MULTIGROUP ANALYSIS IN PARTIAL LEAST SQUARES (PLS) PATH MODELING: ALTERNATIVE METHODS AND EMPIRICAL RESULTS

Marko Sarstedt, Jörg Henseler and Christian M. Ringle

ABSTRACT

Purpose – *Partial least squares (PLS) path modeling has become a pivotal empirical research method in international marketing. Owing to group comparisons' important role in research on international marketing, we provide researchers with recommendations on how to conduct multigroup analyses in PLS path modeling.*

Methodology/approach – *We review available multigroup analysis methods in PLS path modeling and introduce a novel confidence set approach. A characterization of each method's strengths and limitations and a comparison of their outcomes by means of an empirical example extend the existing knowledge of multigroup analysis methods. Moreover, we provide an omnibus test of group differences (OTG), which allows testing the differences across more than two groups.*

Measurement and Research Methods in International Marketing
Advances in International Marketing, Volume 22, 195–218

ISSN: 1474-7979/doi:10.1108/S1474-7979(2011)0000022012

Findings – *The empirical comparison results suggest that Keil et al.'s (2000) parametric approach can generally be considered more liberal in terms of rendering a certain difference significant. Conversely, the novel confidence set approach and Henseler's (2007) approach are more conservative.*

Originality/value of paper – *This study is the first to deliver an in-depth analysis and a comparison of the available procedures with which to statistically assess differences between group-specific parameters in PLS path modeling. Moreover, we offer two important methodological extensions of existing research (i.e., the confidence set approach and OTG). This contribution is particularly valuable for international marketing researchers, as it offers recommendations regarding empirical applications and paves the way for future research studies aimed at comparing the approaches' properties on the basis of simulated data.*

INTRODUCTION

Studies on international marketing have frequently made use of partial least squares (PLS) path modeling (Hair, Ringle, & Sarstedt, 2011; Hair, Sarstedt, Ringle, & Mena, 2012; Lohmöller, 1989; Wold, 1975, 1982) to empirically test theoretical models (for an overview, see Henseler, Ringle, & Sinkovics, 2009). As part of international marketing researchers' toolbox, PLS path modeling has become a pivotal instrument for estimating and analyzing complex path relationships between latent variables. This method belongs to a family of alternating least squares algorithms that extend principal component analysis and canonical correlation analysis to estimate (mainly linear) relationships between latent variables (Lohmöller, 1989).

As with any other statistical method, PLS path modeling applications are usually based on the assumption that the analyzed data stem from a single population (i.e., a unique global model represents all the observations well). However, in many real-world applications, such as in international marketing, this assumption of homogeneity is unrealistic, because individuals are likely to be heterogeneous in their perceptions and evaluations of latent constructs (e.g., Jedidi, Jagpal, & DeSarbo, 1997; Sarstedt & Ringle, 2010). This notion holds specifically for research on international marketing, which often analyzes differences in parameters in respect of different subpopulations such as countries and cultures (Brettel, Engelen, Heinemann, & Vadhanasindhu, 2008; Graham, Mintu, & Rodgers, 1994; Grewal, Chakravarty, Ding, & Liechty, 2008; Rodríguez & Wilson, 2002). Although several studies explicitly broach the issue of group-specific effects in their research questions, ignoring population heterogeneity – when performing PLS path modeling on an aggregate data level – can seriously bias the results and, thereby, yield inaccurate management conclusions (Sarstedt, Schwaiger, & Ringle, 2009).

Although cross-national or cross-cultural differences are related to observed heterogeneity, there can also be unobserved heterogeneity that cannot be attributed to one (or more) pre-specified variable(s). Similar to ignoring observed heterogeneity, unobserved heterogeneity is a serious problem in respect of interpreting PLS path modeling results if it is not considered in the analysis. Various response-based segmentation approaches have recently been developed to deal with unobserved heterogeneity. These segmentation approaches generalize, for example, genetic algorithm (Ringle, Sarstedt, & Schlittgen, 2010), and typological regression approaches (Esposito Vinzi, Ringle, Squillacciotti, & Trinchera, 2007; Esposito Vinzi, Trinchera, Squillacciotti, & Tenenhaus, 2008) to PLS path modeling. Finite mixture PLS (FIMIX-PLS; Sarstedt & Ringle, 2010; Hahn, Johnson, Herrmann, & Huber, 2002; Sarstedt, Becker, Ringle, & Schwaiger, 2011) is currently regarded the primary approach of all these segmentation techniques, and has become mandatory for evaluating PLS path modeling results (Sarstedt, 2008; Hair et al., 2012). Hair et al. (2011, p. 147), for example, point out that ''using this technique, researchers can either confirm that their results are not distorted by unobserved heterogeneity or they can identify thus far neglected variables that describe the uncovered data segments.'' Although these response-based segmentation approaches rely on different statistical concepts, they all share the same final analysis step: A comparison of the PLS parameter estimates across the identified latent segments (e.g., Rigdon, Ringle, & Sarstedt, 2010; Ringle, Sarstedt, & Mooi, 2010). Therefore, no matter whether heterogeneity is observed or unobserved, there is a need for PLS-based approaches to multigroup analysis.

Despite its obvious importance for the international marketing discipline, research on multigroup analysis is a rather new field. Only a small number of methodologically oriented articles have to date been dedicated to the discussion of available approaches (e.g., Chin & Dibbern, 2010; Rigdon et al., 2010). Researchers' discussions, for example, on internet forums like http://www.smartpls.de, show that there is a strong need to clarify how multigroup analysis can be carried out within a PLS path modeling framework. Given this background, the purpose of this chapter is to illustrate the use of multigroup analysis procedures in PLS path modeling. Specifically, we describe available multigroup analysis approaches, comment on their strengths and limitations, and illustrate their use by means of an empirical example. We also propose a novel nonparametric approach based on a comparison of bootstrap confidence intervals. This method has been designed as a more conservative approach to PLS multigroup analysis.

Prior approaches to PLS multigroup analysis are restricted in that they only allow testing the differences in two groups' parameters. However, researchers in international marketing and other cross-cultural research fields frequently encounter situations in which they would like to compare more than two groups. A naive approach would be to conduct all possible pairwise group comparisons, which would, however, quickly boost the familywise error rate beyond any prespecified acceptable Type-I error level (Mooi & Sarstedt, 2011). To overcome this problem, we introduce a permutation-based analysis of variance approach, which maintains the familywise error rate, does not rely on distributional assumptions, and exhibits an acceptable level of statistical power.

