



**UNIVERSITY OF LEEDS**

This is a repository copy of *Addressing Endogeneity in International Marketing Applications of Partial Least Squares Structural Equation Modeling*.

White Rose Research Online URL for this paper:  
<http://eprints.whiterose.ac.uk/142302/>

Version: Accepted Version

---

**Article:**

Hult, GTM, Hair, JF, Proksch, D et al. (3 more authors) (2018) Addressing Endogeneity in International Marketing Applications of Partial Least Squares Structural Equation Modeling. *Journal of International Marketing*, 26 (3). pp. 1-21. ISSN 1069-031X

<https://doi.org/10.1509/jim.17.0151>

---

© 2018, American Marketing Association. This is an author produced version of a paper published in *Journal of International Marketing*. Uploaded in accordance with the publisher's self-archiving policy.

**Reuse**

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.



[eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk)  
<https://eprints.whiterose.ac.uk/>

# **Addressing Endogeneity in International Marketing Applications of Partial Least Squares Structural Equation Modeling**

## **ABSTRACT**

Partial least squares structural equation modeling (PLS-SEM) has become a key method in international marketing research. Users of PLS-SEM have, however, largely overlooked the issue of endogeneity, which has become an integral component of regression analysis applications. This lack of attention is surprising, given that the PLS-SEM method is grounded in regression analysis for which numerous approaches for handling endogeneity have been proposed. To identify and treat endogeneity, and create awareness of how to deal with this issue, this study introduces a systematic procedure that translates control variables, instrumental variables, and Gaussian copulas into a PLS-SEM framework. We illustrate the procedure's efficacy by means of empirical data, and offer recommendations to guide international marketing researchers on how to effectively address endogeneity concerns in their PLS-SEM analyses.

## **KEYWORDS**

Control variable, endogeneity, Gaussian copula, instrumental variable, omitted variable, PLS-SEM, partial least squares (PLS), structural equation modeling (SEM)

## INTRODUCTION

The ambition to make causal claims is fundamental to international marketing research. Consider, for example, a sampling of recent research questions addressed in articles published in *Journal of International Marketing*: Is coopetition between foreign and local partners good or bad for international joint venture performance (Shu et al. 2017)? How do strategic fit between host-home country similarity and exploration exploitation strategies impact firm performance (Peng Cui et al. 2014)? Do macroenvironmental factors such as cultural dimensions and globalization forces moderate the relationship between product diversification and market value of large international firms (Qiu et al. 2015; Tianjiao 2014)?

Analyzing such research questions requires a research design that randomly assigns corporations or customers to different experimental conditions, and then tests for differences among the groups (Antonakis et al. 2010). However, like in many other fields of scientific inquiry (e.g., Clougherty et al. 2016; Sande and Ghosh 2018), researchers in international marketing seldom undertake randomized controlled experiments. This is not surprising, given that such experiments are prohibitively expensive, or simply impossible to conduct, due to managers and firms' unwillingness to be randomly assigned to the treatment and control groups (Bascle 2008). As a consequence, researchers typically rely on survey-based primary data or secondary data (e.g., Nakata and Huang 2005) in which independent variables cannot be exogenously manipulated (e.g., Clougherty et al. 2016).

At the same time, opportunities for endogeneity to be present are widespread throughout international marketing empirical research designs and data contexts. As Deng (2016, p. 1) notes, "endogeneity is pervasive and sometimes inevitable in international business and marketing

studies.“ Numerous researchers have, therefore, called for a routine consideration of endogeneity in these and related fields (e.g., Hult et al. 2008; Sande and Ghosh 2018).

Endogeneity arises, for example, when a firm’s international marketing strategies (e.g., channel selection, entry modes, and advertising spending) are nonrandom and influenced by firm and industry-level factors. These factors may include a firm’s corporate culture and organizational capabilities or market forces such as industry growth and competition intensity (Jin et al. 2016). But these factors are typically difficult to measure and include in statistical models (e.g., Hult et al. 2008). Similarly, a multitude of factors, which are difficult to consider in full, potentially influence a firm’s international marketing performance (i.e., the outcome of a firm’s decision-making regarding internationalization activities; Souchon et al. 2016). In these situations, if an omitted construct or variable affects both the dependent variable and one or more independent variable(s) in the regression model, the condition will induce a correlation between the independent variable(s) and the error term (Wooldridge 2013). Hence, variation in the omitted construct or variable will manifest in the error term, thereby violating a key causal modeling assumption of regression analysis—the error term is assumed to have an expected value of zero (Wooldridge 2013). As a consequence, the coefficient estimates from standard regression analyses are biased and inconsistent, thereby becoming causally uninterpretable and potentially triggering type I and type II errors (Papies et al. 2016).

Dealing with endogeneity has been extensively discussed in the marketing literature, especially with respect to different forms of regression and panel models (e.g., Ebbes et al. 2011; Park and Gupta 2012; Rossi 2014), as well as conjoint analysis (e.g., Liu et al. 2007). But little research addresses endogeneity in structural equation modeling (SEM). While several studies have discussed endogeneity in the context of factor-based SEM (e.g., Bollen et al. 2014; Kirby and

Bollen 2009), there is a paucity of research on this topic in partial least squares structural equation modeling (PLS-SEM; Wold 1982) whose usage has recently gained momentum in international marketing (Richter et al. 2016) and business research in general (e.g., Ali et al. 2018; Nitzl 2016; Ringle et al. 2018). Some researchers even claim that PLS-SEM does not allow for addressing endogeneity at all (e.g., Antonakis et al. 2010; McIntosh et al. 2014; Rönkkö and Evermann 2013). This assertion is astonishing and inaccurate given that PLS-SEM is grounded in regression analysis (Lohmöller 1989, Chapter 2), for which numerous approaches for handling endogeneity exist (e.g., Ebbes et al. 2005; Park and Gupta 2012; Staiger and Stock 1997). Indeed, Benitez et al. (2016) recently made an advance in this direction by combining the standard PLS-SEM algorithm with the two-stage least squares (2SLS) method, but did not consider variables that control for endogeneity's sources.

Overall, research has not yet developed a comprehensive framework that (1) considers the multitude of approaches applicable in PLS-SEM and (2) offers guidelines for assessing endogeneity in this context. It is therefore not surprising that PLS-SEM applications do not use approaches for assessing endogeneity. Specifically, our review of articles published in the ten-year period between 2008 and 2017 in the *Journal of International Marketing*, *International Marketing Review* and the *Journal of International Business Studies*—the three highest ranked international marketing journals in Hult et al.'s (2009) journal ranking—shows that none of the 43 reviewed PLS-SEM studies tested for endogeneity (Table A1 in the Appendix).

Addressing this gap in research, this study explains and illustrates how to deal with endogeneity in PLS-SEM-based international marketing studies. Based on a discussion of endogeneity's roots and consequences, we focus on omitted variables as the principal, and most commonly examined, source of endogeneity (Bascle 2008; Clougherty et al. 2016).<sup>1</sup> We then describe the

different methods of (statistically) responding to endogeneity in regression models and translate these to a PLS-SEM context. Based on the above-mentioned discussion, we offer a systematic procedure for identifying and treating endogeneity in PLS-SEM and illustrate its use in an empirical example of corporate reputation, a concept frequently examined in international marketing and business research.

## **TREATING ENDOGENEITY IN REGRESSION MODELS**

While endogeneity can have various roots, such as measurement errors, simultaneous causality, common method variance, and (un)observed heterogeneity (Bascle 2008; McIntosh et al. 2014; Papies et al. 2016; Sande and Ghosh 2018), endogeneity problems most often arise from omitted variables that correlate with one or more independent variable(s) and the dependent variable(s) in the regression model (e.g., Bijmolt et al. 2005; Rossi 2014). Omitting such variables induces a correlation between the corresponding independent variables and the dependent variables' error term (Wooldridge 2010). That is, the independent variables then not only explain the dependent variable, but also the error in the model.

Consider the following regression model where  $y$  represents the dependent variable,  $x_1$  and  $x_2$  are independent variables,  $\beta_0$  the intercept,  $\beta_1$  and  $\beta_2$  the regression coefficients of  $x_1$  and  $x_2$ , and  $\varepsilon$  the error term:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon, \quad (1)$$

Let us assume that the independent variable  $x_2$  is uncorrelated with  $\varepsilon$  (i.e.,  $x_2$  is exogenous), whereas  $x_1$  is endogenous since it is correlated with the error term  $\varepsilon$  (i.e.,  $Cov(x_1, \varepsilon) \neq 0$ ).

Hence, endogeneity occurs in this regression model example. Note that this terminology, which is used in the endogeneity literature (Wooldridge 2010), should not be confused with the terms

used for the explanation of PLS path models with exogenous (i.e., independent) and endogenous (i.e., dependent) constructs (Chin 1998). In the following, and in line with the literature, we refer to endogenous and exogenous to identify variables that endogeneity does (not) impact; we use dependent and independent to identify constructs that explain other constructs in the PLS path model, or are explained by them.

A straightforward way of handling, or at least reducing, endogeneity is to specify a set of control variables (Bernerth and Aguinis 2016) accounting for a part of the dependent variable's variance (Ebbes et al. 2017). While carefully chosen control variables often alleviate endogeneity problems substantially (Papies et al. 2016), the success of this endeavor depends on the researcher's ability to (1) identify relevant control variables, and (2) collect corresponding data. Furthermore, specifying a great number of control variables does not necessarily account for all of endogeneity in the model (Bascle 2008; Reeb et al. 2012) and contradicts the basic philosophy of Occam's razor (i.e., the principle of parsimony) upon which the scientific discipline rests.

Despite their careful selection of control variables, researchers also need to apply a statistical approach to address endogeneity, if there could be a potential endogeneity problem. Two broad statistical approaches have been developed to examine the presence of endogeneity: instrumental variable and instrumental variable-free approaches (Papies et al. 2016).

### **Instrumental Variable Approaches**

In econometrics and other areas of applied research, the instrumental variable (IV) approach is the technique of choice when dealing with endogeneity (e.g., Bound et al. 1995; Staiger and Stock 1997; Villas-Boas and Winer 1999). This approach uses information in additionally specified independent variables (called IVs) to decompose the endogenous independent

variables' variance into two parts (Bascle 2008): The exogenous part, which does not correlate with the error term, and the endogenous part, which still correlates with the error term. The estimation of the regression model then only uses the exogenous part, which endogeneity does not affect, to provide unbiased and consistent regression estimates. The IV approach is usually implemented by means of the 2SLS algorithm. This algorithm's first stage involves regressing an independent variable on an IV, followed by computing the predicted values from this equation and using them as input for re-estimating the original model in the second stage.

