CHAPTER 4

Modeling Observed Heterogeneity

LEARNING OUTCOMES

1. Appreciate the importance of considering heterogeneity in PLS-SEM.
2. Explain the difference between observed and unobserved heterogeneity.
3. Comprehend the concept of measurement model invariance and its assessment in PLS-SEM.
4. Understand multigroup analysis and execute it in PLS-SEM.

CHAPTER PREVIEW

Relationships in PLS path models imply that exogenous latent variables explain endogenous latent variables without any systematic influences of other variables. In many instances, however, this assumption does not hold. For example, respondents are likely to be heterogeneous in their perceptions and evaluations of latent variables, yielding significant differences in, for example, path coefficients across two or more groups of respondents (e.g., females vs. males). Recognizing that heterogeneous data structures are often present, researchers are increasingly interested in identifying and understanding such differences. In fact, failure to consider heterogeneity can be a threat to the validity of partial least squares structural equation modelling (PLS-SEM) results (Becker, Rai, Ringle, & Völckner, 2013; Hair, Sarstedt, Ringle, & Mena, 2012).
As a solution, we learn about different concepts that enable researchers to model heterogeneous data. This chapter first provides an overview of observed and unobserved heterogeneity, showing how disregarding heterogeneous data structures can provoke biased results. Next, we discuss measurement invariance, which is a primary concern before comparing groups of data. By establishing measurement invariance, researchers can be confident that group differences in model estimates result from neither the distinctive content and/or meanings of the latent variables across groups nor the measurement scale. To assess measurement invariance in a PLS-SEM context, researchers can use the measurement invariance of the composite models (MICOM) procedure. Next, we introduce different types of multigroup analyses that are used to compare parameters (usually path coefficients) between two or more groups of data. A PLS-SEM multigroup analysis is typically applied when researchers want to explore differences that can be traced back to observable characteristics such as gender or country of origin. In this situation, it is assumed that there is a categorical moderator variable (e.g., gender) that influences the relationships in the PLS path model. Yet, while categorical data allow easy specification of groups, other data can also provide a basis to specify groups. The aim of multigroup analysis, therefore, is to disclose the effect of this categorical moderator variable. The chapter closes with an illustration of measurement model invariance assessment using the MICOM procedure and a multigroup analysis.

OBSERVED AND UNOBSERVED HETEROGENEITY

Applications of PLS-SEM usually analyze the full set of data, implicitly assuming that the data stem from a homogeneous population. This assumption of relatively homogeneous data characteristics is often unrealistic. Individuals (e.g., in their behavior), corporations (e.g., in their structure), or environments (e.g., in their dynamism) are frequently different, and pooling data across observations is likely to produce misleading results. Failure to consider such heterogeneity can be a threat to the validity of PLS-SEM results, since it can lead to incorrect conclusions (Sarstedt & Ringle, 2010). Becker, Rai, Ringle, and Völckner (2013) address this problem in detail and provide examples of situations with significantly different positive
and negative path coefficients, which show a nonsignificant value close to 0 on the aggregate data level. Concluding that no relationship exists between the constructs is invalid and misleading.

The model shown in Exhibit 4.1, in which customer satisfaction with a product ($Y_3$) depends on the two perceptual dimensions—satisfaction with the quality ($Y_1$) and satisfaction with the price ($Y_2$)—illustrates the problems stemming from failure to treat heterogeneity in the context of PLS-SEM. Suppose there are two segments of similar sample sizes. Group 1 comprises male customers, whereas Group 2 consists of female customers. Both groups differ regarding their price consciousness as indicated by the different segment-specific path coefficients. More precisely, the effect of perceived quality ($Y_1$) on customer satisfaction ($Y_3$) is much stronger among males ($p_1^{(1)} = 0.50$; the superscript in parentheses indicates the group) than among females ($p_1^{(2)} = 0.10$). In contrast, perceived price ($Y_2$) has a somewhat stronger influence on customer satisfaction ($Y_3$) among females ($p_2^{(2)} = 0.35$) than among males ($p_2^{(1)} = 0.25$). In this example, heterogeneity is the results of one group of customers (females) being more price sensitive but less quality sensitive, whereas it is the opposite for the other group (males). From a technical perspective, there is a categorical moderator variable gender that splits the data set into two customer groups and thus requires estimating two separate models, as indicated in Exhibit 4.1. Importantly, if we fail to recognize the heterogeneity between the groups and analyze the model using the full set of data, the path coefficient estimates offer an incomplete picture of the model relationships. That is, both estimates would equal approximately 0.30 when using the full set of data, thus leading the researcher to conclude that price and quality are equally important for males and females, although they are not. Similarly, group-related differences in model estimates can cancel each other out, yielding nonsignificant effects when analyzing the data on the aggregate level (for an example, see Sarstedt, Schwaiger, & Ringle, 2009). Consequently, it is important to identify, assess, and, if present, treat heterogeneity in the data.

Heterogeneity in data can be observed or unobserved. When differences between two or more groups of data relate to observable characteristics, such as gender, age, or country of origin, this is observed heterogeneity. Researchers can use these observable characteristics to partition the data into separate groups of observations and carry out group-specific PLS-SEM analyses, as illustrated in Exhibit 4.1.
with regard to customers’ gender. On the contrary, **unobserved heterogeneity** implies that differences between two or more groups of data do not emerge a priori from a specific observable characteristic or combinations of several characteristics but becomes apparent in differences in structural path coefficients.

In an attempt to account for unobserved heterogeneity with regard to endogenous but also exogenous variables, researchers have routinely used **cluster analysis** techniques, such as **k-means** (Hair, Black, Babin, & Anderson, 2010; Sarstedt & Mooi, 2014) on the indicator data, or latent variable scores derived from a preceding analysis of the entire data set. The partition that this analysis produces is then used as input for group-specific PLS-SEM estimations. While easy to apply, such an approach is conceptually flawed because traditional clustering techniques ignore the path model relationships that researchers specified prior to the analysis. But it is exactly these relationships that are likely responsible for some of the group differences. Therefore, it is not surprising that prior research has shown that traditional clustering approaches perform very poorly in identifying group differences in PLS-SEM (e.g., Sarstedt & Ringle, 2010).

Recognizing the limitations of sequential approaches, methodological research in PLS-SEM has proposed several specific methods.
to identify and treat unobserved heterogeneity, commonly referred to as **latent class techniques**. These techniques have a long tradition in covariance-based SEM (CB-SEM) research (e.g., Jedidi, Jagpal, & DeSarbo, 1997; Masyn, 2013; Muthén, 1989) and have proven very useful in identifying unobserved heterogeneity and partitioning the data accordingly. The resulting partition can then be analyzed for significant differences using multigroup analysis approaches. Alternatively, latent class techniques may ascertain that unobserved heterogeneity does not influence the results, supporting an analysis of a single model based on the aggregate-level data.

The different techniques for identifying unobserved heterogeneity in PLS-SEM include genetic algorithms (Ringle, Sarstedt, & Schlittgen, 2014; Ringle, Sarstedt, Schlittgen, & Taylor, 2013), weighted least squares (Schlittgen, Ringle, Sarstedt, & Becker, 2015), and several other approaches (e.g., Becker, Rai, Ringle, & Völkner, 2013; Esposito Vinzi, Trinchera, Squillacciotti, & Tenenhaus, 2008). See Hair, Sarstedt, Matthews, and Ringle (2016) for a review. In Chapter 5, we introduce two latent class procedures, **finite mixture partial least squares** (FIMIX-PLS) and **prediction-oriented segmentation in PLS-SEM** (PLS-POS), whose joint use allows for reliably identifying and treating unobserved heterogeneity in PLS path models.

### TESTING MEASUREMENT MODEL INVARIANCE

A primary concern before comparing group-specific parameter estimates for significant differences using a multigroup analysis is ensuring **measurement invariance**, also referred to as **measurement equivalence**. By establishing measurement invariance, researchers can be confident that group differences in model estimates do not result from the distinctive content and/or meanings of the latent variables across groups. Variations in the structural relationships between latent variables could stem from different meanings the groups’ respondents attribute to the phenomena being measured, rather than the true differences in the structural relationships. Reasons for such differences may stem from, for example, (a) respondents embracing different cultural values who interpret a given measure in a conceptually different manner; (b) gender, ethnic, or other individual differences that entail responding to instruments in systematically different ways; and (c) respondents who use the
available options on a scale differently (e.g., tendency to choose or not to choose the extremes). Hult et al. (2008, p. 1028) describe these concerns and conclude that “failure to establish data equivalence is a potential source of measurement error” (i.e., discrepancies between what is intended to be measured and what is actually measured). When measurement invariance is not present, it can reduce the power of statistical tests, influence the precision of estimators, and provide misleading results. In short, when measurement invariance is not demonstrated, any conclusions about model relationships are questionable. Hence, multigroup comparisons require establishing measurement invariance to ensure the validity of outcomes and conclusions.

