post analysis (see the following chapter), an explanatory variable must be identified that allows forming groups of data as indicated by FIMIX-PLS. Then, these a-priori segmented data is used as new inputs for PLS estimations providing group-specific latent variables scores as well as results for the outer and inner measurement models. By this means, concerns on the subject of static utilization of latent variable scores are relaxed and turned into a key advantage of this segmentation approach (Sarstedt and Ringle 2010). FIMIX-PLS is generally applicable for all kinds of PLS path models regardless of whether measurement models for latent variables are operationalized as formative or reflective (see the numerical example in Chap. 8.4).

8.3 Segmentation and Ex Post Analysis

When applying FIMIX-PLS, the number of segments is unknown and the identification of an appropriate number of K classes is not straightforward. A statistically satisfactory solution does not exist for several reasons (Wedel and Kamakura 2000), i.e., mixture models are not asymptotically distributed as chi-square and do not allow for the likelihood ratio statistic. For this reason, Hahn et al. (2002) propose the repeated operation of FIMIX-PLS with consecutive numbers of latent classes K (e.g., 1–10) and to compare the class-specific outcomes for criteria such as the lnL, the Akaike information criterion ($AIC_K = -2lnL + 2N_K$), the consistent AIC ($CAIC_K = -2lnL + (ln(I) + 1)N_K$) or the Bayesian Information Criterion ($BIC_K = -2lnL + ln(I)N_K$). The results of these heuristic measures and their comparison for different numbers of classes provide evidence about an appropriate number of segments. Moreover, an entropy statistic (EN), limited between 0 and 1, indicates the degree of separation in the individually estimated class probabilities (Ramaswamy et al. 1993):

$$EN_K = 1 - \frac{\left[\sum_i \sum_k -P_{ik} ln(P_{ik})\right]}{I ln(K)} \tag{8.14}$$

The quality of separation of the derived classes will improve the higher EN is. Values of EN above 0.5 imply estimates for P_{ik} that permit unambiguous segmentation. Thus, this criterion is especially relevant for identifying and clustering different types of customers in the field of marketing.

Given these assumptions, FIMIX-PLS is only applicable for additional analytic purposes, if an explanatory variable can be identified. An explanatory variable must facilitate both a-priori clustering of data, as indicated by the evaluated FIMIX-PLS results, and interpretability of the distinctive groups. This kind of analysis is essential for exploiting FIMIX-PLS findings for PLS path modeling, and it is the most challenging analytical step to accomplish. Hahn et al. (2002) suggest an ex post analysis of the estimated probabilities of membership using an approach proposed by Ramaswamy et al. (1993). The additional findings can be used to a-priori group data (e.g., into "younger customers" and "older customers") as well as to compute and analyze the LVP for each segment. The following numerical examples, which use experimental and empirical data, document this approach.

8.4 Example Using Experimental Data

Suppose that a market researcher has formulated a LVP on theoretically well developed cause-effect relationships. The researcher suspects, however, that an unobserved moderating factor accounts for Heterogeneity or that the data belongs to a finite number of segments. In such situations, theoretical assumptions can be used to identify a-priori moderating factors that account for consumer heterogeneity in PLS path model. This kind of strategy is not feasible in many marketing applications (Jedidi et al. 1997), and it gives rise to analytical techniques like FIMIX-PLS.

SmartPLS 2.0 (Ringle et al. 2005) is the first statistical software application for (graphical) path modeling with latent variables employing both the basic PLS algorithm (Lohmöller 1989) as well as FIMIX-PLS capabilities for the kind of segmentation proposed by Hahn et al. (2002). Applying this statistical software module to experimental data for a marketing-related path model demonstrates the potentials of the methodology for PLS-based research. In terms of heterogeneity in the inner model, it might be desirable to identify and describe price sensitive consumers (Kim et al. 1999) and consumers who have the strongest preference for another particular product attribute (Allenby et al. 1998), e.g., quality. Thus, the path model for our numerical example with experimental data has one endogenous latent variable, *Satisfaction*, and two exogenous latent variables, *Price* and *Quality*, in the inner model (DeSarbo et al. 2001; Dillon et al. 1997). The used experimental set of data consist of the following equally sized segments:

- Price-oriented customers (segment 1) – this segment is characterized by a strong relationship between *Price* and *Satisfaction* and a weak relationship between *Quality* and *Satisfaction*.
- Quality-oriented customers (segment 2) – this segment is characterized by a strong relationship between *Quality* and *Satisfaction* and a weak relationship between *Price* and *Satisfaction*.

Instead of using single item constructs, each exogenous latent variable (*Price* and *Quality*) has five indicators, and the endogenous latent variable (*Satisfaction*) is measured by three manifest variables (Sarstedt and Wilczynski 2009). We use the correlation matrix in Table 8.1 to generate experimental data. This matrix is partially adopted with changed variable names from Albers and Hildebrandt (2006) who compare, among other aspects, results of formative and reflective operationalized PLS path models with experimental data. A Monte Carlo simulation is performed employing the SEPATH module of the software application STATISTICA 7.1 to

