Estimating and assessing second-order constructs using PLS-PM: the case of composites of composites

Florian Schuberth (Faculty of Engineering Technology, Universiteit Twente, Enschede, The Netherlands)

Manuel Elias Rademaker (Faculty of Business Management and Economics, Julius-Maximilians-Universität Würzburg, Würzburg, Germany)

Jörg Henseler (Faculty of Engineering Technology, Universiteit Twente, Enschede, The Netherlands) (Nova Information Management School, Universidade Nova de Lisboa, Lisbon, Portugal)

Estimating and assessing second-order constructs using PLS-PM: the case of composites of composites

Florian Schuberth   Manuel Rademaker   Jörg Henseler

August 5, 2020

Florian Schuberth, Faculty of Engineering Technology, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands
   E-mail: f.schuberth@utwente.nl
Manuel Rademaker, Faculty of Business Management and Economics, University of Würzburg, Sanderring 2, 97072 Würzburg, Germany
   E-mail: manuel.rademaker@uni-wuerzburg.de
Jörg Henseler, Faculty of Engineering Technology, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands
   E-mail: j.henseler@utwente.nl

Abstract

Purpose – The study's purpose is threefold: (i) to propose partial least squares path modeling (PLS-PM) as a way to estimate models containing composites of composites and to compare the performance of the PLS-PM approaches in this context, (ii) to provide and evaluate two testing procedures to assess the overall model fit of such models, and (iii) to introduce user-friendly step-by-step guidelines.

Design/methodology/approach – A simulation is conducted to examine the PLS-PM approaches and the performance of the two proposed testing procedures.

Findings – The simulation results show that the two-stage approach, its combination with the repeated indicators approach and the extended repeated indicators approach perform similarly. However, only the former is Fisher consistent. Moreover, our simulation shows that guidelines that neglect model fit assessment miss an important opportunity to detect misspecified models. Finally, the results show that both testing procedures based on the two-stage approach allow for assessment of the model fit.

Practical implications – Analysts who estimate and assess models containing composites of composites should use our guidelines, since the majority of guidelines neglect model fit assessment and thus omit a crucial step of structural equation modeling.

Originality/value – This study contributes to the understanding of the discussed approaches. Moreover, it highlights the importance of overall model fit assessment and provides insights about testing the fit of models containing composites of composites. Based on these findings, step-by-step guidelines are introduced to estimate and assess models containing composites of composites.

Keywords second-order constructs, composites of composites, model fit assessment, partial least squares path modeling, user-guidelines, Monte Carlo simulation

Research type Research paper

Acknowledgements

1 Introduction

Structural equation modeling (SEM) has become a highly appreciated modeling framework in social and behavioral sciences including information systems (IS) research (Chin and Todd, 1995; Green and Inman, 2007). It allows for operationalizing theoretical concepts by a set of observable variables and connecting constructs, the statistical representations of the concepts, via a structural model (Bollen, 1989). In doing so, random measurement errors and reciprocal relationships among the constructs can be taken into account. Hence, SEM allows researchers to statistically model complex theories. Moreover, it can be used to empirically assess these theories, which makes it a favorable tool in many disciplines.

In SEM, several ways of operationalizing theoretical concepts have been established. Arguably, the most widespread ways are the common factor model and the composite model (Fornell and Bookstein, 1982; Rigdon, 2016; Henseler, 2017). The common factor model – also known as the reflective measurement model – is the prevalent modeling approach in social and behavioral sciences. It assumes that each indicator is a measurement-error-prone consequence of the theoretical concept. Consequently, there is a presumed causal relationship between the common factor, a latent variable representing the theoretical concept, and its belonging indicators. Examples of concepts in the field of IS that have been operationalized by a reflective measurement model are trust in mobile applications (Hajiheydari and Ashkani, 2018) and purchase intention (Hsu, 2017). Next to the reflective measurement model, the composite model can be used to operationalize concepts (Rigdon, 2012). In the composite model, the theoretical concept is represented as an emergent variable (Cohen et al., 1990; Cole et al., 1993; Benitez et al., 2020), i.e., a composite of (measurement-error-free) indicators. In contrast to the reflective measurement model, the relationship between the indicators and the construct is a definitional and not a causal one (Henseler, 2017). Hence, the relationship between the indicators and the construct is modeled from the indicators to the construct, rather than from the construct to the indicators. Examples of concepts from IS research that have been operationalized by the composite model are organizational Internet use (Brock and Zhou, 2005) and IT infrastructure capabilities (Benitez et al., 2018b).

Hierarchical constructs – so-called higher-order constructs – have gained increasing popularity in SEM (see, e.g., Law and Wong, 1999; Johnson et al., 2012; Polites et al., 2012). In contrast to unidimensional constructs, a higher-order construct contains several layered structures of constructs and therefore involves several dimensions. Its application allows researchers to match the level of abstraction of predictor and outcome variables and reduces model complexity. As a result, a model's parsimony increases because fewer parameters must be estimated (Edwards, 2001). Although generally any number of levels of abstraction is conceivable, second-order constructs clearly prevail in empirical studies.

The choice between common factors and composites is not limited to unidimensional constructs but rather equally applies to both higher- and lower-order constructs. Consequently, we can distinguish between four types of second-order constructs: common factors of common factors, common factors of composites, composites of common factors, and composites of composites. Figure 1 contrasts the four main types of second-order constructs. In doing so, we follow common symbol usage and use circles to represent common factors and error terms, hexagons to represent composites, and squares to represent observable variables.

– INSERT FIGURE 1 HERE –

The most well-explored type of second-order construct in the SEM literature is the common factor of common factors, also known as the reflective-reflective second-order construct, as displayed in Figure 1a (e.g., Bollen, 1989; Mulaik and Quartetti, 1997). It is used to operationalize a multidimensional concept that is assumed to cause several concepts, each itself measured by a set of observable variables. Examples include green supply chain management (Lee et al., 2012), IT competency (Pérez-López and Alegre, 2012), or a study by Hong et al. (2008) investigating the effect of knowledge on system integration project performance. Furthermore, the common factor of composites as depicted in Figure 1c can be used to operationalize a unidimensional concept that causes other concepts which are assumed to be composed of their observable variables. However, they are not often found in empirical research and thus are currently of less practical relevance (Becker et al., 2012). In contrast, composites of common factors, as shown in Figure 1b, are often encountered in the literature, e.g., to model personal traits with its big five dimensions (Leong et al., 2017), or tourist engagement (Rasoolimanesh et al., 2019). This kind of second-order construct was investigated quite recently, and guidelines on their estimation and assessment have been proposed (Van Riel et al., 2017). Finally, second-order constructs of the composite of composites type, as displayed in Figure 1d, are assumed to be composed of its first-order constructs, which are in turn built by a weighted linear combination of observable indicators. In doing so, the composite of composites can be used to operationalize concepts which are assumed to consist of other concepts that are again assumed to be composed. The practical relevance of this kind of second-order construct for IS research is highlighted by the study of Becker et al. (2012), who reviewed 25 models containing second-order constructs in the journal 'Management Information Systems Quarterly' and identified that this type of second-order construct is the second most often employed. Concrete empirical examples employing second-order composites are the studies of Gómez-Cedeño et al. (2015); Benitez et al. (2018a) and Benitez et al. (2018b). However, the SEM literature has hardly studied this type of second-order construct.

Various estimators have been proposed to estimate structural equation models. The most well-known one is the maximum-likelihood (ML) estimator which minimizes the discrepancy between the model-implied and sample variance-covariance matrix to obtain the parameter estimates (Jöreskog, 1970). However, in its current form, the ML estimator is hardly suitable to estimate models containing composites, particularly when they are in an endogenous position in the structural model (Cadogan and Lee, 2013; Rigdon, 2014). In contrast, an estimator that is able to cope with this type of model is the composite-based estimator partial least squares path modeling (PLS-PM, Lohmöller, 1989). It first creates proxies, i.e., linear combinations of observable variables, for the constructs, and the model parameters are subsequently estimated based on these proxies. For a recent overview of the methodological research on PLS-PM, we refer to Khan et al. (2019).

The PLS-PM literature provides several approaches to estimate models containing second-order constructs, such as the repeated indicators approach (Wold, 1982), the two-stage approach (Agarwal and Karahanna, 2000), and the hybrid approach (Wilson and Henseler, 2007). Although there are some theoretical research papers investigating these approaches (Wilson and Henseler, 2007; Becker et al., 2012; Van Riel et al., 2017; Duarte and Amaro, 2018; Sarstedt et al., 2019), their statistical evaluation has received only minimal attention in the past. Particularly, the performance of the approaches for models containing composites of composites is largely unexplored because existing studies have mainly focused on composites of common factors. This is not without problems, since choosing the wrong approach can substantially impact on the estimation. For example, Becker et al. (2012) showed that the hybrid approach produces severely biased path coefficient estimates in the case of composites of common factors. Consequently, researchers studying models containing composites of composites are left in the dark about the choice of the approach to estimate this type of model, and thus are not

prevented from an inappropriate choice which may lead to questionable results.

Although several guidelines have been proposed recently to estimate and assess models containing second-order constructs in the context of PLS-PM (e.g., Sarstedt et al., 2019; Cheah et al., 2019), the majority of these guidelines omit a crucial step of SEM, namely the assessment of the overall model fit (Mulaik et al., 1989; Yuan, 2005; Barrett, 2007; Steiger, 2007; Henseler, 2018). Overall model fit assessment is crucial because it investigates whether the proposed model is consistent with the collected data and thus can be used to falsify a researcher's theory. Moreover, it can help to identify potential problems in the estimated model (Henseler et al., 2014). In SEM, the chi-square test is typically used for that purpose, which is closely tied to the ML estimator and its accompanying assumptions (e.g., Lawley and Maxwell, 1962; Jöreskog, 1969). As these assumptions are not necessary for PLS-PM, Dijkstra and Henseler (2015) propose a bootstrap-based test, also known as Bollen-Stine bootstrap (Bollen and Stine, 1992), to assess the exact overall fit of models estimated by PLS-PM. This testing procedure has been adopted by only a single guidelines paper on second-order constructs using PLS-PM (Van Riel et al., 2017). However, these guidelines focus on second-order constructs of the composite of common factors type. So far, no similar guidelines for assessing models containing composites of composites have been introduced. Moreover, the testing procedure proposed by Van Riel et al. (2017) has not yet been subject of a simulation study. As a consequence, researchers who follow current guidelines that omit the overall model fit assessment miss an important opportunity to detect misspecified models. Moreover, researchers have only little guidance with respect to how to assess the exact fit of models containing composites of composites and generally know little about the performance of this testing procedure.

To overcome the previously mentioned gaps in the literature, the study at hand contributes in three ways: first, it shows how PLS-PM can be employed to estimate models containing second-order composites of composites. In doing so, it investigates the performance of various approaches proposed in the context of PLS-PM to estimate this type of model. Second, it proposes two testing procedures to assess the exact overall fit of models containing such second-order constructs and assesses their efficacy. Third, based on the empirical findings, it provides guidelines for analysts to estimate and assess models containing composites of composites using PLS-PM.

The remainder of the paper is organized as follows. The next section, Section 2, presents the extant approaches commonly used to estimate models containing composites of composites by PLS-PM. Section 3 introduces two testing procedures to examine whether the estimated model containing composites of composites is consistent with the collected data. In Section 4, the performance of the commonly used approaches to estimate models containing composites of composites and the efficacy of the two testing procedures are evaluated by means of a Monte Carlo simulation. In doing so, it is demonstrated on the population level that recently proposed guidelines that ignore overall model fit assessment are unable to detect misspecified models containing composites of composites that would have been detected by an overall model fit assessment. Additionally, the results are reported, and the most important findings are discussed. Based on our findings, in Section 5 we provide user-friendly guidelines on estimating and assessing this kind of model. In doing so, we recommend use of the two-stage approach and the proposed two-step testing procedure. The paper closes with a discussion and an outlook for future research in Section 6.

3

2 Extant approaches to estimate models containing composites of composites

In the following section, we present the approaches that are often employed in the context of PLS-PM to estimate models containing second-order constructs. In doing so, two models containing second-order constructs must be distinguished: (i) models where the second-order construct is not embedded in a structural model and (ii) models where the second-order construct is embedded in a structural model. Since the former is generally not identified in the case of a second-order construct specified as a composite of composites (see Section 5), in the following, we mainly focus on models where the second-order construct is embedded in a nomological net.

To review commonly used approaches in the context of PLS-PM to estimate models containing second-order constructs, Figure 2 depicts a model example containing a composite of composites in an endogenous position in the structural model. Such a setting highlights the importance of applying PLS-PM, since models with endogenous composites cannot currently be estimated by commonly employed estimators for SEM, such as ML. Moreover, it allows us to demonstrate how the original PLS-PM approaches need to be modified to overcome their shortcomings, i.e., that the effects of independent constructs on the endogenous composite of composites cannot be estimated.

– INSERT FIGURE 2 HERE –

The example model consists of one second-order composite (C) that is built by two first-order composites (c_1 and c_2). This represents a situation where two composed concepts compose a second concept. To model such a situation in SEM, composites of composites can be employed. In our example model, two sets of indicators (y_{1i} and y_{2i}, $i = 1, 2$) completely define the respective first-order composites which in turn fully determine the second-order composite. Consequently, the correlations between the two sets of indicators and variables not forming the respective first-order composite, and the correlations between the first-order composites and variables not forming the second-order composite (x_1 and x_2 in this example), are fully mediated by the first-order and the second-order composites, respectively. Furthermore, the model contains an exogenous common factor (ξ) that, in contrast to the first-order composites, causally affects the second-order composite (C). The common factor is connected to two indicators (x_1 and x_2). To preserve clarity, the structural error, measurement errors, and correlations among the variables are not displayed.

Originally, the extant PLS-PM literature provides three approaches to estimate models containing second-order constructs: (i) the *repeated indicators approach*, (ii) the *two-stage approach*, and (iii) the *hybrid approach*. Since scholars have recognized the shortcomings of the repeated indicators and the hybrid approach in the case of antecedent constructs affecting the second-order construct (here, common factor ξ, which affects the second-order composite C), two adjustments have been proposed: (i) specification of indirect effects through the first-order constructs (*extended repeated indicators approach* and *extended hybrid approach* respectively, Becker et al., 2012) and (ii) a combination of the repeated indicators and the two-stage approach (*embedded two-stage approach*, Wilson, 2010). According to Ringle et al. (2012) and Sarstedt et al. (2019), the repeated indicators and the two-stage approach, respectively, are currently the most widely applied approaches. Table 1 summarizes and illustrates the approaches by means of the model example illustrated in Figure 2.

4

2.1 *The repeated indicators approach*

The repeated indicators approach was originally proposed by Wold (1982) and is also known as the superblock approach (Tenenhaus et al., 2005). To estimate models containing a second-order construct, it reuses all indicators of the first-order constructs as indicators of the second-order construct (Lohmöller, 1989). Consequently, the indicators of the first-order constructs are used twice in the estimation process. A known caveat is that the first-order construct with the majority of indicators is likely to explain the largest share of variation in the second-order construct. Hence, the repeated indicators approach may face problems in situations where an unequal number of indicators is connected to the first-order constructs (Ringle et al., 2012). However, it allows for the estimation of models containing second-order constructs by PLS-PM in one step and can be extended in a straightforward manner to estimate models containing constructs of an even higher order, such as third-order or fourth-order constructs (Wetzels et al., 2009). Moreover, this approach can be used to mimic principal component analysis (PCA) or approaches to generalized canonical correlation analysis (GCCA) when applied to an isolated second-order construct (Lohmöller, 1989, p. 132).

2.2 *The two-stage approach*

As its name suggests, the two-stage approach consists of two stages (Agarwal and Karahanna, 2000; Henseler et al., 2007). In the first stage, the model is estimated without the second-order constructs in order to obtain construct scores for the first-order constructs. Technically, after removing the second-order constructs, researchers may use any structural model specification involving the remaining constructs to obtain construct scores. However, for testing purposes, it offers advantages with respect to specifying a saturated structural model, i.e., a structural model with zero degrees of freedom. In the second stage, the first-order constructs are removed from the structural model, and their construct scores are utilized as indicators of the second-order construct. Moreover, prior to the estimation of the model containing the second-order construct, the indicators of the remaining constructs are replaced by the corresponding construct scores, i.e., the remaining constructs become single-indicator constructs. Consequently, in contrast to the repeated indicators approach, two estimation steps are required to estimate the parameters of models containing composites of composites. If the second-order construct is modeled as a common factor or if any of the first-order constructs are modeled as common factors, a correction for attenuation is necessary to obtain consistent parameter estimates (Van Riel et al., 2017). The authors label this the *three-stage approach.*

2.3 *The hybrid approach*

The hybrid approach (Wilson and Henseler, 2007) is similar to the repeated indicators approach but avoids reusing the indicators of the first-order constructs. In doing so, the sets of indicators of the first-order constructs are split in half: one half of each set of indicators is used as indicators of the second-order construct, while the other half remains as indicators of the respective first-order constructs. The primary intention of this procedure is to avoid artificially correlated errors (Wilson and Henseler, 2007). However, in the case of constructs specified as composites, this is not an issue, and contradicts the concept of a composite that assumes that a composite is

fully defined by a unique set of indicators (Bollen and Bauldry, 2011). Consequently, the hybrid approach is not a viable approach to estimate models containing composites of composites.

2.4 *Adjustments to the original approaches*

A drawback of both the repeated indicators approach and the hybrid approach is that almost all of the variation in the second-order constructs is explained by the first-order constructs (Ringle et al., 2012). Therefore, path coefficients of potential predictors are estimated towards zero, which likely leads to incorrect conclusions. Two modified versions of the original approaches overcome this shortcoming: Becker et al. (2012) propose to additionally specify indirect effects of potential predictors through the first-order constructs, which we call the *extended repeated indicators* approach and the *extended hybrid approach*, respectively, in the following. Hence, the effect of a predictor construct on the second-order construct can be assessed by its total effect on the second-order construct. Furthermore, a combination of the repeated indicators approach and the two-stage approach was proposed (Ringle et al., 2012; Wilson, 2010), which is also known as the *embedded two-stage approach* (Sarstedt et al., 2019). In doing so, the *embedded two-stage approach* applies the repeated indicators approach in the first stage to obtain the first-order construct scores. In contrast, for the *two-stage approach*, we employ a model with a saturated structural model in the first stage. Subsequently, in the second stage, both approaches utilize these construct scores as indicators of the second-order construct to estimate the structural model.

3 Overall fit assessment of models containing composites of composites

In general, rigorous research demands researchers to maintain a critical distance to their own work and constantly question their proposals and findings. Of course, this is also true for models proposed in the context of explanatory and confirmatory research using SEM, in which the model reflects the underlying mechanisms of a studied population. For our specific case, this means that researchers should investigate whether a proposed model containing a second-order construct of the composite of composites type adequately reflects the mechanisms of the studied population.