Researchers have suggested a variety of methods to assess measurement invariance for CB-SEM. Multigroup confirmatory factor analysis based on the guidelines of Steenkamp and Baumgartner (1998) and Vandenberg and Lance (2000) is by far the most common approach to invariance assessment. However, the well-established measurement invariance techniques used to assess CB-SEM’s common factor models and related extensions to formative measurement models (Diamantopoulos & Papadopoulos, 2010) cannot be readily transferred to PLS-SEM’s composite models. For this reason, Henseler, Ringle, and Sarstedt (2016) developed the **measurement invariance of composite models (MICOM)** procedure. The MICOM procedure builds on the scores of the latent variables. In PLS-SEM, these latent variables are represented as composites, that is, linear combinations of indicators and the indicator weights as estimated by the PLS-SEM algorithm. Therefore, in the description of the MICOM procedure, we talk about composites when referring to the entities (scores) the PLS-SEM algorithm uses to refer to the latent variables as specified by the researcher.

The MICOM procedure involves three steps: (1) **configural invariance**, (2) **compositional invariance**, and (3) **equality of composite mean values and variances**. The three steps are hierarchically interrelated, as displayed in Exhibit 4.2. This means that configural invariance is a precondition for compositional invariance, which is again a precondition for meaningfully assessing the equality of composite mean values and variances.

If configural invariance (Step 1) and compositional invariance (Step 2) are established, **partial measurement invariance** is confirmed. When partial measurement invariance is confirmed for all latent
Exhibit 4.2 The MICOM Procedure

Configural invariance assessment

- Configural invariance not established
  - Analyze groups separately. No multigroup analysis feasible.
- Configural invariance established
  - Compositional invariance assessment
    - Compositional invariance established
      - Assess equality of composite means and variances
        - Composite means and variances are equal
          - Full measurement invariance established. Multigroup analysis and pooling of data feasible.
        - Composite means and variances are unequal
          - Partial measurement invariance established. Multigroup analysis feasible.
variables in the PLS path model, researchers can compare the path coefficients by means of a multigroup analysis. If partial measurement invariance is established and, additionally, the composites have equal mean values and variances across the groups, full measurement invariance is confirmed. From a measurement model perspective, researchers can then pool the data of the different groups, if they come from separate groups initially, and benefit from the increase in statistical power. However, full measurement invariance does not imply that there are no differences in the structural models (i.e., the path coefficients). The latter need to be tested by means of a multigroup analysis. Only if the multigroup analysis indicates that the structural models are also invariant (i.e., there are no significant differences in the path coefficients) can researchers pool the data and exclusively focus on the aggregate-level analysis. In the following sections, we discuss each step in greater detail (for an application of MICOM on data from five countries, see Schlägel & Sarstedt, 2016).

**Step 1: Configural Invariance**

Step 1 addresses the establishment of configural invariance to ensure that each latent variable in the PLS path model has been specified equally for all the groups. Configural variance exists when constructs are equally parameterized and estimated across groups. An initial qualitative assessment of the latent variables’ specification across all the groups must ensure that the following three requirements have been met:

1. **Identical indicators per measurement model.** Each measurement model must employ the same indicators and scale across the groups. Checking whether exactly the same indicators apply to all the groups seems rather simple. However, when conducting surveys using different languages, the application of good empirical research practices (e.g., translation, back-translation) is of the utmost importance to establish the indicators’ equivalence. In this context, an assessment of face and/or expert validity can help verify whether the researcher(s) used the same set of indicators across the groups.

2. **Identical data treatment.** The indicators’ data treatment must be identical across all the groups, which includes the coding (e.g., dummy coding), reverse coding, and other
forms of recoding as well as the data handling (e.g., standardization or missing value treatment). Outliers should be detected and treated similarly.

3. **Identical algorithm settings or optimization criteria.** Variance-based model estimation methods such as PLS consist of many variants with different target functions and algorithm settings (e.g., choice of initial outer weights and the inner model weighting scheme; Hair, Ringle, & Sarstedt, 2011; Henseler, Ringle, & Sinkovics, 2009). Researchers must ensure that differences in the group-specific model estimations do not result from dissimilar algorithm settings.

Configural invariance is a necessary but not sufficient condition for drawing valid conclusions from multigroup analyses. Researchers also must ensure that differences in path coefficients do not result from differences in the way a latent variable formed across the groups. The next step, compositional invariance, focuses on this aspect.

**Step 2: Compositional Invariance**

Compositional invariance exists when the composite scores are the same across the groups, despite possible differences in the group-specific weights used to compute the scores. Step 2 of the MICOM procedure applies a statistical test to assess whether the composite scores differ significantly across the groups. For this purpose, the MICOM procedure examines \( c \), which is the correlation between the composite scores \( Y^{(1)} \) and \( Y^{(2)} \):

\[
c = \text{cor} \left( Y^{(1)}, Y^{(2)} \right).
\]

Compositional invariance requires that \( c \) equals 1. Technically, the procedure tests the null hypothesis that \( c \) is 1. In order to establish compositional invariance, we must not reject this null hypothesis. That is, if the test yields a \( p \) value larger than 0.05 (in case we assume a significance level of 5%), we can assume compositional invariance. On the contrary, if we reject the null hypothesis, we cannot establish compositional invariance for the specific construct under consideration. Note that this test does not work for single-item constructs since their (single) outer relationship always is 1. Hence, the latent
variable scores \( Y^{(1)} \) and \( Y^{(2)} \) of a single item construct are identical entailing a \( c \) value of 1.

For hypothesis testing, the MICOM procedure draws on the concept of **permutation testing** (Fisher, 1935). Similar to bootstrapping, permutation tests generate a reference distribution from the actual data. However, instead of sampling observations from the original data set with replacement—as is the case in bootstrapping—permutation tests randomly exchange observations between the groups multiple times. Permutation testing has routinely been used in a multitude of contexts as it provides an efficient approach to nonparametric testing, also when the sample size is small (e.g., Ernst, 2004; Good, 2000).

The compositional invariance assessment in MICOM follows a five-step approach, which Exhibit 4.3 illustrates:

1. Group-specific estimation of the PLS path models to obtain composite scores of group 1 with \( n^{(1)} \) observations and group 2 with \( n^{(2)} \) observations. Note that in the example of two groups the sum of \( n^{(1)} \) and \( n^{(2)} \) equals \( n \), which is the total number of observations in the full set of data.

2. Computation of the correlations \( c \) between the composite scores of groups 1 and 2.

3. Random permutation of the data, meaning that the observations are randomly assigned to the groups. In formal terms, \( n^{(1)} \) observations are randomly drawn without replacement from the aggregate data set and assigned to group 1. The remaining \( n^{(2)} \) observations are assigned to group 2. Researchers should use a minimum of 1,000 permutations.

4. For each permutation run \( u \) computations of the correlation \( c_u \) between the composite scores of group 1 and group 2.

5. For hypothesis testing, the procedure sorts the permutations’ correlation results \( c_u \) in descending order. The cutoff value is the 95% point (or, in our example, the 950th of 1,000 permutations in the sorted list) and the corresponding correlation value. If \( c \) is smaller than that value, it falls out of the 95% permutation-based confidence interval and is significantly different from 1 (at the \( p < 0.05 \) level), which entails a rejection of the null hypothesis that \( c \) equals 1.
Exhibit 4.3 Compositional Invariance Testing in the MICOM Procedure

Permutation sample 1 (nine observations)

Group 1 ($n^{(1)} = 4$)

Latent variable scores (group 1)

Randomly permute observations

Latent variable scores (sample 1, group 1)

Latent variable scores (sample 1, group 2)

Group 2 ($n^{(2)} = 4$)

Latent variable scores (group 2)

Latent variable scores (sample 5,000, group 1)

Latent variable scores (sample 5,000, group 2)
Consequently, the composite scores are significantly different, and compositional invariance has not been established. In the opposite situation, if $c$ does not fall into the 5% extreme tail, but into the 95% confidence interval, the composite scores are not significantly different, which substantiates compositional invariance. Based on the permutation results, it is also possible to return the $p$ value of $c$. In this analysis, the $p$ value represents the percentage of the permutations’ correlation $c_u$ results that have a lower value than $c$. A $p$ value above 0.05 indicates that $c$ is not significantly different from 1, which means that compositional invariance has been established.