Table 8.1 Correlation of manifest variables

	Price1	Price2	Price3	Price4	Price5	Quality1	Quality2	Quality3	Quality4	Quality5	Satisfaction1	Satisfaction2	Satisfaction3
Price1	1.00												
Price2	0.03	1.00											
Price3	−0.01	0.15	1.00										
Price4	0.06	−0.05	0.10	1.00									
Price5	−0.01	0.08	0.06	0.56	1.00								
Quality1	−0.02	0.07	0.10	−0.05	−0.06	1.00							
Quality2	−0.03	−0.05	−0.02	0.06	−0.01	0.12	1.00						
Quality3	0.07	0.02	0.01	0.01	−0.04	0.24	0.57	1.00					
Quality4	−0.02	−0.04	0.02	0.00	−0.05	0.29	0.49	0.53	1.00				
Quality5	0.05	−0.02	0.00	−0.02	−0.01	0.13	0.20	0.29	0.27	1.00			
Satisfaction1	0.15	0.14	0.19	−0.02	0.01	0.08	0.08	0.03	0.06	−0.02	1.00		
Satisfaction2	0.19	0.11	0.16	0.04	0.01	0.10	0.04	0.02	0.04	−0.01	0.85	1.00	
Satisfaction3	0.09	0.14	0.15	0.01	0.04	0.11	0.01	0.00	0.03	0.00	0.89	0.83	1.00

generate manifest variable scores. The first one hundred case values are computed for a strong relationship of 0.9 between *Price* and *Satisfaction* and a week relationship of 0.1 between *Quality* and *Satisfaction* in the inner path model (segment 1). Correspondingly, another one hundred cases reflect the characteristics of the quality-oriented segment 2 so that the full set of experimental data includes 200 cases.

PLS path modelling permits both, formative as well as reflective operationalization of latent variables' measurement model with manifest variables (Lohmöller 1989; Ringle et al. 2009). The choice depends on the theoretical foundation and interpretation of cause-effect relationships (Diamantopoulos and Winklhofer 2001; Jarvis et al. 2003; Gudergan et al. 2008; Rossiter 2002). Consequently, FIMIX-PLS must properly perform for this experimental set of data using three different examples of outer measurement models:

- Reflective case – all latent variables have reflective indicators.
- Formative case – all latent variables have formative indicators.
- Mixed case – the exogenous latent variables have a formative while the latent endogenous variable has a reflective measurement model.

To begin with, we use reflective measurement model for all three latent variables. FIMIX-PLS employs the estimates of the standard PLS procedure for this numerical example with experimental data in order to process the latent variable scores for $K = 2$ classes. The standard PLS inner model weights in Table 8.2 show that both constructs, *Price* and *Quality*, have a relatively high effect on *Satisfaction* resulting in a substantial R^2 of 0.465. An overview of results is provided by Table 8.8 in the appendix. However, it is quite misleading to instantaneously examine and further interpret these good estimates for a PLS path model.

The application of FIMIX-PLS permits additional analysis that lead to different conclusions. This procedure identifies two equally sized groups of data that exhibit segment-specific path coefficients with the same characteristics as expected for the experimental set of data (see Table 8.2). Attributable to the experimental design, segment-specific regression variances are very low for the latent endogenous variable *Satisfaction* (0.170 for segment 1 and 0.149 for segment 2) resulting in corresponding outcomes for R^2 at a high level for each segment. Among other results, SmartPLS 2.0 provides the final probability of membership P_{ik} of each case to fit into one of the two classes. More than 80% of the cases are assigned to the class they have been intended to belong to in accordance with the design of data generation in this numerical example. An EN above 0.5 indicates a good separation of data.

Table 8.2 Inner model weights

	Price → Satisfaction	Quality → Satisfaction
Standard PLS	0.538	0.450
FIMIX-PLS segment 1	0.899	0.009
FIMIX-PLS segment 2	0.113	0.902

In the second analytical step, we test the FIMIX-PLS results for segment-specific PLS analysis. The FIMIX-PLS probabilities of membership allow splitting the experimental set of data into two groups. These two sets of data are then separately used as input matrices for manifest variables to estimate the path model for each group with PLS. The FIMIX-PLS results for segment-specific relationships in the inner model are essentially re-establish by this supplementary analysis. While the lower relationship in the inner path model for each group of price- or quality-oriented consumers remains at a value around 0.1, the higher relationship is at a value close to 0.9 and R^2 is around 0.8 in both cases. An overview of these result is given by Table 8.8 in the appendix.

FIMIX-PLS reliably identifies two a-priori formed segments in this numerical example with experimental data and reflective operationalization of latent variables in the PLS path model. However, the question remains, if the methodology also properly performs for path models with formative measurement model. For this reason, all three latent variables are measured with formative indicators and, in the mixed case, *Price* and *Quality* have a formative measurement model while *Satisfaction* has reflective indicators. The standard inner PLS path model estimates as well as the FIMIX-PLS results for two segments in these additional analysis (for the formative and the mixed case) are at the same level as indicated for the reflective case. Then, in the second analytical step, we split the experimental set of data according to the FIMIX-PLS probabilities of membership P_{ik} into two sets of data that are then used as new input matrices for groups specific PLS path model estimates. The computations also provide almost the same estimates for the inner path model relationships and the R^2 of *Satisfaction* as described before for the reflective case (see Tables 8.9 and 8.10 in the appendix).

As a result from these numerical examples with experimental data, we further substantiate the earlier stated rationale that FIMIX-PLS is capable to identify and treat heterogeneity of inner path model estimates by segmentation no matter if latent variables have formative or reflective measurement models. The corresponding group-specific PLS analysis are important for marketers to further differentiate interpretations of the path model resulting in more specific recommendations for the use of the marketing-mix instruments to effectively target each group of consumers.