To address this issue, the overall model fit is usually assessed in SEM. Overall model fit assessment is an integral part of SEM (Yuan, 2005). It investigates whether the specified model is consistent with the collected data. This is important because "[i]f a model is consistent with reality, then the data should be consistent with the model" (Bollen, 1989, p. 68). The importance of overall model fit assessment is acknowledged across various disciplines and emphasized in literally every textbook on SEM (e.g., Bollen, 1989; Schumacker and Lomax, 2009; Kline, 2015); in other words, "if SEM is used, then model fit testing and assessment is paramount, indeed crucial, and cannot be fudged for the sake of 'convenience' or simple intellectual laziness on the part of the investigator" (Barrett, 2007, p. 823).

To examine whether the estimated model fits the collected data, the estimated model is compared to the saturated model; in other words, the discrepancy between the estimated model-implied correlation matrix is compared to the empirical correlation matrix of the indicators. Hence, how well the constraints imposed by the model hold is investigated. Because the model fit indicates whether the underlying theory is reflected in the data, an acceptable fit is required before interpreting the estimated model parameters. In general, the overall model fit can be

assessed in two nonexclusive ways: (i) fit indices; and (ii) tests for exact overall model fit.[1]

Instead of merely defining perfect fit as the only desirable objective, fit indices quantify model fit on a continuous scale to assess how well the proposed model corresponds to the collected data. To judge whether the fit of a model is acceptable on the basis of fit indices, the decision typically relies on heuristic rules and not on statistical inferences, i.e., the values of fit indices are compared to threshold values to decide whether the model exhibits an acceptable model fit. The most often used fit index in the context of PLS-PM is the standardized root mean square residual (SRMR, Hu and Bentler, 1999; Henseler et al., 2014). Moreover, indices such as the comparative fit index (CFI, Bentler, 1990), the normed fit index (NFI, Bentler and Bonett, 1980), and the root mean square outer residual covariance (RMS$_\theta$, Lohmöller, 1989) can be applied. However, their efficacy in the context of composite models is not yet well understood. Therefore, we recommend reporting fit indices but not judging the model fit simply based on thresholds proposed for common factor models.

Next to employing fit indices, researchers can rely on statistical testing to assess the exact overall model fit. In the context of PLS-PM and composite models, a bootstrap-based test has been proposed to assess the exact overall model fit (Dijkstra and Henseler, 2015; Schuberth et al., 2018). In doing so, discrepancy measures such as the geodesic distance (d_G), the squared Euclidean distance (d_L), and the standardized root mean square residual (SRMR) are embedded in a bootstrap-based procedure to obtain their reference distributions (Beran and Srivastava, 1985; Bollen and Stine, 1992). This allows for statistical inferences of the exact fit of the estimated model, i.e., whether the collected data are consistent with the postulated correlation structure of the model.[2] For decision making, the calculated discrepancy measure based on the original dataset is usually compared to the 95% or 99% quantiles of the reference distribution (Henseler et al., 2016). The null hypothesis assuming that the population variance-covariance matrix of the indicators equals the population variance-covariance matrix implied by the model ($H_0 : \mathbf{\Sigma} = \mathbf{\Sigma}(\boldsymbol{\theta})$) is rejected when the calculated discrepancy measure is larger than the quantile of the reference distribution.

The overall fit assessment of models containing composites of composites in the context of PLS-PM depends on the estimation approach. Compared to the other approaches, the two-stage approach has the following conceptual and practical advantages: (i) the dimension of the indicator correlation matrix remains untouched, as no indicators are added, and (ii) due to its way of estimation, it allows for a more in-depth model fit assessment. In the case of the (extended) repeated indicators and (extended) hybrid approach, the originally specified model is modified, i.e., indicators are added or differently assigned. Consequently, it is not clear how to test the originally specified model. The same applies to the embedded two-stage approach, which modifies the model of the first stage. Hence, in terms of model fit assessment, the two-stage approach can be regarded as superior. Therefore, in the following, overall model fit assessment is only considered for models estimated by the two-stage approach based on a saturated structural model.

To assess the exact overall fit of models containing second-order composites estimated by the two-stage approach, we propose two strategies: (i) following a two-step testing procedure, and (ii) assessing the complete postulated model all at once. The two-step testing procedure

---

[1]For an elaborate discussion about model fit testing and fit indices, we refer to the special issue on SEM in the journal 'Personality and Individual Differences' (Vernon and Eysenck, 2007).

[2]Although doubts were recently raised about the applicability of the bootstrap-based test in the context of PLS-PM (Hair et al., 2019a;b; 2020), they are unfounded, as can be determined by the mathematical proofs derived by Beran and Srivastava (1985) on the testing procedure. The bootstrap-based test does not require an estimator that minimizes the discrepancy between the model-implied and the empirical correlation matrix to obtain the estimates; instead, it requires a consistent estimate of the model-implied indicator correlation matrix, which typically requires consistent parameter estimates. For composite models, PLS-PM using Mode B provides such consistent estimates (Dijkstra, 2017).

has already been suggested in the PLS-PM literature to assess models containing composites of common factors (Van Riel et al., 2017). In the first step, the fit of the model without the second-order composite and a saturated structural model is assessed. In doing so, the indicator correlation matrix implied by the model of the first stage is compared to the empirical indicator correlation matrix. In this step, potential misspecifications of the structural model are ignored, and the model assessment solely focuses on problems in the operationalization of the concepts. Since all concepts of the first stage are operationalized by composites that are embedded in a saturated structural model, the first stage resembles a confirmatory composite analysis (CCA, Schuberth et al., 2018). If no empirical evidence against the operationalization was found, in the second step, the structural model containing the second-order composites with its first-order composites is assessed. Hence, the estimated implied correlation matrix of the model from the second stage is compared to the construct correlation matrix of the first stage. It is noted that the latter does not contain any constraints, as the structural model of the first stage is saturated. In contrast, in the case of the complete model assessment, the estimated model-implied indicator correlation matrix based on the originally specified model, i.e., the model in the researcher's mind, is compared to the sample counterpart. To construct the model-implied indicator correlation matrix of the originally specified model containing the second-order composite, we slightly modified the implied correlation matrix for linear composite models containing no second-order constructs (Dijkstra, 2017).[3] As parameters, the estimated weights and indicator intra-block correlations of the first stage and the first-order composite weights, and the correlations among the first-order composites and the path coefficient estimates of the second stage, are used to build the model-implied indicator correlation matrix.

Although both strategies of model fit assessment are conceptually sound, in the case of a significant test result, the complete model assessment gives the researcher no indication as to which part of the model is misspecified, i.e., the operationalization of the concepts or the structural model containing the second-order composites. Moreover, no simulation study has yet been conducted to assess their performance.

4 A Monte Carlo simulation

A Monte Carlo simulation is conducted to further contribute to the understanding of the approaches used to estimate models containing second-order composites and to investigate the performance of two proposed ways of model fit assessment. Due to the well-known drawback of the original repeated indicators approach in the case of potential predictors of the second-order constructs (Becker et al., 2012; Ringle et al., 2012) and the inherent inappropriateness of the hybrid approach for concepts modeled as composites, in the following, we only investigate the two-stage approach based on a saturated structural model in the first stage, the embedded two-stage approach that combines the repeated indicators and the two-stage approach, and the extended repeated indicators approach that specifies additional effects.

---

[3]To construct the model-implied indicator matrix of the originally specified model, we first construct the model-implied correlation matrix for the structural model containing the second-order composite with its first-order composites. In line with the composite model, this matrix contains no constraints on the correlations among the first-order composites, but constraints are contained for the correlations between the first-order composites and the other constructs, namely, that this correlation matrix is of rank one. Second, the model-implied indicator correlation matrix is obtained. In doing so, the main diagonal blocks containing the correlations among the indicators forming a composite are set to the corresponding sample correlations of the indicators. The off-diagonal blocks, which contain the correlations between the indicators of two different blocks, equal the product of the correlations between the indicators and their respective composite and the corresponding construct correlation. Finally, the model-implied indicator correlation matrix is constructed. To determine how the model-implied indicator variance-covariance matrix is constructed for linear composite models without second-order constructs, we refer to Dijkstra (2017).

4.1 *Simulation design*

To assess the performance of the various approaches, we choose a population model that consists of one second-order composite (η_3) built by three first-order composites $c_i = \boldsymbol{w}_i' \boldsymbol{x}_i$, $i = 1, 2, 3$. In selecting the population weights to form the second-order composite and the population correlations among the first-order composites, we choose parameters that conform with recent guidelines about the assessment of models containing this type of second-order construct (Sarstedt et al., 2019), i.e., each first-order composite contributes substantially to the second-order composite, and the correlations among the first-order composites are below the recommended threshold of 0.5 (Hair et al., 2017b). The population weights of the first-order composites are $\boldsymbol{w}_c' = \begin{pmatrix} 0.4 & 0.4 & 0.5 \end{pmatrix}$, and their population correlations are set to: $r_{c_1 c_2} = 0.49$, $r_{c_1 c_3} = 0.27$, and $r_{c_2 c_3} = 0.413$.

The first-order composites c_1 to c_3 are built by two (y_{11} and y_{12}), four (y_{21} to y_{24}), and six (y_{31} to y_{36}) indicators, respectively. The number of indicators per first-order composite intentionally differs in order to investigate the claim that an unequal number of indicators exerts an adverse effect on the parameter estimates of the repeated indicators approach.[4] The weights of the indicators forming the first-order composites c_1 to c_3 and the correlations among these indicators are again chosen to conform with the previously mentioned guidelines, i.e., all indicators contribute substantially to their composites, and the correlations among the indicators forming a composite are below 0.5. The employed population weights are as follows: $\boldsymbol{w}_1' = \begin{pmatrix} 0.8 & 0.4 \end{pmatrix}$, $\boldsymbol{w}_2' = \begin{pmatrix} 0.5 & 0.3 & 0.2 & 0.4 \end{pmatrix}$, and $\boldsymbol{w}_3' = \begin{pmatrix} 0.3 & 0.3 & 0.2 & 0.2 & 0.4 & 0.3 \end{pmatrix}$. The population correlation between the two indicators of c_1 is set to 0.3125, and the population correlation matrices of the indicators forming c_2 and c_3, respectively, are given in Equations 1 and 2. To preserve clarity, the correlations are rounded to the second decimal place.

$$\boldsymbol{\Sigma}_{c_2} = \begin{matrix} \begin{matrix} y_{21} & y_{22} & y_{23} & y_{24} \end{matrix} \\ \begin{pmatrix} 1.00 & 0.40 & 0.30 & 0.31 \\ & 1.00 & 0.28 & 0.31 \\ & & 1.00 & 0.30 \\ & & & 1.00 \end{pmatrix} \end{matrix} \tag{1}$$

$$\boldsymbol{\Sigma}_{c_3} = \begin{matrix} \begin{matrix} y_{31} & y_{32} & y_{33} & y_{34} & y_{35} & y_{36} \end{matrix} \\ \begin{pmatrix} 1.00 & 0.10 & 0.25 & 0.13 & 0.10 & 0.30 \\ & 1.00 & 0.20 & 0.40 & 0.30 & 0.20 \\ & & 1.00 & 0.30 & 0.10 & 0.30 \\ & & & 1.00 & 0.20 & 0.20 \\ & & & & 1.00 & 0.10 \\ & & & & & 1.00 \end{pmatrix} \end{matrix} \tag{2}$$

In addition to the second-order composite η_3, the structural model consists of three composites (ξ, η_1, and η_2) that are antecedents of this second-order composite. The second-order composite is deliberately in an endogenous position to investigate the efficacy of the extended repeated indicators approach that was proposed to overcome the drawbacks of the original repeated indicators approach in such a situation. In designing the structural model, we choose a complexity that can be similarly found in the empirical literature, such as Ainin et al. (2015);

---

[4]Additionally, we considered a population model with three indicators per first-order composite. The results were almost identical, except that the bias in the weights of the first-order composite slightly differed. This is in line with recent findings based on an empirical example which shows that an unequal number of indicators of the first-order composites does not substantially influence the estimates (Cheah et al., 2019).

Yim and Leem (2013); Hsieh et al. (2006). To ensure a sufficient amount of degrees of freedom of the structural model, which is required for the assessment of the structural model, we opt for a multiple mediation structure. In doing so, we choose the following population path coefficients: $\gamma_1 = 0.2$, $\gamma_2 = -0.4$, $\gamma_3 = 0.35$, $\beta_1 = 0.4$, and $\beta_2 = 0.2$. As a consequence, the effect sizes (f^2) range from 0.04 to 0.22, indicating small and medium effects (Cohen, 1992). Moreover, the structural error terms are mutually independent. The whole model is depicted in Figure 3. For clarity, we conceal error terms and correlations among variables forming a composite.[5]

– INSERT FIGURE 3 HERE –

To assess the performance of the various approaches, we consider the estimated bias and the estimated root mean square error (RSME), which are defined as follows:

$$\widehat{\text{Bias}} = \frac{1}{N} \sum_{i=1}^{N} \hat{\theta}_i - \theta \quad \text{and} \quad \widehat{\text{RMSE}} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (\hat{\theta}_i - \theta)^2} \tag{3}$$

where θ represents a generic population parameter, $\hat{\theta}_i$ is its corresponding estimate from the i-th Monte Carlo simulation run, and N denotes the total number of all Monte Carlo simulation runs. While the estimated bias indicates how much an estimate differs on average from its population counterpart, the RMSE combines the bias and the uncertainty involved in an estimate, namely, its standard error. While a value close to 0 is desired for the estimated bias, small values are preferred for the estimated RMSE.

To estimate the bias and the RMSE, a model that equals the baseline population model from Figure 3 is estimated by each approach 1,000 times using sample sizes of 100, 300, 500, and 1,000. We check each estimation for admissibility and replace inadmissible results with admissible ones. Hence, each condition is based on exactly 1,000 admissible estimations. Additionally, we assess whether the approaches are Fisher consistent, i.e., whether they are able to retrieve the population parameters when the indicator population correlation matrix is provided. Finally, it is noted that all specified models, i.e., the two models of the two-stage approach and the model of the repeated indicators approach, are identified. A necessary condition to achieve identification of composite models is that the scale of each composite is fixed (Schuberth et al., 2018). As is common in PLS-PM, this is achieved by scaling the weights to obtain composites with a unit variance. Moreover, each composite must be formed by at least one indicator and no composite is allowed to be isolated in the structural model. Similarly, each second-order composite must be composed of at least one first-order composite, and no second-order composite must be isolated in the structural model. Both are the cases here. Finally, the structural model must also be identified. According to the 'recursive rule', recursive structural models with uncorrelated error terms, such as ours, are always identified (Bollen, 1989, p. 104). As a consequence, all specified models are identified.

To study the performance of the two presented testing procedures, we assess their type I error rate, i.e., the probability of falsely rejecting the null hypothesis, and their power, i.e., the probability of correctly rejecting the null hypothesis, under various conditions. In doing so, the two-stage approach with a saturated structural model in the first stage is employed for the

---

[5]The complete indicator population correlation matrix is provided in Table 5 of the Online Supplementary Material.

estimation. For the assessment of the type I error rate, we estimate the model that matches the population model from Figure 3. Hence, the model is correctly specified in this situation. In the case of the complete model assessment, the estimated model-implied indicator correlation matrix of the originally specified model, i.e., the specification that matches the population model from Figure 3, is compared to the sample correlation matrix of the indicators. In the case of the two-step testing procedure, in the first step, the fit of the model from the first stage of the two-stage approach is assessed, i.e., the model without the second-order composite and a saturated structural model. Subsequently, in the second step, the fit of the structural model containing the second-order composite and the corresponding first-order composites is tested. In doing so, only the estimations that exhibit an acceptable model fit in the first step are tested in the second step. In this scenario, we expect rejection rates close to the predefined significance level.

To assess the tests' power, we consider population models that differ from the specified model containing the second-order composite.[6] In doing so, the same population weights are chosen as in the baseline population model depicted in Figure 3; however, the structural model is different. As displayed in Figure 4, it does not contain a second-order composite, and ξ, η_1, and η_2 directly affect c_1 to c_3. In contrast to the baseline population model, the composites c_1 to c_3 are now in an endogenous position. The variance that cannot be explained by their antecedents is captured by the structural error terms ζ_3 to ζ_5, which are uncorrelated with all other variables in the model. As a consequence, the researcher's model is misspecified because he/she wrongly assumes that a second-order composite exists, although the world functions according to a different model.

– INSERT FIGURE 4 HERE –

To determine various degrees of model misspecification, we generate approximately 110,000 population models with a structural model as displayed in Figure 4. To obtain these models, only the population path coefficients are varied. The number of indicators attached to each composite, the magnitude of the weights and the indicator correlations remain identical to the baseline population model. Subsequently, we estimate each population by the two-stage approach using the specification that contains the second-order composite and calculate the SRMR as a measure of model misfit. Finally, we determine the 5%, 50%, and 95% quantiles of the SRMR values to obtain relatively small, medium, and large misspecifications, and we choose sets of population path coefficients that produce SRMR values which are close to these quantiles. The sets of path coefficients and the corresponding SRMR values are displayed in Table 2. The three indicator population correlation matrices are provided in Table 6 to 8 of the Online Supplementary Material, which also contains the R code used to obtain the three population models. As highlighted by the positive SRMR values, there exists no population model with the same structure as the baseline model but different values of the parameters, which implies one of the three indicator population correlation matrices. This is important for a testing procedure that compares the sample and model-implied indicator variance-covariance matrix because otherwise a model specification matching the baseline population model would

---

[6]As asked by an anonymous reviewer, we also use a second model specification that matches the population of Figure 4 and assess the test's performance for the 4 population models, i.e., the original baseline population model and the three population models that were originally used to assess the tests' power. The results are provided in the Online Supplementary Material: see Figures 10 and 11.

still be incorrectly specified, but it would not be detected by the testing procedure. This problem is well-known in the SEM literature as 'equivalent models' (Raykov and Penev, 1999).

Additionally, Table 2 reports the SRMR for a model without the second-order composite and a saturated structural model, i.e., the model from the first stage of the two-stage approach ($SRMR_{1st}$) and the SRMR for the structural model containing the second-order composites with its first-order composites, i.e., the model from the second stage of the two-stage approach ($SRMR_{2nd}$). As seen, the misfit is fully caused by the structural model, since the SRMR for the model from the first stage is zero, while the SRMR for the model from the second stage is larger than zero. Consequently, we expect rejection rates close to the predefined significance level in the first step of the two-step testing procedure. In contrast, for the second step and the complete model testing, we expect rejection rates which are larger than the significance level. In these situations, rejection rates above 80% are desired (Cohen, 1988), and we expect that rejection rates will increase for larger misspecifications and larger sample sizes.