MULTIGROUP ANALYSIS IN PLS PATH MODELING

Conceptually, the comparison of group-specific effects entails the consideration of a categorical moderator variable which, in line with Baron and Kenny (1986, p. 1174), "affects the direction and/or strength of the relation between an independent or predictor variable and a dependent or criterion variable." Following this concept, group effects are nothing more than a variable's moderating effect whereby the categorical moderator variable expresses each observation's group membership (Henseler et al., 2009). As a consequence, multigroup analysis is generally regarded as a special case of modeling continuous moderating effects (Henseler & Chin, 2010; Henseler & Fassott, 2010). Fig. 1 illustrates the categorical moderator variable concept graphically. Here, x_1 to x_3 represent (reflective) indicator variables of an exogenous latent variable ξ, y_1 to y_3 represent (reflective) indicator variables of an endogenous latent variable η, and θ is the parameter of the relationship between ξ and η. Lastly, m represents a categorical moderating variable, which potentially exerts an influence on all

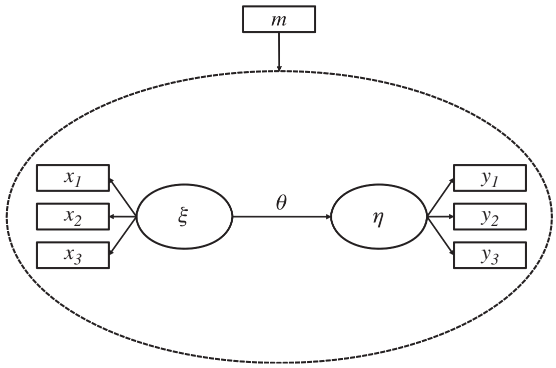

Fig. 1. Moderator Modeling Framework.

model relations. Researchers are usually interested in analyzing group effects related to structural model relations. More precisely, a population parameter θ is hypothesized as different across two subpopulations (i.e., $\theta^{(1)}$ and $\theta^{(2)}$), which are expressed by different modalities in m.

A primary concern when comparing model estimates across groups is ensuring that the construct measures are invariant across the groups. Amongst other criteria, as described by Steenkamp and Baumgartner (1998), this entails, for example, that the estimates satisfy the requirement of measurement invariance. With reference to Fig. 1, this requirement implies that the moderator variable's effect is restricted to the parameter θ and does not entail group-related differences in the item loadings.

Three approaches to multigroup analysis have been proposed within a PLS path modeling framework thus far. The first approach, introduced by Keil et al. (2000), involves estimating model parameters for each group separately, and using the standard errors obtained from bootstrapping as the input for a parametric test. This method is generally labeled the parametric approach (Henseler, 2007). Chin (2003b) proposed and further described a distribution-free data permutation test (Chin & Dibbern, 2010; Dibbern & Chin, 2005), because the parametric approach's distributional assumptions do not fit PLS path modeling's distribution-free character. This test seeks to scale the observed differences between groups by comparing these differences to those between groups randomly assembled from the data. Henseler (2007) proposed and described another nonparametric procedure, which directly compares group-specific bootstrap estimates from each bootstrap sample (see also Henseler et al., 2009).

Parametric Approach

The parametric approach was initially applied by Keil et al. (2000) (see also Chin, 2000) and depicts a modified version of the two independent samples t-test. As such, this approach requires the data (i.e., the PLS estimations of a certain path coefficient across all bootstrapping subsamples) to be normally distributed, which runs contrary to PLS path modeling's distribution-free character. Consequently, researchers should run a Kolmogorov–Smirnov test with Lilliefors correction – or, in the case of small sample sizes below 50, the Shapiro–Wilk test – to assess whether the data follow a normal distribution (Mooi & Sarstedt, 2011). In addition to carrying out these tests, researchers should also visually inspect the theoretical and empirical probability distributions by means of q–q plots (Chambers, Cleveland, Kleiner, & Tukey, 1983).

Executing the parametric test requires researchers to first run the standard PLS path modeling algorithm for each group, followed by the bootstrapping procedure (e.g., Hair et al., 2011; Henseler et al., 2009) to obtain the standard errors of the group-specific parameter estimates (Keil et al., 2000). The choice of test statistic depends on whether the parameter estimates' standard deviations differ significantly across the groups, which can be assessed by means of Levene's test. If the parameter estimates' standard deviations are equal, the test statistic is computed as follows (Keil et al., 2000; the equation provided by these authors has a flaw that we corrected):

$$t = \frac{\tilde{\theta}^{(1)} - \tilde{\theta}^{(2)}}{\sqrt{((n^{(1)}-1)^2/(n^{(1)}+n^{(2)}-2)) \cdot se^2_{\theta^{(1)}} + ((n^{(2)}-1)^2/(n^{(1)}+n^{(2)}-2)) \cdot se^2_{\theta^{(2)}}} \cdot \sqrt{(1/n^{(1)}) + (1/n^{(2)})}} \tag{1}$$

Here, $\tilde{\theta}^{(1)}$ ($\tilde{\theta}^{(2)}$) denote the original parameter estimate for a path relationship in group one (two), $n^{(1)}$ ($n^{(2)}$) the number of observations in group one (two), and $se_{\theta^{(1)}}$ ($se_{\theta^{(2)}}$) the path coefficient's standard error in group one (two) obtained from the bootstrapping procedure. Moreover, t represents the empirical t-value that must be larger than the critical value from a t-distribution with $n^{(1)}+n^{(2)}-2$ degrees of freedom.[1] In cases where Levene's test indicates that the standard errors are unequal, the test statistic takes the following form (Chin, 2000):

$$t = \frac{\tilde{\theta}^{(1)} - \tilde{\theta}^{(2)}}{\sqrt{((n^{(1)}-1)/n^{(1)})se^2_{\theta^{(1)}} + ((n^{(2)}-1)/n^{(2)})se^2_{\theta^{(2)}}}} \tag{2}$$

This test statistic is asymptotically t-distributed and the degrees of freedom (df) are determined by means of the Welch–Satterthwaite equation. The equation below was derived by Nitzl (2010) for use in combination with bootstrapping (note that the first draft by Chin (2000) is not entirely correct):

$$df = \left\| \frac{\left((n^{(1)}-1)/n^{(1)} \cdot se^2_{\theta^{(1)}} + (n^{(2)}-1)/n^{(2)} \cdot se^2_{\theta^{(2)}} \right)^2}{(n^{(1)}-1)/n^{(1)^2}) \cdot se^4_{\theta^{(1)}} + (n^{(2)}-1)/n^{(2)^2} \cdot se^4_{\theta^{(2)}}} - 2 \right\| \tag{3}$$

Permutation-Based Approach

The permutation-based approach was developed by Chin (2003b) and subsequently further described by Chin and Dibbern (2010), as well as Dibbern and Chin (2005). Analogous to Edgington and Onghena (2007), the permutation-based test procedure builds on the observations' random assignment to groups. The procedure is as follows:

1. Run the PLS path modeling algorithm separately for each group.
2. Randomly permute the data; that is, the observations are randomly exchanged between the two groups. More precisely, $n^{(1)}$ observations are drawn without replacement and assigned to the first group; all remaining observations are assigned to the second group. Thus, in each permutation run u ($u \in \{1, \ldots, U\}$), the group-specific sample size remains constant (i.e., $n_u^{(1)} = n^{(1)}$ and $n_u^{(2)} = n^{(2)}, \forall u$). In accordance with commonly suggested rules of thumb for bootstrapping sample sizes (Hair et al., 2012), the minimum number of permutation runs should be 5,000.
3. Run the PLS path modeling algorithm for each group per permutation run u to obtain the group-specific parameter estimates $\tilde{\theta}_u^{(1)}$ and $\tilde{\theta}_u^{(2)}$.
4. Compute the differences in the permutation run-specific parameter estimates $d_u = \tilde{\theta}_u^{(1)} - \tilde{\theta}_u^{(2)}$.
5. Test the null hypothesis that the population parameters are equal across the two groups ($H_0 : \theta^{(1)} = \theta^{(2)}$).