To illustrate the approach, consider the regression model example given in Equation 1. Let us assume that the independent variable  $x_1$  is endogenous, whereas  $x_2$  is exogenous. The first stage of the IV approach involves regressing the endogenous variable  $x_1$  on the exogenous variable  $x_2$  and the IV  $z$ :

$$x_1 = \gamma_0 + \gamma_1 z + \gamma_2 x_2 + \zeta. \quad (2)$$

Here,  $\gamma_0$  represents the intercept,  $\gamma_1$  the regression coefficient of  $z$ ,  $\gamma_2$  the regression coefficient of  $x_2$ , and  $\zeta$  the error term. Estimating Equation 2 yields predicted values of  $x_1$  (i.e.,  $\hat{x}_1$ ), which express the exogenous part of the independent variable  $x_1$ . In the second stage, we replace  $x_1$  in the original regression model (Equation 1) with  $\hat{x}_1$  to estimate the following:

$$y = \beta'_0 + \beta'_1 \hat{x}_1 + \beta'_2 x_2 + \varepsilon'. \quad (3)$$

Endogeneity occurs when there is a significant difference between  $\beta_1$  (Equation 1) and  $\beta'_1$  (Equation 3), with the latter representing the influence of the independent variable's exogenous part on  $y$ . To test for this difference, researchers use the Durbin-Wu-Hausman test (Hausman 1978). If the test indicates a significant difference between the  $\beta$  estimates, researchers should revert to the endogeneity-corrected estimates (Equation 3). If the Durbin-Wu-Hausman test does



not indicate a significant degree of endogeneity, researchers should use the non-corrected estimates from the original regression model (Equation 1).

To work effectively, the IVs must meet two criteria: the instrument relevance criterion and the exclusion restriction. First, according to the instrument relevance criterion, the correlation of each IV with the endogenous variable should be strong, which means that the IV's effect  $\gamma_1$  on the dependent variable  $x_1$  in Equation 2 must be pronounced and significant (Papies et al. 2016). If it is not, the corrected  $\beta$  coefficient estimates (Equation 3) may be considerably biased (Rossi 2014). But as Bound et al. (1993) note, "the cure can be worse than the disease." To further substantiate an IV's relevance and adequacy, researchers need to test whether its inclusion in the first stage regression (Equation 2) improves the model's  $R^2$  value significantly by referring to the  $F$ -statistic. As a rule of thumb, an  $F$ -statistic value of 10 and higher is generally viewed as indicative that the IV is adequate (e.g., Staiger and Stock 1997). Alternatively, Stock et al. (2002) provide a table of critical  $F$ -statistic values for IV models based on the number of instruments used. Note that in situations with only one endogenous variable and one IV, researchers must apply a multivariate  $F$ -test (e.g., Angrist and Pischke 2009) to test for the IVs' adequacy. In contrast, situations with multiple endogenous variables require considering at least the same number of IVs to ensure that the approach draws on an identified, or over-identified, model (Murray 2006b).

Second, according to the exclusion restriction for evaluating IVs when examining endogeneity, each IV should not be correlated with the error term in the original regression model (Murray 2006b). This requirement cannot be assessed statistically unless the researcher specifies more IVs than endogenous variables. In this case, the Sargan test and the Hansen's J test allow testing for significant correlations between the IVs and the error term (Bascle 2008; Wooldridge 2010).

Rejecting the null hypothesis of no correlation, however, does not help to verify *which* of the IVs are endogenous (Papiés et al. 2016). While this test provides some empirical substantiation when the first stage model is over-identified, researchers should primarily rely on theoretical arguments when assessing the IVs' adequacy (e.g., Murray 2006a).

Research has proposed the control function approach, a modification of the 2SLS method, which also leverages IVs to control for endogeneity. Specifically, the control function approach draws on the first stage regression (Equation 2) to obtain the fitted residuals ( $\hat{x}_1 - x_1$ ), which are then included in the original model to obtain correct(ed) estimates of the coefficients. Furthermore, to interpreting the significance of the fitted residuals requires running the Durbin-Wu-Hausman test (Papiés et al. 2016). That is, if a coefficient is significant, researchers should revert to the endogeneity-corrected estimates.

The control function approach is particularly advantageous when estimating regression models with discrete dependent variables (Petrin and Train 2010). However, such model types currently cannot be estimated in PLS-SEM since the standard algorithm requires the dependent constructs to be continuous (Lohmöller 1989, Chapter 4). For linear models as typically estimated in PLS-SEM, the control function approach yields exactly the same results as the 2SLS approach (Papiés et al. 2016).

### **Instrumental Variable-Free Approaches**

When considering the IV approach, identifying conceptually fitting IVs is very challenging (Rossi 2014), especially in light of the data constraints characterizing a typical research project. Research has suggested instrument-free approaches that do not require IVs to be specified as a solution. For example, Park and Gupta (2012) introduced the Gaussian copula approach, which

controls for endogeneity by directly modeling the correlation between the endogenous variable and the error term by means of a copula. Another instrument-free approach, proposed by Ebbes et al. (2005), is conceptually similar to the IV approach in that it splits an endogenous variable into an exogenous and an endogenous part. However, this latent instrumental variables (LIVs) approach does not draw on an observed variable, but generates an unobserved (i.e., latent) IV. Ebbes et al. (2005) assume that the latent instrument is discrete with at least two categories, separating the sample into  $L$  latent categories. As such, the LIV approach belongs to the class of normal mixture models with  $L$  components (Papies et al. 2016) whose model parameters can be identified for  $L > 1$  and non-normality of the endogenous regressor (Ebbes et al. 2005).

The LIV approach can be used in a PLS-SEM context, but a fundamental challenge in its application is the selection of the number of categories  $L$ . A priori, this number is unknown but has a substantial effect on the results. While Papies et al. (2016, p. 609) suggest that “the researcher fits at least  $L = 2, 3, 4$  and  $5$ ,” there is no clear guidance regarding the desired stability of the estimates in order to vary  $L$  and, hence, regarding this parameter’s final choice. At the same time, while the statistical software R facilitates the use of the Gaussian copula approach, current software support for the LIV approach is limited because models can only include one independent and one dependent variable (REndo package version 1.3). As these two issues limit the LIV approach’s applicability, we focus on the Gaussian Copula approach, which allows correcting for endogeneity in the absence of IVs.

Table 1 provides an overview of the advantages and disadvantages of (1) the control variable approach, (2) the IV approach, and (3) two types of IV-free approaches (i.e., Gaussian copula and LIV) when dealing with endogeneity in PLS-SEM.

=== Insert Table 1 about here ===

## ADDRESSING ENDOGENEITY IN PLS-SEM

### Procedure

PLS-SEM is a statistical method to estimate path models with latent variables (Lohmöller 1989, Chapter 2; Wold 1982). Unlike factor-based SEM, which considers constructs as common factors that explain the covariation between the associated indicators (Jöreskog 1978), PLS-SEM is a composite-based approach to SEM that uses linear combinations of indicator variables to represent conceptual variables in a statistical model (Henseler et al. 2014; Sarstedt et al. 2016). PLS-SEM estimates the parameters of a set of equations in a structural equation model by combining an approach similar to principal components analysis with regression-based path analysis. Specifically, the PLS-SEM method uses partial regressions to estimate the path coefficients between the latent variables and their indicators in the measurement models, as well as between the latent variables in the structural model; for instance, Lohmöller (1989, Chapter 2) and give an introduction to the PLS-SEM algorithm. Approaches for dealing with endogeneity in regression models can therefore be readily transferred to PLS-SEM. However, contrary to regression analysis, in which the approaches use the (observed) indicator variables as input, their implementation in PLS-SEM draws on the composite scores obtained after the algorithm's convergence (Lohmöller 1989, Chapter 2).<sup>2</sup> Figure 1 illustrates our systematic procedure for dealing with endogeneity in PLS-SEM. The following sections discuss each step in greater detail.

=== Insert Figure 1 about here ===

Stage 1 of the procedure requires checking if testing for endogeneity is necessary. Specifically, the meaningfulness of considering endogeneity in PLS-SEM depends on whether researchers use the method for predictive or explanatory modeling purposes. In predictive modeling, the purpose of the analysis is to predict new or future observations (Shmueli 2010). Here, the main concern is predicting the case values of the dependent variables, rather than (causally) interpreting the structural paths. Controlling for endogeneity is not useful in this case, as it may reduce the model's predictive power (Ebbes et al. 2005). Conversely, in explanatory modeling, which involves “the use of statistical models for testing causal explanations,” (Shmueli 2010, p. 290) controlling for endogeneity is crucial in order to adequately test hypotheses (Papies et al. 2016). In practice, however, researchers using PLS-SEM generally balance prediction and explanation in that they expect their model to have high predictive accuracy, while also being grounded in well-developed causal explanations (Sarstedt et al. 2017a). Gregor (2006; p. 626) refers to this interplay as explanation and prediction theory, noting that this approach “implies both understanding of underlying causes and prediction, as well as description of theoretical constructs and the relationships among them.” Such a setting *may* require the consideration of endogeneity, depending on whether explanation (i.e., testing theory) or prediction (e.g., to derive managerial recommendations) prevail in a certain research project.

Stage 2 of the procedure (Figure 1) uses existing theory and prior research findings to identify potential endogeneity issues a priori. If prior research adequately theorized, detected, and treated endogeneity issues in a specific research context, this knowledge should be applied when proposing a PLS path model and collecting data. This typically involves including control variables in the model, or gathering data on IVs for an ex post assessment, using the IV approach (e.g., Ebbes et al. 2017; Reeb et al. 2012). Thus, if prior research suggested that endogeneity is

an issue, or the results cast doubt that all the endogeneity sources have been detected, researchers should also test whether endogeneity substantially affects the model estimates.