To assess the tests' performance, we conducted 1,000 simulation runs for each sample size of 100, 300, 500, and 1,000, with the four different population models. As significance levels, we consider $\alpha = 1\%$ and 5%. All bootstrap-based model fit tests are based on the SRMR to measure the discrepancy between the empirical and the model-implied indicator correlation matrix.[7] The number of bootstrap runs per test is set to 1,000 and nonconverged estimations during the bootstrapping are replaced such that each test is based on exactly 1,000 admissible bootstrap runs.

– INSERT TABLE 2 HERE –

Our simulation design highlights the importance of model fit assessment and emphasizes the need for updated guidelines on the assessment of models containing second-order composites. Current guidelines for the assessment of formative measurement models[8] (Sarstedt et al., 2019; Hair et al., 2017a;b; Ramayah et al., 2016; Becker et al., 2012) fail to detect our incorrectly specified models. On the population level, all four estimated models, i.e., the correctly specified and the three misspecified models, are not indicated as problematic by current guidelines on the assessment of models containing composite of composites.[9] The necessary results to apply these guidelines can be found in Tables 6, 7, 8, and 10 of the Online Supplementary Material. In all cases, all indicators and first-order composites contribute substantially to their composite, i.e., all weights are sizable. Moreover, the respective correlations among the indicators and first-order composites forming a composite are all below the recommended threshold of 0.5 (Hair et al., 2017b). As a consequence, researchers are not alerted that their models are misspecified and will draw the wrong conclusions.

The complete simulation is conducted in the statistical programming environment R (R Core Team, 2019). All samples are drawn from a multivariate normal distribution using the `mvrnorm()` function of the MASS package (Venables and Ripley, 2002). In doing so, the means

---

[7]Since the squared Euclidean and the geodesic distances have also been proposed to assess the discrepancy (Dijkstra and Henseler, 2015), we additionally employed these two distance measures. However, the results are hardly affected.

[8]In PLS-SEM parlance, a formative measurement model is a composite model for which the weights are estimated by PLS-PM Mode B (Hair et al., 2017b, Glossary).

[9]It is noted that significance of the weights has not been assessed, since the parameters are retrieved from the population and not estimated for a sample. Thus, statistical inference is not required. Moreover, convergent validity cannot be assessed, since the model does not contain any reflective measurement models.

of the indicators are set to zero, and we use the respective indicator population correlation matrix as variance-covariance matrix. The PLS-PM parameter estimates are obtained by the `csem()` function of the cSEM package (Rademaker and Schuberth, 2020).[10] Outer weighting scheme Mode B is employed for all constructs since it has been shown to produce consistent estimates for composite models (Dijkstra, 2017). As stopping criterion, the absolute change of the weights is considered, i.e., during the PLS-PM algorithm, the largest absolute difference between the current weights and the weights from the previous iteration step are compared. In the case that the largest absolute difference is smaller than 10^{-5}, the algorithm stops. Furthermore, the maximum number of iterations is set to 100. For models without second-order constructs, all inner weighting schemes produce similar results (Noonan and Wold, 1982), and there is only minimal advice regarding which scheme to use for models containing second-order composites (Becker et al., 2012; Sarstedt et al., 2019). Although the path weighting scheme seems to be favored for models containing composites of composites (Becker et al., 2012), we further examine whether the choice of the inner weighting scheme affects the behavior of the different approaches. Therefore, the path, the factorial, and the centroid weighting scheme are employed for every approach. Finally, the exact overall model fit tests are conducted using the `testOMF()` function of the cSEM package. For the complete R code, we refer to the Online Supplementary Material.

4.2 *Simulation results*

4.2.1 **Behavior of the various approaches**

Table 3 displays the results for the estimated bias, the estimated RMSE, Fisher consistency, and the share of converged estimations of the three investigated approaches using the path, factorial, and centroid weighting scheme during PLS-PM's inner estimation. Due to space constraints, we present the results only for the estimates of two path coefficients (γ_1 and β_1), the first weight of the first-order composite c_1 (w_{11}), and the three weights used to build the second-order composite (w_{c_1}, w_{c_2}, and w_{c_3}). The remaining estimates behave similarly to those presented.

– INSERT TABLE 3 HERE –

The two-stage approach based on a saturated structural model in the first stage appears to produce consistent estimates. The average estimates of all parameters converge to their population counterparts for increasing sample size. Moreover, the decrease in the estimated RMSE reflects that the standard errors of the parameter estimates decrease when the sample size increases. This result is hardly affected by the choice of the inner weighting scheme. Among the approaches considered, it produces the smallest number of nonconverged estimations. In none of the conditions have more than 0.6% of the estimations not converged. Finally, it returns the population parameters when provided with the indicator population correlation matrix, showing the Fisher consistency of the two-stage approach.

---

[10]Since the cSEM package has been released on CRAN only recently, we cross-validated the initial estimations with ADANCO (Henseler, 2019), SmartPLS (Ringle et al., 2015) and matrixpls (Rönkkö, 2019).

The embedded two-stage approach behaves similarly to the original two-stage approach, but it produces a larger share of nonconverged estimations. The highest share of 10.9% is observed for small sample sizes ($n = 100$) in combination with the path weighting scheme. However, the number of converged estimations decreases for an increasing sample size and diminishes for sample sizes $n \geq 300$. Considering the estimated bias, it decreases for all parameters with an increasing sample size and almost diminishes for a sample size of 1,000 observations. The estimated RMSE also decreases for all parameters with increasing sample size, as the standard errors of the estimates decrease. In contrast to the two-stage approach, the embedded two-stage approach cannot retrieve all population parameters from the indicator population correlation matrix. The retrieved weights to build the first-order composites slightly differ from their population counterparts.[11] Consequently, the embedded two-stage approach is not Fisher consistent. Considering the three inner weighting schemes, the path weighting scheme produces the smallest deviations from their population counterparts for the weights used to build the first-order composites.

The extended repeated indicators approach, i.e., indirect effects of ξ, η_1, and η_2 on the second-order composite η_3 specified through the first-order composites, behaves similarly to the two other approaches when the centroid or the factorial schemes are employed. In the case of the path weighting scheme, the estimated weights to build the second-order composite considerably deviate on average from their population counterparts. In contrast to the other model parameters, the estimated bias of these weights does not diminish with increasing sample size. The large values of the estimated RMSE for these weights show that the estimates have larger standard errors compared to the other approaches. As expected, the direct effects of constructs on the second-order composites are biased towards zero. However, the results confirm that the total effects can be used to capture these effects. Moreover, in none of the conditions is the share of nonconverged estimations is above 1.8%. However, not all population parameters can be retrieved when the indicator population correlation matrix is provided as input. Hence, the extended repeated indicators approach is not Fisher consistent. This result is unaffected by the employed inner weighting scheme.

4.2.2 Performances of the two testing procedures

In the following, we present the results of our simulation for the two testing procedures. We only report the rejection rates for the two-stage approach in combination with the path weighting scheme. The results using the centroid and the factorial weighting scheme are very similar.

– INSERT FIGURE 5 HERE –

Figure 5 displays the tests' rejection rates in the case of no model misspecification, i.e., the null hypothesis is true. The complete model testing procedure produces rejection slightly below the predefined significance level. Particularly, for a small sample size and a significance level of 5%, the rejection rates are too small. However, the rejection rates converge to the predefined significance level with increasing sample size. Similar results can be observed for

---

[11]Although it is not presented that in the case of the path weighting scheme, the retrieved weights used to build c_3 deviate from their population counterparts. A table containing all retrieved parameters can be found in the Online Supplementary Material: see Table 9.

the first step of the two-step testing procedure. In contrast, in the second step of the two-step testing procedure, the rejection rates are too conservative, i.e., the model is rejected too often. However, with increasing sample size, the rejection rates converge to the predefined significance level.

– INSERT FIGURE 6 HERE –

Figure 6 depicts the tests' rejection rates in cases of the various misspecifications and the two considered significance levels of 1% and 5%. The complete model testing reliably detects medium and large misspecified models, except for small sample sizes, if a significance level of 1% is assumed. In the case of the small misspecification, small sample sizes are not sufficient to reliably reject the null hypothesis, and larger sample sizes are required. Considering the two-step testing procedure, in the first step the rejection rates are close to the predefined significance level. In contrast, in the second step of the two-step testing procedure almost all misspecifications are reliably detected, except small misspecifications in cases of small sample sizes.

4.3 *Simulation insights*

The Monte Carlo simulation shows that, for finite sample sizes, all three approaches perform similarly with respect to estimated bias and RMSE. An exception is the extended repeated indicators approach in combination with the path weighting scheme, which produces considerably biased estimates for those weights that are used to build the second-order composite. Moreover, our simulation confirms that the extended repeated indicators approach is able to estimate the constructs' direct effects on the second-order composite by their total effects. Considering Fisher consistency, only the two-stage approach returns all population parameters when the indicator population correlation matrix is applied. Additionally, its results are hardly affected by the inner weighting scheme.

Considering the results of the two presented testing procedures, the overall results are largely in line with our expectations. In the case that the specified model is consistent with the data, both testing procedures produce rejection rates which are close to the predefined significance levels: even though they slightly deviate for smaller sample sizes, they converge to the significance level with increasing sample size. Moreover, the complete model test reliably detects most of the misspecifications, even though small misspecifications require larger sample sizes. The two-step testing procedure produces rejection rates close to the significance level in the first step. This is not surprising, since a model with a saturated structural model is employed in the first stage of the two-stage approach. Hence, misspecifications in the structural model are ignored and remain undetected. However, in the second step, the two-step test reliably detects the misspecifications of the structural model. In general, and as expected, the power of the tests increases with increasing sample size.

5 Guidelines for estimating and assessing models containing composites of composites

To enable researchers to appropriately estimate and assess models containing composites of composites, particularly in the context of confirmatory and explanatory research, we provide guidelines embracing the two-stage approach and the two-step testing procedure. Although our simulation study shows that both testing procedures detect similarly reliable misspecifications in the structural model, the two-step testing procedure is recommended in the following because it allows for separate assessment of the operationalization of the concepts and the structural model. To fully leverage the strength of the two-step testing procedure, a saturated structural model should be specified in the first stage of the two-stage approach. This allows for localizing the misspecification, i.e., determining whether it occurs in the operationalization of the concepts or in the structural model.

Prior to the estimation of the model and its assessment, the researcher should ensure that the model is identified, guaranteeing that the model parameters can be uniquely retrieved from the system of equations. Although model identification is straightforward in the case of composite models (Dijkstra, 2017; Schuberth et al., 2018), it must not be disregarded. A necessary condition for the identification of a composite model is that each (higher-order) composite is built by at least one (lower-order composite) indicator and that no (higher-order) composite is isolated in the structural model. Apart from all composites being connected via the structural model, it must be additionally ensured that the structural model is also identified. Since the study at hand focuses only on recursive models, we refer to the 'recursive rule', which states that recursive models with uncorrelated structural error terms are always identified (Bollen, 1989, p. 104).

After model identification is ensured, the model can be estimated, and its fit can be assessed. For this purpose, guidelines are developed and depicted in Table 4. In the following, we elaborate each step of our proposed guidelines to estimate and assess linear recursive models containing composites of composites. In doing so, the two-stage approach is employed for model estimation based on a saturated structural model in the first stage. To illustrate our guidelines, we use the example model from Figure 2.

– INSERT TABLE 4 HERE –

Stage 1

In the first stage, as suggested by the two-stage approach, the model is estimated without second-order composites. The goal of this stage is to calculate the construct scores and to assess the operationalization of the concepts represented in the structural model. If all concepts are modeled as composites, this stage largely resembles a CCA (Schuberth et al., 2018).

Step 1: Estimating the model without second-order composites

In Step 1, the model without the second-order composites is estimated. As explained earlier, a model with a saturated structural model should be used. Figure 7 illustrates a potential model

specification with a saturated structural model for our example model.

For concepts that are conceptualized as composites, regression weights (PLS-PM Mode B) should be employed, since Mode B produces consistent estimates (Dijkstra, 2017). If the model contains common factors, the use of correlation weights (PLS-PM Mode A) in combination with a correction for attenuation is recommended for these constructs, i.e., PLSc (Dijkstra and Henseler, 2015), as traditional PLS-PM is known to produce inconsistent estimates for common factors (Dijkstra, 1981).

Step 2: Assessing the overall model fit

Once the model is estimated, it is necessary to evaluate whether it makes sense to continue forming second-order composites from the first-order composites. Thus, the fit of the estimated model from Step 1 needs to be assessed to examine whether the model is consistent with the collected data. For this purpose, and as shown by our simulation, the bootstrap-based test based on a discrepancy measure – such as the SRMR – can be used (see Section 3). For decision making in the context of hypothesis testing, the 95% or 99% quantiles should be used as critical values (Henseler et al., 2016), i.e., the value of the discrepancy measure based on the original dataset is compared to the 95% or 99% quantiles of its reference distribution, to decide whether the null hypothesis that the indicator population correlation matrix equals the model-implied counterpart is rejected. Next to the tests for exact overall model fit, fit indices can be employed to judge the fit of the model. However, fit indices for composite models have not yet been thoroughly examined, and therefore they should be used with caution (Hair et al., 2019b). If researchers encounter a significant inconsistency, they are advised to rethink their operationalization of the concepts. For general guidelines on assessing models estimated by PLS-PM in the context of explanatory and confirmatory research, we refer to Benitez et al. (2020).

Step 3: Extracting construct scores

Once the estimated model from Step 1 has exhibited acceptable fit, the construct scores are extracted for the first-order composites and the constructs not belonging to a second-order composite. The standardized scores are appended as new indicators to the file containing the observations of the original indicators.

Stage 2

In the second stage, the model containing the second-order composite with the originally postulated structural model is specified. Similar to Stage 1, the model is estimated and its overall model fit is assessed.

Step 4: Estimating the structural model containing second-order composites

In Step 4, the second-order composite is included in the structural model and the first-order composites are replaced by their construct scores. In doing so, the construct scores of the first-order composites are utilized as indicators to build the second-order composite. Similarly, the indicators of the other constructs are replaced by their construct scores, i.e., the constructs become single-indicator constructs. It is noted that in the case of factor scores, the reliability of these scores needs to be adjusted. For the estimation of the weights used to build the second-order composite, the use of Mode B is recommended because it is a consistent estimator (Dijkstra, 2017). The model specification of the second stage for our example model is depicted in Figure 8.

– INSERT FIGURE 8 HERE –

Step 5: Assessing model fit

In the last step, which is similar to Step 3, the overall model fit must be assessed. In doing so, we investigate whether the structural model containing the second-order composite is consistent with the data, and thus whether it is at all useful to build a second-order composite from the first-order composites. Supported by our simulation results, the bootstrap-based test of exact overall model fit from Section 3 can be used again for this purpose. As mentioned in Step 2 of our guidelines, the threshold of fit indices as known from common factor analysis should be used cautiously, since they are hardly explored in the context of composite models. In case researchers face misfitting in the second step, they are advised to reconsider their specified structural model, including the specification of the second-order composite.

6   Discussion and future research

Estimating and assessing models containing composites including second-order composites poses a challenge for researchers employing SEM, particularly when composites, including second-order composites, are located in an endogenous position in the structural model. To overcome these difficulties, the study at hand suggests employing PLS-PM and provides step-by-step user-guidelines for estimating and assessing the fit of models containing composites of composites. In doing so, we propose to employ the two-stage approach for estimation and a two-step testing procedure for overall model fit assessment. This enables researchers from the field of IS to appropriately estimate models containing second-order composites of composites and to assess their overall model fit. The latter is particularly important in causal research, i.e., explanatory and confirmatory research, as it can indicate inconsistencies between the model and the collected data and thus provide evidence against a researcher's postulated theory.

Although several guidelines on estimating and assessing models containing second-order constructs have been published in the context of PLS-PM, they either focus on a different type of second-order construct or ignore the assessment of overall model fit – a crucial step in SEM. The need for updated guidelines for models containing composites of composites in the context of explanatory and confirmatory research is additionally emphasized by our simulation design.

Similar to Schuberth (forthcoming), it shows on the population level that existing guidelines, which neglect overall model fit assessment and focus mainly on measurement quality evaluation (e.g., Sarstedt et al., 2019; Hair et al., 2020; 2019a; 2017a;b; Ramayah et al., 2016; Becker et al., 2012), are unable to detect incorrectly specified models that would have been detected through overall model fit assessment. Hence, they miss an important opportunity to detect potential model misspecifications and leave researchers unaware of whether their specified models are (in)consistent with the collected data. Although recently proposed guidelines in the context of PLS-PM recommend evaluating the predictive power of a model (e.g., Hair et al., 2019b), it is emphasized that such an evaluation cannot replace the overall model fit assessment in explanatory and confirmatory research. This is highlighted by the fact that erroneous models can result in accurate prediction and that well-specified models can result in poor prediction (Shmueli, 2010). Therefore, guidelines in the context of causal research that currently ignore overall model fit assessment should be extended.

Considering our proposed guidelines, we recommend employing Mode B to estimate composite models because it produces consistent estimates (Dijkstra, 2017). However, in some situations, e.g., in the case of a large degree of multicollinearity among the indicators forming a composite, it can be beneficial to desist from consistent estimators and choose an estimator that is inconsistent but robust against multicollinearity. In doing so, the use of Mode A might be worthwhile. Although its estimates are inconsistent for composite models, and thus will not converge in probability to the population weights for an increasing sample size, they are unaffected by multicollinearity. An analogous approach is customary in the context of regression analysis, where ridge regression can be employed when researchers face multicollinearity issues (Mason and Brown, 1975). Moreover, the use of predefined weights or weights obtained by other techniques can be beneficial (Dijkstra, 2013; 2016) to overcome issues caused by multicollinearity.

To support our proposed guidelines and to further illuminate the performance of the commonly applied approaches in the context of PLS-PM to estimate models containing composites of composites, a Monte Carlo simulation was conducted. The results of our simulation showed that, for finite sample sizes, none of the approaches considered are superior in terms of estimated bias and RMSE. For most approaches, the average parameter estimates converged to the population counterpart and the standard errors of the estimates decreased with increasing sample size. However, for the extended repeated indicators approach, we observed a slightly different behavior compared with the findings of previous simulation studies. For example Becker et al. (2012) concluded that the path weighting scheme is the preferable option for the repeated indicators approach, and we observed in our simulation that this inner weighting scheme produces more biased first-order composite weight estimates compared with the centroid and the factorial inner weighting scheme. Overall, only the two-stage approach is Fisher consistent, i.e., it retrieves the population parameters when the indicator population correlation matrix is applied. Moreover, its results were hardly affected by the choice of the inner weighting scheme and it produced a low share of nonconverged estimations. Additionally, the way of the two-stage approach to estimate models containing composites of composites offers advantages over the other two approaches because in model testing the source of misspecification can be located. Therefore, for models containing a composite of composites, we recommend use of the two-stage approach based on a saturated structural model in the first stage.