When configural invariance exists and compositional invariance is established for all latent variables in the PLS path model, there is partial measurement invariance. In this case, researchers can compare the path coefficients by means of a multigroup analysis. If partial measurement invariance cannot be established for a latent variable, then group-specific comparisons using a multigroup analysis on any relationships involving this latent variable are not feasible. In this case, researchers may refrain from running a multigroup analysis altogether and analyze each group separately without attempting to compare the results for the two groups. Alternatively, one can eliminate the latent variables that did not achieve compositional invariance, provided theory supports this step. Importantly, when making any changes to the model, MICOM needs to be rerun.

**Step 3: Equality of Composite Mean Values and Variances**

If the results of Step 2 support measurement invariance, the assessment should continue with the equality assessment of the composites’ mean values and variances.

This final step of the MICOM procedure first requires estimating the PLS path model using the pooled (i.e., aggregate) data. We then examine whether the mean values and variances between the composite scores of the first group and the composite scores of the second group differ regarding their means and variances. Note that, different from Step 2, we do not use the composite scores computed by means of separate, group-specific PLS-SEM estimations but, instead, use the
entire data set and assign the scores to each group ex post. For the analysis of the mean values’ equivalence, the null hypothesis is

$$H_0 : \bar{Y}^{(1)}_{\text{pooled}} - \bar{Y}^{(2)}_{\text{pooled}} = 0.$$

Here, the index “pooled” indicates that the composite scores of a certain group stem from the analysis of the pooled data.

Analyzing the equivalence of the variances requires determining the logarithm of the variance ratio of the composite scores of both groups. If the logarithm of this ratio is statistically not different from 0, we conclude that the variances are equal across groups. The corresponding null hypothesis is

$$H_0 : \log \left( \frac{\text{var} \left( Y^{(1)}_{\text{pooled}} \right)}{\text{var} \left( Y^{(2)}_{\text{pooled}} \right)} \right) - \log \left( \frac{\text{var} \left( Y^{(1)}_{\text{pooled}} \right)}{\text{var} \left( Y^{(2)}_{\text{pooled}} \right)} \right) = 0.$$

The testing of these two hypotheses follows a similar approach as in Step 2. MICOM permutes (i.e., rearranges) the observations’ group membership many times and generates the empirical distribution of the differences in mean values and logarithms of variances. Full measurement invariance is established when there are no significant differences in mean values and (logarithms of) variances across the groups. This is the case if the permutation-based confidence intervals of the differences in mean values and (logarithms of) variances include the original differences in mean values and variances as obtained by the original model estimation (i.e., without permutation). In contrast, if one of these differences is significant, we must acknowledge that full measurement invariance cannot be established.

In summary, comparing group-specific model relationships for significant differences using a multigroup analysis requires establishing configural (Step 1) and compositional (Step 2) invariance. If these two steps support measurement invariance, partial measurement invariance is established. As the focus is usually on the structural model relationships, we can then run a multigroup analysis on the path coefficients using one of the approaches described in the following sections. However, an exception in this regard is the comparison of interaction terms, created in the course of a moderator analysis (see Hair et al., 2016). The measurement model of an interaction term
represents an auxiliary measurement that incorporates the interrelationships between the moderator and the exogenous construct in the path model. This characteristic, however, renders any measurement model assessment of the interaction term and related comparisons across different groups in terms of invariance assessment meaningless. Importantly, if partial measurement invariance cannot be established for a latent variable, group-specific comparisons using a multigroup analysis on any relationships involving this latent variable are not feasible.

To reiterate, researchers may analyze the groups separately, if partial measurement invariance is not apparent, without drawing any conclusions about group-related effects or change the model by eliminating the latent variables that did not achieve compositional invariance. Finally, if partial measurement invariance is confirmed and the composites have equal mean values and variances across the groups, full measurement invariance is confirmed, which additionally supports the pooled data analysis, when considered applicable.

**Multigroup Analysis**

Path coefficients generated drawing on different samples are almost always numerically different, but the question is whether the differences are statistically significant. Multigroup analysis helps answer this question. Technically, a multigroup analysis tests the null hypotheses \( H_0 \) that the path coefficients between two groups (e.g., \( p^{(1)} \) in group 1 and \( p^{(2)} \) in group 2) are not significantly different (e.g., \( p^{(1)} = p^{(2)} \)), which amounts to the same as saying that the absolute difference between the path coefficients is 0 (i.e., \( H_0 : |p^{(1)} - p^{(2)}| = 0 \)). The corresponding alternative hypothesis \( H_1 \) is that the path coefficients are different (i.e., \( H_0 : p^{(1)} \neq p^{(2)} \) or, put differently, \( H_0 : |p^{(1)} - p^{(2)}| > 0 \)).

Research has proposed several approaches to multigroup analysis that are illustrated in Exhibit 4.4 (Sarstedt, Henseler, & Ringle, 2011). When comparing two groups of data, researchers need to distinguish between the parametric approach and several nonparametric approaches.

**The Parametric Test**

The parametric test (Keil et al., 2000) was the first approach applied in PLS-SEM studies and has been widely adopted because of its ease of implementation. This approach is a modified version of a standard
two independent samples \( t \) test, which relies on standard errors derived from bootstrapping. As with the standard \( t \) test, the parametric approach has two versions (Sarstedt & Mooi, 2014), depending on whether population variances can be assumed to be equal (homoscedastic) or unequal (heteroscedastic). The latter can be tested by means of the Levene’s test as described in, for example, Sarstedt and Mooi (2014). If the variances are equal, the test statistic (i.e., the empirical \( t \) value) is computed as follows:

\[
t = \frac{\hat{p}^{(1)} - \hat{p}^{(2)}}{\sqrt{\left(\frac{n^{(1)} - 1}{n^{(1)} + n^{(2)} - 2}\right) \cdot \text{se}(\hat{p}^{(1)})^2 + \left(\frac{n^{(2)} - 1}{n^{(1)} + n^{(2)} - 2}\right) \cdot \text{se}(\hat{p}^{(2)})^2 \times \frac{1}{n^{(1)}} + \frac{1}{n^{(2)}}}}
\]

In this formula, \( \hat{p}^{(1)}(\hat{p}^{(2)}) \) describes the path coefficient to be compared in group 1 (group 2), whereas \( n^{(1)}(n^{(2)}) \) stands for the number of observations in group 1 (group 2). Finally, \( \text{se}(\hat{p}^{(1)})(\text{se}(\hat{p}^{(2)})) \) describes the standard errors of the parameter estimates of group 1 (group 2), which can be obtained via bootstrapping. To reject the null hypothesis of equal path coefficients, the empirical \( t \) value must be
larger than the critical value from a $t$ distribution with $n_1^{(1)} + n_2^{(2)} - 2$ degrees of freedom.

On the contrary, if the variances are unequal, researchers have to use a modified version of the Welch-Satterthwaite $t$ test (Satterthwaite, 1946; Welch, 1947), whose test statistic takes the following form:

$$ t = \frac{p_1^{(1)} - p_2^{(2)}}{\sqrt{\frac{(n_1^{(1)} - 1)}{n_1^{(1)}} \cdot \text{se}(p_1^{(1)})^2 + \frac{(n_2^{(2)} - 1)}{n_2^{(2)}} \cdot \text{se}(p_2^{(2)})^2}}. $$

This test statistic is also asymptotically $t$ distributed but with the following degrees of freedom ($df$):

$$ df = \sqrt{\frac{\left(\frac{(n_1^{(1)} - 1)}{n_1^{(1)}} \cdot \text{se}(p_1^{(1)})^2 + \frac{(n_2^{(2)} - 1)}{n_2^{(2)}} \cdot \text{se}(p_2^{(2)})^2\right)^2}{\frac{(n_1^{(1)} - 1)}{n_1^{(1)}} \cdot \text{se}(p_1^{(1)})^4 + \frac{(n_2^{(2)} - 1)}{n_2^{(2)}} \cdot \text{se}(p_2^{(2)})^4}} - 2}. $$

Prior research suggests that the parametric approach is rather liberal and likely subject to Type I errors (Sarstedt, Henseler, & Ringle, 2011). Furthermore, from a conceptual perspective, the parametric approach is problematic since it relies on distributional assumptions, which are inconsistent with PLS-SEM’s nonparametric nature. Against this background, researchers have proposed several nonparametric alternatives to multigroup analysis.