To do so, Stage 3 of the procedure (Figure 1) applies the Gaussian copula approach. If the results do not indicate a problem, researchers can use the original model's estimates. Otherwise, the sources of endogeneity must be further explored and controlled for in the next step of the procedure. Since endogeneity problems in international marketing often arise from omitted variables (Jean et al. 2016), Stage 4 involves conceptually identifying the missing information and adding control variables to the PLS path model (Papies et al. 2016). Including suitable control variables handles, or at least reduces, the impact of endogeneity on the model estimates (Ebbes et al. 2017). In situations where no suitable control variables are available, researchers can draw on the IV approach. If IVs are unavailable, or researchers cannot identify at least one suitable IV for each endogenous variable in a partial regression model, they can use the Gaussian copula results, if available, and report that they cannot identify adequate omitted variables that explain the endogeneity problem. If the Gaussian copula results are not available, because they do not meet the method's requirements, researchers should interpret the original PLS path model estimates and specify the reason(s) for not addressing the potential endogeneity issues as a limitation (Figure 1).

### **The Gaussian copula approach in PLS-SEM**

To illustrate the Gaussian copula approach in PLS-SEM, consider a simple path model with one dependent latent variable  $Y_1$ , explained by two independent latent variables  $X_1$  and  $X_2$ . Each of the latent variables is measured using three indicators. Figure 2 shows this simple path model. The hexagons represent composites formed by linear combinations of the indicator variables, as indicated by the arrows linking the indicators and the composites. These arrows do not

necessarily correspond to the measurement model specification, which can be reflective or formative (Sarstedt et al. 2016). However, we use this illustration because PLS-SEM forms composites from the indicator variables to represent the conceptual variables in the model estimation.

=== Insert Figure 2 about here ===

Model estimation in PLS-SEM draws on a three-stage approach that belongs to the family of (alternating) least squares algorithms (Mateos-Aparicio 2011). Stage 1 of the PLS-SEM algorithm iteratively determines indicator weights, composite scores, and path coefficients. The algorithm converges when the indicator weights change only marginally from one iteration to the next. Stages 2 and 3 use the final composite scores from Stage 1 as input for a series of ordinary least squares regressions. These regressions produce the final indicator loadings, indicator weights, and path coefficients, as well as related elements, such as the indirect and total effects, and  $R^2$  values (e.g., Sarstedt et al. 2017a). With reference to our example, the final stage of the PLS-SEM algorithm entails regressing  $Y$  on  $X_1$  and  $X_2$  to estimate the structural model path coefficients  $\beta_1$  and  $\beta_2$ :

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \varepsilon \quad (4)$$

Let us assume that the independent variable  $X_1$  is thought to be correlated with the error term ( $\varepsilon$ ) of the dependent variable  $Y$ , and may therefore induce an endogeneity problem in the estimation of Equation 4.

The Gaussian copula approach requires the composite scores of the endogenous construct  $X_1$  to be non-normally distributed. If this requirement is met, the Gaussian copula of  $X_1$  can be calculated as follows (Papies et al. 2016):

$$c^* = \Phi^{-1}(H(X_1)), \quad (5)$$

where  $H(x)$  is the empirical cumulative density function and  $\Phi^{-1}$  is the inverse normal cumulative density function. Next, the copula is included as an additional independent variable that controls for the correlation between the error term and the endogenous independent construct in the regression model:

$$y = \beta_0 + \beta'_1 x_1 + \beta'_2 x_2 + \beta_3 c_1^* + \varepsilon' \quad (6)$$

To attend to the endogeneity problem, estimating Equation 6 yields the new path coefficients  $\beta'_1$  and  $\beta'_2$ .

To determine whether endogeneity is at a critical level, researchers need to assess the copula coefficient's ( $\beta_3$ ) significance using bootstrapping (Hausman 1978; Park and Gupta 2012). A significant coefficient indicates a critical level of endogeneity. If the copula coefficient is significant, researchers can assume that an endogeneity issue is present—even though treated by including the copula in the regression model (Equation 6). If the coefficient is not significant, no critical endogeneity issue affects the regression results.

While the standard Gaussian copula approach assumes a single endogenous variable, Park and Gupta (2012) also discuss the case of modeling multiple endogenous variables simultaneously. Doing so requires creating the copula for each individual endogenous variable and adding all copulas to the original model. The procedure separately computes all the Gaussian copulas of the specified endogenous variables (Park and Gupta 2012).



## **Control Variables in PLS-SEM**

If the Gaussian copula approach indicates a significant degree of endogeneity, researchers should use control variables to identify its sources. The introduction of control variables should always be grounded in strong theory (Bernerth and Aguinis 2016). That is, researchers should be explicit (rather than implicit) regarding the role of control variables in the analysis and potentially even match the hypotheses precisely to their choice (Spector and Brannick 2010). Typical examples of control variables used in international marketing models include a company's location, age, size, export experience, years of global experience, percentage of products sold overseas, number of countries entered, and cultural distance (e.g., Peng Cui et al. 2014; Shu et al. 2017).

The use of control variables in PLS-SEM is analogous to regression models. More precisely, control variables are included as single-item constructs in a PLS path model. Each control variable is then linked to the dependent latent variable whose predictor constructs have been identified (or are assumed to be) as possibly exhibiting endogeneity. Note that criticism of the use of single-item constructs (Diamantopoulos et al. 2012) does not apply to control variables, as they usually represent observable characteristics. However, if the control variables represent unobservable concepts (e.g., cultural dimensions), they need to be measured by means of multiple items. Finally, the interpretation of the control variables' effect is also analogous to regression models and depends on each control variable's measurement scale, which can be categorical or metric in PLS-SEM (Hair et al. 2017).

Once specified, researchers need to use bootstrapping to assess the control variables' significance. Similar to the analysis of interaction effects in PLS-SEM (Hair et al. 2017; Chapter 7), the assessment of significant control variables should use the  $f^2$  effect size to also consider

their relevance. According to Chin (1998)  $f^2$  values of 0.02, 0.15, and 0.35 suggest small, moderate, and substantial effects.

### **The Instrumental Variable Approach in PLS-SEM**

Implementing the IV approach in the context of PLS-SEM requires adjusting the final stage of the PLS-SEM algorithm. In its original form, this stage uses the composite scores obtained after convergence as input to run a series of ordinary least squares regressions that estimate the structural model path coefficients. Instead of using ordinary least squares regressions, however, this stage now uses 2SLS (Benitez et al. 2016). When implementing the IV approach in PLS-SEM, researchers need to ensure that they include at least one IV per endogenous variable.

Analogous to regression analysis, the IV approach's implementation requires researchers to find IVs that are highly correlated with the independent variables (instrument relevance), but are uncorrelated with the dependent variable's error term (exclusion restriction). In international marketing, previous regression-based models used IVs such as the organization quality, brand quality, environmental causes, brand strength (Swoboda and Hirschmann 2017), exposure to a multinational companies' marketing (Swoboda and Hirschmann 2016), and export memory (Souchon et al. 2016). However, the identification of suitable IVs is highly context-specific and the most challenging step in the analysis (e.g., Larcker and Rusticus 2010). Angrist and Pischke (2009, p. 17) argue that finding good IVs requires "a combination of institutional knowledge and ideas about processes determining the variable of interest," especially as there is no true test for their quality (Rossi 2014).

To illustrate the implementation of the IV approach in PLS-SEM, consider the simple model in Figure 2 with the construct  $X_I$  possibly exhibiting endogeneity. We now introduce an IV (i.e.,

$Z_I$ ), measured by three indicators, that correlates strongly with the construct  $X_I$ , but not with the error term of  $Y_I$  in the original model. Thereafter, we create and estimate the model shown in Figure 3.

=== Insert Figure 3 about here ===

Next, a 2SLS analysis uses the composite scores of  $X_I$  and  $X_2$ , obtained from the first model (Figure 2), and the composite scores of  $Z_I$ , derived from the second model (Figure 3).

Specifically, we regress  $X_I$  on  $Z_I$  and  $X_2$  to obtain  $\hat{X}_1$ :

$$\hat{X}_1 = \gamma_0 + \gamma_1 Z_1 + \gamma_2 X_2. \quad (7)$$

We first assess the IV's strength by comparing the  $R^2$  values of the first stage regression (Equation 7) with and without the IV. The results of an  $F$ -test indicate whether the  $R^2$  value of the regression model with and without the IVs differ significantly. If there are significant differences, we continue with the analysis by theoretically substantiating that  $Z_I$  is not correlated with the error term of  $Y_I$  (e.g., Semadeni et al. 2014; Staiger and Stock 1997). In the case of an over-identified model, statistical assessments by means of the Sargan test and the Hansen's J test may support this analysis.

Finally, we re-estimate the original model by regressing  $Y$  on  $X_2$  and  $\hat{X}_1$  from Equation 7 to obtain the corrected path coefficient estimates  $\beta'_1$  and  $\beta'_2$ :

$$Y = \beta'_0 + \beta'_1 \hat{X}_1 + \beta'_2 X_2 + \varepsilon'. \quad (8)$$

The Durbin-Wu-Hausman test (Hausman 1978) allows for testing whether  $\beta'_1$  (Equation 8) differs significantly from  $\beta_1$  (Equation 5). If so, researchers should revert to the IV approach's

coefficients (Equation 8). Otherwise, if the difference is not significant researchers should continue using the results of the partial regression model's original PLS-SEM estimation.

## **EMPIRICAL EXAMPLE TO ADDRESS ENDOGENEITY IN PLS-SEM**

### **Model and Data**

To illustrate our procedure for identifying and treating endogeneity (Figure 1), we draw on a corporate reputation model originally presented by Eberl (2010). The goal of this model is to explain the effects of competence (*COMP*) and likeability (*LIKE*), which represent corporate reputation's two dimensions (Sarstedt et al. 2013; Schwaiger 2004), on customer satisfaction (*CUSA*), and ultimately on customer loyalty (*CUSL*). This model has been used to illustrate aspects of the PLS-SEM method (e.g., Hair et al. 2017; Hair et al. 2018) and has also been frequently applied in an international marketing context.<sup>3</sup> For example, Zhang and Schwaiger (2012) use this model to study the differentiating effects of corporate social responsibility and other antecedent dimensions on corporate reputation in three Western countries (Germany, the USA, and the UK) and in China. Similarly, Radomir and Alan (2018) apply this model to contrast the effects of service quality and relationship investment on corporate reputation assessments of Romanian and UK consumers. Furthermore, Walsh et al. (2009) developed and validated a variant of this model using cross-country data, which features prominently in the international marketing literature. For example, several researchers have used this model to analyze the moderating effect of institutional country differences (Swoboda et al. 2016) and cultural dimensions (Swoboda and Hirschmann 2017) on the relationship between corporate reputation and customer loyalty.