Similarly, to investigate the efficacy of the two presented testing procedures, i.e., complete testing and two-step testing, a Monte Carlo simulation was conducted. The results showed that both testing procedures produce rejection rates which are slightly different from the predefined significance, particularly for small sample sizes. However, for an increasing number of observations, the rejection rates converged to the predefined significance level. Moreover, both tests were capable of reliably detecting misspecifications, except for small misspecifications and

in the case of a small sample size. In this case, the null hypothesis was rejected too rarely. Therefore, researchers should ensure a sufficient sample size to avoid underpowered statistical tests. Moreover, it is noted that our proposed testing procedures for exact model fit are not a panacea to detect all types of model misspecifications. As is well-known in the SEM literature, tests for model fit do not sound an alarm if a model is misspecified but equivalent to the correctly specified model, i.e., both produce the same model-implied variance-covariance matrix (Raykov and Penev, 1999). This highlights the importance of the theory from which a model is derived, and the fact that it is not sufficient that models are solely statistical sound: "if the data are consistent with the model, this does not imply that the model corresponds to reality." (Bollen, 1989, p. 68)

A limitation of our guidelines is that they focus solely on linear models, i.e., models without interaction terms; however, there are instances in which researchers want to include a second-order composite of composites in a moderation. In this case, the two-stage approach used for moderation analysis in the context of PLS-PM (Henseler and Fassott, 2009; Dijkstra and Schermelleh-Engel, 2014) needs to be slightly modified. First, identical to the Step 1 of our proposed guidelines, a model with a saturated structural model and without the second-order composite and interaction term is estimated. Second, a model with a saturated structural model is again specified; however, in contrast to the previous model, the constructs' indicators are replaced by the corresponding construct scores from the previous step, and the second-order composite with first-order constructs' scores as indicators is included. This stage resembles Step 4 of our guidelines, except for the specification of the structural model. Finally, the originally proposed structural model is specified. In doing so, all constructs are modeled as single-indicator constructs, including the second-order composite, using their corresponding construct scores, and the interaction term is incorporated as the product of the scores of the involved constructs. It is noted that in all stages, a correction for attenuation is required to obtain consistent estimates if common factors are involved. A software implementation of this approach can be found in the R package cSEM (Rademaker and Schuberth, 2020). Unfortunately, the exact fit of models containing interaction terms, i.e., the model from the last step, can currently not be tested in the context of PLS-PM.

Furthermore, we only focus on the assessment of a single model; however in empirical research it is not uncommon that researchers face several theoretically well-motivated models. In such a situation, the use of model selection criteria such as the Akaike's information criterion (AIC, Akaike, 1998) and the Bayesian information criterion (BIC, Schwarz, 1978) can be helpful to decide between the competing models. With respect to how to apply model selection criteria in the context of PLS-PM, we refer to Sharma et al. (2019); Danks et al. (2020); Liengaard et al. (accepted).

Since a simulation study is always limited by its design, we recommend future research to conduct more complex simulation studies. For instance, researchers should further vary the distributional assumption and the model complexity to broaden the understanding of the behavior of the different approaches. This is particularly important for the repeated indicators approach, since we observed a different behavior with regard to the preferable inner weighting scheme compared with previous simulation studies, e.g., Becker et al. (2012). Future research should study these contradictory findings in more detail. Similarly, future research should strive for explanations as to why the repeated indicators approach and its modifications are not Fisher consistent in order to mathematically support our findings. Moreover, the performances of the PLS-PM approaches to estimate models containing composites of composites should be compared to other estimators such as generalized structured component analysis (GSCA). Considering the testing procedures, our simulation study is limited to misspecifications in the structural model. However, there is a simulation study showing that the test is also capable of detecting misspecifications in the composite models (Schuberth et al., 2018). This provides a

starting point to further investigate the performance of the testing procedures with respect to misspecifications in both, the structural model and the operationalization of the concepts.

Moreover, we only investigate the performances of the approaches for models containing second-order constructs specified as composites of composites. However, second-order constructs specified as mixtures of different constructs are conceivable. To the best of our knowledge, this type of second-order construct is not yet applied in the literature; however, it may open new avenues for modeling theories in IS and beyond. Potential candidates are a second-order construct that is specified as a common factor of a mixture of composites and common factors and a second-order construct that is specified as a composite of a mixture of composites and common factors. Figure 9 displays two possibilities. An investigation into the performances of the commonly employed approaches and the assessment of such models would further contribute to the literature.

– INSERT FIGURE 9 HERE –

The indicators building a composite are typically assumed to be free of measurement errors, which can limit the applicability of the composite model in practice. To overcome this shortcoming, the composite reliability can be adjusted if the researcher doubts this assumption (Henseler, 2017). In doing so, the share of measurement error in the indicator must be determined externally, e.g., by previous studies. Subsequently, the composite reliability can be taken into account during the estimation (Van Riel et al., 2017). Finally, it is straightforward to extend our guidelines to estimate and assess models with third-, fourth-, or even higher-order composites, e.g., Wetzels et al. (2009). However, future research must investigate the behaviors of the approach and testing procedures.

References

Agarwal, R. and Karahanna, E. (2000), "Time flies when you're having fun: Cognitive absorption and beliefs about information technology usage", *MIS Quarterly*, Vol. 24(4), pp. 665–694.

Ainin, S., Parveen, F., Moghavvemi, S., Jaafar, N. I., and Shuib, N. L. M. (2015), "Factors influencing the use of social media by SMEs and its performance outcomes", *Industrial Management & Data Systems*, Vol. 115(3), pp. 570–588.

Akaike, H. (1998), "Information theory and an extension of the maximum likelihood principle", in Parzen, E., Tanabe, K., and Kitagawa, G., (Eds.), "Selected Papers of Hirotugu Akaike. Springer Series in Statistics (Perspectives in Statistics)", Springer, New York, pp. 199–213.

Barrett, P. (2007), "Structural equation modelling: Adjudging model fit", *Personality and Individual Differences*, Vol. 42(5), pp. 815–824.

Becker, J.-M., Klein, K., and Wetzels, M. (2012), "Hierarchical latent variable models in PLS-SEM: Guidelines for using reflective-formative type models", *Long Range Planning*, Vol. 45(5), pp. 359–394.

Benitez, J., Castillo, A., Llorens, J., and Braojos, J. (2018a), "IT-enabled knowledge ambidexterity and innovation performance in small u.s. firms: The moderator role of social media capability", *Information & Management*, Vol. 55(1), pp. 131–143.

Benitez, J., Henseler, J., Castillo, A., and Schuberth, F. (2020), "How to perform and report an impactful analysis using partial least squares: Guidelines for confirmatory and explanatory IS research", *Information & Management*, Vol. 57(2), p. 103168.

Benitez, J., Llorens, J., and Braojos, J. (2018b), "How information technology influences opportunity exploration and exploitation firm's capabilities", *Information & Management (in press)*.

Bentler, P. M. (1990), "Comparative fit indexes in structural models.", *Psychological Bulletin*, Vol. 107(2), pp. 238–246.

Bentler, P. M. and Bonett, D. G. (1980), "Significance tests and goodness of fit in the analysis of covariance structures.", *Psychological Bulletin*, Vol. 88(3), pp. 588–606.

Beran, R. and Srivastava, M. S. (1985), "Bootstrap tests and confidence regions for functions of a covariance matrix", *The Annals of Statistics*, Vol. 13(1), pp. 95–115.

Bollen, K. A. (1989), *Structural Equations with Latent Variables*, John Wiley & Sons, New York.

Bollen, K. A. and Bauldry, S. (2011), "Three Cs in measurement models: Causal indicators, composite indicators, and covariates.", *Psychological Methods*, Vol. 16(3), pp. 265–284.

Bollen, K. A. and Stine, R. A. (1992), "Bootstrapping goodness-of-fit measures in structural equation models", *Sociological Methods & Research*, Vol. 21(2), pp. 205–229.

Brock, J. K.-U. and Zhou, Y. J. (2005), "Organizational use of the Internet: Scale development and validation", *Internet Research*, Vol. 15(1), pp. 67–87.

Cadogan, J. W. and Lee, N. (2013), "Improper use of endogenous formative variables", *Journal of Business Research*, Vol. 66(2), pp. 233–241.

Cheah, J.-H., Ting, H., Ramayah, T., Memon, M. A., Cham, T.-H., and Ciavolino, E. (2019), "A comparison of five reflective–formative estimation approaches: reconsideration and recommendations for tourism research", *Quality & Quantity*, Vol. 53(3), pp. 1421–1458.

Chin, W. W. and Todd, P. A. (1995), "On the use, usefulness, and ease of use of structural equation modeling in MIS research: A note of caution", *MIS Quarterly*, Vol. 19(2), pp. 237–246.

Cohen, J. (1988), *Statistical Power Analysis for the Behavioral Sciences*, Lawrence Erlbaum Associates, Hillsdale.

Cohen, J. (1992), "A power primer", *Psychological Bulletin*, Vol. 112(1), pp. 155–159.

Cohen, P., Cohen, J., Teresi, J., Marchi, M., and Velez, C. N. (1990), "Problems in the measurement of latent variables in structural equations causal models", *Applied Psychological Measurement*, Vol. 14(2), pp. 183–196.

Cole, D. A., Maxwell, S. E., Arvey, R., and Salas, E. (1993), "Multivariate group comparisons of variable systems: MANOVA and structural equation modeling.", *Psychological Bulletin*, Vol. 114(1), pp. 174–184.

Danks, N. P., Sharma, P. N., and Sarstedt, M. (2020), "Model selection uncertainty and multi-model inference in partial least squares structural equation modeling (PLS-SEM)", *Journal of Business Research*, Vol. 113, pp. 13–24.

Dijkstra, T. K. (1981), *Latent variables in linear stochastic models: Reflections on "Maximum Likelihood" and "Partial Least Squares" methods*, PhD thesis, Groningen University.

Dijkstra, T. K. (2013), "On the extraction of weights from pairwise comparison matrices", *Central European Journal of Operations Research*, Vol. 21(1), pp. 103–123.

Dijkstra, T. K. (2016), "High-dimensional concepts and substance driven ways to make them operational, some tools", in "Cluj Economics and Business Seminar", Cluj-Napoca, Babes-Bolyia University, pp. 1–38.

Dijkstra, T. K. (2017), "A perfect match between a model and a mode", in Latan, H. and Noonan, R., (Eds.), "Partial Least Squares Path Modeling: Basic Concepts, Methodological Issues and Applications", Springer, Cham, pp. 55–80.

Dijkstra, T. K. and Henseler, J. (2015), "Consistent and asymptotically normal PLS estimators for linear structural equations", *Computational Statistics & Data Analysis*, Vol. 81(1), pp. 10–23.

Dijkstra, T. K. and Schermelleh-Engel, K. (2014), "Consistent partial least squares for nonlinear structural equation models", *Psychometrika*, Vol. 79(4), pp. 585–604.

Duarte, P. and Amaro, S. (2018), "Methods for modelling reflective-formative second order constructs in PLS", *Journal of Hospitality and Tourism Technology*, Vol. 9(3), pp. 295–313.

Edwards, J. R. (2001), "Multidimensional constructs in organizational behavior research: An integrative analytical framework", *Organizational Research Methods*, Vol. 4(2), pp. 144–192.

Fornell, C. and Bookstein, F. L. (1982), "Two structural equation models: LISREL and PLS applied to consumer exit-voice theory", *Journal of Marketing Research*, Vol. 19(4), pp. 440–452.

Gómez-Cedeño, M., Castán-Farrero, J. M., Guitart-Tarrés, L., and Matute-Vallejo, J. (2015), "Impact of human resources on supply chain management and performance", *Industrial Management & Data Systems*, Vol. 115(1), pp. 129–157.

Green, K. W. and Inman, R. A. (2007), "The impact of JIT-II-selling on organizational performance", *Industrial Management & Data Systems*, Vol. 107(7), pp. 1018–1035.

Hair, J., Hollingsworth, C., Randolph, A., and Chong, A. (2017a), "An updated and expanded assessment of PLS-SEM in information systems research", *Industrial Management & Data Systems*, Vol. 117(3), pp. 442–458.

Hair, J. F., Howard, M. C., and Nitzl, C. (2020), "Assessing measurement model quality in PLS-SEM using confirmatory composite analysis", *Journal of Business Research*, Vol. 109, pp. 101–110.

Hair, J. F., Hult, G. T. M., Ringle, C. M., and Sarstedt, M. (2017b), *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)*, Sage, Thousand Oaks, CA, 2nd ed.

Hair, J. F., Risher, J. J., Sarstedt, M., and Ringle, C. M. (2019a), "When to use and how to report the results of PLS-SEM", *European Business Review*, Vol. 31(1), pp. 2–24.

Hair, J. F., Sarstedt, M., and Ringle, C. M. (2019b), "Rethinking some of the rethinking of partial least squares", *European Journal of Marketing*, Vol. 53(4), pp. 566–584.

Hajiheydari, N. and Ashkani, M. (2018), "Mobile application user behavior in the developing countries: A survey in Iran", *Information Systems*, Vol. 77, pp. 22–33.

Henseler, J. (2017), "Bridging design and behavioral research with variance-based structural equation modeling", *Journal of Advertising*, Vol. 46(1), pp. 178–192.

Henseler, J. (2018), "Partial least squares path modeling: Quo vadis?", *Quality & Quantity*, Vol. 52(1), pp. 1–8.

Henseler, J. (2019), *ADANCO 2.0.1*, Composite Modeling GmbH & Co., Kleve, Germany.

Henseler, J., Dijkstra, T. K., Sarstedt, M., Ringle, C. M., Diamantopoulos, A., Straub, D. W., Ketchen, D. J., Hair, J. F., Hult, G. T. M., and Calantone, R. J. (2014), "Common beliefs and reality about PLS: Comments on Rönkkö and Evermann (2013)", *Organizational Research Methods*, Vol. 17(2), pp. 182–209.

Henseler, J. and Fassott, G. (2009), "Testing moderating effects in PLS path models: An illustration of available procedures", in "Handbook of Partial Least Squares: Concepts, Methods and Applications", Springer, Berlin, pp. 713–735.

Henseler, J., Hubona, G., and Ray, P. A. (2016), "Using PLS path modeling in new technology research: Updated guidelines", *Industrial Management & Data Systems*, Vol. 116(1), pp. 2–20.

Henseler, J., Wilson, B., Götz, O., and Hautvast, C. (2007), "Investigating the moderating role of fit on sports sponsorship and brand equity", *International Journal of Sports Marketing and Sponsorship*, Vol. 8(4), pp. 34–42.

Hong, H.-K., Kim, J.-S., Kim, T., and Leem, B.-H. (2008), "The effect of knowledge on system integration project performance", *Industrial Management & Data Systems*, Vol. 108(3), pp. 385–404.

Hsieh, C.-T., Lai, F., and Shi, W. (2006), "Information orientation and its impacts on information asymmetry and e-business adoption", *Industrial Management & Data Systems*, Vol. 106(6), pp. 825–840.

Hsu, L.-C. (2017), "Investigating community members' purchase intention on Facebook fan page", *Industrial Management & Data Systems*, Vol. 117(5), pp. 766–800.

Hu, L. and Bentler, P. M. (1999), "Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives", *Structural Equation Modeling: A Multidisciplinary Journal*, Vol. 6(1), pp. 1–55.

Johnson, R. E., Rosen, C. C., Djurdjevic, E., Taing, M. U., et al. (2012), "Recommendations for improving the construct clarity of higher-order multidimensional constructs", *Human Resource Management Review*, Vol. 22(2), pp. 62–72.

Jöreskog, K. (1969), "A general approach to confirmatory maximum likelihood factor analysis", *Psychometrika*, Vol. 34(2), pp. 183–202.

Jöreskog, K. G. (1970), "A general method for estimating a linear structural equation system", *ETS Research Bulletin Series*, Vol. 1970(2), pp. i–41.

Khan, G. F., Sarstedt, M., Shiau, W.-L., Hair, J. F., Ringle, C. M., and Fritze, M. P. (2019), "Methodological research on partial least squares structural equation modeling (PLS-SEM): An analysis based on social network approaches", *Internet Research*, Vol. 29(3), pp. 407–429.

Kline, R. B. (2015), *Principles and Practice of Structural Equation Modeling*, Guilford Press, New York, London, 4th ed.

Law, K. S. and Wong, C.-S. (1999), "Multidimensional constructs in structural equation analysis: An illustration using the job perception and job satisfaction constructs", *Journal of Management*, Vol. 25(2), pp. 143–160.

Lawley, D. N. and Maxwell, A. E. (1962), "Factor analysis as a statistical method", *The Statistician*, Vol. 12(3), pp. 209–229.

Lee, S. M., Kim, S. T., and Choi, D. (2012), "Green supply chain management and organizational performance", *Industrial Management & Data Systems*, Vol. 112(8), pp. 1148–1180.

Leong, L.-Y., Jaafar, N. I., and Sulaiman, A. (2017), "Understanding impulse purchase in Facebook commerce: does Big Five matter?", *Internet Research*, Vol. 27(4), pp. 786–818.

Liengaard, B. D., Sharma, P. N., Hult, G. T. M., Jensen, M. B., Sarstedt, M., Hair, J. F., and Ringle, C. M. (accepted), "Prediction: Coveted, yet forsaken? introducing a cross-validated predictive ability test in partial least squares path modeling", *Decision Sciences*, pp. 1–31.

Lohmöller, J.-B. (1989), *Latent Variable Path Modeling with Partial Least Squares*, Physica, Heidelberg.

Mason, R. and Brown, W. G. (1975), "Multicollinearity problems and ridge regression in sociological models", *Social Science Research*, Vol. 4(2), pp. 135–149.

Mulaik, S. A., James, L. R., Van Alstine, J., Bennett, N., Lind, S., and Stilwell, C. D. (1989), "Evaluation of goodness-of-fit indices for structural equation models", *Psychological Bulletin*, Vol. 105(3), p. 430.

Mulaik, S. A. and Quartetti, D. A. (1997), "First order or higher order general factor?", *Structural Equation Modeling: A Multidisciplinary Journal*, Vol. 4(3), pp. 193–211.

Noonan, R. and Wold, H. (1982), "PLS path modeling with indirectly observed variables: A comparison of alternative estimates for the latent variable", in Jöreskog, K. G. and Wold, H., (Eds.), "Systems under Indirect Observation: Causality, Structure, Prediction Part II", North-Holland, Amsterdam, pp. 75–94.