8.5 Marketing Example Using Empirical Data

When researchers work with empirical data and do not have a-priori segmentation assumptions to capture unobserved heterogeneity in the inner PLS path model relationships, FIMIX-PLS is often not as clear-cut as demonstrated in the foregoing example that is based on experimental data. Until now, research efforts to apply FIMIX-PLS and to assess its usefulness for expanding methodological instruments in marketing was restricted by the unavailability of a statistical software application for this kind of analysis. Since such functionalities are provided as presented in Chap. 8.2, extensive use of FIMIX-PLS with empirical data in future research ought to furnish additional findings about the methodology and its applicability. For this reason, we make use of that technique for a marketing-related path model and empirical data from Gruner+Jahr's "Brigitte Communication Analysis 2002".

Gruner+Jahr is one of the leading publishers of printed magazines in Germany. They have been conducting their communication analysis survey every other year since 1984. In the survey, over 5,000 women answer numerous questions on brands in different product categories and questions regarding their personality. The women represent a cross section of the German female population. We choose answers to questions on the Benetton fashion brand name (on a four-point scale from "low" to "high") in order to use the survey as a marketing-related example of FIMIX-PLS-based customer segmentation. We assume that Benetton's aggressive and provocative advertising in the 1990s resulted in a lingering customer heterogeneity that is more distinctive and easier to identify compared with other fashion brands in the Communication Analysis Survey (e.g., Esprit or S.Oliver).

The scope of this paper does not include a presentation of theoretically hypothesized LVP and its PLS-based estimation with empirical data (Bagozzi 1994; Hansmann and Ringle 2005). Consequently, we do not provide a discussion if one ought use CSA or PLS to estimate the cause-effect relationship model with latent variables (Bagozzi and Yi 1994), a line of reasoning if the measurement models of latent variables should be operationalized as formative or reflective (Diamantopoulos and Winklhofer 2001; Rossiter 2002) or an extensive presentation of the survey data. Our goal is to demonstrate the applicability of FIMIX-PLS to empirical data for a reduced cause-effect relationship model on branding (Yoo et al. 2000) that principally guides all kinds of LVP analysis in marketing employing this segmentation technique.

The PLS path model for Benetton's brand preference consists of one latent endogenous *Brand preference* variable, and two exogenous latent variables, *Image* and *Person*, in the inner model. All latent variables are operationalized via a reflective measurement model. Figure 8.2 illustrates the path model with the latent variables and the particular manifest variables from Gruner+Jahr's "Brigitte Communication Analysis 2002" employed. The basic PLS algorithm (Lohmöller 1989) is applied for estimating that LVP using the SmartPLS 2.0 (Ringle et al. 2005) software application.

We follow the suggestions given by Chin (1998a) and Henseler et al. (2009) for arriving at a brief evaluation of results. All relationships in the reflective measurement model have high factor loadings (the smallest loading has a value of 0.795). Moreover, results for the average variance extracted (AVE) and ρ_c are at good levels (see Table 8.11 in the appendix). The exogenous latent *Image* variable (weight of 0.423) exhibits a strong relationship to the endogenous latent *Brand preference* variable. The influence of the exogenous latent *Person* variable is considerably weaker (weight of 0.177). Both relationships are statistically significant [tested with the bootstrapping procedure using individual sign change (Tenenhaus et al. 2005)]. The endogenous latent variable *Brand preference* has a R^2 of 0.239 and, thus, is at a moderate level for PLS path models.

The FIMIX-PLS module of SmartPLS 2.0 is applied for customer segmentation based on the estimated scores for latent variables. Table 8.3 shows heuristic

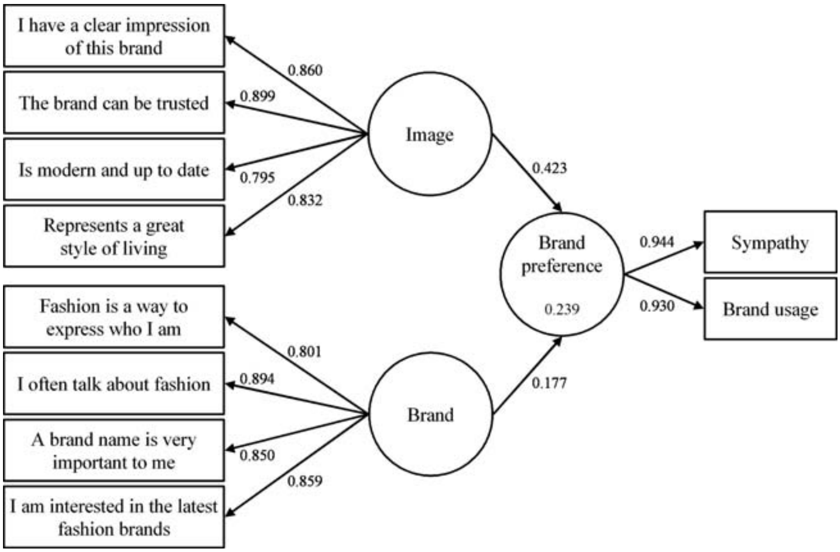

Fig. 8.2 The brand preference model

Table 8.3 Evaluation of FIMIX-PLS results

Number of latent classes	*lnL*	*AIC*	*BIC*	*CAIC*	*EN*
$K = 2$	−713.233	1448.466	1493.520	1493.545	0.501
$K = 3$	−942.215	1954.431	2097.784	2097.863	0.216
$K = 4$	−1053.389	2192.793	2450.830	2450.972	0.230
$K = 5$	−1117.976	2441.388	2846.874	2847.097	0.214

FIMIX-PLS evaluation criteria for alternative numbers of classes K. According to these results, the choice of two latent classes seems to be appropriate for customer segmentation purposes, especially in terms of EN. Compared to EN of 0.43 arrived at in the only other proficient FIMIX-PLS segmentation presented thus far in literature by Hahn et al. (2002), our EN result of 0.501 also reaches a proper level indicating well separable groups of data.

