Measurement error correlation within blocks of indicators in consistent partial least squares
Issues and remedies

Manuel E. Rademaker
Faculty of Business Management and Economics,
Julius-Maximilians-Universität Würzburg, Würzburg, Germany
Florian Schuberth
Faculty of Engineering Technology, University of Twente, Enschede,
The Netherlands, and
Theo K. Dijkstra
Faculty of Economics and Business, Rijksuniversiteit Groningen,
Groningen, The Netherlands

Abstract

Purpose The purpose of this paper is to enhance consistent partial least squares (PLSc) to yield consistent parameter estimates for population models whose indicator blocks contain a subset of correlated measurement errors.

Design/methodology/approach Correction for attenuation as originally applied by PLSc is modified to include *a priori* assumptions on the structure of the measurement error correlations within blocks of indicators. To assess the efficacy of the modification, a Monte Carlo simulation is conducted.

Findings In the presence of population measurement error correlation, estimated parameter bias is generally small for original and modified PLSc, with the latter outperforming the former for large sample sizes. In terms of the root mean squared error, the results are virtually identical for both original and modified PLSc. Only for relatively large sample sizes, high population measurement error correlation, and low population composite reliability are the increased standard errors associated with the modification outweighed by a smaller bias. These findings are regarded as initial evidence that original PLSc is comparatively robust with respect to misspecification of the structure of measurement error correlations within blocks of indicators.

Originality/value Introducing and investigating a new approach to address measurement error correlation within blocks of indicators in PLSc, this paper contributes to the ongoing development and assessment of recent advancements in partial least squares path modeling.

Keywords Monte Carlo simulation, Structural equation modelling, Consistent partial least squares, Measurement error correlation, Model specification

Paper type Research paper

1. Introduction

Structural equation modeling (SEM) is a versatile, widely used analytical technique to statistically examine relationships between theoretical concepts. In SEM these concepts are predominantly operationalized by latent variables, the so-called common factors assumed to be measured by a set of observable indicators within the measurement model framework.

To estimate the measurement model parameters as well as the postulated structural relationship between latent variables, two conceptually different estimation approaches have been established: covariance-based (CB) estimation (e.g. Jöreskog, 1978) and variance-based (VB) estimation (e.g. Lohmöller, 1989). CB parameter estimates are obtained by minimizing a distance measure of the empirical covariance matrix of the indicators and its theoretical counterpart implied by the model. VB estimators, on the other hand, use linear combinations of the indicators to build proxies as stand-ins for the constructs and, subsequently, estimate the model parameters based on these proxies.

Among VB estimators, partial least squares path modeling (PLS) is arguably most wide-spread. It has been used for research in numerous fields, including strategic management (e.g. Hair, Sarstedt, Pieper and Ringle, 2012), marketing (e.g. Hair, Sarstedt, Ringle and Mena, 2012), information systems (e.g. Ringle *et al.*, 2012), tourism research (e.g. Müller *et al.*, 2018) and internet research (e.g. Chiang and Hsiao, 2015; Yan *et al.*, 2017; Wu and Li, 2018). For a recent overview of the methodological research on PLS see Khan *et al.* (forthcoming).

However, despite its popularity, PLS has been subject to intense debate in recent years (see e.g. Rigdon *et al.*, 2017, for a recent stocktaking of the debate) that helped show its limitations. Most notably, PLS is only consistent at large (e.g. Dijkstra, 1981; Schneeweiss, 1993), hence yielding generally inconsistent parameter estimates for common factor models. In fact, unless all measurement errors are zero in the population, proxies cannot generally be expected to be a perfect substitute for the underlying common factor. As a consequence, the probability limit of the estimated correlation between proxies is smaller than the population correlation between their corresponding common factors. Hence, path coefficients and factor loadings based on estimated proxy correlations are inconsistent estimates for their underlying latent variable counterpart (Dijkstra and Henseler, 2015a).