Pérez-López, S. and Alegre, J. (2012), "Information technology competency, knowledge processes and firm performance", *Industrial Management & Data Systems*, Vol. 112(4), pp. 644–662.

Polites, G. L., Roberts, N., and Thatcher, J. (2012), "Conceptualizing models using multidimensional constructs: A review and guidelines for their use", *European Journal of Information Systems*, Vol. 21(1), pp. 22–48.

R Core Team (2019), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria.

Rademaker, M. and Schuberth, F. (2020), *cSEM: Composite-Based Structural Equation Modeling*, R package version 0.2.0.9000.
URL: *https://github.com/M-E-Rademaker/cSEM*

Ramayah, T., Cheah, J., Chuah, F., Ting, H., and Memon, M. A. (2016), *Partial Least Squares Structural Equation Modeling (PLS-SEM) using SmartPLS 3.0: An Updated Guide and Practical Guide to Statistical Analysis*, Pearson, Singapore, 2nd ed.

Rasoolimanesh, S. M., Noor, S. M., Schuberth, F., and Jaafar, M. (2019), "Investigating the effects of tourist engagement on satisfaction and loyalty", *The Service Industries Journal*, Vol. 39(7-8), pp. 559–574.

Raykov, T. and Penev, S. (1999), "On structural equation model equivalence", *Multivariate Behavioral Research*, Vol. 34(2), pp. 199–244.

Rigdon, E. E. (2012), "Rethinking partial least squares path modeling: In praise of simple methods", *Long Range Planning*, Vol. 45(5), pp. 341–358.

Rigdon, E. E. (2014), "Comment on 'Improper use of endogenous formative variables'", *Journal of Business Research*, Vol. 67(1), pp. 2800–2802.

Rigdon, E. E. (2016), "Choosing PLS path modeling as analytical method in European management research: A realist perspective", *European Management Journal*, Vol. 34(6), pp. 598–605.

Ringle, C. M., Sarstedt, M., and Straub, D. (2012), "A critical look at the use of PLS-SEM in MIS Quarterly", *MIS Quarterly*, Vol. 36(1), pp. iii–xiv.

Ringle, C. M., Wende, S., and Becker, J.-M. (2015), *SmartPLS 3*, http://www.smartpls.com, Bönningstedt.

Rönkkö, M. (2019), *matrixpls: Matrix-based Partial Least Squares Estimation*, R package version 1.0.6.
URL: *https://github.com/mronkko/matrixpls*

Sarstedt, M., Hair, J. F., Cheah, J.-H., Becker, J.-M., and Ringle, C. M. (2019), "How to specify, estimate, and validate higher-order constructs in PLS-SEM", *Australasian Marketing Journal*, Vol. 27(3), pp. 197–211.

Schuberth, F. (forthcoming), "Confirmatory composite analysis using partial least squares: Setting the record straight", *Review of Managerial Science.*

Schuberth, F., Henseler, J., and Dijkstra, T. K. (2018), "Confirmatory composite analysis", *Frontiers in Psychology*, Vol. 9(2541).

Schumacker, R. E. and Lomax, R. G. (2009), *A Beginner's Guide to Structural Equation Modeling*, Routledge, New York, 3rd ed.

Schwarz, G. (1978), "Estimating the dimension of a model", *The Annals of Statistics*, Vol. 6(2), pp. 461–464.

Sharma, P., Sarstedt, M., Shmueli, G., Kim, K., and Thiele, K. (2019), "PLS-based model selection: The role of alternative explanations in information systems research", *Journal of the Association for Information Systems*, Vol. 20(4), pp. 346–397.

Shmueli, G. (2010), "To explain or to predict?", *Statistical Science*, Vol. 25(3), pp. 289–310.

Steiger, J. H. (2007), "Understanding the limitations of global fit assessment in structural equation modeling", *Personality and Individual Differences*, Vol. 42(5), pp. 893–898.

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

Van Riel, A. C. R., Henseler, J., Kemény, I., and Sasovova, Z. (2017), "Estimating hierarchical constructs using partial least squares: The case of second order composites of factors", *Industrial Management & Data Systems*, Vol. 117(3), pp. 459–477.

Venables, W. N. and Ripley, B. D. (2002), *Modern Applied Statistics with S*, Springer, New York, 4th ed.

Vernon, T. and Eysenck, S. (2007), "Introduction", *Personality and Individual Differences*, Vol. 42(5), p. 813.

Wetzels, M., Odekerken-Schröder, G., and Van Oppen, C. (2009), "Using PLS path modeling for assessing hierarchical construct models: Guidelines and empirical illustration", *MIS Quarterly*, Vol. 33(1), pp. 177–195.

Wilson, B. (2010), "Using PLS to investigate interaction effects between higher order branding constructs", in Esposito Vinzi, V., Chin, W. W., Henseler, J., and Wang, H., (Eds.), "Handbook of Partial Least Squares: Concepts, Methods and Applications", Springer, Berlin, pp. 621–652.

Wilson, B. and Henseler, J. (2007), "Modeling reflective higher-order constructs using three approaches with PLS path modeling: A Monte Carlo comparison", in Thyne, M. and Deans, K. R., (Eds.), "Conference Proceedings ANZMAC 2007", pp. 791–800.

Wold, H. (1982), "Soft modeling: The basic design and some extensions", in Jöreskog, K. G. and Wold, H., (Eds.), "Systems under Indirect Observation: Causality, Structure, Prediction Part II", North-Holland, Amsterdam, pp. 1–54.

Yim, B. and Leem, B. (2013), "The effect of the supply chain social capital", *Industrial Management & Data Systems*, Vol. 113(3), pp. 324–349.

Yuan, K.-H. (2005), "Fit indices versus test statistics", *Multivariate Behavioral Research*, Vol. 40(1), pp. 115–148.

Figures
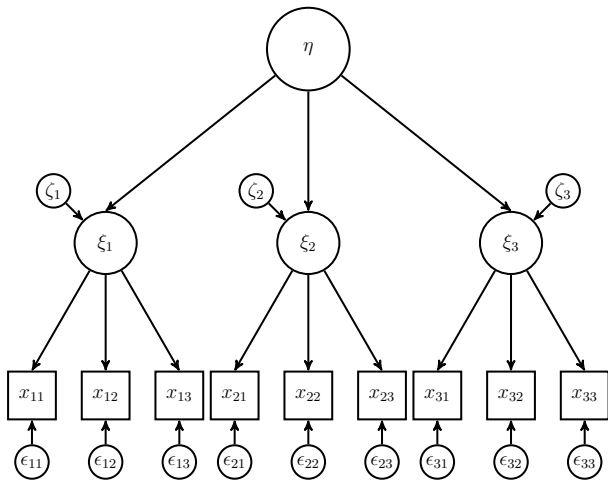
(a) Type I: Common factor of common factors
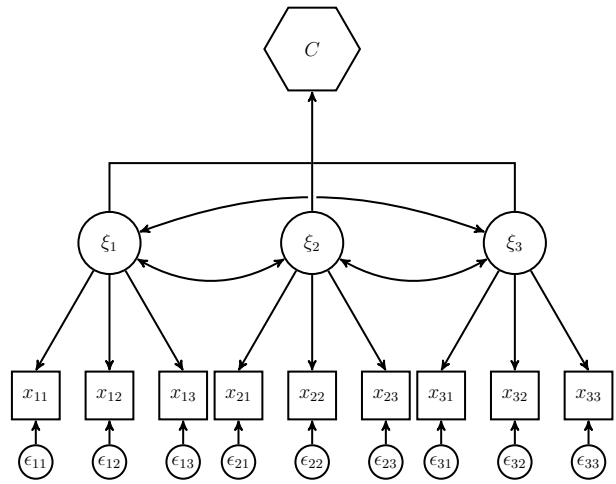
(b) Type II: Composite of common factors
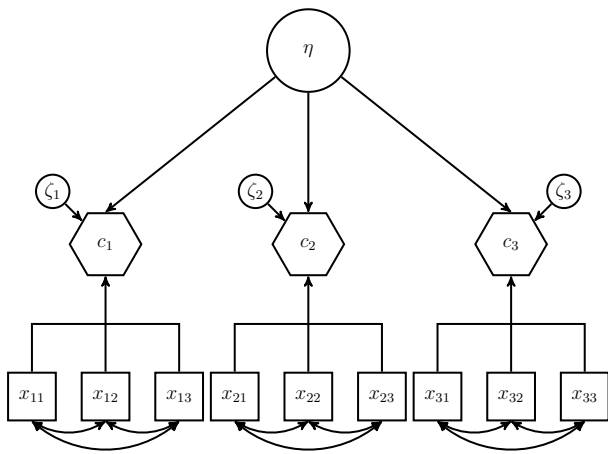
(c) Type III: Common factor of composites
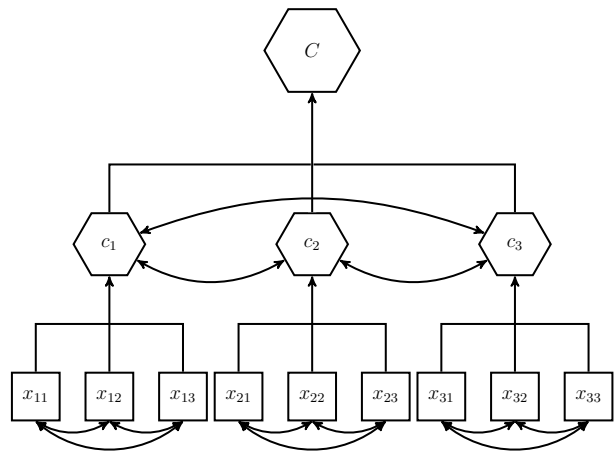
(d) Type IV: Composite of composites

Figure 1: The four main types of second-order constructs
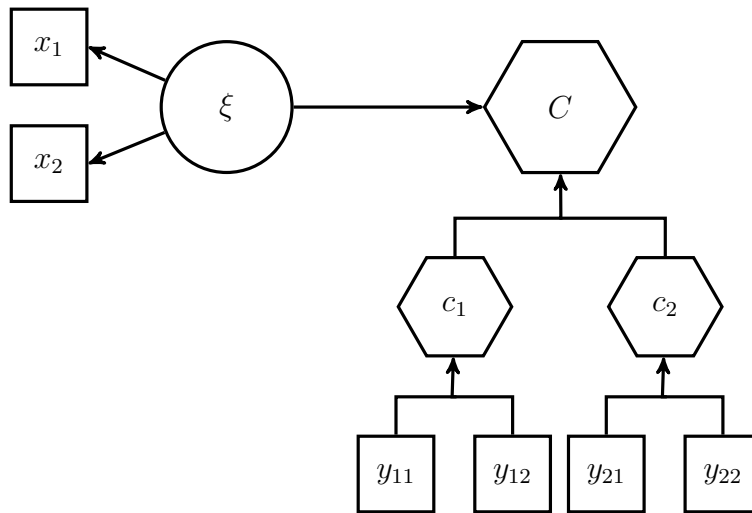
Figure 2: Example of a model containing a second-order composite formed by composites