Table 8.4 presents the FIMIX-PLS results for two latent classes. In a large segment (relative size of 0.809), the explained variance of the endogenous latent *Brand preference* variable is at a relatively weak level for PLS models ($R^2 = 0.108$). The variance is explained by the exogenous latent *Image* variable, with its weight of 0.343, and the exogenous latent *Person* variable, with its weight of 0.177. A smaller segment (relative size of 0.191) has a relatively high R^2 for *Brand preference* (value of 0.930). The influence of the *Person* variable does not change much for this segment. However, the weight of the *Image* variable is more than twice as high and has a value of 0.759. This result reveals that the preference for Benetton is explained to a high degree whenever the image of this brand is far more important than the individuals' personality.

Table 8.4 FIMIX-PLS disaggregate results for two latent classes

	$K = 1$	$K = 2$
Relative segment size	0.809	0.191
R^2 (for *Brand preference*)	0.108	0.930
Path *Image* to *Brand preference*	0.343	0.759
Path *Person* to *Brand preference*	0.177	0.170

Table 8.5 A-priori segmentation based on *I like to buy fashion designers' perfumes*

	Segment 1	Segment 2
R^2 (for *Brand preference*)	0.204	0.323
Image → *Brand preference*	0.394	0.562
Person → *Brand preference*	0.164	0.104

The next step of FIMIX-PLS involves the identification of a certain variable to form and characterize the two uncovered customer segments. For this reason, we conducted an ex post analysis for finite mixture models according to the approach proposed by Ramaswamy et al. (1993). Among several possible indicators examined, the most significant explanatory variable are: *I am very interested in the latest fashion trends*, *I get information about current fashion from magazines for women*, *Brand names are very important for sports wear* and *I like to buy fashion designers' perfumes* (t-statistics ranging from 1.462 to 2.177). These variables may be appropriate for explaining the segmentation of customers into two classes.

Table 8.5 shows PLS results using the *I like to buy fashion designers' perfumes* variable for an a-priori customer segmentation into two classes. Both corresponding outcomes for segment-specific LVP estimations (see Table 8.12 in the appendix) satisfy the relevant criteria for model evaluation (Chin 1998a; Henseler et al. 2009). Segment 1 represents customers that are not interested in fashion designers' perfumes (relative size of 0.777). By contrast, segment 2 (relative size of 0.223) is characterized by female consumers that are attracted to Benetton and who would enjoy using Benetton products in other product categories, such as perfumes. From a marketing viewpoint, these customers are very important to fashion designers who want to plan for brand extensions.

Except for the *I like to buy fashion designers' perfumes* variable, the other four variables identified in the ex post analysis to explain the two classes (with reasonable t-statistics) do not offer much potential for a meaningful a-priori separation of data into two groups and segment-specific PLS path modeling. The corresponding results are at similar levels as the estimates for the full set of data. We therefore consider reasonable alternatives and test the *Customers' age* variable for an a-priori segmentation of Benetton's brand preference LVP. The ex post analysis of FIMIX-PLS results does not furnish evidence for the relevance of this variable (t-statistic of 0.690). Yet, when creating a customer segment for females over age 28 (segment 1; relative segment size: 0.793) and for younger women (segment 2; relative segment size: 0.207), we do achieve a result (see Table 8.6) that is nearly identical to the a-priori segmentation using *I like to buy fashion designers' perfumes*. The evaluation of results (Chin 1998a; Henseler et al. 2009) substantiates that the PLS path model estimates are acceptable for each segment (see Table 8.13 in the appendix).

Table 8.6 A-priori segmentation based on *Customers' age*

	Segment 1	Segment 2
R^2 (for *Brand preference*)	0.172	0.356
Image → *Brand preference*	0.364	0.559
Person → *Brand preference*	0.158	0.110

The findings that we present for the technique to uncover explanatory variables proposed by Ramaswamy et al. (1993) depict indistinct outcomes for PLS path modeling. Consequently, reliable procedures for the identification of fitting explanatory variables in the ex post analysis are required and future research must advance on this essential issue for the applicability of FIMIX-PLS.

Another implication addresses the FIMIX-PLS segment-specific estimates for relationships in the inner model and R^2 of endogenous latent variables. The procedure must be executed for successively increased numbers of classes and the outcomes for evaluation criteria must be compared in order to determine an appropriate number of segments. However, segment-specific FIMIX-PLS results are often improper for interpretation when a certain number of classes is exceeded. In most cases, the standardized weights in the inner model are at values higher than one and/or the unexplained variance of endogenous latent variables exceeds the value of one (or becomes negative). These kinds of outcomes indicate that the heterogeneity in the inner path model cannot appropriately be segmented by FIMIX-PLS for the chosen number of classes and that the analysis of additional classes may be stopped. Thus, these findings allow to further improve this methodology. Hahn et al. (2002) suggest limiting segment-specific FIMIX-PLS estimates between reasonable bounds. Future research must determine if such bounds for FIMIX-PLS computation impart useful improvements of the methodology regarding the identification of an adequate number of segments.