**NONPARAMETRIC TESTS**

**PLS-MGA**

Henseler et al. (2009) proposed the nonparametric PLS-MGA approach that builds on bootstrapping results of each data group. For a specific relationship in the PLS path model, their approach compares each bootstrap estimate of one group with all other bootstrap estimates of the same parameter in the other group. By counting the
number of occurrences where the bootstrap estimate of the first group is larger than those of the second group, the approach derives a \( p \) value for a one-tailed test. This approach can be described in formal terms as follows:

\[
\text{Prop.}(p^{(1)} \geq p^{(2)} | \beta^{(1)} < \beta^{(2)}) = 1 - \frac{1}{B} \sum \sum \Theta\left(\left(p_j^{(1)} + p^{(1)} - \bar{p}^{(1)}\right) - \left(p_j^{(2)} + p^{(2)} - \bar{p}^{(2)}\right)\right).
\]

With regard to the \( \beta^{(1)} \) and \( \beta^{(2)} \) true population parameters of each data group, this approach determines the conditional probability \( \text{Prop.}(p^{(1)} \geq p^{(2)} | \beta^{(1)} < \beta^{(2)}) \) that the \( p^{(1)} \) estimation of a certain path coefficient in group 1 is larger than or equal to its estimated \( p^{(2)} \) coefficient in group 2, whereby the superscript in parentheses marks the respective group. By using the PLS-MGA, a researcher would like to ensure that this conditional probability is below a specified \( \alpha \)-level before concluding that \( \beta^{(1)} \) is greater than \( \beta^{(2)} \). In the above equation, \( B \) is the number of bootstrap runs, \( p_j^{(1)} \) \((j = 1, \cdots, B) \) and \( p_j^{(2)} \) \((j = 1, \cdots, B) \) are the group-specific bootstrap results of the path coefficient, \( \bar{p}^{(1)} \) and \( \bar{p}^{(2)} \) denote the mean values of the group-specific bootstrap samples, and \( \Theta \) (i.e., the Greek letter \( \theta \)) stands for the unit step function, which has a value of 1 if its argument exceeds 0, otherwise 0. Note that in accordance with Henseler et al. (2009), all group-specific bootstrap results need to be corrected by the difference between the original group-specific parameter estimate and the average over the group-specific bootstrap values.

To illustrate the working principle of the PLS-MGA, consider a simple model with two constructs, estimated in two groups yielding the path coefficients \( p^{(1)} = 0.501 \) and \( p^{(2)} = 0.336 \). We want to test the hypothesis that the path coefficient is larger in the first group compared to the second group. To test this hypothesis, we draw 10 bootstrap samples for \( p^{(1)} \) and \( p^{(2)} \) and contrast the estimates as shown in Exhibit 4.5. The first column represents the 10 bootstrap samples of the path coefficient in the first group, whereas the first row represents the 10 bootstrap samples in the second group. We now compare each bootstrap estimate of \( p^{(1)} \) (e.g., the first bootstrap sample's estimate 0.494) with each bootstrap estimate of \( p^{(2)} \) (i.e., 0.357, 0.226, ..., 0.272) and count the number of cases where
$p^{(1)} \geq p^{(2)}$, indicated by an X in Exhibit 4.5. Dividing this number (i.e., 11) by the total number of comparisons (i.e., 100) yields the $p$ value, which is 0.11 in our case. Therefore, we cannot conclude that the path coefficient $p$ is significantly larger in the first group compared to the second group.

PLS-MGA involves a great number of comparisons of bootstrap estimates. For an initial analysis of group differences, we suggest drawing on 500 bootstrap samples, using the algorithm settings (e.g., no sign change option) as recommended in Hair et al. (2016). For the final analysis, use 5,000 bootstrap samples, which results in 25,000,000 comparisons for each parameter.

Finally, it is important to emphasize that the PLS-MGA approach allows for testing only one-sided hypotheses. Specifically, the SmartPLS software always tests the hypothesis that $p^{(1)}$ is larger than $p^{(2)}$. In case you want to test the opposite direction (i.e., $p^{(2)} > p^{(1)}$), you therefore need to subtract the resulting $p$ value from 1 to obtain the actual $p$ value for our hypothesis. Using the PLS-MGA approach to test two-sided hypotheses is not possible without limitations as the bootstrap-based distribution is not necessarily symmetric. This characteristic clearly limits its applicability as researchers routinely draw on two-tailed tests, despite frequent concerns (e.g., Cho & Abe, 2012).

**Permutation Test**

The permutation test was originally developed by Chin (2003) and further substantiated by Chin and Dibbern (2010). As its name implies and analogous to its role in Step 2 of the MICOM procedure, the permutation test randomly exchanges observations between the data groups and re-estimates the model for each permutation. Computing the differences between the group-specific path coefficients per permutation enables testing whether these also differ in the population.

The permutation test follows a six-step process:

1. Carry out group-specific estimations of the PLS path models to obtain path coefficient estimates for groups 1 and 2.

2. Compute the difference between the group-specific path coefficient estimates: $d = p^{(1)} - p^{(2)}$. 

Copyright ©2018 by SAGE Publications, Inc. This work may not be reproduced or distributed in any form or by any means without express written permission of the publisher.
### Exhibit 4.5 Data Matrix for PLS-MGA

<table>
<thead>
<tr>
<th>Bootstrap samples $p^{(1)}$</th>
<th>0.357</th>
<th>0.226</th>
<th>0.318</th>
<th>0.281</th>
<th>0.372</th>
<th>0.318</th>
<th>0.296</th>
<th>0.308</th>
<th>0.415</th>
<th>0.272</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.494</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.423</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.324</td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>0.591</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.698</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.291</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>0.509</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.400</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>0.526</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.538</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: X indicates situations when $p^{(1)} \geq p^{(2)}$. 

Copyright ©2018 by SAGE Publications, Inc.

This work may not be reproduced or distributed in any form or by any means without express written permission of the publisher.
3. Produce random permutation of the data, meaning that the observations are randomly assigned to the groups. In formal terms, $n^{(1)}$ observations are drawn without replacement from the aggregate data set and assigned to group 1. The remaining $n^{(2)}$ observations are assigned to group 2. Researchers should use a minimum number of 1,000 permutations.

4. Conduct group-specific estimations of the PLS path models for each permutation run $u$. For example, with 1,000 permutations, we obtain 1,000 model estimates for group 1 (i.e., $p_u^{(1)}$) and for group 2 (i.e., $p_u^{(2)}$).

5. Compute the differences in the permutation run-specific path coefficient estimates: $d_u = p_u^{(1)} - p_u^{(2)}$.

6. Create the two-tailed 95% permutation-based confidence interval. More specifically, sort the resulting $d_u$ values in ascending order and determine the two values that separate (1) the 2.5% lowest from the 97.5% highest values (lower boundary) and (2) the 97.5% lowest from the 2.5% highest values (upper boundary). If the original difference $d$ of the group-specific path coefficient estimates does not fall into the confidence interval, it is statistically significant. In addition, based on the permutation results, SmartPLS provides the $p$ value of the difference $d$. A $p$ value smaller than 0.05 suggests that the difference $d$ between the group-specific path coefficients does not fall into the 95% permutation-based confidence interval and, thus, is statistically significant.

Permutation tests in general have been shown to perform very well across a broad range of conditions (e.g., Ernst, 2004; Good, 2000). Most importantly, permutation tests reliably control for Type I errors when the assignment of observations occurs randomly, as is the case in the permutation test’s application in PLS-SEM. Correspondingly, prior research indicates that the permutation test performs more conservatively than the parametric test in terms of rendering differences significant (Sarstedt, Henseler, & Ringle, 2011). Against this background and because of the test’s nonparametric character, we generally recommend using the permutation test. However, large differences in the group-specific sample sizes have adverse consequences for the permutation test’s performance. Therefore, when one group’s sample is more than double the size of the other group’s, researchers should
draw on the PLS-MGA approach (in case of testing a one-sided hypothesis) or the parametric approach. As an alternative approach, researchers can randomly withdraw another sample from the large group that is comparable in size to the smaller group and compare the two samples using the permutation test.