y_{51} .2617

y_{52} .3926

y_{53} .6543

η_1

y_{41} .4887

y_{42} .3665

y_{43} .3665

ξ

$\gamma_1 = .2$

$\gamma_3 = .35$

$\beta_1 = .4$

η_3

$\gamma_2 = -.4$

$\beta_2 = .2$

y_{61} .6696

y_{62} .3348

y_{63} .3348

η_2

y_{11} y_{12}

.80   .40

c_1

$w_{c_1} = .4$

$w_{c_2} = .4$

c_2

$w_{c_3} = .5$

c_3

.50 y_{21}

.30 y_{22}

.20 y_{23}

.40 y_{24}

.30 .30 .20 .20 .40 .30

y_{31} y_{32} y_{33} y_{34} y_{35} y_{36}
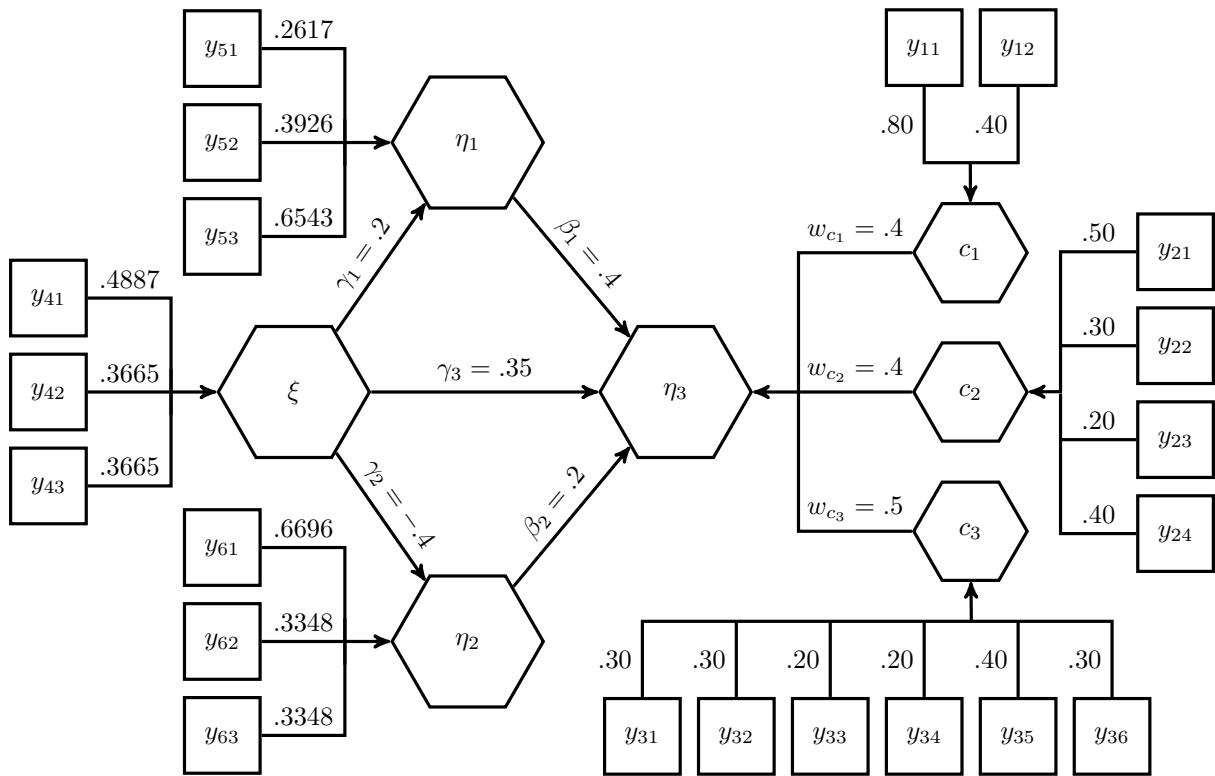
Figure 3: Baseline population model
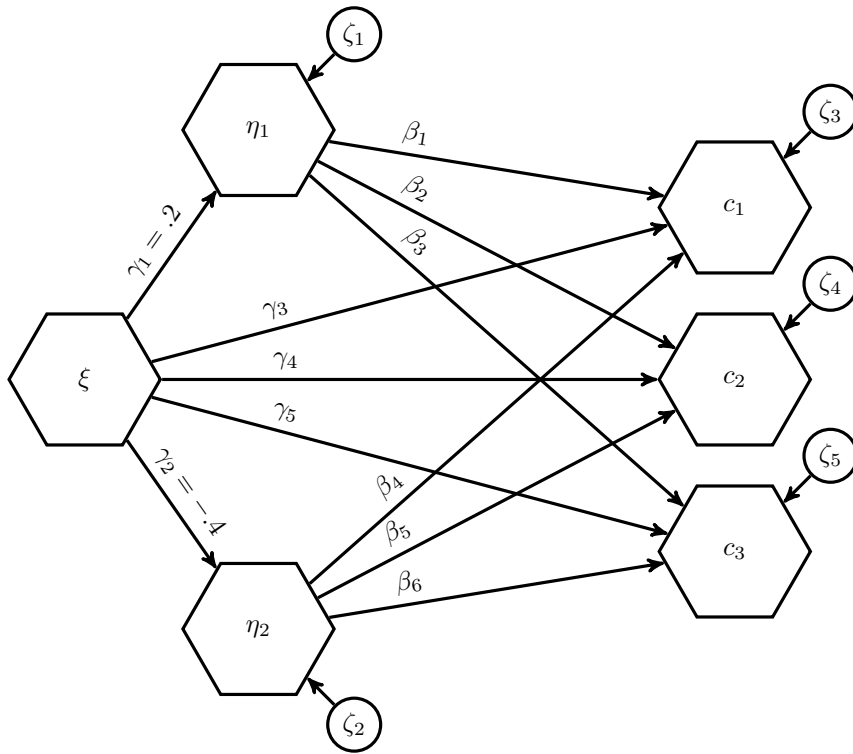
Figure 4: Population structural model in the case of misspecification
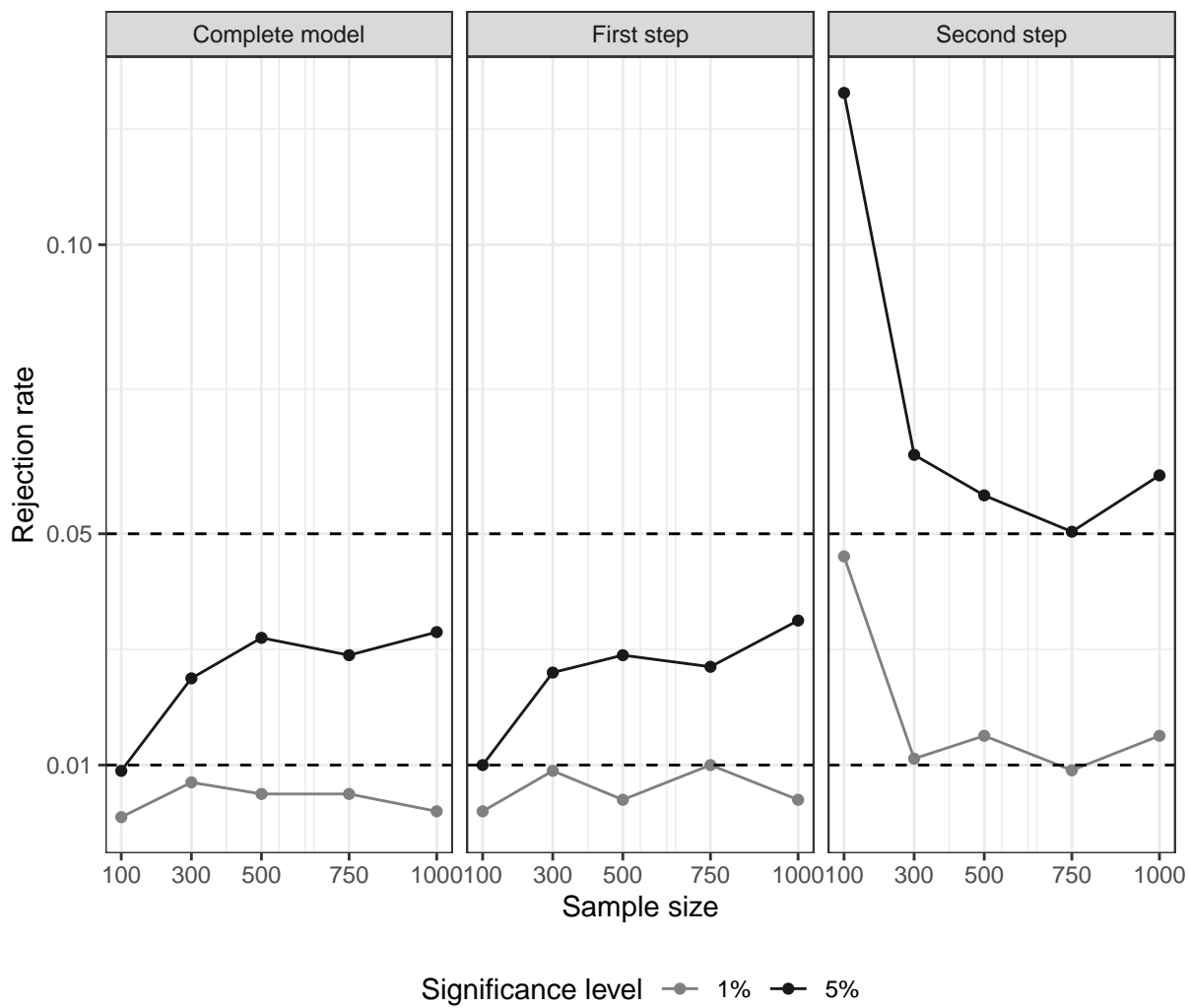
Figure 5: Tests' rejection rates in the case of no misspecification
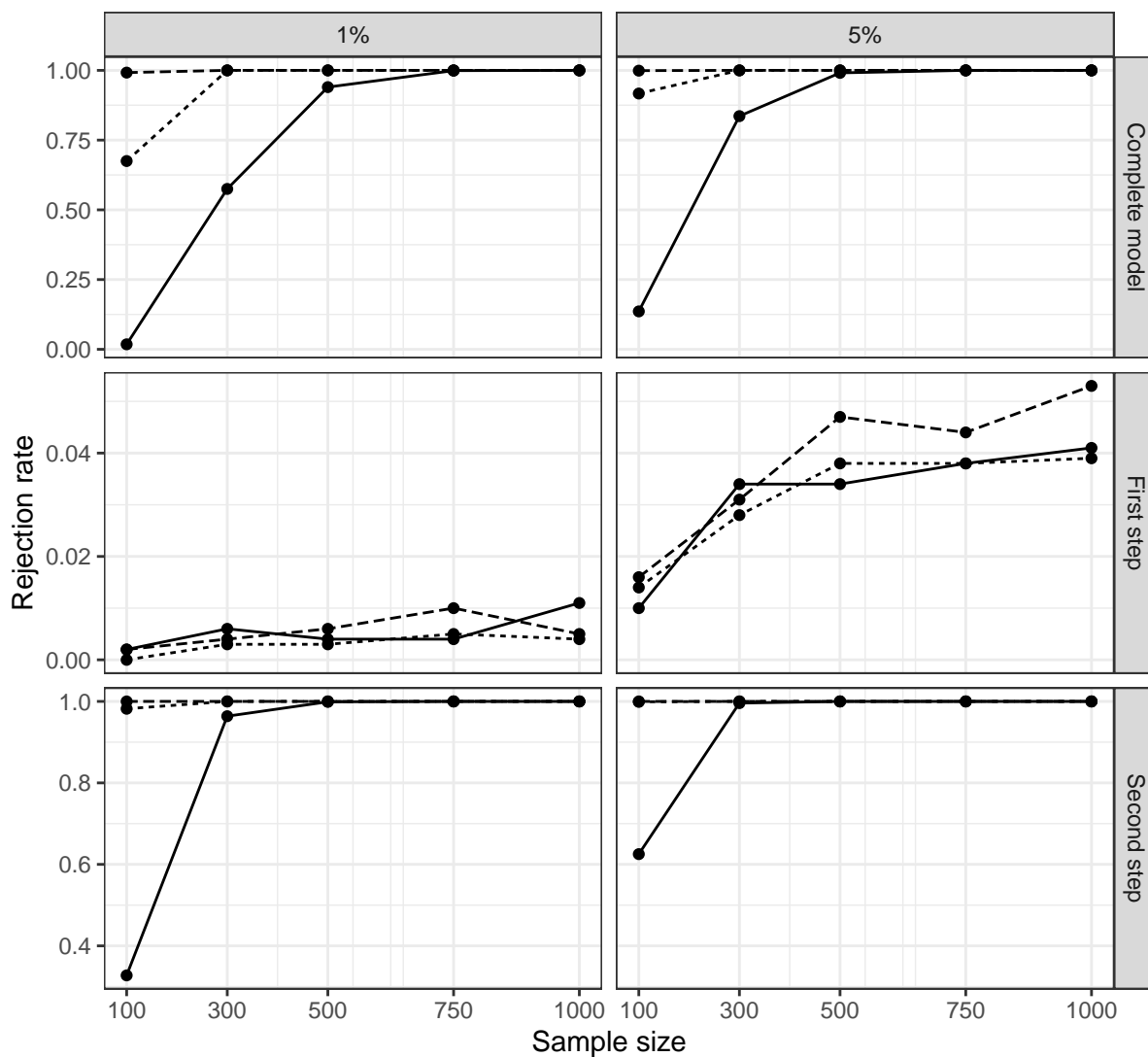
Figure 6: Tests' rejection rates for various degrees of misspecification
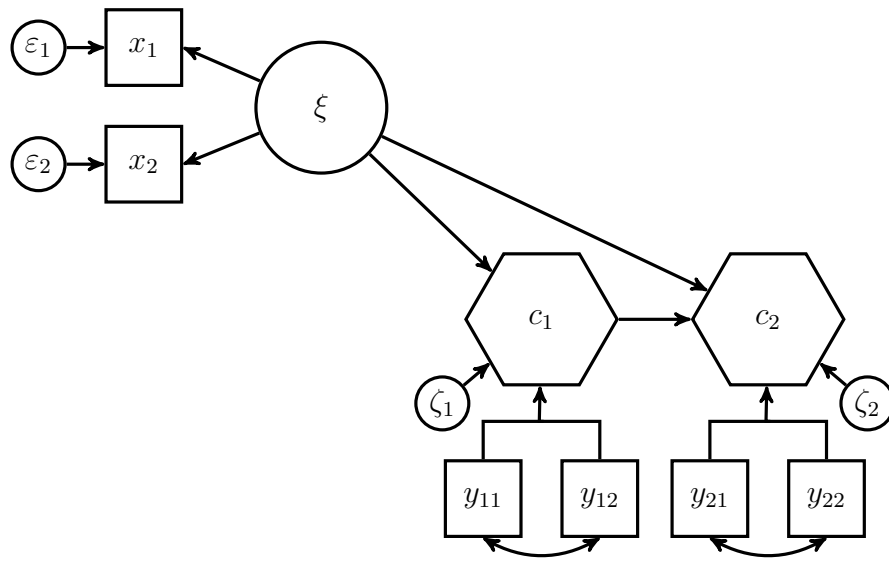
Figure 7: Model specification of the first stage for our example model
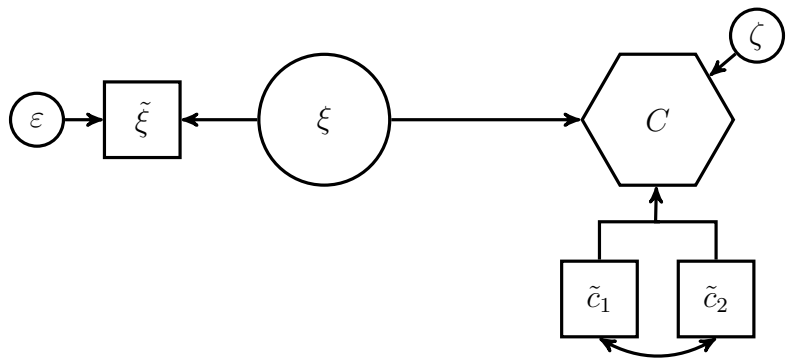
Figure 8: Model specification of the second stage for our example model
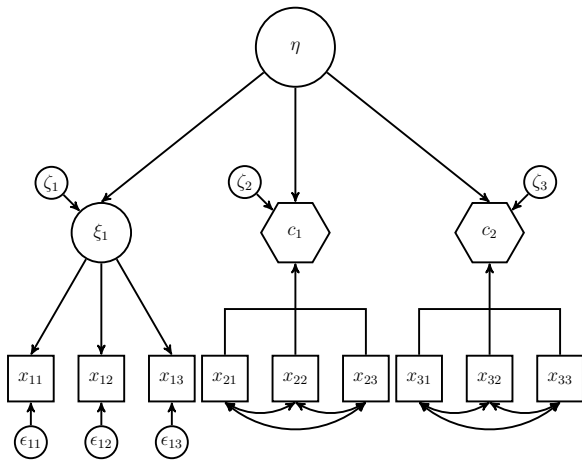
Type V: Common factor of composites and common factors
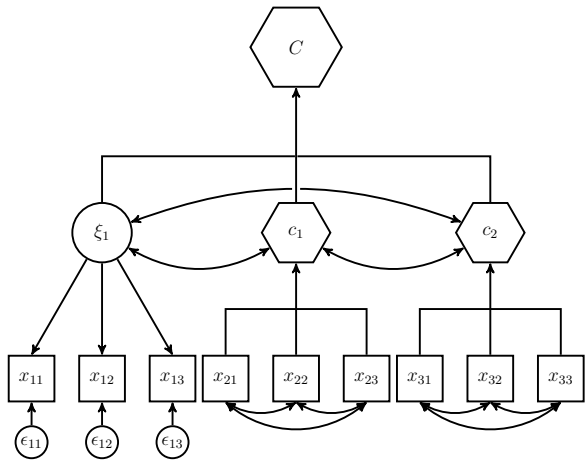
Type VI: Composite of composites and common factors

Figure 9: Specifications of mixed second-order constructs

Tables

Table 1: PLS-PM approaches to estimate models containing composites of composites

	(Extended) Repeated indicators approach (Wold, 1982)	Two-stage approach (Agarwal and Karahanna, 2000)	Embedded two-stage approach (Wilson, 2010)	(Extended) Hybrid approach (Wilson and Henseler, 2007)
Approach:				
Illustration:		First stage: Second stage:	First stage: Second stage:	
Description:	All indicators of the first-order constructs building the second-order construct are used as indicators for the second-order construct. Originally, it was proposed to specify only the direct effect of the antecedent construct on the second-order construct. As a consequence, the second-order construct is fully explained by the first-order constructs c_1 and c_2, and all path from antecedent constructs (here, the path from ξ on C) are estimated towards zero. To address this shortcoming, it was proposed to additionally specify the indirect effects of the antecedent constructs on the second-order construct through the first-order constructs (dotted lines) to estimate the effect of the respective antecedent construct on the second-order construct as the total effects (Becker et al., 2012).	Stage 1: Construct scores ($\tilde{c}_1$, $\tilde{c}_2$, and $\tilde{\xi}$) are calculated based on a (saturated) structural model without the second-order construct. Stage 2: Construct scores are used as indicators to estimate the structural model containing the second-order construct.	Stage 1: Construct scores ($\tilde{c}_1$, $\tilde{c}_2$, and $\tilde{\xi}$) are calculated by the repeated indicators approach. Stage 2: Construct scores are used as indicators to estimate the structural model containing the second-order construct.	Split the sets of original indicators of the first-order constructs in half, such that one half of each set is used as indicators of the first-order constructs, while the other half of each set is used as indicators of the second-order construct. Originally, it was proposed to specify only the direct effect of the antecedent construct on the hierarchical construct. As a consequence, the second-order construct is fully explained by the first-order constructs c_1 and c_2, and all path from antecedent constructs (here, the path from ξ on C) are estimated towards zero. To address this shortcoming, it was proposed to additionally specify the indirect effects of the antecedent constructs on the second-order construct through the first-order constructs (dotted lines) to estimate the effect of the respective antecedent construct on the second-order construct as the total effects (Becker et al., 2012).

Table 2: Population path coefficients in the case of misspecifications

Misspecification	SRMR	SRMR$_{1st}$	SRMR$_{2nd}$	γ_3	γ_4	γ_5	β_1	β_2	β_3	β_4	β_5	β_6
Small	0.0302	0.000	0.0626	0.4	0.1	0.5	0.1	0.5	0.6	0.1	0.2	0.5
Medium	0.0629	0.000	0.1262	-0.4	0.1	0.5	0.1	0.5	0.6	0.2	0.2	0.5
Large	0.0949	0.000	0.1718	0.4	0.1	0.5	0.5	0.4	0.6	-0.4	0.2	0.5

Table 3: Results for the estimated bias, estimated RMSE, Fisher consistency, and number of converged estimations

Inner weighting scheme	Sample size	Approach	% converged estimations	Estimated bias and Fisher consistency							Estimated RMSE						
				γ_1	β_1	w_{11}	w_{c_1}	w_{c_2}	w_{c_3}	κ_2	γ_1	β_1	w_{11}	w_{c_1}	w_{c_2}	w_{c_3}	κ_2
Centroid scheme	100	two-stage	100.00	0.037	0.003	-0.024	-0.055	0.008	-0.016	0.003	0.101	0.081	0.147	0.197	0.209	0.181	0.081
		embedded TS	99.80	0.048	-0.002	-0.021	-0.061	0.000	0.004	-0.002	0.108	0.083	0.149	0.194	0.193	0.158	0.083
		extended RI	100.00	0.033	-0.389	-0.031	-0.000	0.022	-0.045	0.017	0.099	0.389	0.171	0.054	0.063	0.069	0.081
	500	two-stage	100.00	0.005	0.000	-0.000	-0.014	0.009	-0.008	0.000	0.042	0.035	0.057	0.097	0.096	0.085	0.035
		embedded TS	100.00	0.007	-0.002	0.001	-0.015	0.006	-0.002	-0.002	0.042	0.035	0.056	0.090	0.091	0.076	0.035
		extended RI	100.00	0.003	-0.398	-0.001	0.019	0.025	-0.048	0.002	0.042	0.398	0.067	0.032	0.036	0.054	0.035
	1,000	two-stage	100.00	0.004	0.003	-0.005	-0.005	-0.002	0.000	0.003	0.030	0.026	0.042	0.064	0.070	0.060	0.026
		embedded TS	100.00	0.005	0.002	-0.005	-0.006	-0.001	0.001	0.002	0.030	0.026	0.042	0.061	0.067	0.056	0.026
		extended RI	100.00	0.003	-0.399	-0.006	0.022	0.023	-0.048	0.004	0.030	0.399	0.049	0.028	0.030	0.051	0.026
	Σ	two-stage	—	-3.89e-16	5.00e-16	7.77e-16	-1.55e-15	2.11e-15	-1.22e-15	5.00e-16	—	—	—	—	—	—	—
		embedded TS	—	-3.33e-16	-2.27e-11	-1.24e-05	-6.66e-11	1.39e-10	-3.03e-11	-2.27e-11	—	—	—	—	—	—	—
		extended RI	—	-3.61e-16	-4.00e-01	-6.21e-06	2.41e-02	2.41e-02	-4.94e-02	-5.09e-04	—	—	—	—	—	—	—
Factorial scheme	100	two-stage	99.80	0.031	0.014	-0.005	-0.056	-0.001	-0.004	0.014	0.110	0.086	0.134	0.193	0.196	0.180	0.086
		embedded TS	99.50	0.035	0.008	-0.003	-0.054	0.011	-0.021	0.008	0.112	0.086	0.133	0.189	0.193	0.172	0.086
		extended RI	98.70	0.027	-0.388	-0.007	-0.004	0.045	-0.075	0.022	0.107	0.388	0.145	0.047	0.064	0.090	0.086
	500	two-stage	100.00	0.005	0.005	-0.002	-0.010	-0.000	-0.001	0.005	0.042	0.036	0.053	0.079	0.085	0.074	0.036
		embedded TS	100.00	0.006	0.003	-0.002	-0.009	0.002	-0.005	0.003	0.043	0.036	0.055	0.079	0.085	0.073	0.036
		extended RI	100.00	0.004	-0.398	-0.002	0.015	0.055	-0.081	0.005	0.042	0.398	0.056	0.025	0.059	0.083	0.036
	1,000	two-stage	100.00	0.004	0.002	-0.000	-0.005	-0.000	-0.001	0.002	0.030	0.025	0.037	0.054	0.056	0.051	0.025
		embedded TS	100.00	0.004	0.001	0.000	-0.005	0.001	-0.004	0.001	0.030	0.025	0.038	0.054	0.056	0.052	0.025
		extended RI	100.00	0.003	-0.399	-0.000	0.017	0.056	-0.082	0.001	0.030	0.399	0.041	0.022	0.057	0.083	0.025
	Σ	two-stage	—	-6.66e-16	2.78e-16	5.55e-16	6.66e-16	1.44e-15	-2.33e-15	2.78e-16	—	—	—	—	—	—	—
		embedded TS	—	-6.94e-16	-1.44e-11	-8.52e-06	-1.75e-12	3.05e-11	5.66e-12	-1.44e-11	—	—	—	—	—	—	—
		extended RI	—	-6.94e-16	-4.00e-01	-5.33e-06	1.97e-02	5.69e-02	-8.20e-02	-1.34e-03	—	—	—	—	—	—	—
Path scheme	100	two-stage	99.40	0.022	0.031	-0.008	-0.045	-0.024	0.009	0.031	0.113	0.093	0.128	0.181	0.196	0.178	0.093
		embedded TS	89.10	0.031	-0.002	-0.016	-0.022	0.003	-0.051	-0.002	0.113	0.087	0.152	0.179	0.193	0.181	0.087
		extended RI	98.20	0.023	-0.400	-0.010	-0.153	0.045	0.058	0.012	0.105	0.400	0.137	0.156	0.057	0.064	0.085
	500	two-stage	100.00	0.005	0.008	-0.005	-0.011	-0.001	0.002	0.008	0.045	0.037	0.057	0.077	0.075	0.065	0.037
		embedded TS	100.00	0.007	0.000	-0.005	-0.006	0.005	-0.011	0.000	0.045	0.036	0.068	0.077	0.076	0.067	0.036
		extended RI	100.00	0.005	-0.400	-0.005	-0.144	0.050	0.069	0.001	0.045	0.400	0.060	0.144	0.051	0.070	0.036
	1,000	two-stage	100.00	0.000	0.004	-0.003	-0.001	-0.003	-0.001	0.004	0.030	0.025	0.039	0.052	0.056	0.048	0.025
		embedded TS	100.00	0.001	0.000	-0.002	0.001	0.000	-0.007	0.000	0.030	0.025	0.047	0.052	0.056	0.049	0.025
		extended RI	100.00	0.000	-0.400	-0.003	-0.143	0.050	0.069	-0.001	0.030	0.400	0.042	0.143	0.051	0.070	0.025
	Σ	two-stage	—	-6.94e-16	2.78e-16	0.00e+00	9.44e-16	-1.67e-16	-1.11e-15	2.78e-16	—	—	—	—	—	—	—
		embedded TS	—	-6.94e-16	-1.05e-10	-4.58e-10	2.18e-10	5.74e-10	-6.17e-10	-1.05e-10	—	—	—	—	—	—	—
		extended RI	—	-6.94e-16	-4.00e-01	-1.35e-09	-1.42e-01	5.05e-02	7.01e-02	-3.63e-03	—	—	—	—	—	—	—

Note: embedded TS: embedded two-stage approach; extended RI: extended repeated indicators approach; two-stage: two-stage approach based on a saturated structural model in the first stage; Σ: the indicator population correlation matrix is provided as input; κ_2: the total effect of η_1 on η_3.