By not relying on distributional assumptions, the permutation-based approach overcomes a key disadvantage of the parametric approach and, thus, fits the PLS path modeling method's characteristics. However, the permutation-based approach requires group-specific sample sizes to be fairly similar (Chin & Dibbern, 2010), which is a central limitation.

Henseler's PLS Multigroup Analysis

From a procedural perspective, the approach proposed by Henseler (2007) closely resembles the parametric approach. Initially, the subsamples are exposed to separate bootstrap analyses, and the bootstrap outcomes serve as a basis for testing the potential group differences. However, Henseler's (2007) approach differs in the way the bootstrap estimates are used to assess the robustness of the group-specific parameter estimates. Instead of relying on distributional assumptions, the new approach evaluates the bootstrap outcomes' observed distribution. Given two subsamples with different parameter estimates $\tilde{\theta}^{(1)}$ and $\tilde{\theta}^{(2)}$, groups can be indexed – without any loss of generality – so that $\tilde{\theta}^{(1)} > \tilde{\theta}^{(2)}$. In order to assess the significance of a group effect, the conditional probability $p(\theta^{(1)} \leq \theta^{(2)} | \tilde{\theta}^{(1)}, \tilde{\theta}^{(2)}, CDF(\theta^{(1)}), CDF(\theta^{(2)}))$ has to be determined on the basis of the group-specific parameter estimates $\tilde{\theta}^{(g)} (g \in \{1, 2\})$ and the empirical cumulative distribution functions (CDFs).

In an initial step, the centered bootstrap estimates ($\tilde{\theta}_i^{(g)\bar{*}}$) have to be computed as follows:

$$\tilde{\theta}_i^{(g)\bar{*}} = \tilde{\theta}_i^{(g)*} - \frac{1}{B}\sum_{i=1}^{B} \tilde{\theta}_i^{(g)*} + \tilde{\theta}^{(g)} \tag{4}$$

where $\tilde{\theta}_i^{(g)*}$ represents the bootstrap estimate in group g ($g \in \{1, 2\}$) and bootstrap sample i ($i \in \{1, \ldots, B\}$). By using the Heaviside step function $H(x^*)$, as defined by

$$H(x^*) = \frac{1 + \operatorname{sgn}(x^*)}{2} \tag{5}$$

and the bootstrap estimates as discrete manifestations of the CDFs, the conditional probability is computed as follows:

$$p(\theta^{(1)} \leq \theta^{(2)} | \tilde{\theta}^{(1)}, \tilde{\theta}^{(2)}, CDF(\theta^{(1)}), CDF(\theta^{(2)})) = \frac{1}{B^2}\sum_{i=1}^{B}\sum_{j=1}^{B} H\left(\tilde{\theta}_j^{(2)\bar{*}} - \tilde{\theta}_i^{(1)\bar{*}}\right) \tag{6}$$

The idea behind Henseler's (2007) approach is simple. Each centered bootstrap estimate of the second group is compared with each centered bootstrap of the first group across all the bootstrap samples. The number of positive differences divided by the total number of comparisons (i.e., B^2) indicates the probability that the second group's population parameter will be greater than that of the first group.

Henseler's (2007) approach does not build on any distributional assumptions and is simple to apply by using the bootstrap outputs generated by established PLS path modeling software packages such as SmartPLS (Ringle, Wende, & Will, 2005) and PLS-graph (Chin, 2003a). Researchers can easily make the final calculations with available spreadsheet software applications. However, Henseler's (2007) approach only allows testing the one-sided hypotheses. As the bootstrap-based distribution is not necessarily symmetric, it cannot be used to test two-sided hypotheses.

Nonparametric Confidence Set Approach

As an answer to prior methods' deficiencies, we propose the confidence set approach, which builds conceptually on Keil et al.'s (2000) parametric test. Keil et al.'s (2000) approach is a modified version of the two independent samples *t*-test, which accounts for the fact that the standard deviation is obtained through bootstrapping. As such, the test indirectly compares two bootstrap confidence intervals, assuming that the data are normally distributed.

In accordance with this test, researchers can directly compare the group-specific bootstrap confidence intervals, regardless of whether the data are normally distributed or not. The procedure is as follows:

1. Run the PLS path modeling algorithm separately for each group.
2. Construct the bias-corrected α%-bootstrap confidence intervals (preferably 95% in order to avoid Type-II error inflation) for groups one and two, $(\tilde{\theta}^{(1)}_{low}, \tilde{\theta}^{(1)}_{up})$, and $(\tilde{\theta}^{(2)}_{low}, \tilde{\theta}^{(2)}_{up})$.
3. If the parameter estimate for a path relationship of group one $\tilde{\theta}^{(1)}$ sided falls within the corresponding confidence interval of group two $(\tilde{\theta}^{(2)}_{low}, \tilde{\theta}^{(2)}_{up})$, *or* if the parameter estimate of group two $\tilde{\theta}^{(2)}$ falls within the corresponding confidence interval of group one $(\tilde{\theta}^{(1)}_{low}, \tilde{\theta}^{(1)}_{up})$, it can be assumed that there are *no* significant differences between the group-specific path coefficients with regard to a significance level α. Conversely, if there is no overlap, one can assume that group-specific path coefficients are significantly different.

An important element of the confidence set approach is the bootstrap confidence interval. Several methods for constructing bootstrap confidence intervals have been proposed in the literature (e.g., Davison & Hinkley, 1997; Efron & Tibshirani, 1993). An obvious way to construct a confidence interval for a parameter based on bootstrap estimates is to use a set of B bootstrap samples $x^*_i (i \in \{1, \ldots, B\})$ and calculate the bootstrap-specific parameters $\tilde{\theta}_i^*$. Similar to random subsampling, it is presumed that an interval containing 90% of the $\tilde{\theta}_i^*$ is a 90% confidence interval for θ if the estimates are sorted in ascending sequence. Although this so-called percentile method (Efron, 1981) is appealing due to its easy implementation, prior research has shown that – in the case of small samples (especially regarding asymmetric distributions) – the percentile method does not work well (Chernick, 2008). In addition, this method has a clear tendency to underestimate the upper confidence limit, leading to severe under-coverage (Shi, 1992).

The double bootstrap is an alternative approach which generally provides more accurate bootstrap confidence intervals (i.e., bootstrap the bootstrap; McCullough & Vinod, 1998). Articles on double bootstrap methods appear regularly in the statistical literature (e.g., Davidson & MacKinnon, 2007; McKnight, McKean, & Huitema, 2000), but this technique has not yet found its way into methodological research on PLS path modeling. The double bootstrap's basic principle is to take resamples from each bootstrap resample; that is, for each element of $x_i^* = (x_1^*, x_2^*, \ldots, x_B^*)$ (i.e., the first-level bootstrap), further resamples $x_{ij}^{**} = (x_{11}^{**}, \ldots, x_{1M}^{**}, \ldots, x_{B1}^{**}, \ldots, x_{BM}^{**})(j \in \{1, \ldots, M\})$ are drawn from the second level. Both types of bootstrap samples are used to estimate path coefficients on the two levels; that is, $\tilde{\theta}_i^*$ (first level) and $\tilde{\theta}_{ij}^{**}$ (second level). Fig. 2 illustrates the general concept.