**Comparing More Than Two Groups**

Standard approaches to multigroup comparison in PLS-SEM have in common that they test the difference in the parameters between two groups. However, researchers frequently encounter situations in which they would like to compare a parameter (e.g., a path coefficient) across more than two groups. In PLS-SEM, the complexity of this analysis—determined by the number of group-specific path coefficient comparisons—increases with the path model complexity (i.e., the number of relationships) and the number of groups. For example, a model with 10 path relationships in the structural model and five groups requires 100 comparisons of group-specific path coefficients. If no significant differences exist across groups, the likelihood to find at least one significant difference is larger than 95% when conducting the 100 comparisons. Not only do such complex analyses become increasingly time-consuming, but there is a more severe problem associated with this approach, called **alpha inflation** (also referred to as the **multiple testing problem**). This refers to the fact that the more tests you conduct at a certain significance level, the more likely you are to claim a significant result when this is not so (i.e., a Type I error). For example, assuming a significance level of 5% and making all possible pairwise comparisons as specified above, the overall probability of a Type I error (also referred to as the **familywise error rate**) is not 5% but 99.41% (Sarstedt & Mooi, 2014). In other words, when conducting all possible pairwise group comparisons, the familywise error rate quickly increases beyond the acceptable Type I error level (i.e., the acceptance of false positives) of, for example, 5%.

Against this background, when comparing more than two groups, two issues warrant attention. First, we need to test whether the path coefficient differs from the others in at least one of the groups (i.e., testing for the overall difference). For this purpose, we suggest using a permutation-based analysis of variance (ANOVA), which maintains the familywise error rate, does not rely on distributional assumptions, and exhibits an acceptable level of statistical power. In case we find a significant difference, the second step is to assess and determine
between which groups the path coefficient differs (i.e., pairwise comparisons) by using the standard procedures (e.g., the permutation test). However, the results become adjusted to account for the multiple testing problem.

In the following section, we explain these two steps in more detail. As a note of caution, you must keep in mind that the more you search, the more likely it is that you will find something. For this reason, you should generally focus on revealing substantial differences across groups in PLS-SEM.

**Testing for the Overall Difference**

The standard approach to testing for the overall difference in a parameter across multiple groups involves running an ANOVA \( F \) test. However, in light of the test’s parametric character and its asymptotic properties when used on bootstrapping results, comparing the different groups by means of an ANOVA \( F \) test is not meaningful in a PLS-SEM context.

As a remedy, Sarstedt, Henseler, and Ringle (2011) developed the omnibus test of group differences (OTG), which combines the bootstrapping procedure with permutation testing to mimic an overall \( F \) test. This approach maintains the Type I error level as specified by the researcher (i.e., the familywise error rate) and delivers an acceptable level of statistical power while not relying on distributional assumptions. The OTG approach consists of the following steps:

1. Run the bootstrapping procedure on each group separately to derive reference distributions of the group-specific parameters.
2. Compute the variance ratio. This is the ratio of variance explained by the grouping variable and the overall variance:

\[
F_k = \frac{K \cdot B \cdot (1 / K - 1) \sum_{i=1}^{K} (\overline{p}^{(i)} - \overline{p})^2}{1 / (B - 1) \sum_{i=1}^{B} \sum_{j=1}^{K} (p_i^{(j)} - \overline{p}^{(j)})^2}
\]

In this equation \( p_i^{(j)} \) is the parameter estimate from the \( j \)th bootstrap sample (\( j = 1, \ldots, B \)) in the \( i \)th group (\( i = 1, \ldots, K \)), \( \overline{p}^{(i)} \) the average over all bootstrap parameter estimates of group \( i \), and \( \overline{p} \) the grand mean of the bootstrap parameter estimates across all the groups.
3. Produce permutations of the parameter estimates $p_{ij}^{(i)}$ between the groups, separately for each bootstrap sample. The process starts with the first bootstrap sample by randomly exchanging the parameter estimates $p_{ij}^{(i)}$ between the groups. Next, the same is done for the second and all further bootstrap samples until all parameter estimates have been permuted for all $B$ bootstrap samples. This process yields a total of $(K!)^{B-1}$ different permutations, which is difficult to complete computationally—for example, for $B=5,000$ bootstrap samples and $K=3$ groups, $(3!)^{4,899} = 9.508 \cdot 10^{1,889}$ permutations are required! Therefore, the OTG approach randomly engages in, for example, 5,000 permutations and computes the variance ratio $F_{R_{l}}$ for each permutation run $l (l = 1, \ldots, U)$.

4. Compute the $p$ value using the Heaviside step function. The Heaviside step function is a discontinuous function whose value is 0 for the negative argument and 1 for the positive argument. In the context of the OTG approach, it returns 1 when $F_{R} - F_{R_{l}}$ yields a positive value and 0 for a negative value:

$$ p = \frac{1}{U} \sum_{l=1}^{U} H \left( F_{R} - F_{R_{l}} \right) $$

If the OTG approach indicates a significant effect, we can conclude that at least one group’s path coefficient differs significantly from those of the other groups. This does not mean, however, that all the groups’ path coefficients differ significantly. To assess which groups differ from each other, we need to apply a pairwise comparisons test.

The OTG approach has not yet been included in SmartPLS. However, researchers can download the corresponding R code at http://www.pls-sem.com/downloads.

**Pairwise Comparisons**

The OTG approach allows identifying path candidate path coefficients for a pairwise comparison using, for example, the permutation test. Since such tests compare only two groups at a time, you need to run this test multiple times until you have conducted all relevant comparisons. For example, if you like to compare a specific path relationship $p$ across five groups, you must conduct 10 comparisons.
For this example, Exhibit 4.6 shows how to report the comparisons results of a specific path coefficient across multiple groups (e.g., by reporting the difference and their significance in each cell).

As indicated earlier, a primary concern of conducting pairwise comparisons is the alpha inflation, according to which the overall probability of a Type I error (i.e., the familywise error rate) increases exponentially with the number of comparisons. A standard approach for controlling for the familywise error rate is the Bonferroni correction. Instead of using a specific significance level in the test decision, the Bonferroni correction tests each individual hypothesis at a significance level of $\alpha / m$. For example, in case of five groups, there would be $m = 10$ comparisons, yielding a significance level of $0.05 / 10 = 0.005$ instead of 0.05. Another popular approach is the Šidák procedure, which uses $1 - \left(1 - \alpha\right)^{1/m}$ in each of the hypothesis tests. Here, in the example of five groups (i.e., 10 comparisons), one would use a significance level of $1 - \left(1 - 0.05\right)^{1/10} = 0.005116$ instead of 0.05.

Both approaches counteract the increase in the familywise error rate when performing multiple comparisons across several groups of data. However, the maintenance of the familywise error rate comes to the expense of statistical power. For example, with $m = 10$, for a single comparison to be considered significant, with the Bonferroni correction the effect needs to be 43% larger than that with an alpha level of 5%.

<table>
<thead>
<tr>
<th>Exhibit 4.6</th>
<th>Pairwise Comparison of a Specific Path Coefficient Across Multiple Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Group 1</td>
</tr>
<tr>
<td>Group 1</td>
<td></td>
</tr>
<tr>
<td>Group 2</td>
<td>$p^{(1)} - p^{(2)}$</td>
</tr>
<tr>
<td>Group 3</td>
<td>$p^{(1)} - p^{(3)}$</td>
</tr>
<tr>
<td>Group 4</td>
<td>$p^{(1)} - p^{(4)}$</td>
</tr>
<tr>
<td>Group 5</td>
<td>$p^{(1)} - p^{(5)}$</td>
</tr>
</tbody>
</table>

**Note:** $p_1$ represents the specific path coefficient in the structural model; the comparison is across five groups; the superscript numbers in brackets indicate the group for which the specific coefficient was obtained.
Exhibit 4.7  Rules of Thumb for Invariance Assessment and Multigroup Analyses

Invariance assessment:

- Running a multigroup analysis requires partial measurement invariance, which is given when configural invariance and compositional invariance hold. If partial measurement invariance is confirmed and the composites have equal mean values and variances across the groups, full measurement invariance is confirmed, which supports the pooled data analysis.

- For configural invariance assessment, check whether the measurement models employ the same indicators across the groups in terms of numbers, item content, and coding. Any type of data treatment (e.g., outliers, missing values) must occur identically for the groups. Group-specific model estimations must draw on the same algorithm settings.