Figure 4 shows the original model, which includes two partial regression models: (1) *CUSL* is regressed on *COMP*, *CUSA*, and *LIKE*, while (2) *CUSA* is regressed on *COMP* and *LIKE*. The following illustrations focus on the first, more complex, regression model. The measurement models of *COMP*, *LIKE*, and *CUSL* draw on three reflective items each, whereas *CUSA* is measured with a single item. The model estimation draws on data from two major German mobile communications network providers and two smaller competitors. A total of 344 respondents rated each item on a seven-point Likert scale. Observations with missing values were deleted, leaving a total sample size of 336. We used the statistical software R (R Core Team 2018), STATA (StataCorp 2015), and SmartPLS 3 (Ringle et al. 2015) for the computations. For the PLS-SEM estimations, we use Mode A for all composites (Sarstedt et al. 2017a). The results of the bootstrap routine with 5,000 samples, applying the no sign changes option, enable us to test for the coefficients' significance. We find that the measurement models meet all the relevant evaluation criteria. In the structural model assessment, all the path coefficient estimates, except *COMP* to *CUSL*, are significant and the model's  $R^2$  values are satisfactory (for this application, Hair et al. 2017 provide further details of the PLS-SEM results and their assessment).

=== Insert Figure 4 about here ===

Most of the corporate reputation model's analyses take a predictive perspective (e.g., Schwaiger 2004). However, in line with Eberl (2010), who elaborates on the causal effects of corporate reputation's two components (i.e., *COMP* and *LIKE*) on *CUSA* and *CUSL*, we consider our analysis primarily explanatory in nature. Hence, following Stage 1 of the procedure (Figure 1), we continue by applying the Gaussian copula approach.

## **Application of the Gaussian copula approach in PLS-SEM**

Before initiating the Gaussian copula approach, to meet its assumptions, we first verify if the variables, which potentially exhibit endogeneity, are non-normally distributed (Figure 1). We do so by running the Kolmogorov–Smirnov test with Lilliefors correction (Sarstedt and Mooi 2014) on the standardized composite scores of *COMP*, *LIKE*, and *CUSA*, which the PLS path model estimation provides. The results show that none of the constructs has normally distributed scores, which allows us to consider them endogenous in the Gaussian copula analysis.

Next, we create three regression models in which we consider the independent constructs *COMP* (model 1), *LIKE* (model 2), and *CUSA* (model 3) as possibly exhibiting endogeneity. In addition, to simultaneously account for multiple endogenous variables when applying the Gaussian copula approach, we create four regression models that include all possible combinations of multiple endogenous variables in this PLS path model example: *LIKE* and *COMP* (model 4), *LIKE* and *CUSA* (model 5), *COMP* and *CUSA* (model 6), and *LIKE*, *COMP*, and *CUSA* (model 7).

We use the constructs' standardized composite scores to compute the Gaussian copula of the partial regressions in the structural model. To cross-validate the results, we used two different statistical software packages, which returned the same results: STATA 14 (StataCorp 2015) and the REndo package of the R program (Gui et al. 2017). To implement STATA, we used specific commands (1) to compute the Gaussian cumulative distribution function, (2) to create the copula as the inverse of the Gaussian normal cumulative distribution function, and (3) to calculate the model that includes the Gaussian copula. Since the REndo package does not currently support the computation bootstrap standard errors, we used our own R code and the boot package (Canty and Ripley 2017) to run this part of the analyses.

=== Insert Table 2 about here ===

The results in Table 2 show that only one Gaussian copula (i.e.,  $c_{CUSA}$ ) is significant ( $p < 0.1$ ) when treating one endogenous variable, which points to a potential endogeneity issue. Including the significant Gaussian copula in the model changes the effect of *CUSA* on *CUSL* by 0.073 units (from 0.509 to 0.582), which points to a potential endogeneity problem for *CUSA* (model 3). Similarly,  $c_{CUSA}$  is also significant in the *CUSA* models in combination with *LIKE* (model 5,  $p < 0.1$ ), *COMP* (model 6,  $p < 0.05$ ), and *LIKE* and *COMP* (model 7,  $p < 0.1$ ). This confirms the possibility of *CUSA* being endogenous.<sup>4</sup>

### **Application of the control variable approach in PLS-SEM**

Following Stage 3 of the procedure for identifying and treating endogeneity (Figure 1), we next include control variables in the model, which from a theoretical perspective, are likely to influence customers' loyalty (Eberl 2010). The first control variable (*serviceprovider*) distinguishes between large, globally operating mobile phone services providers (*serviceprovider*=1;  $n_{(1)} = 251$ ) and smaller, locally operating competitors (*serviceprovider*=2;  $n_{(2)} = 85$ ). In addition, we include a second control variable (*servicetype*), which is relevant from a service marketing perspective and distinguishes between customers with a contract and those with a prepaid plan.

Analyzing the impact of the two control variables, both separately and jointly, shows that neither has a significant influence on the *CUSL* construct. Specifically, the *servicetype* variable has a -0.003 effect (sig.=0.917;  $f^2=0.000$ ), and the *serviceprovider* variable has a -0.038 effect (sig.=0.319;  $f^2=0.003$ ) on *CUSL* when included separately. Including both control variables simultaneously changes the results only marginally. As none of the control variables identifies a

source of endogeneity, we continue with the procedure (Figure 1) by running the IV approach on the data.

### **Application of the IV approach in PLS-SEM**

The application of the IV approach in the context of our model requires selecting an IV that is correlated with *COMP*, *CUSA*, and *LIKE*. At the same time, the IV needs to be uncorrelated with the omitted determinants of customer loyalty and, hence, be uncorrelated with *CUSL*'s error term. Our implementation of the IV approach uses customers' perceived switching costs as an IV. Switching costs comprise the time, money, and effort that consumers expect to expend when changing the service provider. In the mobile phone market, switching costs come primarily in the form of transfer costs (Klemperer 1987) when ending a relationship with one provider and starting a new one with another. Specifically, the implementation and configuration of the new relationship require the customer to invest time, effort, and other assets, while ending an old service may also require procedural tasks (Ray et al. 2012). Hence, in terms of the instrument relevance criterion, it is reasonable to assume that the perceived switching costs correlate with the customers' satisfaction and reputation assessments, because significant transfer costs limit their freedom to configure the business relationship. Given the manifold factors that influence loyalty, arguing that perceived switching costs are uncorrelated with omitted variables that affect customer loyalty (exclusion restriction) is more difficult. However, research has identified customers' variety seeking tendency—that is, the “tendency of individuals to seek diversity in their choices of services or goods” (Kahn 1995, p. 139)—as a major determinant of customer loyalty (e.g., Berné et al. 2001). Since this tendency is a psychological trait that is independent of individuals' preference structures and is stable over time (Menon and Kahn 1995), variety seeking is probably unrelated to customers' perception of the switching costs. Network quality is



another key determinant of customer loyalty that is relevant for the context of our study (e.g., in terms of coverage, stability, and voice quality; Lim et al. 2006). Network quality can differ substantially, especially in rural areas, and cannot be fully observed a priori. Poor network coverage will have a strong negative effect on customer loyalty, independent of the customers' perception of the switching costs.<sup>5</sup>

We use an extended version of Jones et al.'s (2000) scale and measure the switching costs by using four reflective indicators (*switch<sub>1</sub>* to *switch<sub>4</sub>*), each measured on a five-point Likert scale. For our analysis, we extract all (standardized) composite scores from the SmartPLS 3 software. We subsequently run the IV approach by using STATA's *ivregress* method (StataCorp 2017), as well as the statistical software R with its AER package (Kleiber and Zeileis 2017), to validate the results. Both programs provide the same results. Table 3 shows the results of the analysis for the IVs.

To assess if the IV is strong (Figure 1), we use the weakID test's *F*-value, which compares the  $R^2$  values of the first stage regression with and without the IV. We consider three models, each of which considers another predictor construct of *CUSL* as likely to exhibit endogeneity (i.e., *LIKE*, *COMP*, and *CUSA*). The results in Table 3 show that including the IV significantly increases the 1<sup>st</sup> stage  $R^2$  value when considering *COMP* and *LIKE* as endogenous, as evidenced by the weakID test values well above 10. Conversely, the model's  $R^2$  value only increases marginally when considering *CUSA* as likely to exhibit endogeneity. Hence, in the following step, we use the Wu-Hausman test to assess whether the coefficients of the IV approach differ significantly from the original coefficients. The results in Table 3 show that, in our analysis, the original and IV coefficients do not differ significantly. When considering *LIKE* and *COMP* as endogenous, the Wu-Hausman test specifically returns *p*-values of 0.216.<sup>6</sup>

=== Insert Table 3 about here ===

Including switching costs as an IV does not explain the endogeneity problem's origin. More specifically, the IV correction of *CUSA* increases the corresponding point estimate to above one, further indicating that the switching costs construct is not a suitable IV.

Following the procedure in Figure 3, we should therefore examine the results of the Gaussian copula approach. Specifically, researchers should reinterpret the results that assume *CUSA* is endogenous, since including *CUSA*'s significant Gaussian copula in the model changes the relationship of *CUSA* on *CUSL* from 0.509 to 0.582 (Table 3). This finding is in line with prior research, which has theoretically and empirically substantiated endogeneity problems with customer satisfaction, especially its relationship to customer loyalty (Ivanov et al. 2013). Nevertheless, this analysis also illustrates that, in the context of our empirical example, considering endogeneity does not, from a theory testing perspective, change any conclusions drawn from the model estimation.

To summarize, by carrying out the procedure shown in Figure 1, we identified a potential endogeneity issue based on the Gaussian Copula results. However, we have not been able to explain the problem by the available control and instrumental variables. Thus, in this kind of situation the researcher should report the Gaussian Copula approach's solution, acknowledging that the findings do not reveal the origins of the endogeneity.

## **SUMMARY AND CONCLUSIONS**

While addressing endogeneity has become an integral component of regression-based analyses in international marketing (Jean et al. 2016), users of the PLS-SEM method have overlooked this issue. A potential reason for this disregard is the predictive nature of PLS-SEM, which, if taken

into full consideration when using the method, renders endogeneity irrelevant (Ebbes et al. 2011). However, researchers conventionally use PLS-SEM to test formal hypotheses, which corresponds to an explanation perspective, while simultaneously deriving managerial implications, which follows a prediction perspective. Endogeneity becomes an important concern, therefore, depending on whether researchers emphasize explanation or prediction in their analyses.