However, this approach is computationally demanding. Specifically, the second-level bootstrap generates M bootstrap samples for each first-level bootstrap, leading to an overall number of $B \cdot M + B$ bootstrap samples. For example, following Hair et al.'s (2011) recommendation to use at least 5,000 bootstrap samples would require drawing more than 25×10^6 bootstrap samples.

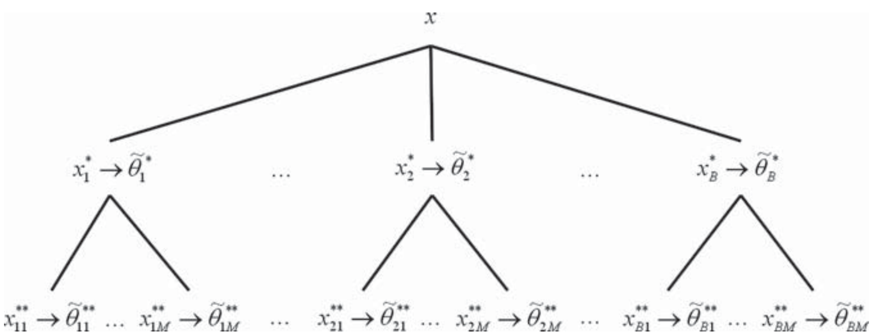

Fig. 2. The Double Bootstrap Method.

Based on this principle, Shi (1992) proposed an accurate and efficient double bootstrap method to estimate bootstrap confidence intervals. In this method, the bootstrap distribution is estimated using

$$Q_i^* = \frac{1}{M}\sum_{j=1}^{M} H(\tilde{\theta}_{ij}^{**} - \tilde{\theta}) \tag{7}$$

where $Q_i^* \in \{0; 1\}$ is random under the empirical distribution $\tilde{F}_i$. Its values are sorted in ascending sequence ($Q_{(1)}^* \leq Q_{(2)}^* \leq \cdots \leq Q_{(B)}^*$) and are used to determine the lower and upper confidence limits:

$$(\tilde{\theta}_{low}^{(g)}, \tilde{\theta}_{up}^{(g)}) = (\tilde{\theta}_{[l]}^{(g)}, \tilde{\theta}_{[u]}^{(g)}) \tag{8}$$

where $[\cdot]$ is a nearest integer function with the arguments given by

$$l = (B+1) \cdot Q_{(\alpha/2)}, \text{ and} \tag{9}$$

$$u = (B+1) \cdot Q_{(1-\alpha/2)} \tag{10}$$

Since estimating the bootstrap confidence interval (Efron & Tibshirani, 1993) entails potential systematic errors, Davison and Hinkley (1997) proposed a bias correction, which should be considered when constructing the interval. The use of bias-corrected confidence intervals was introduced to PLS path modeling in the context of the confirmatory tetrad analysis (Gudergan, Ringle, Wende, & Will, 2008) and bootstrapping-based significance testing (Henseler et al., 2009). The bias correction is as follows:

$$\begin{aligned} bias &= \frac{1}{B}\sum_{i=1}^{B}\tilde{\theta}_i^* - \tilde{\theta} - \left(\frac{1}{BM}\sum_{i=1}^{B}\sum_{j=1}^{M}\tilde{\theta}_{ij}^{**} - \frac{2}{B}\sum_{i=1}^{B}\tilde{\theta}_i^* + \tilde{\theta}\right) \\ &= \frac{3}{B}\sum_{i=1}^{B}\tilde{\theta}_i^* - \frac{1}{BM}\sum_{i=1}^{B}\sum_{j=1}^{M}\tilde{\theta}_{ij}^{**} - 2\tilde{\theta} \end{aligned} \tag{11}$$

This bias correction is used to estimate the confidence interval's lower and upper limits:

$$(\tilde{\theta}_{low}^{(g)}, \tilde{\theta}_{up}^{(g)}) = (\tilde{\theta}_{[l]}^{(g)} - bias, \tilde{\theta}_{[u]}^{(g)} - bias) \tag{12}$$

Although Shi's (1992) method for estimating double bootstrap-based confidence intervals has proven to be accurate in various data constellations, the improvement in accuracy comes at the expense of computational demand.

MULTIGROUP ANALYSIS WITH MORE THAN TWO GROUPS

All previously presented approaches to group comparison in PLS path modeling have in common that they test the difference in the parameters between two groups. As previously mentioned, researchers in international marketing and other cross-cultural research fields frequently encounter situations in which they would like to compare more than two groups. As soon as there are more than two groups, two questions arise: Does a parameter differ between groups? And, if so, between which groups does it differ? Although the second question can be answered by means of pairwise group comparisons, the first question demands more attention. Again, as mentioned, a naive approach would be to conduct all possible pairwise group comparisons, which would lead to the well-known multiple testing problem; that is, the familywise error rate quickly exceeds any prespecified acceptable Type-I error level.

There are, however, several ways of controlling the familywise error rate. A standard remedy is the Bonferroni correction, which aims at retaining the familywise error rate by dividing each comparison's error-rate by the overall number of comparisons. The Bonferroni correction tends to be conservative; that is, it sacrifices statistical power for the sake of a predefined level of Type-I error. An alternative would be to conduct an ANOVA (i.e., an overall *F*-test), comparing the different groups' bootstrap outputs. However, using an ANOVA would mean relying on distributional assumptions (e.g., Hair, Black, Babin, & Anderson, 2010; Mooi & Sarstedt, 2011), which Chin and Dibbern (2010) criticize. An optimal test for the differences between multiple groups in a PLS path modeling framework should (1) maintain the familywise error rate, (2) deliver an acceptable level of statistical power, and (3) not rely on distributional assumptions. Another desirable feature is that such a test should be available in PLS path modeling software packages. In this section, we propose such an omnibus test of group differences (OTG).

Our OTG approach uses bootstrapping, permutation, and random selection's asymptotic properties. The underlying idea of this nonparametric OTG dates back to Pitman (1938) – although the concrete implementation is inspired by Bortz, Lienert, and Boehnke (2003), who proposed a "randomized ANOVA" method. Their method tests the hypothesis that *G* samples are drawn from populations with identical means. Applied to PLS path modeling, the OTG approach consists of the following steps:

1. The first step encompasses groupwise bootstrapping. Per group, a large number of bootstrap samples are drawn and estimated in order to obtain

an empirical distribution of the group-specific model parameters. The number of bootstrap samples should be equal across the groups. The presentation of the bootstrap estimates may be structured as shown in Table 1.