Our numerical example that uses empirical data demonstrates that FIMIX-PLS reliably identifies distinctive groups of customers. The larger segment tendencyally exhibits comparable results to the overall PLS path model estimates. Thus, this group of individuals is not subject for obtaining additional conclusions. In contrast, the smaller segment with a substantial relationship between *Image* and *Brand preference* is of high relevance from a marketing perspective. For these women, *Brand preference* of Benetton is foremost explained by aspects that are potentially under control of marketing activities that aim at creating an exclusive *Image* for the brand. Characteristics of the individual *Person* that are more difficult to influence by marketers are not an important issue for Benetton's brand preference in this segment of consumers. Furthermore, two kinds of explanatory variables are uncovered to form and characterize these two groups of data. Females who would like to buy Benetton's perfume or, alternatively, younger female consumers account for the smaller group of data. Hence, the specific PLS path model outcomes for the a priori formed smaller group of customers are particularly important for originating marketing strategies with regard to potential brand extensions or Benetton's target group of customers.

8.6 Summary

FIMIX-PLS allows us to capture unobserved heterogeneity in the estimated scores for latent variables in path models by grouping data. This is advantageous to a priori segmentation because homogeneous segments are explicitly generated for the inner path model relationships. The procedure is broadly applicable in business research. For example, marketing-related path modeling can exploit this approach for distinguishing certain groups of customers.

In the first numerical example involving experimental data, FIMIX-PLS reliably identifies and separates the two a-priori created segments of price- and quality-oriented customers no matter what kind of outer measurement model, reflective or formative, is employed. The second numerical example of a marketing-related path model for Benetton's brand preference is based on empirical data, and it also demonstrates that FIMIX-PLS reliably identifies an appropriate number of customer segments if distinctive groups of customers exist that cause heterogeneity within the inner model. In this case, FIMIX-PLS enables us to identify and characterize: (1) a large segment of customers that shows similar results when compared to the original model estimation as well as (2) a smaller segment of customers that is highly important for marketing programs revealing a strong relationship between *Image* and *Brand preference*.

We accordingly conclude that the methodology offers valuable capabilities to extend and further differentiate PLS-based analysis of LVP in order to develop targeted marketing strategies (Rigdon et al. 2010; Ringle et al. 2010a). Under extreme circumstances, poor standard PLS results for the overall set of data, caused by the heterogeneity of estimates in the inner model, may result in significant estimates of the inner relationships and substantial values for R^2 of endogenous latent variables for at least one group after segmentation. (Sarstedt and Ringle 2010; Sarstedt et al. 2009). Researchers and practitioners should employ FIMIX-PLS as a standard procedure to evaluate their PLS path modeling results. They thereby assure that outcomes on the aggregate data level are not affected by unobserved heterogeneity in the inner path model estimates. Otherwise, the analysis provides further indications on how to treat that problem by forming groups of data. Significantly distinctive group-specific path model estimations impart further differentiated interpretations of PLS modeling results and may foster the origination of more effective (marketing) strategies.

The initial application and critical review of this new segmentation technique for partial least squares path modeling finally allows us to unveil and discuss some of the problematic aspects (Ringle 2006) and to address significant areas of future research. As pointed out in the foregoing chapters, advances on the problem of local optimum solutions, not interpretable FIMIX-PLS estimates as well as a reliable procedure to identify explanatory variables in the ex post analysis are crucial for the applicability of this approach. In addition, extensive simulations with experimental data and broad use of empirical data are required to further exemplify how FIMIX-PLS provides additional findings for PLS path modeling.

8.7 Appendix

8.7.1 Description of Symbols

Table 8.7 Table of symbols

A_m	number of exogenous variables as regressors in regression m
a_m	exogenous variable a_m with $a_m = 1, \ldots, A_m$
B_m	number of endogenous variables as regressors in regression m
b_m	endogenous variable b_m with $b_m = 1, \ldots, B_m$
$\gamma_{a_m mk}$	regression coefficient of a_m in regression m for class k
$\beta_{b_m mk}$	regression coefficient of b_m in regression m for class k
τ_{mk}	$((\gamma_{a_m mk}), (\beta_{b_m mk}))'$ vector of the regression coefficients
ω_{mk}	cell $(m \times m)$ of Ψ_k
c	constant factor
$f_{i\|k}(\cdot)$	probability for case i given a class k and parameters $(\cdot)$
I	number of cases or observations
i	case or observation i with $i = 1, \ldots, I$
J	number of exogenous variables
j	exogenous variable j with $j = 1, \ldots, J$
K	number of classes
k	class or segment k with $k = 1, \ldots, K$
M	number of endogenous variables
m	endogenous variable m with $m = 1, \ldots, M$
N_k	number of free parameters defined as $(K - 1) + KR + KM$
P_{ik}	probability of membership of case i to class k
R	number of predictor variables of all regressions in the inner model
S	stop or convergence criterion
V	large negative number
X_{mi}	case values of the regressors for regression m of individual i
Y_{mi}	case values of the regressant for regression m of individual i
z_{ik}	$z_{ik} = 1$, if the case i belongs to class k; $z_{ik} = 0$ otherwise
ζ_i	random vector of residuals in the inner model for case i
η_i	vector of endogenous variables in the inner model for case i
ξ_i	vector of exogenous variables in the inner model for case i
B	$M \times M$ path coefficient matrix of the inner model
Γ	$M \times J$ path coefficient matrix of the inner model
Δ	difference of $current_{lnL}$ and $last_{lnL}$
B_k	$M \times M$ path coefficient matrix of the inner model for latent class k
Γ_k	$M \times J$ path coefficient matrix of the inner model for latent class k
Ψ_k	$M \times M$ matrix for latent class k containing the regression variances
ρ	$(\rho_1, \ldots, \rho_K)$, vector of the K mixing proportions of the finite mixture
ρ_k	mixing proportion of latent class k