Table 4: Guidelines for estimating and assessing models containing second-order composites made of composites

Stage 1
- Step 1: Estimate the model without second-order composites
- Step 2: Validate the first-order constructs by means of overall model fit
- Step 3: Extract construct scores

Stage 2
- Step 4: Estimate the structural model containing second-order composites
- Step 5: Validate the second-order composites by means of overall model fit

Online Supplementary Material

Table 5: Indicator population correlation matrix of the baseline model

	y_{11}	y_{12}	y_{21}	y_{22}	y_{23}	y_{24}	y_{31}	y_{32}	y_{33}	y_{34}	y_{35}	y_{36}	y_{41}	y_{42}	y_{43}	y_{51}	y_{52}	y_{53}	y_{61}	y_{62}	y_{63}
	1.00	0.31	0.36	0.31	0.25	0.32	0.13	0.16	0.13	0.14	0.15	0.15	0.19	0.18	0.18	0.16	0.22	0.27	0.02	0.01	0.01
	0.31	1.00	0.26	0.22	0.18	0.23	0.09	0.11	0.09	0.10	0.11	0.10	0.14	0.12	0.12	0.12	0.15	0.19	0.01	0.01	0.01
	0.36	0.26	1.00	0.40	0.30	0.31	0.18	0.21	0.18	0.19	0.20	0.20	0.19	0.17	0.17	0.16	0.21	0.26	0.02	0.01	0.01
	0.31	0.22	0.40	1.00	0.28	0.31	0.15	0.18	0.15	0.16	0.17	0.17	0.16	0.14	0.14	0.13	0.18	0.22	0.01	0.01	0.01
	0.25	0.18	0.30	0.28	1.00	0.30	0.12	0.15	0.12	0.13	0.14	0.14	0.13	0.12	0.12	0.11	0.14	0.18	0.01	0.01	0.01
	0.32	0.23	0.31	0.31	0.30	1.00	0.16	0.19	0.15	0.16	0.18	0.17	0.16	0.15	0.15	0.14	0.18	0.23	0.01	0.01	0.01
	0.13	0.09	0.18	0.15	0.12	0.16	1.00	0.10	0.25	0.13	0.10	0.30	0.12	0.11	0.11	0.10	0.13	0.17	0.01	0.01	0.01
	0.16	0.11	0.21	0.18	0.15	0.19	0.10	1.00	0.20	0.40	0.30	0.20	0.14	0.13	0.13	0.12	0.16	0.20	0.01	0.01	0.01
	0.13	0.09	0.18	0.15	0.12	0.15	0.25	0.20	1.00	0.30	0.10	0.30	0.12	0.11	0.11	0.10	0.13	0.17	0.01	0.01	0.01
	0.14	0.10	0.19	0.16	0.13	0.16	0.13	0.40	0.30	1.00	0.20	0.20	0.13	0.11	0.11	0.11	0.14	0.18	0.01	0.01	0.01
	0.15	0.11	0.20	0.17	0.14	0.18	0.10	0.30	0.10	0.20	1.00	0.10	0.14	0.12	0.12	0.12	0.15	0.19	0.01	0.01	0.01
	0.15	0.10	0.20	0.17	0.14	0.17	0.30	0.20	0.30	0.20	0.10	1.00	0.13	0.12	0.12	0.11	0.15	0.19	0.01	0.01	0.01
	0.19	0.14	0.19	0.16	0.13	0.16	0.12	0.14	0.12	0.13	0.14	0.13	1.00	0.40	0.40	0.09	0.12	0.15	−0.29	−0.18	−0.22
	0.18	0.12	0.17	0.14	0.12	0.15	0.11	0.13	0.11	0.11	0.12	0.12	0.40	1.00	0.40	0.08	0.11	0.13	−0.26	−0.17	−0.20
	0.18	0.12	0.17	0.14	0.12	0.15	0.11	0.13	0.11	0.11	0.12	0.12	0.40	0.40	1.00	0.08	0.11	0.13	−0.26	−0.17	−0.20
	0.16	0.12	0.16	0.13	0.11	0.14	0.10	0.12	0.10	0.11	0.12	0.11	0.09	0.08	0.08	1.00	0.20	0.30	−0.04	−0.02	−0.03
	0.22	0.15	0.21	0.18	0.14	0.18	0.13	0.16	0.13	0.14	0.15	0.15	0.12	0.11	0.11	0.20	1.00	0.40	−0.05	−0.03	−0.04
	0.27	0.19	0.26	0.22	0.18	0.23	0.17	0.20	0.17	0.18	0.19	0.19	0.15	0.13	0.13	0.30	0.40	1.00	−0.06	−0.04	−0.05
	0.02	0.01	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	−0.29	−0.26	−0.26	−0.04	−0.05	−0.06	1.00	0.25	0.40
	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	−0.18	−0.17	−0.17	−0.02	−0.03	−0.04	0.25	1.00	0.16
	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	−0.22	−0.20	−0.20	−0.03	−0.04	−0.05	0.40	0.16	1.00

Table 6: Indicator population correlation matrix in the case of the small model misspecification

	y_{11}	y_{12}	y_{21}	y_{22}	y_{23}	y_{24}	y_{31}	y_{32}	y_{33}	y_{34}	y_{35}	y_{36}	y_{41}	y_{42}	y_{43}	y_{51}	y_{52}	y_{53}	y_{61}	y_{62}	y_{63}
	1.00	0.31	0.22	0.18	0.15	0.19	0.11	0.13	0.11	0.11	0.12	0.12	0.17	0.15	0.15	0.07	0.09	0.12	0.27	0.17	0.20
	0.31	1.00	0.15	0.13	0.10	0.13	0.08	0.09	0.07	0.08	0.09	0.08	0.12	0.11	0.11	0.05	0.06	0.08	0.19	0.12	0.14
	0.22	0.15	1.00	0.40	0.30	0.31	0.15	0.18	0.15	0.16	0.17	0.17	−0.03	−0.02	−0.02	0.20	0.27	0.34	0.37	0.23	0.27
	0.18	0.13	0.40	1.00	0.28	0.31	0.13	0.15	0.13	0.13	0.15	0.14	−0.02	−0.02	−0.02	0.17	0.23	0.29	0.31	0.20	0.23
	0.15	0.10	0.30	0.28	1.00	0.30	0.10	0.12	0.10	0.11	0.12	0.12	−0.02	−0.02	−0.02	0.14	0.18	0.23	0.26	0.16	0.19
	0.19	0.13	0.31	0.31	0.30	1.00	0.13	0.16	0.13	0.14	0.15	0.15	−0.02	−0.02	−0.02	0.18	0.24	0.30	0.33	0.20	0.24
	0.11	0.08	0.15	0.13	0.10	0.13	1.00	0.10	0.25	0.13	0.10	0.30	−0.03	−0.02	−0.02	0.05	0.07	0.09	0.21	0.13	0.16
	0.13	0.09	0.18	0.15	0.12	0.16	0.10	1.00	0.20	0.40	0.30	0.20	−0.03	−0.03	−0.03	0.06	0.08	0.10	0.25	0.16	0.18
	0.11	0.07	0.15	0.13	0.10	0.13	0.25	0.20	1.00	0.30	0.10	0.30	−0.03	−0.02	−0.02	0.05	0.07	0.08	0.21	0.13	0.15
	0.11	0.08	0.16	0.13	0.11	0.14	0.13	0.40	0.30	1.00	0.20	0.20	−0.03	−0.03	−0.03	0.05	0.07	0.09	0.22	0.14	0.16
	0.12	0.09	0.17	0.15	0.12	0.15	0.10	0.30	0.10	0.20	1.00	0.10	−0.03	−0.03	−0.03	0.06	0.08	0.10	0.24	0.15	0.18
	0.12	0.08	0.17	0.14	0.12	0.15	0.30	0.20	0.30	0.20	0.10	1.00	−0.03	−0.03	−0.03	0.06	0.08	0.10	0.23	0.15	0.17
	0.17	0.12	−0.03	−0.02	−0.02	−0.02	−0.03	−0.03	−0.03	−0.03	−0.03	−0.03	1.00	0.40	0.40	0.09	0.12	0.15	−0.29	−0.18	−0.22
	0.15	0.11	−0.02	−0.02	−0.02	−0.02	−0.02	−0.03	−0.02	−0.03	−0.03	−0.03	0.40	1.00	0.40	0.08	0.11	0.13	−0.26	−0.17	−0.20
	0.15	0.11	−0.02	−0.02	−0.02	−0.02	−0.02	−0.03	−0.02	−0.03	−0.03	−0.03	0.40	0.40	1.00	0.08	0.11	0.13	−0.26	−0.17	−0.20
	0.07	0.05	0.20	0.17	0.14	0.18	0.05	0.06	0.05	0.05	0.06	0.06	0.09	0.08	0.08	1.00	0.20	0.30	−0.04	−0.02	−0.03
	0.09	0.06	0.27	0.23	0.18	0.24	0.07	0.08	0.07	0.07	0.08	0.08	0.12	0.11	0.11	0.20	1.00	0.40	−0.05	−0.03	−0.04
	0.12	0.08	0.34	0.29	0.23	0.30	0.09	0.10	0.08	0.09	0.10	0.10	0.15	0.13	0.13	0.30	0.40	1.00	−0.06	−0.04	−0.05
	0.27	0.19	0.37	0.31	0.26	0.33	0.21	0.25	0.21	0.22	0.24	0.23	−0.29	−0.26	−0.26	−0.04	−0.05	−0.06	1.00	0.25	0.40
	0.17	0.12	0.23	0.20	0.16	0.20	0.13	0.16	0.13	0.14	0.15	0.15	−0.18	−0.17	−0.17	−0.02	−0.03	−0.04	0.25	1.00	0.16
	0.20	0.14	0.27	0.23	0.19	0.24	0.16	0.18	0.15	0.16	0.18	0.17	−0.22	−0.20	−0.20	−0.03	−0.04	−0.05	0.40	0.16	1.00

Table 7: Indicator population correlation matrix in the case of the medium model misspecification

	y_{11}	y_{12}	y_{21}	y_{22}	y_{23}	y_{24}	y_{31}	y_{32}	y_{33}	y_{34}	y_{35}	y_{36}	y_{41}	y_{42}	y_{43}	y_{51}	y_{52}	y_{53}	y_{61}	y_{62}	y_{63}
	1.00	0.31	0.24	0.20	0.17	0.21	0.10	0.12	0.10	0.11	0.12	0.11	−0.44	−0.40	−0.40	−0.01	−0.01	−0.02	0.54	0.34	0.40
	0.31	1.00	0.17	0.14	0.12	0.15	0.07	0.08	0.07	0.08	0.08	0.08	−0.31	−0.28	−0.28	−0.01	−0.01	−0.01	0.38	0.24	0.28
	0.24	0.17	1.00	0.40	0.30	0.31	0.15	0.18	0.15	0.16	0.17	0.17	−0.03	−0.02	−0.02	0.20	0.27	0.34	0.37	0.23	0.27
	0.20	0.14	0.40	1.00	0.28	0.31	0.13	0.15	0.12	0.13	0.14	0.14	−0.02	−0.02	−0.02	0.17	0.23	0.29	0.31	0.20	0.23
	0.17	0.12	0.30	0.28	1.00	0.30	0.10	0.12	0.10	0.11	0.12	0.11	−0.02	−0.02	−0.02	0.14	0.18	0.23	0.26	0.16	0.19
	0.21	0.15	0.31	0.31	0.30	1.00	0.13	0.16	0.13	0.14	0.15	0.15	−0.02	−0.02	−0.02	0.18	0.24	0.30	0.33	0.20	0.24
	0.10	0.07	0.15	0.13	0.10	0.13	1.00	0.10	0.25	0.13	0.10	0.30	0.02	0.02	0.02	0.06	0.08	0.10	0.19	0.12	0.14
	0.12	0.08	0.18	0.15	0.12	0.16	0.10	1.00	0.20	0.40	0.30	0.20	0.02	0.02	0.02	0.07	0.09	0.11	0.23	0.14	0.17
	0.10	0.07	0.15	0.12	0.10	0.13	0.25	0.20	1.00	0.30	0.10	0.30	0.02	0.02	0.02	0.06	0.07	0.09	0.19	0.12	0.14
	0.11	0.08	0.16	0.13	0.11	0.14	0.13	0.40	0.30	1.00	0.20	0.20	0.02	0.02	0.02	0.06	0.08	0.10	0.20	0.13	0.15
	0.12	0.08	0.17	0.14	0.12	0.15	0.10	0.30	0.10	0.20	1.00	0.10	0.02	0.02	0.02	0.07	0.09	0.11	0.22	0.14	0.16
	0.11	0.08	0.17	0.14	0.11	0.15	0.30	0.20	0.30	0.20	0.10	1.00	0.02	0.02	0.02	0.06	0.08	0.11	0.21	0.13	0.16
	−0.44	−0.31	−0.03	−0.02	−0.02	−0.02	0.02	0.02	0.02	0.02	0.02	0.02	1.00	0.40	0.40	0.09	0.12	0.15	−0.29	−0.18	−0.22
	−0.40	−0.28	−0.02	−0.02	−0.02	−0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.40	1.00	0.40	0.08	0.11	0.13	−0.26	−0.17	−0.20
	−0.40	−0.28	−0.02	−0.02	−0.02	−0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.40	0.40	1.00	0.08	0.11	0.13	−0.26	−0.17	−0.20
	−0.01	−0.01	0.20	0.17	0.14	0.18	0.06	0.07	0.06	0.06	0.07	0.06	0.09	0.08	0.08	1.00	0.20	0.30	−0.04	−0.02	−0.03
	−0.01	−0.01	0.27	0.23	0.18	0.24	0.08	0.09	0.07	0.08	0.09	0.08	0.12	0.11	0.11	0.20	1.00	0.40	−0.05	−0.03	−0.04
	−0.02	−0.01	0.34	0.29	0.23	0.30	0.10	0.11	0.09	0.10	0.11	0.11	0.15	0.13	0.13	0.30	0.40	1.00	−0.06	−0.04	−0.05
	0.54	0.38	0.37	0.31	0.26	0.33	0.19	0.23	0.19	0.20	0.22	0.21	−0.29	−0.26	−0.26	−0.04	−0.05	−0.06	1.00	0.25	0.40
	0.34	0.24	0.23	0.20	0.16	0.20	0.12	0.14	0.12	0.13	0.14	0.13	−0.18	−0.17	−0.17	−0.02	−0.03	−0.04	0.25	1.00	0.16
	0.40	0.28	0.27	0.23	0.19	0.24	0.14	0.17	0.14	0.15	0.16	0.16	−0.22	−0.20	−0.20	−0.03	−0.04	−0.05	0.40	0.16	1.00