2. The bootstrap results of the previous step facilitate the variance ratio's computation. Analogous to a one-way ANOVA (e.g., Mooi & Sarstedt, 2011), the variance explained by the grouping variable is evaluated relatively to the overall variance:

$$F_R = \frac{s^2_{\text{between}}}{s^2_{\text{within}}} = \frac{G \cdot B \cdot (1/(G-1)) \cdot \sum_{g=1}^{G} (\bar{A}_g - \bar{A})^2}{1/(B-1) \cdot \sum_{g=1}^{G} \sum_{i=1}^{B} (\tilde{\theta}_i^{(g)*} - \bar{A}_g)^2} \tag{13}$$

In this equation, $\tilde{\theta}_i^{(g)*}$ is the parameter estimate from the i^{th} bootstrap sample ($i = 1, \ldots, B$) of group g ($g = 1, \ldots, G$), $\bar{A}_g$ the average over the bootstrap parameter estimates of group g, and $\bar{A}$ the grand mean of all the bootstrap values.

3. This permutation step uses the previously generated bootstrap estimates (e.g., as displayed in Table 1). The elements of the first row – the outcomes of the first bootstrap estimation in each group – can be permuted in $G!$ different ways, whereby each permutation has the same likelihood of occurrence. If this idea is extended to all B rows, this results in $(G!)^B$ permutations. Since the test outcomes are independent of the group index, there are only $(G!)^{B-1}$ different permutations.

For many bootstrap samples, the associated number of permutations becomes extremely high (e.g., in the case of $B = 5{,}000$ bootstraps and $G = 3$ groups, $(3!)^{4{,}999} = 9.508 \times 10^{3{,}889}$ permutations are required). Such extensive computations are not feasible within a reasonable time.

Table 1. Arranging the Groupwise Bootstrap Estimates of a Specific Model Parameter.

Bootstrap Estimation	Groups			
	1	2	…	G
1	$\tilde{\theta}_1^{(1)}$	$\tilde{\theta}_1^{(2)}$	…	$\tilde{\theta}_1^{(G)}$
2	$\tilde{\theta}_2^{(1)}$	$\tilde{\theta}_2^{(2)}$	…	$\tilde{\theta}_2^{(G)}$
…	…	…	…	…
B	$\tilde{\theta}_B^{(1)}$	$\tilde{\theta}_B^{(2)}$	…	$\tilde{\theta}_B^{(G)}$

Consequently, we draw on the random selection (i.e., Monte Carlo) concept. A reasonably high number of permutations (e.g., 5,000) are sufficient to obtain an outcome that approximates the results for $(G!)^{B-1}$ different permutations. Subsequently, the variance ratio F_R can be computed for each randomly selected permutation (e.g., one obtains 5,000 F_R values).

4. The error probability is computed in the final step. As is usual with regard to randomization tests, one has to examine whether the empirical F_R value from Step 1 is among the α% largest values of the empirical F_R value distribution obtained from the previous step. The error probability p can be determined as follows:

$$p = \frac{1}{U}\sum_{u=1}^{U} H(F_R - F_{R_u}) \tag{14}$$

In this equation, $H(\cdot)$ is again the Heaviside step function, U denotes the number of permutations, and F_{R_u} the empirical F_R-value obtained in permutation run u.

The proposed OTG approach offers a possibility to control the familywise error rate. This approach does not rely on distributional assumptions, nor is it as conservative as the Bonferroni correction. The OTG approach can be applied to the regular bootstrap output of standard PLS path modeling software implementations, such as SmartPLS (Ringle et al., 2005).[2]

EMPIRICAL EXAMPLE

Overview

In this section, we use a well-established PLS path model and empirical data to illustrate and compare the different multigroup analysis approaches. The selected PLS path model draws on prior studies by Homburg and Rudolph (1997), as well as by Festge and Schwaiger (2007), and examines the effects of customer satisfaction drivers on customer loyalty in industrial markets.[3] Since the focus of this section is not centered on the substantive model as such, but on an illustration of the multigroup analysis approaches, we only provide a brief description of the data and model set-up.

Measures and Data

The data originate from a survey – by means of standardized mail questionnaires – of a major industrial firm's customers in three countries (Germany, $n = 65$; the United Kingdom, $n = 115$; and France, $n = 170$). All the respondents rated their satisfaction with the different performance features related to the firm's products and services. Our model includes the following three performance features, which have been shown to significantly affect customer satisfaction in industrial markets (e.g., Festge & Schwaiger, 2007; Sarstedt et al., 2009): (1) satisfaction with products, (2) satisfaction with services, and (3) satisfaction with pricing. The corresponding construct measures were adapted from Homburg and Rudolph (1997), as well as from Festge and Schwaiger (2007), using a five-item scale to measure satisfaction with products, and two three-item scales to measure satisfaction with services and pricing. Loyalty was measured with three well-known items (intention to repurchase, word-of-mouth recommendation, and intention to remain a customer in the long run) from prior research (Zeithaml, Berry, & Parasuraman, 1996). We used reflective indicator variables measured on seven-point Likert-type scales.

Results

Principal components analysis supports the scales' unidimensionality. In addition, we computed coefficient β values which range from 0.59 (satisfaction with services; German subsample) to 0.83 (satisfaction with products; German subsample), and are thus above the commonly suggested threshold of 0.50 (Revelle, 1979). Subsequent PLS path model analyses reveal that all measures meet the commonly suggested criteria for measurement model assessment as described, for example, by Chin (1998), Henseler et al. (2009), and Hair et al. (2012). Specifically, the analyses per country show that all indicators exhibit loadings above 0.70, and that the constructs' average variance extracted (AVE) values are above 0.50. Likewise, all constructs achieve high composite reliability values of 0.80 and higher (Table 2).

We used two approaches to assess the constructs' discriminant validity. First, we examined the indicators' cross loading, which revealed that no indicator loads higher on an opposing construct (Hair et al., 2012). Second, we applied the Fornell and Larcker (1981) criterion and tested whether each construct's AVE is greater than its squared correlation with the remaining constructs. Both analyses clearly indicate that the constructs exhibit

Table 2. Country-Specific Results.

		Germany	United Kingdom	France
Latent variables				
Satisfaction with services	Composite reliability	0.829	0.860	0.848
	AVE	0.619	0.672	0.650
Satisfaction with products	Composite reliability	0.910	0.889	0.895
	AVE	0.669	0.616	0.630
Satisfaction with prices	Composite reliability	0.829	0.833	0.895
	AVE	0.619	0.625	0.623
Loyalty	Composite reliability	0.869	0.836	0.846
	AVE	0.689	0.631	0.646
n		65	115	170
Path relationships				
Satisfaction with services → Loyalty		0.040	0.238***	0.195***
Satisfaction with products → Loyalty		0.669***	0.130*	0.289***
Satisfaction with prices → Loyalty		0.163*	0.500***	0.398***
R^2		0.690	0.600	0.609

Notes: *Significance at 0.10, **significance at 0.05, ***significance at 0.01.

discriminant validity. Overall, these results provide clear support for the measures' reliability and convergent validity.