8.7.2 PLS Path Modeling Results for Experimental Data

Table 8.8 Overview of experimental PLS modeling results (reflective case)

Analysis with reflective measurement models	PLS results for the full set of experimental data			Group 1: PLS results for a-priori segmented experimental data			Group 2: PLS results for a-priori segmented experimental data		
	Price (loadings)	Quality (loadings)	Satisfaction (loadings)	Price (loadings)	Quality (loadings)	Satisfaction (loadings)	Price (loadings)	Quality (loadings)	Satisfaction (loadings)
Price 1	0.854			0.848			0.730		
Price 2	0.872			0.888			0.740		
Price 3	0.843			0.829			0.712		
Price 4	0.860			0.882			0.871		
Price 5	0.890			0.898			0.941		
Quality 1		0.862			0.738			0.847	
Quality 2		0.864			0.639			0.843	
Quality 3		0.892			0.599			0.893	
Quality 4		0.880			0.977			0.862	
Quality 5		0.903			0.684			0.893	
Satisfaction 1			0.877			0.892			0.855
Satisfaction 2			0.861			0.867			0.872
Satisfaction 3			0.907			0.919			0.882
Price -> Satisfaction			0.538			0.873			0.169
Quality -> Satisfaction			0.450			0.077			0.898
AVE	0.746	0.775	0.777	0.756	0.547	0.797	0.646	0.754	0.756
ρ_c	0.936	0.945	0.913	0.939	0.854	0.777	0.900	0.939	0.903
R^2			0.465			0.777			0.843

Table 8.9 Overview of experimental PLS modeling results (formative case)

Analysis with formative measurement models	PLS results for the full set of experimental data			Group 1: PLS results for a-priori segmented experimental data			Group 2: PLS results for a-priori segmented experimental data		
	Price (weights)	Quality (weights)	Satisfaction (weights)	Price (weights)	Quality (weights)	Satisfaction (weights)	Price (weights)	Quality (weights)	Satisfaction (weights)
Price 1	−0.116			0.127			−0.654		
Price 2	0.079			0.090			−0.115		
Price 3	0.017			0.257			−0.778		
Price 4	0.458			0.326			0.531		
Price 5	0.641			0.342			0.928		
Quality 1		−0.005			−0.074			−0.001	
Quality 2		−0.137			−0.584			0.160	
Quality 3		0.059			−0.907			0.320	
Quality 4		0.779			0.948			0.362	
Quality 5		0.354			0.422			0.289	
Satisfaction 1			0.502			0.348			0.382
Satisfaction 2			0.167			0.279			0.411
Satisfaction 3			0.449			0.485			0.357
Price − > Satisfaction			0.527			0.875			0.065
Quality − > Satisfaction			0.406			0.052			0.892
R^2			0.470			0.784			0.835

Table 8.10 Overview of experimental PLS modeling results (mixed case)

Analysis with formative and reflective measurement models	PLS results for the full set of experimental data			Group 1: PLS results for a-priori segmented experimental data			Group 2: PLS results for a-priori segmented experimental data		
	Price (weights)	Quality (weights)	Satisfaction (loadings)	Price (weights)	Quality (weights)	Satisfaction (loadings)	Price (weights)	Quality (weights)	Satisfaction (loadings)
Price 1	−0.066			0.168			0.647		
Price 2	0.037			0.071			−0.102		
Price 3	0.016			0.248			−0.787		
Price 4	0.436			0.303			0.539		
Price 5	0.659			0.352			0.927		
Quality 1		−0.098			−0.102			0.005	
Quality 2		−0.109			−0.602			0.158	
Quality 3		0.078			−0.901			0.322	
Quality 4		0.785			0.952			0.359	
Quality 5		0.384			0.330			0.287	
Satisfaction 1			0.881			0.892			0.852
Satisfaction 2			0.862			0.867			0.876
Satisfaction 3			0.902			0.919			0.880
Price -> Satisfaction			0.513			0.873			0.063
Quality -> Satisfaction			0.418			0.053			0.893
AVE			0.777			0.797			0.756
ρ_c			0.913			0.922			0.903
R^2			0.464			0.781			0.835

8.7.3 PLS Path Modeling Results for the Example with Empirical Data

Table 8.11 Overview of empirical PLS path modeling results

	PLS results for the full set of empirical data		
	Image	Person	Brand Preference
I have a clear impression of this brand	0.860		
This brand can be trusted	0.899		
Is modern and up to date	0.795		
Represents a great style of living	0.832		
Fashion is a way to express who I am		0.801	
I often talk about fashion		0.894	
A brand name is very important to me		0.850	
I am interested in the latest trends		0.859	
Sympathy			0.944
Brand usage			0.930
AVE	0.718	0.725	0.881
ρ_c	0.910	0.913	0.937
R^2			0.239
Image − > Brand preference			0.423
Person − > Brand preference			0.177
Relative segment size			1.000