Table 8: Indicator population correlation matrix in the case of the large model misspecification

y_{11}	y_{12}	y_{21}	y_{22}	y_{23}	y_{24}	y_{31}	y_{32}	y_{33}	y_{34}	y_{35}	y_{36}	y_{41}	y_{42}	y_{43}	y_{51}	y_{52}	y_{53}	y_{61}	y_{62}	y_{63}
1.00	0.31	0.27	0.23	0.19	0.24	0.05	0.06	0.05	0.06	0.06	0.06	0.17	0.15	0.15	0.07	0.09	0.12	0.27	0.17	0.20
0.31	1.00	0.19	0.16	0.13	0.17	0.04	0.04	0.04	0.04	0.04	0.04	0.12	0.11	0.11	0.05	0.06	0.08	0.19	0.12	0.14
0.27	0.19	1.00	0.40	0.30	0.31	0.06	0.07	0.06	0.06	0.07	0.07	0.23	0.20	0.20	0.19	0.26	0.32	0.26	0.16	0.19
0.23	0.16	0.40	1.00	0.28	0.31	0.05	0.06	0.05	0.05	0.06	0.06	0.19	0.17	0.17	0.16	0.22	0.27	0.22	0.14	0.16
0.19	0.13	0.30	0.28	1.00	0.30	0.04	0.05	0.04	0.04	0.05	0.05	0.16	0.14	0.14	0.13	0.18	0.22	0.18	0.11	0.13
0.24	0.17	0.31	0.31	0.30	1.00	0.05	0.06	0.05	0.06	0.06	0.06	0.20	0.18	0.18	0.17	0.23	0.28	0.23	0.14	0.17
0.05	0.04	0.06	0.05	0.04	0.05	1.00	0.10	0.25	0.13	0.10	0.30	−0.25	−0.23	−0.23	0.02	0.03	0.04	0.31	0.19	0.23
0.06	0.04	0.07	0.06	0.05	0.06	0.10	1.00	0.20	0.40	0.30	0.20	−0.29	−0.26	−0.26	0.03	0.04	0.05	0.36	0.23	0.27
0.05	0.04	0.06	0.05	0.04	0.05	0.25	0.20	1.00	0.30	0.10	0.30	−0.24	−0.22	−0.22	0.02	0.03	0.04	0.30	0.19	0.22
0.06	0.04	0.06	0.05	0.04	0.06	0.13	0.40	0.30	1.00	0.20	0.20	−0.26	−0.24	−0.24	0.02	0.03	0.04	0.32	0.20	0.24
0.06	0.04	0.07	0.06	0.05	0.06	0.10	0.30	0.10	0.20	1.00	0.10	−0.28	−0.26	−0.26	0.03	0.03	0.04	0.35	0.22	0.26
0.06	0.04	0.07	0.06	0.05	0.06	0.30	0.20	0.30	0.20	0.10	1.00	−0.27	−0.25	−0.25	0.03	0.03	0.04	0.34	0.21	0.25
0.17	0.12	0.23	0.19	0.16	0.20	−0.25	−0.29	−0.24	−0.26	−0.28	−0.27	1.00	0.40	0.40	0.09	0.12	0.15	−0.29	−0.18	−0.22
0.15	0.11	0.20	0.17	0.14	0.18	−0.23	−0.26	−0.22	−0.24	−0.26	−0.25	0.40	1.00	0.40	0.08	0.11	0.13	−0.26	−0.17	−0.20
0.15	0.11	0.20	0.17	0.14	0.18	−0.23	−0.26	−0.22	−0.24	−0.26	−0.25	0.40	0.40	1.00	0.08	0.11	0.13	−0.26	−0.17	−0.20
0.07	0.05	0.19	0.16	0.13	0.17	0.02	0.03	0.02	0.02	0.03	0.03	0.09	0.08	0.08	1.00	0.20	0.30	−0.04	−0.02	−0.03
0.09	0.06	0.26	0.22	0.18	0.23	0.03	0.04	0.03	0.03	0.03	0.03	0.12	0.11	0.11	0.20	1.00	0.40	−0.05	−0.03	−0.04
0.12	0.08	0.32	0.27	0.22	0.28	0.04	0.05	0.04	0.04	0.04	0.04	0.15	0.13	0.13	0.30	0.40	1.00	−0.06	−0.04	−0.05
0.27	0.19	0.26	0.22	0.18	0.23	0.31	0.36	0.30	0.32	0.35	0.34	−0.29	−0.26	−0.26	−0.04	−0.05	−0.06	1.00	0.25	0.40
0.17	0.12	0.16	0.14	0.11	0.14	0.19	0.23	0.19	0.20	0.22	0.21	−0.18	−0.17	−0.17	−0.02	−0.03	−0.04	0.25	1.00	0.16
0.20	0.14	0.19	0.16	0.13	0.17	0.23	0.27	0.22	0.24	0.26	0.25	−0.22	−0.20	−0.20	−0.03	−0.04	−0.05	0.40	0.16	1.00

Table 9: Approaches' Fisher consistency

Parameter	Centroid scheme			Factorial scheme			Path scheme		
	two-stage	embedded TS	extended RI	two-stage	embedded TS	extended RI	two-stage	embedded TS	extended RI
γ_1	-3.89e-16	-3.33e-16	-3.61e-16	-6.66e-16	-6.94e-16	-6.94e-16	-6.94e-16	-6.94e-16	-6.94e-16
γ_2	6.66e-16	6.11e-16	5.55e-16	-2.22e-16	-1.11e-16	-2.22e-16	-1.11e-16	-1.11e-16	-1.67e-16
γ_3	-7.77e-16	-1.98e-11	-3.50e-01	-1.67e-16	-1.26e-11	-3.50e-01	0.00e+00	-9.16e-11	-3.50e-01
β_1	5.00e-16	-2.27e-11	-4.00e-01	2.78e-16	-1.44e-11	-4.00e-01	2.78e-16	-1.05e-10	-4.00e-01
β_2	-1.39e-16	-1.13e-11	-2.00e-01	-8.33e-17	-7.22e-12	-2.00e-01	5.55e-17	-5.24e-11	-2.00e-01
w_{11}	7.77e-16	-1.24e-05	-6.21e-06	5.55e-16	-8.52e-06	-5.33e-06	0.00e+00	-4.58e-10	-1.35e-09
w_{12}	-8.33e-16	1.77e-05	8.83e-06	-4.44e-16	1.21e-05	7.59e-06	0.00e+00	6.52e-10	1.92e-09
w_{21}	-3.33e-16	-4.20e-06	-2.80e-06	3.89e-16	-7.99e-06	-5.29e-06	9.99e-16	-3.14e-07	-3.77e-07
w_{22}	5.00e-16	1.98e-06	1.32e-06	4.44e-16	3.78e-06	2.50e-06	1.67e-16	1.48e-07	1.78e-07
w_{23}	-5.55e-16	5.07e-06	3.39e-06	-5.83e-16	9.66e-06	6.39e-06	-1.30e-15	3.80e-07	4.56e-07
w_{24}	-5.55e-17	-1.11e-06	-7.40e-07	-3.89e-16	-2.11e-06	-1.39e-06	-3.33e-16	-8.29e-08	-9.94e-08
w_{31}	-3.33e-16	-1.28e-06	-4.71e-07	-2.22e-15	-9.20e-07	-3.60e-07	-8.33e-16	-3.96e-06	-1.46e-06
w_{32}	-1.67e-16	-1.28e-06	-4.71e-07	1.61e-15	-9.20e-07	-3.60e-07	1.11e-15	-3.96e-06	-1.46e-06
w_{33}	1.11e-16	8.07e-06	2.96e-06	-8.33e-16	5.78e-06	2.26e-06	-1.30e-15	2.48e-05	9.19e-06
w_{34}	2.66e-15	8.07e-06	2.96e-06	3.61e-15	5.78e-06	2.26e-06	1.17e-15	2.48e-05	9.19e-06
w_{35}	-1.17e-15	-1.06e-05	-3.90e-06	-1.39e-15	-7.62e-06	-2.98e-06	2.78e-16	-3.28e-05	-1.21e-05
w_{36}	-1.17e-15	-1.28e-06	-4.71e-07	-8.88e-16	-9.20e-07	-3.60e-07	-4.44e-16	-3.96e-06	-1.46e-06
w_{41}	4.44e-16	4.44e-16	2.22e-16	3.33e-16	6.66e-16	0.00e+00	-5.55e-16	-3.33e-16	-6.66e-16
w_{42}	-5.00e-16	-3.33e-16	-6.11e-16	6.11e-16	5.55e-17	5.00e-16	-6.11e-16	-9.44e-16	-7.22e-16
w_{43}	1.11e-16	1.67e-16	6.66e-16	-8.88e-16	-8.88e-16	-5.55e-16	1.11e-15	1.28e-15	1.39e-15
w_{51}	2.00e-15	2.16e-15	1.67e-15	1.05e-15	1.33e-15	1.28e-15	7.77e-16	1.17e-15	1.11e-15
w_{52}	-2.16e-15	-1.28e-15	-1.94e-15	-1.44e-15	-1.28e-15	-1.94e-15	-2.11e-15	-1.89e-15	-2.05e-15
w_{53}	3.33e-16	-3.33e-16	2.22e-16	3.33e-16	-1.11e-16	6.66e-16	1.33e-15	7.77e-16	9.99e-16
w_{61}	-3.11e-15	-2.89e-15	-7.88e-15	1.11e-15	7.77e-16	9.99e-16	-1.11e-16	-6.66e-16	-3.33e-16
w_{62}	3.66e-15	3.89e-15	6.16e-15	-7.22e-16	-1.11e-16	-5.00e-16	-1.72e-15	-1.44e-15	-1.78e-15
w_{63}	1.61e-15	1.17e-15	6.16e-15	-7.77e-16	-4.44e-16	-5.55e-16	2.22e-15	2.61e-15	2.61e-15
w_{c1}	-1.55e-15	-6.66e-11	—	6.66e-16	-1.75e-12	—	9.44e-16	2.18e-10	—
w_{c2}	2.11e-15	1.39e-10	—	1.44e-15	3.05e-11	—	-1.67e-16	5.74e-10	—
w_{c3}	-1.22e-15	-3.03e-11	—	-2.33e-15	5.66e-12	—	-1.11e-15	-6.17e-10	—
κ_2	5.00e-16	-2.27e-11	-5.09e-04	2.78e-16	-1.44e-11	-1.34e-03	2.78e-16	-1.05e-10	-3.63e-03
κ_3	-1.39e-16	-1.13e-11	-2.55e-04	-8.33e-17	-7.22e-12	-6.72e-04	5.55e-17	-5.24e-11	-1.81e-03
κ_1	-3.89e-16	-1.98e-11	-4.46e-04	-1.67e-16	-1.26e-11	-1.18e-03	0.00e+00	-9.16e-11	-3.17e-03

Note: embedded TS: embedded two-stage approach; extended RI: extended repeated indicators approach; two-stage: two-stage approach based on a saturated structural model in the first stage;

κ_1: the total effect of ξ on η_3; κ_2: the total effect of η_1 on η_3; κ_3: the total effect of η_2 on η_3.

Table 10: Smallest weight and loading, and largest VIF for the population models used in the context of model misspecification

	small misspecification									medium misspecification									large misspecification								
	TS			Emb. TS			Ext. RI			TS			Emb. TS			Ext. RI			TS			Emb. TS			Ext. RI		
	w	VIF	λ	w	VIF	λ	w	VIF	λ	w	VIF	λ	w	VIF	λ	w	VIF	λ	w	VIF	λ	w	VIF	λ	w	VIF	λ
ξ	.39	1.9	.75	.39	1.9	.75	.39	1.9	.75	.39	1.9	.75	.39	1.9	.75	.39	1.9	.75	.39	1.9	.75	.39	1.9	.75	.39	1.9	.75
η_1	.26	1.8	.54	.26	1.8	.54	.26	1.8	.54	.26	1.8	.54	.26	1.8	.54	.26	1.8	.54	.26	1.8	.54	.26	1.8	.54	.26	1.8	.54
η_2	.33	1.8	.56	.33	1.8	.56	.33	1.8	.56	.33	1.8	.56	.33	1.8	.56	.33	1.8	.56	.33	1.8	.56	.33	1.8	.56	.33	1.8	.56
c_1	.40	1.5	.65	.40	1.5	.65	.40	1.5	.65	.40	1.5	.65	.40	1.5	.65	.40	1.5	.65	.40	1.5	.65	.40	1.5	.65	.40	1.5	.65
c_2	.20	1.9	.55	.20	1.9	.55	.20	1.9	.55	.20	1.9	.55	.20	1.9	.55	.20	1.9	.55	.20	1.9	.55	.20	1.9	.55	.20	1.9	.55
c_3	.20	1.9	.53	.20	1.9	.53	.20	1.9	.53	.20	1.9	.53	.20	1.9	.53	.20	1.9	.53	.20	1.9	.53	.20	1.9	.53	.20	1.9	.53
η_3	.27	1.7	.55	.27	1.7	.55	.27	2.7	.54	.22	1.8	.53	.22	1.8	.53	.27	2.9	.55	.27	1.6	.57	.27	1.6	.57	.29	2.9	.55

Note: values in columns w and λ, resepctively, show the smallest weight and loading, respectively, among the indicators attached to the corresponding composite; Values in column VIF show the largest VIF value of the indicators' weight used to build the corresponding composite.
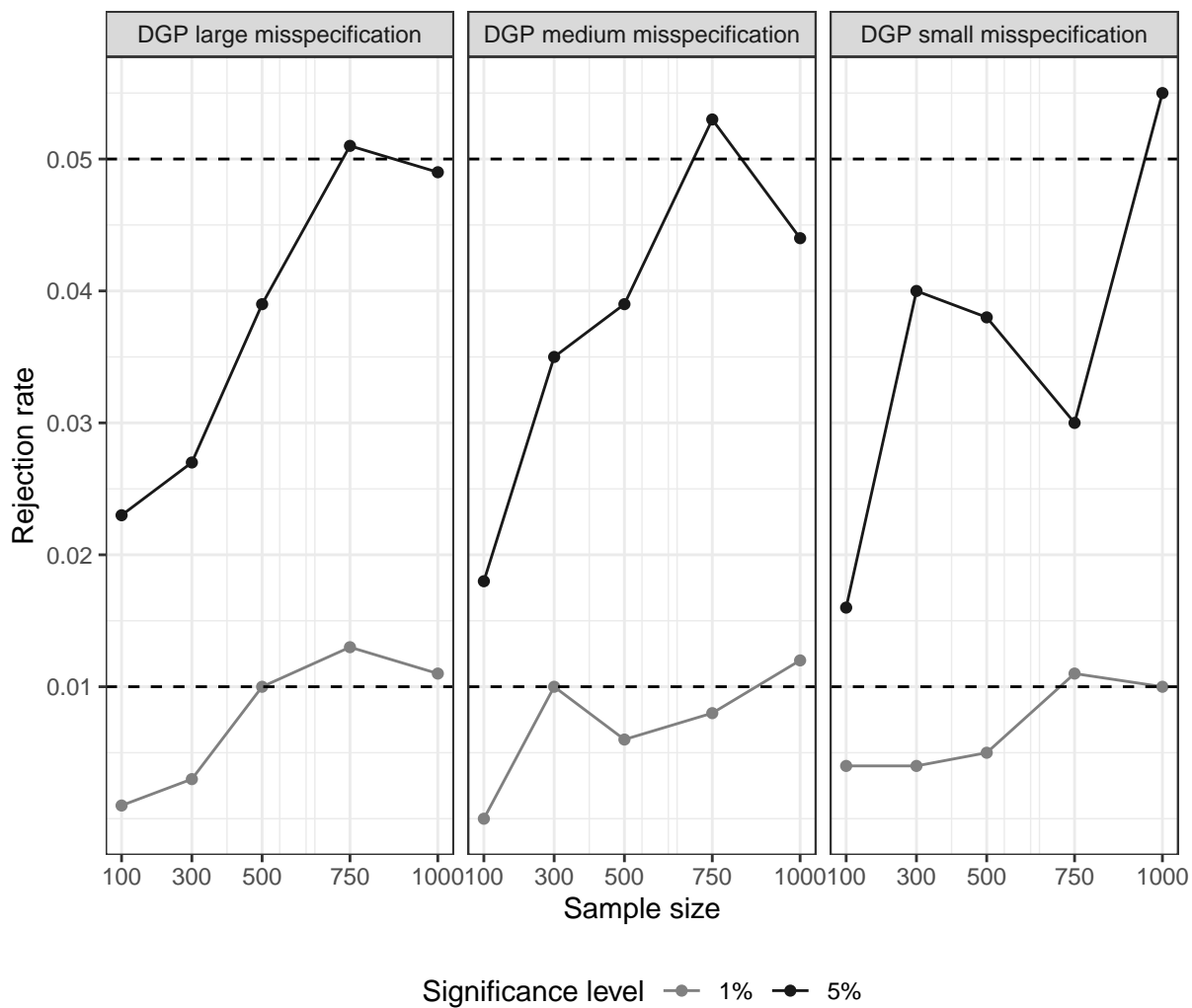
Figure 10: Tests' rejection rates for the alternative population models obtained for a correctly specified model
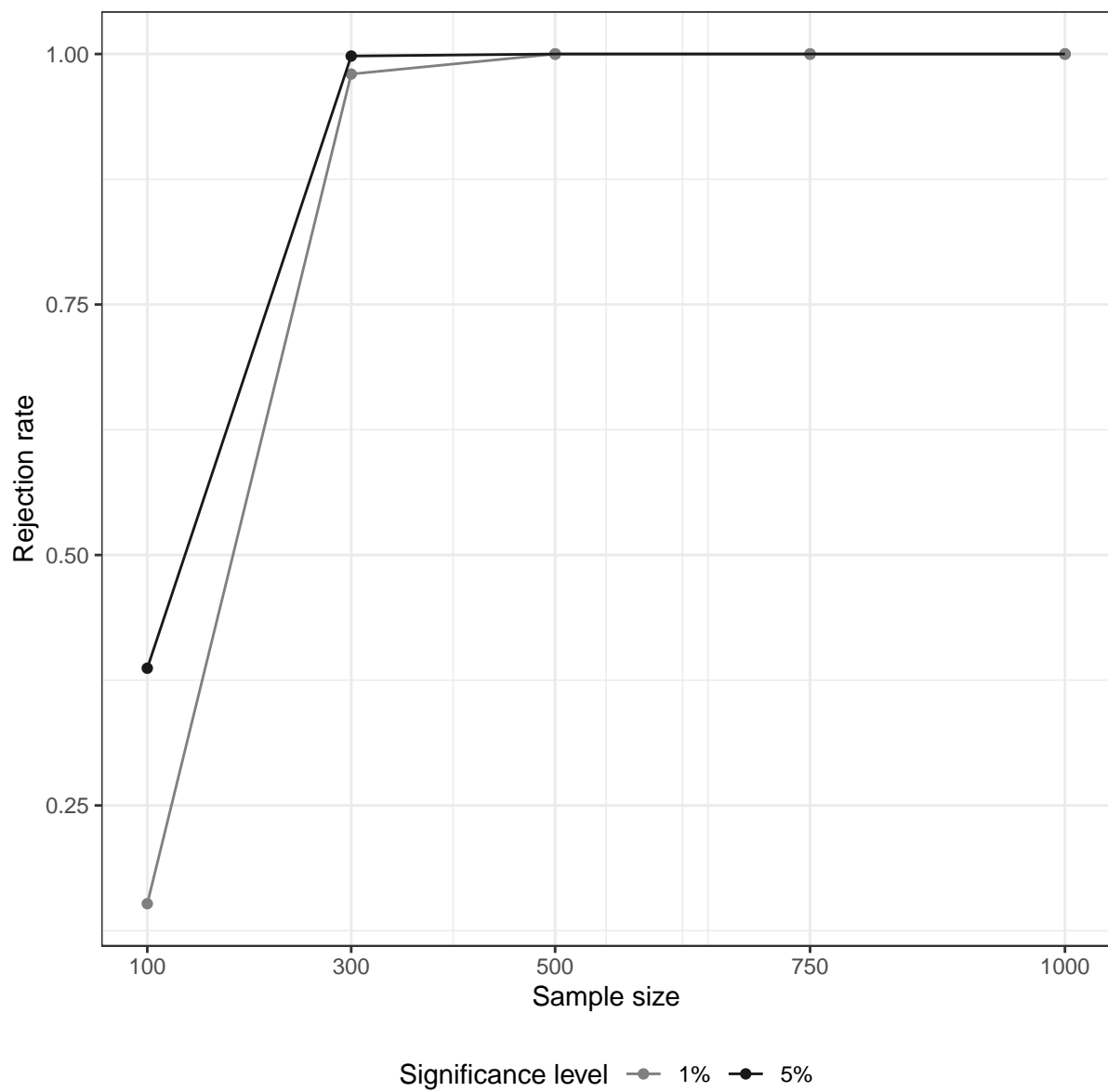
Figure 11: Tests' rejection rates for the baseline population model obtained for a specification that matches the alternative population models