- Use 1,000 permutations (or more) for the assessment of compositional invariance and the equality of composite mean values and variances in Steps 2 and 3 of the MICOM procedure.

- Compositional invariance is established when the original correlations between the composite scores of groups 1 and 2 are larger than the 5%-quantile of the empirical distribution of $c_u$. In exploratory research and when the sample size is small, reverting to the 10%-quantile is more reasonable.

- To test for the equality of the composites’ mean values and variances across the groups, check whether differences in means and variances are significant.

Multigroup analyses:

- Use the permutation test to multigroup analysis using 1,000 permutations (or more).

- When one group’s sample is more than double the size of the other group’s, randomly withdraw another sample from the large group that is comparable in size to the smaller group, and compare the two samples using the permutation test. Alternatively, use the PLS-MGA approach (in case of testing a one-sided hypothesis) or the parametric approach.

- When using the PLS-MGA approach, use 500 bootstrap samples for the initial assessment and 5,000 bootstrap samples for the final assessment of group differences. Use standard bootstrapping algorithm settings (e.g., no sign change option).

- When using the parametric approach, test for differences in the group variances. In case variances differ significantly, apply the Welch-Satterthwaite $t$ test.

- To analyze more than two groups, use the OTG approach.
Therefore, researchers generally prefer the Šidák procedure over the Bonferroni correction as this approach achieves an acceptable level of statistical power. Research has brought forward a range of alternative (and more complex) approaches for maintaining the familywise error rate (for details, see Hochberg & Tamhane, 2011).

Exhibit 4.7 summarizes the rules of thumb for a combined multigroup analysis and invariance assessment in PLS-SEM.

**CASE STUDY ILLUSTRATION—INVARIANCE ASSESSMENT AND MULTIGROUP ANALYSIS**

To illustrate the use of multigroup analysis in conjunction with the assessment of measurement model invariance by means of the MICOM procedure, we draw on the extended corporate reputation model. Rather than analyzing the aggregate data set with 344 observations, we are interested in analyzing whether the effects in the model differ significantly for customers with prepaid cell phone plans from those with contract plans. The research study obtained data on customers with a contract plan (servicetype = 1; n₁ = 219) versus those with a prepaid plan (servicetype = 2; n₂ = 125), and we will compare those two groups. When engaging in a multigroup analysis, we need to ensure that the number of observations in each group meets the rules of thumb for minimum sample size requirements. As the maximum number of arrows pointing at a latent variable is eight, we would need at least 8 \cdot 10 = 80 observations per group, according to the 10 times rule. Following the more rigorous recommendations from a power analysis (Hair et al., 2016), 54 observations per group are needed to detect \( R^2 \) values of around 0.25 at a significance level of 5% and a power level of 80%. Therefore, the group-specific sample sizes can be considered sufficiently large.

In the first step, we have to define the grouping variable that SmartPLS uses to split up the data set. To do so, double-click on the data set Corporate Reputation Data [344 records] in the Corporate Reputation project file. SmartPLS will open a new tab (Exhibit 4.8), which provides information on the data set and its format (data view).

When opening the data view, three options appear in the menu above the modeling window: Add Data Group, Generate Data Groups, and Clear Data Groups. To define a new grouping variable, click on Generate Data Groups and a new dialogue box will open (Exhibit 4.9). Next to Name prefix, you can specify a prefix that
SmartPLS uses in the naming of the groups in the results report. In our example, we maintain the default prefix `Group_`. Under `Group Column`, we can specify one or more grouping variables. When left-clicking on the box next to `Group column 0`, a list of all variables in the dataset will appear. Behind each variable, SmartPLS will show the number of unique values. Scroll down and locate the variable `servicetype`, which has two unique values. Left-click on `servicetype` and close the dialogue box by left-clicking on `OK`.

Back in the data view, SmartPLS will show an additional tab labeled `Data Groups`, which offers information on the group-specific sample sizes. You can now close the data view and go back to the modeling window showing the extended corporate reputation model. Before continuing, make sure that the correct data set (i.e., `Corporate Reputation Data [344 records]`) is activated. If the corresponding font does not appear in green, right-click on `Corporate Reputation Data [344 records]` and click on `Select Active Data File`.

To run the MICOM procedure, you need to navigate to `Calculate → Permutation`, which you can find at the top of the SmartPLS screen. Alternatively, you can left-click on the wheel symbol with the label `Calculate` in the tool bar. A menu with alternative
analysis options opens from which you can select running Permutation procedure.

In the dialogue box that follows (Exhibit 4.10), we first need to specify the groups to be compared. To do so click on the menu next to **Group A** and select **GROUP_servicetype(1.0)**. Next, click on the menu next to **Group B** and select **GROUP_servicetype(2.0)**. Specify at least a number of 1,000 permutations and two-tailed testing. Because service type is known to affect the way customers relate to experiences, we assume a significance level of 0.05. Next, click on Start Calculation. SmartPLS now estimates the model for each group separately, using the same algorithm settings.
The PLS path models as well as the data treatment used in both groups are identical, which is a necessary requirement for the establishment of configural invariance in Step 1 of the MICOM procedure. Furthermore, as our group-specific model estimations also draw on the identical algorithm settings, configural invariance (Step 1 in MICOM) is established. The following analyses address Steps 2 and 3 of the MICOM procedure. In the results report that opens, go to Quality Criteria → MICOM. The first tab labeled Step 2 (see Exhibit 4.11) shows the results of the compositional invariance assessment. Note that as permutation is a random process, your results will slightly differ from those presented here. Similarly, the results will differ every time you rerun the MICOM procedure.

The column 5% shows the 5% quantile of the empirical distribution of $c_u$. Comparing the correlations $c$ between the composite scores of the first and second group (column Original Correlations) with the 5% quantile reveals that the quantile is always smaller than (or equal to) the correlation $c$ for all the constructs. This result is also supported by the $p$ values that are higher than 0.05, indicating the correlation is not significantly lower than 1. For example, the original correlation value of ATTR is 0.993. This result is within the corresponding permutation-based confidence interval with a lower boundary of 0.945; in accordance, ATTR’s $p$ value of 0.673, as displayed in the column Permutation $p$-Values, is considerably larger than 0.05 (Exhibit 4.11). Hence, the original correlation of ATTR is not significantly different from 1, which supports the conclusion.
that compositional invariance has been established for this construct. Similarly, we substantiate that compositional invariance has been established for all multi-item constructs in the model. Note that we cannot run Step 2 for the single item construct CUSA since its single outer relationship is 1 by design. Thus, we can ignore the $p$ value of CUSA (Exhibit 4.11), which is not perfectly 0 in this example since the lower boundary value is 0.999999999999998 due to rounding of certain permutation results. To summarize, the results from Step 2 support partial measurement invariance. Thus, we can compare the standardized path coefficients across the groups by means of a multigroup analysis with confidence. To check whether even full measurement invariance holds, click on the tab Step 3 in the results report (see Exhibit 4.12).

The first two columns in Exhibit 4.12 show the mean differences between the composite scores as resulting from the original model estimation and the permutation procedure, respectively (note that due to space constraints, the screenshot does not show the entire column label). The next two columns show the lower (2.5%) and upper (97.5%) boundaries of the 95% confidence interval of the scores’ mean differences. As we can see, every confidence interval includes the original difference in mean values, indicating that there are no significant differences in the mean values of latent variables across the two groups. For example, for ATTR, the original difference in mean values of the latent variable scores is −0.064, which is within the corresponding confidence interval with a lower boundary of −0.225 and an upper boundary of 0.221. The result in the column Permutation $p$-Values further supports this finding for ATTR and every other construct in the PLS path model as all the $p$ values are considerably larger than 0.05. The next columns show the analogous results for the composite variances. Again, all the
confidence intervals include the original value and all the $p$ values are clearly larger than 0.05. We therefore conclude that all the composite mean values and variances are equal, providing support for full measurement invariance. Exhibit 4.13 summarizes the results of this MICOM analysis. In light of these results, we continue by examining the results of the multigroup analysis. Our assessment focuses on the permutation test, which we can access by going to Final Results → Path Coefficients in the SmartPLS permutation output.