Table 8.12 Segment-specific PLS results for *I like to buy…*

	Group 1: PLS results for a-priori segmented empirical data based on the explanatory variable *I like to buy fashion designers' perfumes*			Group 2: PLS results for a-priori segmented empirical data based on the explanatory variable *I like to buy fashion designers' perfumes*		
	Image	Person	Brand Preference	Image	Person	Brand Preference
I have a clear impression of this brand	0.851			0.873		
This brand can be trusted	0.894			0.906		
Is modern and up to date	0.783			0.832		
Represents a great style of living	0.821			0.870		
Fashion is a way to express who I am		0.761			0.788	
I often talk about fashion		0.878			0.645	
A brand name is very important to me		0.851			0.747	
I am interested in the latest trends		0.823			0.718	
Sympathy			0.949			0.906
Brand usage			0.930			0.934
AVE	0.703	0.687	0.882	0.758	0.528	0.847
ρ_c	0.904	0.898	0.938	0.926	0.816	0.917
R^2			0.204			0.323
Image − > Brand preference			0.394			0.562
Person − > Brand preference			0.164			0.104
Relative segment size			0.777			0.223

Table 8.13 Segment-specific PLS results for *Customers' age*

	Group 1: PLS results for a-priori segmented empirical data based on the explanatory variable *Customers' age*			Group 1: PLS results for a-priori segmented empirical data based on the explanatory variable *Customers' age*		
	Image	Person	Brand Preference	Image	Person	Brand Preference
I have a clear impression of this brand	0.860			0.850		
This brand can be trusted	0.879			0.928		
Is modern and up to date	0.777			0.888		
Represents a great style of living	0.807			0.897		
Fashion is a way to express who I am		0.769			0.849	
I often talk about fashion		0.894			0.834	
A brand name is very important to me		0.867			0.752	
I am interested in the latest trends		0.831			0.818	
Sympathy			0.936			0.958
Brand usage			0.925			0.940
AVE	0.692	0.708	0.866	0.794	0.663	0.900
ρc	0.900	0.906	0.928	0.939	0.887	0.948
R^2			0.172			0.356
Image -> Brand preference			0.364			0.559
Person -> Brand preference			0.158			0.110
Relative segment size			0.793			0.207

References

Albers, S., & Hildebrandt, L. (2006). Methodische Probleme bei der Erfolgsfaktorenforschung – Messfehler, versus reflektive Indikatoren und die Wahl des Strukturgleichungs-Modells. *Schmalenbachs Zeitschrift fr betriebswirtschaftliche Forschung*, *58*(1), 2–33.

Allenby, G. M., Arora, N., & Ginter, J. L. (1998). On the heterogeneity of demand. *Journal of Marketing Research*, *35*(3), 384–389.

Bagozzi, R. P. (1994). Structural equation models in marketing research: basic principles. In R. P. Bagozzi (Ed.), *Principles of marketing research* (pp. 317–385). Oxford: Blackwell.

Bagozzi, R. P., & Yi, Y. (1994). Advanced topics in structural equation models. In R. P. Bagozzi, (Ed.), *Principles of marketing research* (pp. 1–52). Oxford: Blackwell.

Chin, W. W. (1998). The partial least squares approach to structural equation modeling. In G. A. Marcoulides, (Ed.), *Modern methods for business research* (pp. 295–358). Mahwah: Lawrence Erlbaum Associates.

Chin, W. W., & Dibbern, J. (2010). A permutation based procedure for multi-group PLS analysis: results of tests of differences on simulated data. In replace V. Esposito Vinzi, W. W. Chin, J. Henseler, & H. Wang, (Eds.), *Handbook of computational statistics (Vol. II): Partial Least Squares: Concepts, Methods and Applications*, 171–193. Berlin: Springer.

DeSarbo, W. S., Anthony Benedetto, C., Jedidi, K., & Song, M. (2006). Identifying sources of heterogeneity for empirical deriving strategic types: a constrained finite-mixture structural-equation methodology. *Management Science*, *62*(6), 909–924.

DeSarbo, W. S., Jedidi, K., & Sinha, I. (2001). Customer value analysis in a heterogeneous market. *Strategic Management Journal*, *22*(9), 845–857.

Diamantopoulos, A., & Winkelhofer, H. (2001). Index construction with formative indicators: an alternative to scale development. *Journal of Marketing Research*, *38*(2), 269–277.

Dillon, W. R., White, J. B., Rao, V. R., & Filak, D. (1997). Good science: use structural equation models to decipher complex customer relationships. *Marketing Research*, *9*(4), 22–31.

Esposito Vinzi, V., Ringle, C.M., Squillacciotti, S., Trinchera, L. (2007): Capturing and Treating Unobserved Heterogeneity by Response Based Segmentation in PLS Path Modeling: A Comparison of Alternative Methods by Computational Experiments, ESSEC Research Center, Working Paper No. 07019, ESSEC Business School Paris-Singapore, Cergy Pontoise Cedex 2007.

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

Fornell, C., & Larcker, D. F. (1981). Structural equation models with unobservable variables and measurement error: algebra and statistics. *Journal of Marketing Research*, *18*(3), 328–388.

Gudergan, S., Ringle, C.M., Wende, S., Will, A. (2008): Confirmatory tetrad analysis in PLS path modeling, in: Journal of Business Research (JBR), Volume 61 (2008), Issue 12, pp. 1238–1249.