Table 2 shows the results of the structural model evaluation. The bootstrap analyses using 5,000 samples and a number of cases equal to the country-specific sample size (using the individual sign change option) show that all the satisfaction features – with the exception of satisfaction with services in the German subsample – have a significant ($p \leq 0.10$) effect on customer loyalty. A comparison of the country-specific path coefficients reveals several differences in the effects. For example, whereas satisfaction with products has the strongest effect on loyalty in the German subsample, it has a much weaker effect in the UK subsample. Instead, satisfaction with prices exerts the strongest influence on loyalty in the UK subsample. In respect of the French subsample, the effects are somewhat balanced across the three satisfaction types. However, the question emerges whether these numeric differences between country-specific path coefficients are statistically significant.

In a first step, we applied the OTG approach to assess if the path coefficients are equal across the three groups. The analysis reveals that in respect of all three structural model relations, the null hypothesis that the three path coefficients are equal across the three groups can be rejected.

Specifically, the analysis yields F_R values of 579.93 (Services → Loyalty), 3,393.36 (Products → Loyalty), and 1,504.48 (Prices → Loyalty), rendering all differences significant at $p \leq 0.01$. These results suggest that, in respect of all three relationships, at least one path coefficient differs from the remaining two across the three countries.

Table 3 shows the differences in three comparisons' path coefficient estimates (Germany vs. the United Kingdom, Germany vs. France, and the United Kingdom vs. France), and provides the results of multigroup comparisons based on the parametric approach, the permutation test, and Henseler's (2007) approach. The analysis shows that, generally, the multigroup comparison test results correspond very closely. However, differences emerge in respect of Keil et al.'s (2000) parametric tests, which, in most cases, yields higher *t*-values than the permutation test. For example, in the comparison of the German and the UK subsamples, Keil et al.'s (2000) test renders the relationship between satisfaction with services and loyalty significant ($p \leq 0.10$), whereas this result does not occur in the permutation test. Consequently, the parametric approach can generally be considered more liberal in terms of rendering a certain difference significant. Conversely, Henseler's (2007) approach appears to be rather conservative in this respect. Although the approach indicates several significant differences, one has to bear in mind that it only allows testing a one-sided hypothesis. Comparing its results with, for example, the critical *t*-values of a one-sided

Table 3. Multigroup Comparison Test Results.

Relationship	Comparison	$\lvert\text{diff}\rvert$	$t_{\text{Parametric}}$	$t_{\text{Permutation}}$	p_{Henseler}
Services → Loyalty	Germany vs. United Kingdom	0.198	1.930*	1.632	0.095
	Germany vs. France	0.155	1.530	1.351	0.130
	United Kingdom vs. France	0.043	0.410	0.441	0.363
Products → Loyalty	Germany vs. United Kingdom	0.539	4.285***	3.285***	0.005
	Germany vs. France	0.270	2.662***	2.614***	0.013
	United Kingdom vs. France	0.159	1.503	1.367	0.107
Prices → Loyalty	Germany vs. United Kingdom	0.338	2.156**	2.052**	0.021
	Germany vs. France	0.235	1.967**	1.802*	0.063
	United Kingdom vs. France	0.102	0.930	0.959	0.193

Notes: *Significant at 0.10, **significant at 0.05, ***significant at 0.01.
Results for Henseler (2007) eligible for a one-sided test.

parametric test (e.g., 1.28 for $\alpha = 0.10$), clearly shows that Henseler's (2007) approach reveals fewer significant effects.

Table 4 shows the bias-corrected 95% confidence intervals according to Shi's (1992) approach, as well as the corresponding multigroup analysis results. Again, if the parameter estimate for a path relationship of one group (Table 2) does not fall within the corresponding confidence interval of another group (Table 4) and vice versa, there exists no overlap and we can assume that the group-specific path coefficients are significantly different with regard to a significance level α.

Comparing the confidence set approach's results with those of prior tests shows that the former is more conservative than Keil et al.'s (2000) test. Whereas the parametric approach indicates a significant difference ($p \leq 0.05$) between the German and French subsamples in terms of the satisfaction with prices and loyalty relationship, this is not the case with the confidence set approach. Overall, in terms of significant differences, the approach closely resembles the permutation test's results.

Table 4. Bias-corrected 95% Confidence Intervals (Shi 1992) and Multigroup Comparison Results.

Relationship	Confidence Intervals			Comparison	Significance
	Germany	United Kingdom	France		
Services → Loyalty	[−0.206,0.250]	[0.035,0.380]	[0.065,0.325]	Germany vs. United Kingdom	Nsig.
				Germany vs. France	Nsig.
				United Kingdom vs. France	Nsig.
Products → Loyalty	[0.329,0.991]	[−0.021,0.275]	[0.115,0.469]	Germany vs. United Kingdom	Sig.
				Germany vs. France	Sig.
				United Kingdom vs. France	Nsig.
Prices → Loyalty	[−0.158,0.447]	[0.303,0.658]	[0.239,0.551]	Germany vs. United Kingdom	Sig.
				Germany vs. France	Nsig.
				United Kingdom vs. France	Nsig.

Notes: Sig. denotes a significant difference at 0.05; Nsig. denotes a nonsignificant difference at 0.05.

SUMMARY, CONCLUSIONS, AND FUTURE RESEARCH

PLS path modeling is a key multivariate analysis method for empirical research in international marketing (e.g., Henseler et al., 2009), and multigroup analyses are of primary interest in this field (e.g., Hoffmann, Mai, & Smirnova, 2011). This research contributes to the literature on PLS path modeling in several ways: First, we present and compare the procedures available for multigroup analysis in PLS path modeling. Second, we introduce the novel nonparametric confidence set approach based on the comparison of parameter estimates and bootstrap confidence intervals. Third, we address the issue of simultaneously comparing more than two groups by providing a permutation-based analysis of variance approach that maintains the familywise error rate, does not rely on distributional assumptions, and exhibits an acceptable level of statistical power.

The results of the empirical example suggest that Keil et al.'s (2000) parametric approach is the most liberal of the procedures as it, compared to the permutation test, generally yields higher *t*-values. Furthermore, Keil et al.'s (2000) approach renders more differences significant than the novel confidence set approach does. The confidence set approach, just like Henseler's (2007) procedure, appears to be very conservative, as they indicate fewer significant differences vis-à-vis alternative multigroup comparison tests.

International marketing research often deals with relatively small sample sizes and a relatively large number of groups (i.e., data from different cultures or countries). Our novel confidence set approach specifically provides researchers with certain advantageous functions in these kinds of situations. The confidence set approach is nonparametric, can handle relatively small sizes, and is more conservative than the other approaches and, thus, is less prone to Type-II errors. These aspects are particularly relevant when conducting multigroup analysis in international marketing research.

Overall, our findings suggest that if researchers need to compare more than two groups (e.g., countries or cultures), they should first conduct the OTG in order to test the hypothesis that a model parameter differs across groups. If this hypothesis is supported, or if there are only two groups, researchers should subsequently apply the novel confidence set approach to multigroup analysis with regard to comparing two groups of data.