### Exhibit 4.13 Summary of the MICOM Results

<table>
<thead>
<tr>
<th>MICOM Step 1</th>
<th>Configural variance established? Yes</th>
</tr>
</thead>
<tbody>
<tr>
<td>MICOM Step 2</td>
<td></td>
</tr>
<tr>
<td>Composite</td>
<td>Correlation $c$</td>
</tr>
<tr>
<td>ATTR</td>
<td>0.993</td>
</tr>
<tr>
<td>COMP</td>
<td>0.998</td>
</tr>
<tr>
<td>CSOR</td>
<td>0.985</td>
</tr>
<tr>
<td>CUSA</td>
<td>1.000</td>
</tr>
<tr>
<td>CUSL</td>
<td>1.000</td>
</tr>
<tr>
<td>LIKE</td>
<td>1.000</td>
</tr>
<tr>
<td>PERF</td>
<td>0.979</td>
</tr>
<tr>
<td>QUAL</td>
<td>0.961</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>MICOM Step 3</th>
<th>Difference of the composite's mean value ($= 0$)</th>
<th>95% confidence interval</th>
<th>$p$ value</th>
<th>Equal mean values?</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATTR</td>
<td>$-0.064$</td>
<td>$[-0.225; 0.221]$</td>
<td>0.601</td>
<td>Yes</td>
</tr>
<tr>
<td>COMP</td>
<td>0.060</td>
<td>$[-0.238; 0.236]$</td>
<td>0.611</td>
<td>Yes</td>
</tr>
<tr>
<td>CSOR</td>
<td>0.072</td>
<td>$[-0.216; 0.223]$</td>
<td>0.539</td>
<td>Yes</td>
</tr>
<tr>
<td>CUSA</td>
<td>$-0.097$</td>
<td>$[-0.226; 0.214]$</td>
<td>0.420</td>
<td>Yes</td>
</tr>
</tbody>
</table>

(Continued)
The first two columns Exhibit 4.14 show the original path coefficients in group 1 and group 2, respectively, followed by their differences in the original data set and the permutation testing, respectively (note that due to space constraints, the screenshot does not show the entire column label). As can be seen, most structural model relationships do not differ between the two groups. The only exceptions are the relationships between COMP and CUSL, CUSA and CUSL, as well as LIKE and CUSL, which differ significantly on a 10% level. More precisely, the effect between COMP and CUSL is significantly ($p \leq 0.10$) different between customers with a contract plan ($p^{(1)} = 0.136$) and those with a prepaid plan ($p^{(2)} = -0.062$). Similarly, the relationship between CUSA and CUSL is significantly ($p \leq 0.10$) different among customers with a contract plan ($p^{(1)} = 0.599$) versus those with a

<table>
<thead>
<tr>
<th>Composite</th>
<th>Difference of the composite’s mean value ($= 0$)</th>
<th>95% confidence interval</th>
<th>$p$ value</th>
<th>Equal mean values?</th>
</tr>
</thead>
<tbody>
<tr>
<td>CUSL</td>
<td>-0.060</td>
<td>[-0.229; 0.219]</td>
<td>0.587</td>
<td>Yes</td>
</tr>
<tr>
<td>LIKE</td>
<td>-0.054</td>
<td>[-0.227; 0.229]</td>
<td>0.653</td>
<td>Yes</td>
</tr>
<tr>
<td>PERF</td>
<td>-0.023</td>
<td>[-0.222; 0.221]</td>
<td>0.897</td>
<td>Yes</td>
</tr>
<tr>
<td>QUAL</td>
<td>-0.010</td>
<td>[-0.229; 0.244]</td>
<td>0.949</td>
<td>Yes</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Composite</th>
<th>Logarithm of the composite’s variances ratio ($= 0$)</th>
<th>95% confidence interval</th>
<th>$p$ value</th>
<th>Equal variances?</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATTR</td>
<td>0.190</td>
<td>[-0.297; 0.254]</td>
<td>0.691</td>
<td>Yes</td>
</tr>
<tr>
<td>COMP</td>
<td>0.135</td>
<td>[-0.298; 0.266]</td>
<td>0.639</td>
<td>Yes</td>
</tr>
<tr>
<td>CSOR</td>
<td>-0.106</td>
<td>[-0.294; 0.262]</td>
<td>0.581</td>
<td>Yes</td>
</tr>
<tr>
<td>CUSA</td>
<td>0.230</td>
<td>[-0.374; 0.351]</td>
<td>0.613</td>
<td>Yes</td>
</tr>
<tr>
<td>CUSL</td>
<td>-0.070</td>
<td>[-0.34; 0.351]</td>
<td>0.721</td>
<td>Yes</td>
</tr>
<tr>
<td>LIKE</td>
<td>0.012</td>
<td>[-0.258; 0.248]</td>
<td>0.747</td>
<td>Yes</td>
</tr>
<tr>
<td>PERF</td>
<td>0.038</td>
<td>[-0.306; 0.293]</td>
<td>0.927</td>
<td>Yes</td>
</tr>
<tr>
<td>QUAL</td>
<td>0.154</td>
<td>[-0.309; 0.271]</td>
<td>0.981</td>
<td>Yes</td>
</tr>
</tbody>
</table>
Exhibit 4.14 Permutation Test in SmartPLS

<table>
<thead>
<tr>
<th>Path Coefficients</th>
<th>Path Coefficients</th>
<th>Path Coefficients Original Difference</th>
<th>Path Coefficients Permutation</th>
<th>2.5%</th>
<th>97.5%</th>
<th>Permutation p-Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATTR -&gt; COMP</td>
<td>0.185</td>
<td>0.060</td>
<td>0.114</td>
<td>-0.082</td>
<td>-0.217</td>
<td>0.028</td>
</tr>
<tr>
<td>ATTR -&gt; LIKE</td>
<td>0.241</td>
<td>0.105</td>
<td>0.132</td>
<td>-0.084</td>
<td>-0.245</td>
<td>0.225</td>
</tr>
<tr>
<td>COMP -&gt; CUSL</td>
<td>0.214</td>
<td>0.188</td>
<td>0.106</td>
<td>-0.003</td>
<td>-0.293</td>
<td>0.274</td>
</tr>
<tr>
<td>COMP -&gt; CUSL</td>
<td>0.136</td>
<td>0.056</td>
<td>0.187</td>
<td>0.002</td>
<td>-0.214</td>
<td>0.226</td>
</tr>
<tr>
<td>CUSL -&gt; COMP</td>
<td>-0.025</td>
<td>0.081</td>
<td>-0.118</td>
<td>0.001</td>
<td>-0.213</td>
<td>0.228</td>
</tr>
<tr>
<td>CUSL -&gt; LIKE</td>
<td>0.096</td>
<td>0.184</td>
<td>-0.100</td>
<td>0.001</td>
<td>-0.201</td>
<td>0.213</td>
</tr>
<tr>
<td>CUSL -&gt; CUSL</td>
<td>0.399</td>
<td>0.440</td>
<td>0.151</td>
<td>-0.091</td>
<td>-0.166</td>
<td>0.169</td>
</tr>
<tr>
<td>LIKE -&gt; CUSL</td>
<td>0.065</td>
<td>0.076</td>
<td>-0.111</td>
<td>0.003</td>
<td>-0.234</td>
<td>0.253</td>
</tr>
<tr>
<td>LIKE -&gt; CUSL</td>
<td>0.225</td>
<td>0.404</td>
<td>-0.220</td>
<td>-0.002</td>
<td>-0.235</td>
<td>0.235</td>
</tr>
<tr>
<td>PERS -&gt; COMP</td>
<td>0.210</td>
<td>0.359</td>
<td>-0.090</td>
<td>0.007</td>
<td>-0.278</td>
<td>0.287</td>
</tr>
<tr>
<td>PERS -&gt; LIKE</td>
<td>-0.002</td>
<td>0.172</td>
<td>-0.174</td>
<td>0.094</td>
<td>-0.209</td>
<td>0.276</td>
</tr>
<tr>
<td>RELQUAL -&gt; COMP</td>
<td>0.353</td>
<td>0.418</td>
<td>0.165</td>
<td>0.006</td>
<td>-0.246</td>
<td>0.271</td>
</tr>
</tbody>
</table>

prepaid plan ($p^{(2)} = 0.440$). Finally, the effect of LIKE on CUSL is significantly ($p < 0.10$) different among customers with a contract plan ($p^{(1)} = 0.205$) versus those with a prepaid plan ($p^{(2)} = 0.424$).