Similar to scholars who have applied PLS-SEM in international marketing and other disciplines, methodological researchers have also failed to consider endogeneity in PLS-SEM, with some authors suggesting that endogeneity cannot be accounted for in a PLS-SEM framework (e.g., Antonakis et al. 2010; Rönkkö and Evermann 2013). Such notions are astonishing, since PLS-SEM is a regression-based technique, which makes common approaches for identifying and treating endogeneity accessible in this specific methodological context. Benitez et al. (2016) took a first step in this direction by combining the standard PLS-SEM algorithm with the 2SLS method without, however, considering actual IVs that allow for decomposing the independent variables' variance into endogenous and exogenous parts. Broadening the scope of their research, this study introduces a systematic procedure for identifying and treating endogeneity in PLS-SEM that takes into account the most recent statistical developments in the literature on endogeneity. Specifically, implementing the Gaussian copula approach addresses the significant concerns associated with identifying strong and valid IVs. Moreover, combining this novel approach with the inclusion of control variables and the classic IV approach enables a comprehensive assessment of endogeneity when applying PLS-SEM for explanatory purposes.

Our application of the procedure to empirical data discloses endogeneity in Eberl's (2010) corporate reputation model, which has been frequently used in different variations in

international marketing and international business research (e.g., Swoboda et al. 2016; Zhang and Schwaiger 2012). Our analysis suggests that the *CUSA* construct is subject to endogeneity but control variables or IVs were not able to handle this problem. Correcting for endogeneity using the Gaussian copula approach leads to a pronounced change in the model estimates, while nevertheless leaving the general conclusions intact.

We hope this article will be the starting point for future PLS-SEM research on endogeneity.

Researchers may endeavor to introduce bootstrapping-based approaches to test the strength of an IV, its validity, and relevance in PLS-SEM. Further, the weakID test and Wu-Hausman test are parametric tests and therefore violate the nonparametric nature of PLS-SEM, thus requiring other tests to be identified. Moreover, further research should advance the understanding and applicability of IV-free approaches, such as Gaussian copulas or Ebbes et al.'s (2005) LIV method. For example, simulation study results could provide researchers with thresholds for deciding if a significant Gaussian copula entails a rather negligible, or a substantial, endogeneity problem. Future research should also identify ways to parametrize the LIV approach, which our study did not consider due to its strong dependency on the selection of the number of latent instrument categories (Papiés et al. 2016) and lack of software support. Such a step would open ways for a benchmarking study of the Gaussian copula and LIV approaches in the context of PLS-SEM. In these regards, a Monte Carlo simulation study, in which prespecified levels of endogeneity bias the parameter estimates, would allow for comparing the IV and IV-free approaches' efficacy regarding identifying and treating endogeneity.

Recent PLS-SEM developments, such as the consistent PLS-SEM approach (Dijkstra and Henseler 2015), which allows for mimicking factor-based SEM results, and the introduction of model fit criteria into PLS-SEM (e.g., Henseler et al. 2014) may help researchers with specific

explanatory research goals. The next logical step is to extend PLS-SEM's capabilities to address endogeneity issues in empirical international marketing research when using consistent PLS-SEM to estimate models on the grounds of common factor model data (Rigdon et al. 2017; Sarstedt et al. 2016). This kind of advancement should be straightforward in terms of our research findings, since the consistent PLS-SEM method uses the composite scores obtained from a standard PLS-SEM-based estimation as input.

Future research should also explore other approaches to detect endogeneity problems in PLS-SEM. Research has brought forward a number of latent class techniques such as finite mixture PLS (Hair et al. 2018), PLS prediction oriented segmentation (Becker et al. 2013), and PLS iterative reweighted regression segmentation (Schlittgen et al. 2016) to uncover unobserved heterogeneity. Sarstedt et al. (2017b) recently introduced a systematic procedure for identifying and treating unobserved heterogeneity in PLS path models. Combining several latent class techniques, the procedure enables researchers to (1) check if unobserved heterogeneity has a critical impact on the results, (2) decide on the number of latent segments to retain from the data, and (3) recover the segment-specific model estimates. However, future research needs to assess the interplay between the various sources of heterogeneity on the one hand and endogeneity on the other and develop guidelines to assess if endogeneity is at a non-critical level in the segment-specific solutions.

It would be particularly interesting to also consider the fuzzy-set qualitative comparative analysis of regression models (fsQCA; Woodside 2013). A joint consideration of these techniques may allow researchers to better assess and control for endogeneity embedded in unobserved heterogeneity (Ivanov et al. 2013; Schlittgen et al. 2016). Moreover, in this study we did not address how to deal with endogeneity issues in PLS path models that include interaction terms,

which are used, for example, in PLS-SEM-based moderator analyses (Hair et al. 2017; Henseler and Chin 2010). As international marketing researchers frequently apply such analyses (e.g., Jin et al. 2016; Leonidou et al. 2017), future studies should use insights from research on regression analysis (e.g., by using special 2SLS analyses to account for interaction terms and IVs or the control function approach; see, for example, Sande and Ghosh 2018) to furnish recommendations on how to adequately address this issue in PLS-SEM.

Finally, in its standard form PLS-SEM does not support estimating longitudinal or panel data without limitations (Roemer 2016). Once the PLS-SEM method advances in these directions, assessing the efficacy of other approaches for handling endogeneity, such as the lagged variable approach (Jean et al. 2016), the cross-lagged correlation procedure (Tyagi and Wotruba 1993), the Arellano–Bond (1991) estimator (Kumar and Pansari 2016), and the Granger causality test (Filipescu et al. 2013), would be particularly interesting. Relatedly, extending the PLS-SEM algorithm to accommodate binary dependent variables would provide an opportunity to consider endogeneity methods specific to such model settings, such as the control function approach and its extensions (e.g., Garen 1984; Wooldridge 2015). These kinds of advances would be particularly beneficial for international marketing research, which frequently relies on longitudinal and panel data (e.g., Kumar and Pansari 2016), as well as binary outcome variables (e.g., Obadia et al. 2017; Westjohn et al. 2017).

## **REFERENCES**

Ali, Faizan, S. Mostafa Rasoolimanesh, Marko Sarstedt, Christian M. Ringle, and Kisang Ryu (2018), "An Assessment of the Use of Partial Least Squares Structural Equation Modeling (PLS-

SEM) in Hospitality Research," *International Journal of Contemporary Hospitality Management*, 30 (1), 514-38.

Angrist, Joshua D. and Jörn-Steffen Pischke (2009), *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, NY: Princeton University Press.

Antonakis, John, Samuel Bendahan, Philippe Jacquart, and Rafael Lalive (2010), "On Making Causal Claims: A Review and Recommendations," *The Leadership Quarterly*, 21 (6), 1086-120.

Arellano, Manuel and Stephen Bond (1991), "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equation," *The Review of Economic Studies*, 58 (2), 277-97.

Bascle, Guilhem (2008), "Controlling for Endogeneity with Instrumental Variables in Strategic Management Research," *Strategic Organization*, 6 (3), 285-327.

Becker, Jan-Michael, Arun Rai, Christian M. Ringle, and Franziska Völckner (2013), "Discovering Unobserved Heterogeneity in Structural Equation Models to Avert Validity Threats," *MIS Quarterly*, 37 (3), 665-94.

Benitez, Jose, Jörg Henseler, and José L. Roldán (2016), "How to Address Endogeneity in Partial Least Squares Path Modeling," in *22nd Americas Conference on Information Systems (ACIS)*. San Diego, CA.

Berné, Carmen, José M. Múgica, and M. Jesús Yagüe (2001), "The Effect of Variety-seeking on Customer Retention in Services," *Journal of Retailing and Consumer Services*, 8 (6), 335-45.

Bernerth, Jeremy B. and Herman Aguinis (2016), "A Critical Review and Best-Practice Recommendations for Control Variable Usage," *Personnel Psychology*, 69 (1), 229-83.

Bijmolt, Tammo H. A., Harald J. van Heerde, and Rik G. M. Pieters (2005), "New Empirical Generalizations on the Determinants of Price Elasticity," *Journal of Marketing Research*, 42 (2), 141-56.

Bollen, Kenneth A., Stanislav Kolenikov, and Shawn Bauldry (2014), "Model-Implied Instrumental Variable—Generalized Method of Moments (MIIV-GMM) Estimators for Latent Variable Models," *Psychometrika*, 79 (1), 20-50.

Bound, John, David A. Jaeger, and Regina M. Baker (1993), "The Cure Can Be Worse than the Disease: A Cautionary Tale Regarding Instrumental Variables," in NBER Technical Working Paper No. 137. Cambridge, MA: National Bureau of Economic Research.

---- (1995), "Problems with Instrumental Variables Estimation: When the Correlation Between the Instruments and the Endogeneous Explanatory Variable is Weak," *Journal of the American Statistical Association*, 90 (430), 443-50.

Canty, Angelo and Brian Ripley (2017), "R Package Boot: Bootstrap Functions (Version 1.3-20)," <https://cran.r-project.org/web/packages/boot/>.

Chin, Wynne W. (1998), "The Partial Least Squares Approach to Structural Equation Modeling," in *Modern Methods for Business Research*, George A. Marcoulides, ed. Mahwah: Erlbaum.

Clougherty, Joseph A., Tomaso Duso, and Johannes Muck (2016), "Correcting for Self-selection Based Endogeneity in Management Research: Review, Recommendations and Simulations," *Organizational Research Methods*, 19 (2), 286-347.

Deng, Ziliang (2016), "Endogeneity in International Marketing Research, 1995-2014," (accessed 20 March 2018, [available at [http://en.paper.edu.cn/en\\_releasepaper/content/4687265](http://en.paper.edu.cn/en_releasepaper/content/4687265)]).



Diamantopoulos, Adamantios, Marko Sarstedt, Christoph Fuchs, Petra Wilczynski, and Sebastian Kaiser (2012), "Guidelines for choosing between multi-item and single-item scales for construct measurement: a predictive validity perspective," *Journal of the Academy of Marketing Science*, 40 (3), 434-49.

Dijkstra, Theo K. and Jörg Henseler (2015), "Consistent Partial Least Squares Path Modeling," *MIS Quarterly*, 39 (2), 297-316.