Obviously, our empirical illustration using satisfaction data can only be a first step toward understanding the different multigroup analysis approaches' adequacy. For example, with regard to the confidence set approach's conservative performance, it is unclear if the corresponding path relationships are truly identical in the population, or if the approach's potential lack of statistical power biases the outcome. The approaches may perform differently, depending on the model set-up and sample at hand. It is therefore necessary to compare the approaches' point estimation accuracy, and their statistical power in systematically changed data constellations by conducting a Monte Carlo experiment. In a related context, Qureshi and Compeau (2009) evaluate the ability of variance and covariance-based approaches to structural equation modeling to detect between-group differences and to accurately estimate the moderating effects' strength. However, the authors only consider Keil et al.'s (2000) parametric approach, rather than comparing the performance of different approaches within a PLS path modeling framework.

Another important issue of, and avenue for future research on, multi-group comparisons is exploring ways to test for measurement invariance (e.g., Steenkamp & Baumgartner, 1998; Vandenberg & Lance, 2000) in a PLS path modeling context. If measurement invariance cannot be established, the differences in path coefficients cannot be fully attributed to true relationships, because respondents from different groups might have systematically interpreted a given measure in conceptually different ways. Although measurement invariance should be added to the well-established criteria reliability, homogeneity, and validity when performing multigroup analysis, prior research on PLS path modeling has largely neglected this issue. Haenlein and Kaplan (2011) proposed an approach to control for gamma change, which occurs when the construct's domain (i.e., its meaning) differs in each group. Specifically, the authors propose a combination of Box's M test and ordinary least squares regressions, which can help assess this bias's magnitude and, hence, support researchers when they have to decide whether parameter estimates can be trusted or not. However, Rigdon et al. (2010, p. 269) provide a different perspective on measurement invariance in PLS path modeling, stating that "an insistence on measurement invariance across groups carries its own assumption that the impact of group membership is limited to the structural parameters of the structural model. In many cases, this assumption is questionable or even implausible, and researchers should consider group membership effects on both structural and measurement parameters." Furthermore, the authors point out that PLS path modeling is a method based on approximation and designed for situations with a less firmly established theoretical base (Wold, 1982). Therefore, researchers should interpret the results from PLS path modeling involving multiple groups with the necessary caution.

NOTES

1. Sarstedt and Wilczynski (2009) describe a complementary approach for paired samples.
2. A code file for R (R-Development-Core-Team, 2011), which performs the approach, can be obtained from the second author upon request.
3. Sarstedt et al. (2009) applied the FIMIX-PLS method (Hahn et al., 2002; Rigdon et al., 2010; Ringle, Wende, & Will, 2010; Ringle, Sarstedt, & Mooi, 2010; Sarstedt et al., 2011; Sarstedt & Ringle, 2010) to the original study by Festge and Schwaiger (2007) to uncover unobserved heterogeneity.

REFERENCES

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic and statistical considerations. *Journal of Personality and Social Psychology*, *51*(6), 1173–1182.

Bortz, J., Lienert, G. A., & Boehnke, K. (2003). *Verteilungsfreie Methoden in der Biostatistik* (3 ed.). Berlin: Springer.

Brettel, M., Engelen, A., Heinemann, F., & Vadhanasindhu, P. (2008). Antecedents of market orientation: A cross-cultural comparison. *Journal of International Marketing*, *16*(2), 84–119.

Chambers, J. M., Cleveland, W. S., Kleiner, B., & Tukey, P. A. (1983). *Graphical methods for data analysis*. Belmont: Wadsworth.

Chernick, M. R. (2008). *Bootstrap methods. A guide for practitioners and researchers* (Wiley Series in Probability and Statistics, 2: Wiley).

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: Erlbaum.

Chin, W. W. (2000). Multi-group analysis with PLS. Retrieved from http://disc-nt.cba.uh.edu/chin/plsfaq/multigroup.htm.

Chin, W. W. (2003a). *PLS graph 3.0*. Houston: Soft Modeling Inc.

Chin, W. W. (2003b). A permutation procedure for multi-group comparison of PLS models. In: M. Vilares, M. Tenenhaus, P. Coelho, V. Esposito Vinzi, A. Morineau (Eds.), *PLS and Related Methods: Proceedings of the International Symposium PLS'03* (pp. 33–43). Lisbon: Decisia.

Chin, W. W., & Dibbern, J. (2010). An introduction to a permutation based procedure for multi-group PLS analysis: Results of tests of differences on simulated data and a cross cultural analysis of the sourcing of information system services between Germany and the USA. In: V. Esposito Vinzi, W. W. Chin, J. Henseler & H. Wang, et al. (Eds.), *Handbook of partial least squares. Concepts, methods and applications* (pp. 171–193). Berlin, Heidelberg: Springer-Verlag.

Davidson, R., & MacKinnon, J. G. (2007). Improving the reliability of bootstrap tests with the fast double bootstrap. *Computational Statistics & Data Analysis*, *51*(7), 3259–3281.

Davison, A. C., & Hinkley, D. V. (1997). *Bootstrap methods and their application*. Cambridge: Cambridge Univiversity Press.

Dibbern, J., & Chin, W. W. (2005). Multi-group comparison: Testing a PLS model on the sourcing of application software services across Germany and the USA using a permutation based algorithm. In: F. Bliemel, A. Eggert, G. Fassott & J. Henseler, et al. (Eds.), *Handbuch PLS-Pfadmodellierung. Methode, Anwendung, Praxisbeispiele* (pp. 135–160). Stuttgart: Schäffer-Poeschel.

Edgington, E., & Onghena, P. (2007). *Randomization tests* (4th ed.). London: Chapman & Hall.

Efron, B. (1981). Nonparametric standard erroros and confidence intervals. *Canadian Journal of Statistics*, *9*(2), 139–172.

Efron, B., & Tibshirani, R. J. (1993). *An introduction to the bootstrap*. New York: Chapman & Hall.

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. Cergy Pontoise Cedex: ESSEC Research Center, Working Paper No. 07019.

Esposito Vinzi, V., Trinchera, L., Squillacciotti, S., & Tenenhaus, M. (2008). REBUS-PLS: A response-based procedure for detecting unit segments in PLS path modelling. *Applied Stochastic Models in Business and Industry*, *24*(5), 439–458.

Festge, F., & Schwaiger, M. (2007). The drivers of customer satisfaction with industrial goods: An international study. *Advances in International Marketing*, *18*, 179–207.

Fornell, C. G., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, *18*(1), 39–50.

Graham, J. L., Mintu, A. T., & Rodgers, W. (1994). Explorations of negotiation behaviors in ten foreign cultures using a model developed in the United States. *Management Science*, *40*(1), 72–95.

Grewal, R., Chakravarty, A., Ding, M., & Liechty, J. (2008). Counting chickens before the eggs hatch: Associating new product development portfolios with shareholder expectations in the pharmaceutical sector. *International Journal of Research in Marketing*, *25*(3), 261–272.

Gudergan, S. P., Ringle, C. M., Wende, S., & Will, A. (2008). Confirmatory tetrad analysis in PLS path modeling. *Journal of Business Research*, *61*(12), 1238–1249.

Haenlein, M., & Kaplan, A. M. (2011). The influence of observed heterogeneity on path coefficient significance: Technology acceptance within the marketing discipline. *Journal of Marketing Theory and Practice*, *19*(2), 153–168.