To further analyze group-specific effects, we can run another multigroup analysis approach on the data. To do so, go to Calculate → Multi-Group Analysis (MGA), which you can find at the top of the SmartPLS screen. Alternatively, you can left-click on the wheel symbol in the tool bar and click the Multi-Group Analysis (MGA) option.
In the dialogue box that follows (Exhibit 4.15), we again need to specify the groups to be compared. To do so tick the box next to `GROUP_servicetype(1.0)` under Groups A. Next, select `GROUP_servicetype(2.0)` under Groups B. In the other tabs, we can specify settings related to the PLS-SEM algorithm, the bootstrapping procedure, missing value treatment, and variable weighting. Use the standard settings for all the options as described in, for example, Hair et al. (2016).

In the results report that opens, go to Final Results → Path Coefficients and select the PLS-MGA tab to access the results of this multigroup analysis approach (Exhibit 4.16).

The results support those of the permutation test. The vast majority of structural model relationships do not differ between the two groups, with the exceptions of `CUSA → CUSL`, `LIKE → CUSL`, and `COMP → CUSL`. As PLS-MGA represents a one-tailed test, both differences are significant at a 5% level. Note that the PLS-MGA approach relies on the bootstrapping procedure, which is a random process. Therefore, your test results will vary from those presented here, but the differences should not be substantial. Finally, Exhibit 4.17 presents the results of the parametric tests, which you can access by clicking the corresponding tab in the results report of the Multi-Group Analysis (MGA). We show the results of the parametric test since the MICOM procedure revealed that the variances do not differ across groups. However, the Welch-Satterthwaite $t$ test results, which you can find in the corresponding tab of the results report, are very similar.
Exhibit 4.17  Parametric PLS Multigroup Tests Results

<table>
<thead>
<tr>
<th>Path Coefficients</th>
<th>Bootstrapping Results</th>
<th>Confidence Intervals (Bias Corrected)</th>
<th>PLS-MGA</th>
<th>Parametric Test</th>
<th>Welch-Satterthwaite Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATTR → COMP</td>
<td>0.111</td>
<td>1.027</td>
<td>0.305</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATTR → LIKE</td>
<td>0.112</td>
<td>0.817</td>
<td>0.415</td>
<td></td>
<td></td>
</tr>
<tr>
<td>COMP → CUSA</td>
<td>0.106</td>
<td>0.740</td>
<td>0.455</td>
<td></td>
<td></td>
</tr>
<tr>
<td>COMP → CUSL</td>
<td>0.197</td>
<td>1.732</td>
<td>0.084</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CSOR → COMP</td>
<td>0.118</td>
<td>1.074</td>
<td>0.283</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CSOR → LIKE</td>
<td>0.103</td>
<td>0.860</td>
<td>0.390</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CUSL → CUSL</td>
<td>0.139</td>
<td>1.062</td>
<td>0.063</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LIKE → CUSA</td>
<td>0.111</td>
<td>0.878</td>
<td>0.381</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LIKE → CUSL</td>
<td>0.220</td>
<td>1.921</td>
<td>0.056</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PERF → COMP</td>
<td>0.095</td>
<td>0.721</td>
<td>0.472</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PERF → LIKE</td>
<td>0.174</td>
<td>1.197</td>
<td>0.334</td>
<td></td>
<td></td>
</tr>
<tr>
<td>QUAL → COMP</td>
<td>0.117</td>
<td>0.543</td>
<td>0.397</td>
<td></td>
<td></td>
</tr>
<tr>
<td>QUAL → LIKE</td>
<td>0.165</td>
<td>1.160</td>
<td>0.247</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: X indicates significant difference (p < 0.10) of path coefficients across the groups.

Exhibit 4.18  PLS Multigroup Results Across methods

<table>
<thead>
<tr>
<th>Path Coefficient</th>
<th>Permutation Test</th>
<th>PLS-MGA</th>
<th>Parametric Test</th>
<th>Welch-Satterthwaite t Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATTR → COMP</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATTR → LIKE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COMP → CUSA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COMP → CUSL</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>CSOR → COMP</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CSOR → LIKE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CUSL → CUSL</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>LIKE → CUSL</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LIKE → CUSA</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>LIKE → CUSL</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>PERF → COMP</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PERF → LIKE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>QUAL → COMP</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>QUAL → LIKE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: X indicates significant difference (p < 0.10) of path coefficients across the groups.
In comparison, we find the PLS multigroup analysis results do not entail different outcomes across methods, as shown in Exhibit 4.18. Even though we recommend using the permutation test, the results of such an additional multimethods approach provide additional confidence in the final results obtained.

**SUMMARY**

- **Appreciate the importance of considering heterogeneity in PLS-SEM.** Applications of PLS-SEM are usually based on the assumption that the data represent a homogeneous population. In such cases, the estimation of a unique global PLS path model represents all observations. In many real-world applications, however, the assumption of sample homogeneity is unrealistic because respondents are likely to be heterogeneous in their perceptions and evaluations of latent phenomena. While the consideration of heterogeneity is promising from a practical and theoretical perspective to learn about differences between groups of respondents, it is oftentimes necessary, but also challenging, to obtain valid results. For example, when there are significant differences between path coefficients across groups, an analysis on the aggregate data level could cancel out group-specific effects. The results of such an analysis would likely be seriously misleading and render erroneous recommendations to managers and researchers. Researchers are well advised, therefore, to consider the issue of heterogeneous data structures in their modeling efforts.

- **Explain the difference between observed and unobserved heterogeneity.** Heterogeneity can be observed in that differences between two or more groups of data relate to observable characteristics, such as gender, age, or country of origin. Researchers can use these observable characteristics to partition the data into separate groups of observations and carry out group-specific PLS-SEM analyses. Second, heterogeneity can be unobserved in that it does not depend on one specific observable characteristic or combinations of several characteristics. To identify and treat unobserved heterogeneity, researchers can draw on a range of latent class techniques.
• **Comprehend the concept of measurement model invariance and its assessment in PLS-SEM.** Group comparisons are valid if measurement invariance has been established. Thereby, researchers ensure that group differences in model estimates do not result from the distinctive content and/or meanings of the latent variables across groups. When measurement invariance is not demonstrated, any conclusions about model relationships are questionable. The measurement invariances of composites (MICOM) procedure represents a useful tool in PLS-SEM. The procedure comprises three steps that test different aspects of measurement invariance: (1) configural invariance (i.e., equal parameterization and way of estimation), (2) compositional invariance (i.e., equal indicator weights), and (3) equality of composite mean values and variances.

• **Understand multigroup analysis and execute it in PLS-SEM.** Multigroup analysis allows testing whether differences between group-specific path coefficients are statistically significant. Researchers have proposed different approaches to multigroup analysis, with the t test–based parametric approach being most frequently applied. Alternatives such as the permutation test and the PLS-MGA do not rely on distributional assumptions. In comparison, the permutation test has particularly advantageous statistical properties, and we recommend this method when analyzing the difference of parameters across two groups. In the case of more than two groups, the use of the OTG is appropriate.

**REVIEW QUESTIONS**

1. What is the difference between observed and unobserved heterogeneity?
2. Explain the three steps of the MICOM procedure.
3. Why would you run a multigroup analysis in PLS-SEM?
4. Which approaches to multigroup analysis are available in PLS-SEM? Discuss their advantages and disadvantages.
CRITICAL THINKING QUESTIONS

1. Why is the consideration of heterogeneity so important when analyzing PLS path models?

2. Critically comment on the following statement: Measurement invariance is not an issue in PLS-SEM because of the method’s focus on prediction and exploration.

3. What are the implications for your analysis when full measurement invariance holds?

4. Explain the conceptual differences between a multigroup analysis and an analysis of interaction effects.

KEY TERMS

Alpha inflation of composite models (MICOM) procedure
Bonferroni correction Multigroup analysis
Bootstrap samples Multiple testing problem
Bootstrap samples Observed heterogeneity
Bootstrapping Omnibus test of group differences (OTG)
Cluster analysis OTG
Confocal analysis Parametric test
Common factor model Partial measurement invariance
Compositional invariance Permutation test
Confidential invariance Permutation testing
Equality of composite mean values and variances PLSTMGA
Familywise error rate PLS-POS
FIMIX-PLS Prediction-oriented segmentation in PLS-SEM (PLS-POS)
Finite mixture partial least squares (FIMIX-PLS) Šidák procedure
Full measurement invariance Unobserved heterogeneity
k-means clustering
SUGGESTED READINGS