Ebbes, P., D. Papies, and H. J. van Heerde (2017), "Dealing with Endogeneity: A Nontechnical Guide for Marketing Researchers," in *Handbook of Market Research*, Christian Homburg and Martin Klarmann and Arnd Vomberg, eds. Cham: Springer International Publishing.

Ebbes, Peter, Dominik Papies, and Harald J. van Heerde (2011), "The Sense and Non-Sense of Holdout Sample Validation in the Presence of Endogeneity," *Marketing Science*, 30 (6), 1115-22.

Ebbes, Peter, Michel Wedel, Ulf Böckenholt, and Ton Steerneman (2005), "Solving and Testing for Regressor-Error (in)Dependence When no Instrumental Variables are Available: With New Evidence for the Effect of Education on Income," *Quantitative Marketing and Economics*, 3 (4), 365-92.

Eberl, Markus (2010), "An Application of PLS in Multi-Group Analysis: The Need for Differentiated Corporate-Level Marketing in the Mobile Communications Industry," in *Handbook of Partial Least Squares: Concepts, Methods and Applications (Springer Handbooks of Computational Statistics Series, vol. II)*, Vincenzo Esposito Vinzi and Wynne W. Chin and Jörg Henseler and Huiwen Wang, eds. Heidelberg, Dordrecht, London, New York: Springer.

Filipescu, Diana A., Shameen Prashantham, Alex Rialp, and Josep Rialp (2013), "Technological Innovation and Exports: Unpacking Their Reciprocal Causality," *Journal of International Marketing*, 21 (1), 23-38.

Garen, John (1984), "The Returns to Schooling: A Selectivity Bias Approach with a Continuous Choice Variable " *Econometrica*, 52 (5), 1199-218.

Gregor, Shirley (2006), "The Nature of Theory in Information Systems," *MIS Quarterly*, 30 (3), 611-42.

Gui, Raluca, Markus Meierer, and Rene Algesheimer (2017), "R Package REndo: Fitting Linear Models with Endogenous Regressors using Latent Instrumental Variables (Version 1.2)," <https://cran.r-project.org/web/packages/REndo/>.

Hair, Joseph F., G. Tomas M. Hult, Christian M. Ringle, and Marko Sarstedt (2017), *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)* (2 ed.). Thousand Oaks, CA: Sage.

Hair, Joseph F., Marko Sarstedt, Christian M. Ringle, and Siegfried P. Gudergan (2018), *Advanced Issues in Partial Least Squares Structural Equation Modeling (PLS-SEM)*. Thousand Oaks, CA: Sage.

Hausman, Jerry A. (1978), "Specification Tests in Econometrics," *Econometrica*, 46 (6), 1251-71.

Henseler, Jörg and Wynne W. Chin (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, Jörg, Theo K. Dijkstra, Marko Sarstedt, Christian M. Ringle, Adamantios Diamantopoulos, Detmar W. Straub, David J. Ketchen, Joseph F. Hair, G. Tomas M. Hult, and Roger J. Calantone (2014), "Common Beliefs and Reality about Partial Least Squares: Comments on Rönkkö & Evermann (2013)," *Organizational Research Methods*, 17 (2), 182-209.

Hult, G. Tomas M., David J. Ketchen, David A. Griffith, Brian R. Chabowski, Mary K. Hamman, Bernadine Johnson Dykes, Wesley A. Pollitte, and S. Tamer Cavusgil (2008), "An Assessment of the Measurement of Performance in International Business Research," *Journal of International Business Studies*, 39 (6), 1064-80.

Hult, G. Tomas M., Martin Reimann, and Oliver Schilke (2009), "Worldwide Faculty Perceptions of Marketing Journals: Rankings, Trends, Comparisons, and Segmentations," *globalEDGE Business Review*, 3 (3), 1-10.

Ivanov, Vladimir, Kissan Joseph, and M. Babajide Wintoki (2013), "Disentangling the Market Value of Customer Satisfaction: Evidence from Market Reaction to the Unanticipated Component of ACSI Announcements," *International Journal of Research in Marketing*, 30 (2), 168-78.

Jean, Ruey-Jer "Bryan", Ziliang Deng, Daekwan Kim, and Xiaohui Yuan (2016), "Assessing Endogeneity Issues in International Marketing Research," *International Marketing Review*, 33 (3), 483-512.

Jin, Jason Lu, Kevin Zheng Zhou, and Yonggui Wang (2016), "Exploitation and Exploration in International Joint Ventures: Moderating Effects of Partner Control Imbalance and Product Similarity," *Journal of International Marketing*, 24 (4), 20-38.

Jones, Michael A., David L. Mothersbaugh, and Sharon E. Beatty (2000), "Switching barriers and repurchase intentions in services," *Journal of Retailing*, 76 (2), 259-74.

Jöreskog, Karl G. (1978), "Structural Analysis of Covariance and Correlation Matrices," *Psychometrika*, 43 (4), 443-77.

Kahn, Barbara E. (1995), "Consumer Variety-seeking Among Goods and Services: An Integrative Review," *Journal of Retailing and Consumer Services*, 2 (3), 139-48.

Kirby, James B. and Kenneth A. Bollen (2009), "Using Instrumental Variable Tests to Evaluate Model Specification in Latent Variable Structural Equation Models," *Sociological Methodology*, 39, 327-55.

Kleiber, Christian and Achim Zeileis (2017), "R Package AER: Applied Econometrics with R (Version 1.2-5)," <https://cran.r-project.org/web/packages/AER/>.

Klemperer, Paul (1987), "Markets with Consumer Switching Costs," *The Quarterly Journal of Economics*, 102 (2), 375-94.

Kumar, V. and Anita Pansari (2016), "National Culture, Economy, and Customer Lifetime Value: Assessing the Relative Impact of the Drivers of Customer Lifetime Value for a Global Retailer," *Journal of International Marketing*, 24 (1), 1-110.

Larcker, David F. and Tjomme O. Rusticus (2010), "On the Use of Instrumental Variables in Accounting Research," *Journal of Accounting and Economics*, 49 (3), 186-205.

Leonidou, Leonidas C., Bilge Aykol, Thomas A. Fotiadis, and Paul Christodoulides (2017), "Antecedents and Consequences of Infidelity in Cross-Border Business Relationships," *Journal of International Marketing*, 25 (1), 46-71.

Lim, Heejin, Richard Widdows, and Jungkun Park (2006), "M-loyalty: Winning Strategies for Mobile Carriers," *Journal of Consumer Marketing*, 23 (4), 208-18.

Liu, Qing, Thomas Otter, and Greg M. Allenby (2007), "Investigating Endogeneity Bias in Marketing," *Marketing Science*, 26 (5), 642-50.

Lohmöller, Jan-Bernd (1989), *Latent Variable Path Modeling with Partial Least Squares*. Heidelberg: Physica.

Mateos-Aparicio, Gregoria (2011), "Partial Least Squares (PLS) Methods: Origins, Evolution, and Application to Social Sciences," *Communications in Statistics - Theory and Methods* 40 (13), 2305-17.

McIntosh, Cameron. N., Jeffery R. Edwards, and John Antonakis (2014), "Reflections on Partial Least Squares Path Modeling," *Organizational Research Methods*, 17 (2), 210-51.

Menon, Satya and Barbara E. Kahn (1995), "The Impact of Context on Variety Seeking in Product Choices," *Journal of Consumer Research*, 22 (3), 285-95.

Murray, Michael P. (2006a), "Avoiding Invalid Instruments and Coping with Weak Instruments," *The Journal of Economic Perspectives*, 20 (4), 111-32.

---- (2006b), "The Bad, the Weak, and the Ugly: Avoiding the Pitfalls of Instrumental Variables Estimation." SSRN: <http://dx.doi.org/10.2139/ssrn.843185>.

Nakata, Cheryl and Yili Huang (2005), "Progress and Promise: The Last Decade of International Marketing Research," *Journal of Business Research*, 58 (5), 611-18.

Nitzl, Christian (2016), "The Use of Partial Least Squares Structural Equation Modelling (PLS-SEM) in Management Accounting Research: Directions for Future Theory Development," *Journal of Accounting Literature*, 37 (December), 19-35.

Obadia, Claude, Irena Vida, and José Pla-Barber (2017), "Differential Effects of Bilateral Norms on SMEs' Export Relationships: A Dynamic Perspective," *Journal of International Marketing*, 25 (3), 21-41.

Papies, Dominik, Peter Ebbes, and Harald J. van Heerde (2016), "Addressing Endogeneity in Marketing Models," in *Advanced Methods in Modeling Markets*, Peter S.H. Leeflang and Jaap E. Wieringa and Tammo H.A. Bijmolt and Koen H. Pauwels, eds.

Park, Sungho and Sachin Gupta (2012), "Handling Endogenous Regressors by Joint Estimation Using Copulas," *Marketing Science*, 31 (4), 567-86.

Peng Cui, Annie, Michael F. Walsh, and Zou Shaoming (2014), "The Importance of Strategic Fit Between Host: Home Country Similarity and Exploration Exploitation Strategies on Small and Medium-Sized Enterprises' Performance: A Contingency Perspective," *Journal of International Marketing*, 22 (4), 67-85.

Petrin, Amil and Kenneth Train (2010), "A Control Function Approach to Endogeneity in Consumer Choice Models," *Journal of Marketing Research*, 47 (1), 3-13.

Qiu, Hanqin, Ben Haobin Ye, Billy Bai, and Wei Heng Wang (2015), "Do the Roles of Switching Barriers on Customer Loyalty Vary for Different Types of Hotels?," *International Journal of Hospitality Management*, 46, 89-98.

R Core Team (2018), "R: A Language and Environment for Statistical Computing." Vienna, Austria: R Foundation for Statistical Computing (<https://www.R-project.org>).

Radomir, Lacramioara and Wilson Alan (2018), "Corporate Reputation: The Importance of Service Quality and Relationship Investment," in *Partial Least Squares Structural Equation*

*Modeling: Recent Advances in Banking and Finance*, Necmi K. Avkiran and Christian M. Ringle, eds. Heidelberg: Springer.

Raithel, Sascha and Manfred Schwaiger (2015), "The Effects of Corporate Reputation Perceptions of the General Public on Shareholder Value," *Strategic Management Journal*, 36 (6), 945-56.

Ray, Soumya, Sung S. Kim, and James G. Morris (2012), "Research Note: Online Users' Switching Costs: Their Nature and Formation," *Information Systems Research*, 23 (1), 197-213.