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.

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis* (7 ed.). Englewood Cliffs: Prentice Hall.

Hair, J. F., Ringle, C. M., & Sarstedt, M. (2011). PLS-SEM: Indeed a silver bullet. *Journal of Marketing Theory and Practice*, *19*(2), 139–151.

Hair, J. F., Sarstedt, M., Ringle, C, M., & Mena, J. A. (2012). An assessment of the use of partial least squares structural equation modeling in marketing research. *Journal of the Academy of Marketing Science*, forthcoming (online available).

Henseler, J. (2007). A new and simple approach to multi-group analysis in partial least squares path modeling. In: H. Martens & T. Næs (Eds.), *Causalities explored by indirect observation: Proceedings of the 5th international symposium on PLS and related methods* (PLS'07) (pp. 104–107). Oslo.

Henseler, J., & Chin, W. W. (2010). A comparison of approaches for the analysis of interaction effects between latent variables using partial least squares path modeling. *Structural Equation Modeling*, *17*(1), 82–109.

Henseler, J., & Fassott, G. (2010). Testing moderating effects in PLS path models: An illustration of available procedures. In: V. Esposito Vinzi, W. W. Chin, J. Henseler & H. Wang (Eds.), *Handbook of partial least squares. Concepts, methods and applications* (pp. 713–735). Berlin, Heidelberg: Springer-Verlag.

Henseler, J., Ringle, C. M., & Sinkovics, R. R. (2009). The use of partial least squares path modeling in international marketing. *Advances in International Marketing*, *20*, 277–320.

Hoffmann, S., Mai, R., & Smirnova, M. (2011). Development and validation of a cross-nationally stable scale of consumer animosity. *Journal of Marketing Theory and Practice*, *19*(2), 235–251.

Homburg, C., & Rudolph, B. (1997). Customer satisfaction in industrial markets: Dimensional and multiple role issues. *Journal of Business Research*, *52*(1), 15–33.

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.

Keil, M., Saarinen, T., Tan, B. C. Y., Tuunainen, V., Wassenaar, A., & Wei, K.-K. (2000). A cross-cultural study on escalation of commitment behavior in software projects. *Management Information Systems Quarterly*, *24*(2), 299–325.

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

McCullough, B. D., & Vinod, H. D. (1998). Implementing the double bootstrap. *Computational Economics*, *12*(1), 79–95.

McKnight, S. D., McKean, J. W., & Huitema, B. E. (2000). A double bootstrap method to analyze linear models with autoregressive error terms. *Psychological Methods*, *5*(1), 87–101.

Mooi, E. A., & Sarstedt, M. (2011). *A concise guide to market research: The process, data, and methods using IBM SPSS statistics*. Berlin: Springer.

Nitzl, C. (2010). Eine anwenderorientierte Einführung in die Partial Least Square (PLS)-Methode. Hamburg: Universität Hamburg: Industrielles Management, Arbeitspapier Nr. 21.

Pitman, E. J. G. (1938). Significance tests which may be applied to samples from any population. *Journal of the Royal Statistical Society Supplement*, *4*(1), 119–130.

Qureshi, I., & Compeau, D. R. (2009). Assessing between-group differences in information systems research: A comparision of covariance- and component-based SEM. *Management Information Systems Quarterly*, *33*(1), 197–214.

R-Development-Core-Team. (2011). R: A language and environment for statistical computing. Vienna.

Revelle, W. (1979). Hierarchical clustering and the internal structure of tests. *Multivariate Behavioral Research*, *14*(1), 57–74.

Rigdon, E. E., Ringle, C. M., & Sarstedt, M. (2010). Structural modeling of heterogeneous data with partial least squares. In: N. K. Malhotra (Ed.), *Review of marketing research* (Vol. 7, pp. 255–296). Armonk: Sharpe.

Ringle, C. M., Sarstedt, M., & Mooi, E. A. (2010). Response-based segmentation using finite mixture partial least squares: Theoretical foundations and an application to American Customer Satisfaction Index data. *Annals of Information Systems*, *8*, 19–49.

Ringle, C. M., Sarstedt, M., & Schlittgen, R. (2010). Finite mixture and genetic algorithm segmentation in partial least squares path modeling: Identification of multiple segments in a complex path model. In: A. Fink, B. Lausen, W. Seidel & A. Ultsch (Eds.), *Advances in data analysis, data handling and business intelligence* (pp. 167–176). Berlin: Springer.

Ringle, C. M., Wende, S., & Will, A. (2005). SmartPLS 2.0 (Beta). Hamburg: SmartPLS. Retrieved from www.smartpls.de.

Ringle, C. M., Wende, S., & Will, A. (2010). Finite mixture partial least squares analysis: Methodology and numerical examples. In: V. Esposito Vinzi, W. W. Chin, J. Henseler & H. Wang (Eds.), *Handbook of partial least squares. Concepts, methods and applications* (pp. 195–218). Berlin, Heidelberg: Springer-Verlag.

Rodríguez, C. M., & Wilson, D. T. (2002). Relationship bonding and trust as a foundation for commitment is U.S.-Mexican strategic alliances: A structural equation modeling approach. *Journal of International Marketing*, *10*(4), 53–76.

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., Becker, J.-M., Ringle, C. M., & Schwaiger, M. (2011). Uncovering and treating unobserved heterogeneity with FIMIX-PLS: Which model selection criterion provides an appropriate number of segments? *Schmalenbach Business Review*, *63*(1), 34–62.

Sarstedt, M., & Ringle, C. M. (2010). Treating unobserved heterogeneity in PLS path modelling: A comparison of FIMIX-PLS with different data analysis strategies. *Journal of Applied Statistics*, *37*(8), 1299–1318.

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 Schwaiger's model to account for unobserved heterogeneity. *Journal of Business Market Management*, *3*(3), 185–206.

Sarstedt, M., & Wilczynski, P. (2009). More for less? A comparison of single-item and multi-item measures. *Die Betriebswirtschaft*, *69*(2), 211–227.

Shi, S. G. (1992). Accurate and efficient double-bootstrap confidence limit method. *Computational Statistics and Data Analaysis*, *13*(1), 21–32.

Steenkamp, J. B. E. M., & Baumgartner, H. (1998). Assessing measurement invariance in cross national consumer research. *Journal of Consumer Research*, *25*(1), 78–107.

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, *3*(1), 4–70.

Wold, H. (1975). Path models with latent variables: The NIPALS approach. In: H. M. Blalock, A. Aganbegian, F. M. Borodkin, R. Boudon & V. Capecchi (Eds.), *Quantitative sociology: International perspectives on mathematical and statistical modeling* (pp. 307–357). New York: Academic Press.

Wold, H. (1982). Soft modeling: The basic design and some extensions. In: K. G. Jöreskog & H. Wold (Eds.), *Systems under indirect observations: Part II* (pp. 1–54). Amsterdam: North-Holland.

Zeithaml, V. A., Berry, L. L., & Parasuraman, A. (1996). The behavioral consequences of service quality. *Journal of Marketing*, *60*(2), 31–46.