Hackl, P., & Westlund, A. H. (2000). On structural equation modeling for customer satisfaction measurement. *Total Quality Management*, *11*(4/5/6), 820–825.

Hahn, C., Johnson, M. D., Herrmann, A., & Huber, F. (2002). Capturing customer heterogeneity using a finite mixture PLS approach. *Schmalenbach Business Review*, *54*(3), 243–269.

Henseler, J., Ringle, C. M., Sinkovics, R. R.: The use of partial least squares path modeling in international marketing, in: Sinkovics, R. R., Ghauri, P. N. (eds.), Advances in International Marketing, Vol. 20, Bingley 2009, pp. 277–320.

Hansmann, K., & Ringle, C. M. (2005). Strategies for cooperation. In T. Theurl & E. C. Meyer, (Eds.), *Competitive advantage of the cooperatives' networks* (pp. 131–152). Aachen: Shaker.

Jarvis, C. B., MacKenzie, S. B., & Podsakoff, P. M. (2003). A critical review of construct indicators and measurement model misspecification in marketing and consumer research. *Journal of Consumer Research*, *30*(2), 199–218.

Jedidi, K., Jagpal, H. S., & DeSarbo, W. S. (1997). Finite-mixture structural equation models for response-based segmentation and unobserved heterogeneity. *Marketing Science*, *16*(1), 39–59.

Jöreskog, K. G. (1978). Structural analysis of covariance and correlation matrices. *Psychometrika*, *43*(4), 443–477.

Kim, B., Srinivasan, K., & Wilcox, R. T. (1999). Identifying price sensitve consumers: the relative merits of domogrphic vs. purchase pattern information. *Journal of Retailing*, *75*(2), 173–193.

Lohmöller, J. (1989). *Latent variable path modeling with partial least squares*. Heidelberg: Verlag.

McLachlan, G. J., & Krishnan, T. (2004). *The EM algorithm and extensions*. Chichester: Wiley.

Ramaswamy, V., DeSarbo, W. S., Reibstein, D. J., & Robinson, W. T. (1993). An empirical pooling approach for estimating marketing mix elasticities with PIMS data. *Marketing Science*, *12*(1), 103–124.

Rigdon, E.E., Ringle, C.M., Sarstedt, M.: Structural Modeling of Heterogeneous Data with Partial Least Squares, in: Malhotra, N.K. (ed.), Review of Marketing Research, Volume 7, Armonk, 2010, forthcoming.

Ringle, C. M., Wende, S., & Will, A. (2005). SmartPLS 2.0 (beta). http://www.smartpls.de.

Ringle, C. M., Sarstedt, M., Mooi, E. A. (2009a): Response-based segmentation using FIMIX-PLS: Theoretical foundations and an application to American customer satisfaction index data, in: Stahlbock, R., Crone, S. F., Lessmann, S. (eds.): Annals of Information Systems, Special Issue on Data Mining, Vol. 8, Berlin-Heidelberg 2010, 19–49.

Rossiter, J. R. (2002). The C-OAR-SE procedure for scale development in marketing. *International Journal of Research in Marketing*, *19*(4), 305–335.

Sarstedt, M. (2008): A review of recent approaches for capturing heterogeneity in partial least squares path modelling. *Journal of Modelling in Management*, *3*(2): 140–161.

Sarstedt, M., & Wilczynski, P. (2009). More for Less? A Comparison of Single-Item and Multi-Item Measures. Business Administration Review, *69*(2), 211–227.

Sarstedt, M., Schwaiger, M., & Ringle, C. M. (2009). Do we fully understand the critical success factors of customer satisfaction with industrial goods? - Extending Festge and Schwaigers model to account for unobserved heterogeneity, in: Journal of Business Market Management (JBM), Volume 3 (2009), Issue 3, pp. 185–206.

Sarstedt, M., Ringle, C. M. (2010). Treating unobserved heterogeneity in PLS path modelling: A comparison of FIMIX-PLS with different data analysis strategies, in: Journal of Applied Statistics, Vol. 37, 2010, in press.

Schneeweiß, H. (1991). Models with latent variables: LISREL versus PLS. *Statistica Neerlandica*, *45*(2), 145–157.

Squillacciotti, S. (2010). Prediction oriented classification in PLS path modeling. In V. Esposito Vinzi, W. W. Chin, J. Henseler, & H. Wang, (Eds.), *Handbook of Computational Statistics (Vol. II): Partial Least Squares: Concepts, Methods and Applications*, 219–234. Berlin: Springer.

Steenkamp, J., & Baumgartner, H. (2000). On the use of structural equation models for marketing modeling. *International Journal of Research in Marketing*, *17*(2/3), 195–202.

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

Ter Hofstede, F., Steenkamp, J. E. M., & Wedel, M. (1999). International market segmentation based on consumer-product relations. *Journal of Marketing Research*, *36*(1), 1–17.

Wedel, M., & Kamakura, W. (2000). *Market segmentation: conceptual and methodological foundations* (2nd ed.). Dordrecht: Kluwer.

Wu, J., & Desarbo, W. S. (2005). Market segmentation for customer satisfaction studies via a new latent structure multidimensional scaling model. *Applied Stochastic Models in Business and Industry*, *21*(4/5), 303–309.

Yoo, B., Donthu, N., & Lee, S. (2000). An examination of selected marketing mix elements and brand equity. *Academy of Marketing Science Journal*, *28*(2), 195–211.