Reeb, David, Mariko Sakakibara, and Ishtiaq P. Mahmood (2012), "From the Editors: Endogeneity in International Business Research," *Journal of International Business Studies*, 43 (3), 211-18.

Richter, Nicole F., Rudolf R. Sinkovics, Christian M. Ringle, and Christopher Schlägel (2016), "A Critical Look at the Use of SEM in International Business Research " *International Marketing Review*, 33 (3), 376-404.

Rigdon, Edward E., Marko Sarstedt, and Christian M. Ringle (2017), "On Comparing Results from CB-SEM and PLS-SEM. Five Perspectives and Five Recommendations," *Marketing ZFP*, 39 (3), 4-16.

Ringle, Christian M., Marko Sarstedt, Rebecca Mitchell, and Siegfried P. Gudergan (2018), "Partial Least Squares Structural Equation Modeling in HRM Research," *The International Journal of Human Resource Management*, forthcoming.

Ringle, Christian M., Sven Wende, and Jan-Michael Becker (2015), "SmartPLS 3." Bönningstedt: SmartPLS.

Roemer, Ellen (2016), "A Tutorial on the Use of PLS Path Modeling in Longitudinal Studies," *Industrial Management & Data Systems*, 116 (9), 1901-21.

Rönkkö, Mikko and Joerg Evermann (2013), "A Critical Examination of Common Beliefs About Partial Least Squares Path Modeling," *Organizational Research Methods*, 16 (3), 425-48.

Rossi, Peter E. (2014), "Even the Rich Can Make Themselves Poor: A Critical Examination of IV Methods in Marketing Applications," *Marketing Science*, 33 (5), 655-72.

Sande, Jon Bingen and Mrinal Ghosh (2018), "Endogeneity in Survey Research," *International Journal of Research in Marketing*, forthcoming.

Sarstedt, Marko, Joseph F. Hair, Christian M. Ringle, Kai O. Thiele, and Siegfried P. Gudergan (2016), "Estimation Issues with PLS and CBSEM: Where the Bias Lies!," *Journal of Business Research*, 69 (10), 3998-4010.

Sarstedt, Marko and Erik A. Mooi (2014), *A Concise Guide to Market Research: The Process, Data, and Methods Using IBM SPSS Statistics*. Heidelberg et al.: Springer.

Sarstedt, Marko, Christian M. Ringle, and Joseph F. Hair (2017a), "Partial Least Squares Structural Equation Modeling," in *Handbook of Market Research*, Christian Homburg and Martin Klarmann and Arndt Vomberg, eds. Heidelberg: Springer.

---- (2017b), "Treating Unobserved Heterogeneity in PLS-SEM: A Multi-Method Approach," in *Partial Least Squares Structural Equation Modeling: Basic Concepts, Methodological Issues and Applications*, Richard Noonan and Hengky Latan, eds. Heidelberg: Springer.

Sarstedt, Marko, Petra Wilczynski, and T. C. Melewar (2013), "Measuring Reputation in Global Markets: A Comparison of Reputation Measures' Convergent and Criterion Validities " *Journal of World Business*, 48 (3), 329-39.



Schlittgen, Rainer, Christian M. Ringle, Marko Sarstedt, and Jan-Michael Becker (2016), "Segmentation of PLS Path Models by Iterative Reweighted Regressions," *Journal of Business Research*, 69 (10), 4583-92.

Schwaiger, Manfred (2004), "Components and Parameters of Corporate Reputation: An Empirical Study," *Schmalenbach Business Review*, 56 (1), 46-71.

Semadeni, Matthew, Michael C. Withers, and S. Trevis Certo (2014), "The Perils of Endogeneity and Instrumental Variables in Strategy Research: Understanding through Simulations," *Strategic Management Journal*, 35 (7), 1070-79.

Shmueli, Galit (2010), "To Explain or to Predict?," *Statistical Science*, 25 (3), 289-310.

Shu, Chengli, Jason Lu Jin, and Kevin Zheng Zhou (2017), "A Contingent View of Partner Coopetition in International Joint Ventures," *Journal of International Marketing*, 25 (3), 42-60.

Souchon, Anne L., Paul Hughes, Andrew M. Farrell, Ekaterina Nemkova, and João S. Oliveira (2016), "Spontaneity and International Marketing Performance," *International Marketing Review*, 33 (5), 671-90.

Spector, Paul E. and Michael T. Brannick (2010), "Common Method Issues: An Introduction to the Feature Topic in Organizational Research Methods," *Organizational Research Methods*, 13 (3), 403-06.

Staiger, Douglas and James H. Stock (1997), "Instrumental Variables Regression with Weak Instruments," *Econometrica*, 65 (3), 557-86.

StataCorp (2017), "ivregress: Single-equation Instrumental-variables Regression," in Technical Report. College Station, TX: StataCorp LP.

---- (2015), "Stata Statistical Software: Release 14." College Station, TX: StataCorp LP.

Stock, James H., Jonathan H. Wright, and Motohiro Yogo (2002), "A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments," *Journal of Business & Economic Statistics*, 20 (4), 518-29.

Swoboda, Bernhard and Johannes Hirschmann (2016), "Does Being Perceived as Global Pay Off? An Analysis of Leading Foreign and Domestic Multinational Corporations in India, Japan, and the United States," *Journal of International Marketing*, 24 (3), 1-30.

---- (2017), "Perceptions and Effects of Cross-National Corporate Reputation: The Role of Hofstede's Cultural Value Approach," *International Marketing Review*, 34 (6), 909-44.

Swoboda, Bernhard, Cathrin Puchert, and Dirk Morschett (2016), "Explaining the Differing Effects of Corporate Reputation Across Nations: A Multilevel Analysis," *Journal of the Academy of Marketing Science*, 44 (4), 454-73.

Tianjiao, Qiu (2014), "Product Diversification and Market Value of Large International Firms: A Macroenvironmental Perspective," *Journal of International Marketing*, 22 (4), 86-107.

Tyagi, Pradeep K. and Thomas R. Wotruba (1993), "An Exploratory Study of Reverse Causality Relationships Among Sales Force Turnover Variables," *Journal of the Academy of Marketing Science*, 21 (2), 143-53.

Villas-Boas, J. Miguel and Russell S. Winer (1999), "Endogeneity in Brand Choice Models," *Management Science*, 45 (10), 1324-38.

Walsh, Gianfranco, Sharon E. Beatty, and Edward M. K. Shiu (2009), "The Customer-based Corporate Reputation Scale: Replication and Short Form," *Journal of Business Research*, 62 (10), 924-30.

Westjohn, Stanford A., Holger Roschk, and Peter Magnusson (2017), "Eastern Versus Western Culture Pricing Strategy: Superstition, Lucky Numbers, and Localization," *Journal of International Marketing*, 25 (1), 72-90.

Wold, Herman O. A. (1982), "Soft Modeling: The Basic Design and Some Extensions," in *Systems Under Indirect Observations: Part II*, Karl G. Jöreskog and Herman O. A. Wold, eds. Amsterdam: North-Holland.

Woodside, Arch G. (2013), "Moving Beyond Multiple Regression Analysis to Algorithms: Calling for Adoption of a Paradigm Shift from Symmetric to Asymmetric Thinking in Data Analysis and Crafting Theory," *Journal of Business Research*, 66 (4), 463-72.

Wooldridge, Jeffrey M. (2015), "Control Function Methods in Applied Econometrics," *Journal of Human Resources*, 50 (2), 420-45.

---- (2010), *Econometric Analysis of Cross Section and Panel Data* (2 ed.). Cambridge, MA: MIT Press.

---- (2013), *Introductory Econometrics. A Modern Approach* (5 ed.). Mason, OH: South Western, Cengage Learning.

Zhang, Yang and Manfred Schwaiger (2012), "A Comparative Study of Corporate Reputation between China and Developed Western Countries," in *Handbook of Research in International Advertising*, Shintaro Okazaki, ed. Cheltenham: Edward Elgar Publishing.

## APPENDIX

=== Insert Table A1 about here ===

## TABLES

**Table 1:** Comparison of Approaches to Deal with Endogeneity

Criteria \ Methods	Control variable approach	Instrumental variable (IV) approach	Instrumental variable-free approaches	
			Gaussian copula	Latent instrumental variable (LIV)
Number of variables	Data on additional variables must be collected	Instrumental variables have to be identified and data has to be collected	No additional variables needed	No additional variables needed
Distribution of variables	No assumptions required	No assumptions required	Endogenous variables have to be non-normally distributed	Endogenous variables have to be non-normally distributed
Nature of dependent variable	Discrete or continuous	Continuous	Discrete or continuous	Continuous
Statistical tests	Not necessary	Test for significance and relevance	Test for significance	Test for significance
Acceptance in scientific community	Widely accepted and commonly used	Widely accepted and commonly used	Relatively new and therefore rarely used	Relatively new and therefore rarely used
Implementation in software	No additional implementation necessary	Supported by, for example, SPSS, STATA, and R software packages	The REndo (Gui et al. 2017) package supports the Gaussian copula approach	The REndo (Gui et al. 2017) package supports LIV model with one dependent and one independent variables

**Table 2: Results of the Gaussian Copula Approach**

	Original model		Gaussian copula model 1 (endogenous variable: <i>COMP</i> )		Gaussian copula model 2 (endogenous variable: <i>LIKE</i> )		Gaussian copula model 3 (endogenous variable: <i>CUSA</i> )	
Variable	Value	<i>p</i> -value	Value	<i>p</i> -value	Value	<i>p</i> -value	Value	<i>p</i> -value
<i>COMP</i>	0.016	0.746	0.014	0.848	0.017	0.763	0.021	0.707
<i>LIKE</i>	0.331	< 0.01	0.331	< 0.01	0.370	< 0.01	0.331	< 0.01
<i>CUSA</i>	0.509	< 0.01	0.509	< 0.01	0.511	< 0.01	0.582	< 0.01
$c_{COMP}$			0.002	0.974				
$c_{LIKE}$					-0.033	0.245		
$c_{CUSA}$							-0.041	0.063
	Gaussian copula model 4 (endogenous variables: <i>LIKE, COMP</i> )		Gaussian copula model 5 (endogenous variables: <i>LIKE, CUSA</i> )		Gaussian copula model 6 (endogenous variables: <i>COMP, CUSA</i> )		Gaussian copula model 7 (endogenous variables: <i>LIKE, COMP, CUSA</i> )	
Variable	Value	<i>p</i> -value	Value	<i>p</i> -value	Value	<i>p</i> -value	Value	<i>p</i> -value
<i>COMP</i>	-0.006	0.939	0.021	0.705	-0.027	0.717	-0.037	0.653
<i>LIKE</i>	0.381	0.000	0.341	0.000	0.333	0.000	0.362	0.000
<i>CUSA</i>	0.509	0.000	0.580	0.000	0.592	0.000	0.589	0.000
$c_{COMP}$	0.019	0.737			0.039	0.432	0.047	0.405
$c_{LIKE}$	-0.041	0.283	-0.008	0.790			-0.024	0.509
$c_{CUSA}$			-0.039	0.084	-0.049	0.045	-0.047	0.061

**Table 3: Results of the IV Approach**

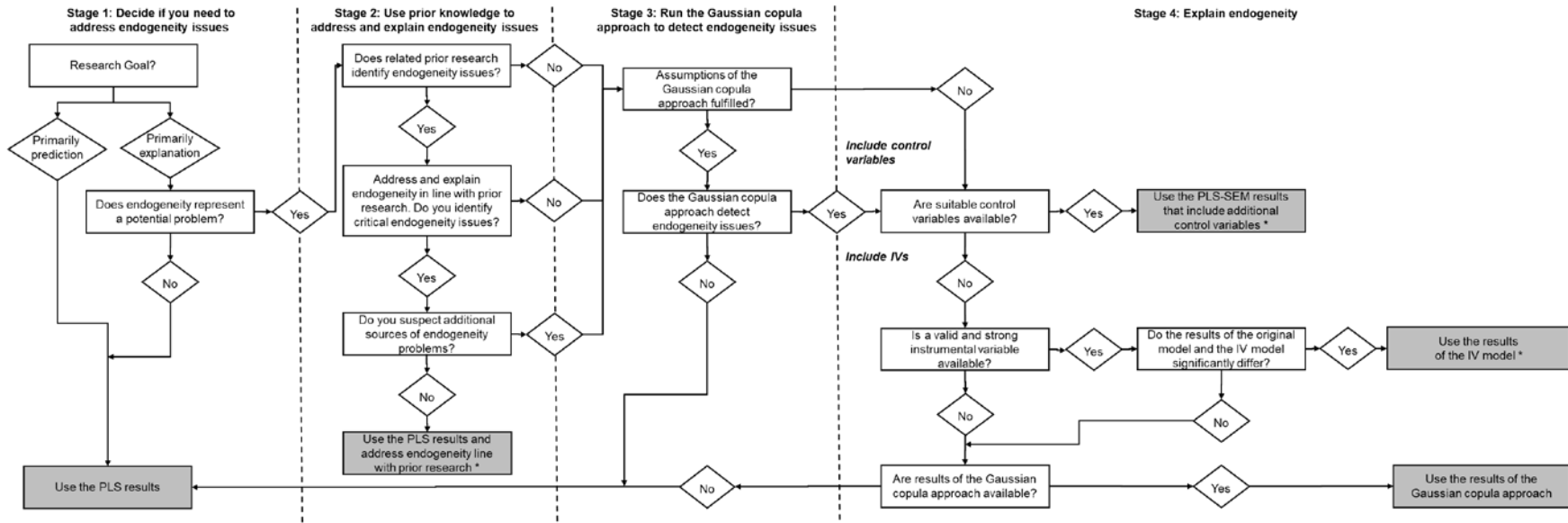
Endogenous Variable	Coefficients			$R^2$ values		Weakid test		Wu-Hausman test
	<i>COMP</i>	<i>LIKE</i>	<i>CUSA</i>	First stage without IV	First stage with IV	<i>F</i> -value	Significant?	<i>p</i> -value
<i>LIKE</i>	-0.075	0.506	0.454	0.496	0.560	49.478	YES	0.216
<i>COMP</i>	0.120	0.278	0.490	0.433	0.560	97.647	YES	0.216
<i>CUSA</i>	-0.087	0.081	1.113	0.295	0.300	2.575	NO	0.216

**Table A1: International marketing studies that apply PLS-SEM**

<b>Journal</b> <b>Year</b>	<b>International Marketing Review</b>	<b>Journal of International Business Studies</b>	<b>Journal of International Marketing</b>
<b>2017</b>	Gomes et al. (2017); Moon and Oh (2017); Pinho and Thompson (2017); Rahman et al. (2017)		Ashraf et al. (2017); Obadia et al. (2017)
<b>2016</b>	Andéhn and L'Espoir Decosta (2016); Henseler et al. (2016b); Jean et al. (2016); Richter et al. (2016); Schlegelmilch et al. (2016)	Lew et al. (2016)	
<b>2015</b>	Rippé et al. (2015); Sinkovics et al. (2015)	Khan et al. (2015); Obadia et al. (2015)	Griffith and Yanhui (2015); Morgeson et al. (2015); Sheng et al. (2015)
<b>2014</b>	Freeman and Styles (2014); Griffith et al. (2014)		Ashraf et al. (2014); Diamantopoulos et al. (2014)
<b>2013</b>	Sinkovics et al. (2013)	Schotter and Beamish (2013)	Magnusson et al. (2013); Obadia (2013); Swoboda and Elsner (2013)
<b>2012</b>		Lam et al. (2012)	Hilmersson and Jansson (2012); Johnston et al. (2012)
<b>2011</b>			Hortinha et al. (2011); Nijssen and Douglas (2011); Sichtmann et al. (2011)
<b>2010</b>	Duque and Lado (2010)	Shi et al. (2010)	Navarro et al. (2010); Sichtmann and von Selasinsky (2010)
<b>2009</b>		Fey et al. (2009)	Lages et al. (2009); Nijssen and Herk (2009)
<b>2008</b>			Brettel et al. (2008); Nijssen and Douglas (2008)
<b>Number of articles (2008-2017)</b>	15	7	21

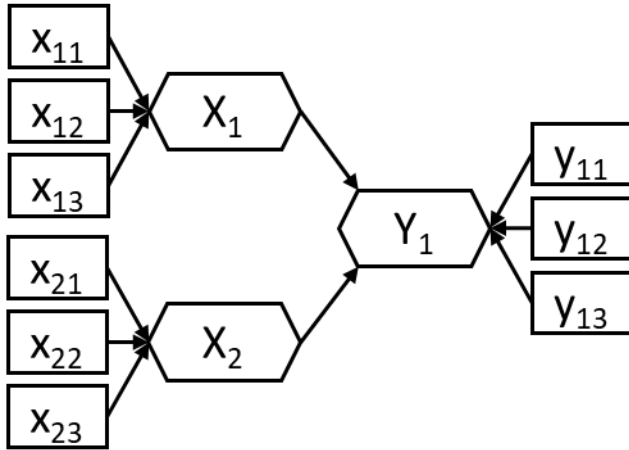


# FIGURES

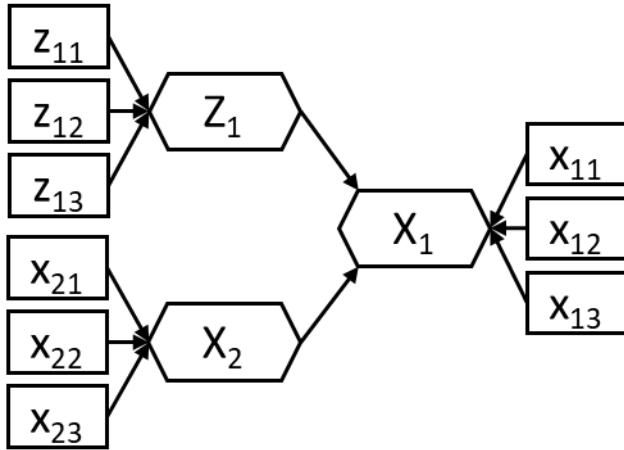


\* Consider continuing at Stage 3 and use the Gaussian copula approach to assess to what extent the endogeneity problem has been successfully addressed by the model.

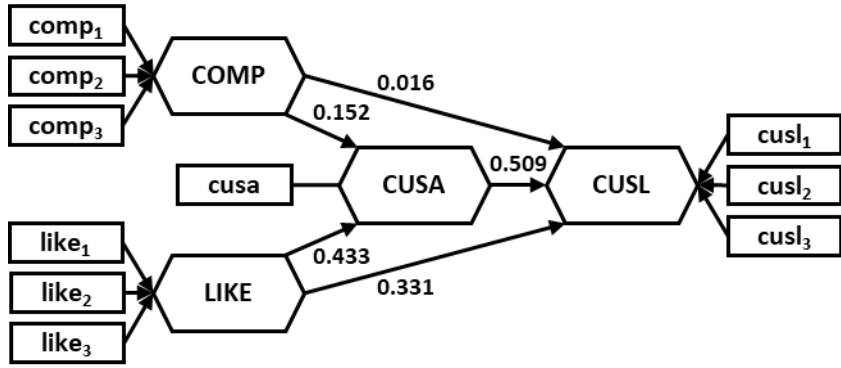
**Figure 1: Procedure to Address Endogeneity in PLS-SEM**



**Figure 2: Simple PLS Path Model**



**Figure 3: First Stage of the IV Approach**



**Figure 4: The Simple Corporate Reputation Model**

## ENDNOTES

<sup>1</sup> Other sources of endogeneity include measurement error and nonrecursive relationships between constructs (Wooldridge 2010).

<sup>2</sup> This implementation is similar to other methodological extensions, such as finite mixture PLS (Hair et al. 2018), which uses composite scores as input for mixture regressions.

<sup>3</sup> The model and data can be downloaded from <https://www.pls-sem.net/>.

<sup>4</sup> The parallel findings from the application of the Gaussian copula approach to single and multiple endogenous variables is not surprising, since the technique computes the copulas separately when considering multiple endogenous variables in the PLS path model.

<sup>5</sup> While these arguments support the use of perceived switching costs, we note that this is not an ideal IV, as one could expect this variable to correlate with other service perceptions (e.g., perceived service quality) that are relevant for customer loyalty and which the two reputation dimensions only capture indirectly (e.g., Raithel and Schwaiger 2015).

<sup>6</sup> We also ran the control function approach on the model using switching costs as the IV. We used the statistical software R for the calculation and the boot package to bootstrap the standard errors (Canty and Ripley 2017). As expected, the results parallel those of the IV approach.