To correct for these shortcomings, consistent partial least squares (PLSc) has been introduced as an enhancement of PLS that essentially maintains all the advantages of PLS while yielding consistent and asymptotically normally distributed parameter estimates for common factor models in line with Wold's (1975) basic design (Dijkstra, 1981; Dijkstra and Henseler, 2015a, b). As one of the defining assumptions of the basic design, uncorrelated measurement errors within and across blocks of indicators are thus necessary in theory for PLSc to maintain consistency.

Practically, however, there are a number of cases in empirical research in which uncorrelatedness of measurement errors may not hold (e.g. Gerbing and Anderson, 1984; Rubio and Gillespie, 1995; Chin *et al.*, 2003; Saris and Aalberts, 2003; Henseler and Chin, 2010; Brown, 2015). Depending on the magnitude of the unobserved correlation between measurement errors, the number of indicators and their quality, ignoring measurement error correlations leads to inconsistent structural parameter estimates and, therefore, to potentially erroneous conclusions (e.g. Podsakoff *et al.*, 2012; Westfall *et al.*, 2012; Gu *et al.*, 2017).

Different remedies have been proposed to prevent correlated measurement errors through a careful study design (e.g. MacKenzie and Podsakoff, 2012; Podsakoff *et al.*, 2012). However, in practice, aspects such as study design, item quality and wording are often beyond the researchers' control, essentially leaving modeling approaches as the only alternative. Several researchers therefore suggest addressing the problem indirectly, e.g., by means of bifactor models and associated hierarchical reliability indices (e.g. McNeish, 2018). Others propose explicitly specifying the measurement error correlation structure in the model (e.g. Rubio and Gillespie, 1995; Brown, 2015, pp. 162–175) – although there is some controversy as to the conceptual justification of such an approach (e.g. Landis *et al.*, 2009; Hermida, 2015).

Against this background, we follow Sarstedt *et al.*'s (2014) call for a continuous improvement of PLS and contribute to the literature by extending PLSc to yield consistent parameter estimates for population models whose indicator blocks contain a subset of

correlated measurement errors. Based on an idea outlined in Dijkstra (2013) and mentioned in Dijkstra and Henseler (2015a), this is achieved by modifying the calculation of the correction factors as defined by PLSc to include *a priori* assumptions on the structure of the within-block measurement error correlations.

The remainder of the paper is structured as follows: Section 2 briefly reviews the PLS algorithm and its consistent version PLSc. Section 3 presents the methodological contribution to obtain consistent and asymptotically normally distributed parameter estimates if within-block measurement error correlation is present. The design and results of a Monte Carlo simulation to assess the approach are described in Sections 4 and 5. The paper closes with a discussion and an outline for potential future research in Section 6.

2. PLS path modeling
PLS was developed by Herman O.A. Wold (1975) for the analysis of high-dimensional data in a low-structure environment but has been extended and modified in recent years to accommodate a wide variety of analytical needs. PLS, which may be regarded as similar to generalized canonical correlation analysis, is capable of emulating several of Kettenring's (1971) techniques for the canonical correlation analysis of several sets of indicators (Tenenhaus *et al.*, 2005). In its most developed form, known as PLSc, it may best be understood as a fully developed SEM approach that includes a global goodness-of-fit test for linear models and the ability to consistently estimate recursive, non-recursive and non-linear common factor models (Dijkstra, 2011; Dijkstra and Schermelleh-Engel, 2014; Dijkstra and Henseler, 2015a, b).

The following section briefly reviews the notation and main aspects of PLS and PLSc as well as their underlying model setup, known as the basic design.

Consider a model with J latent variables $\eta_1, \eta_2, \ldots, \eta_J$ with unit variance related via a set of structural equations and the existence of corresponding vectors of indicators $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_J$ defined as measurement error-prone manifestations of their respective latent variable:

$$\boldsymbol{x}_j = \boldsymbol{\lambda}_j \eta_j + \boldsymbol{\varepsilon}_j \quad \forall j = 1, \ldots, J, \tag{1}$$

where the vector of loadings $\boldsymbol{\lambda}_j$ contains as many components as there are indicators in $\boldsymbol{x}_j$. All variables involved are centered at their mean, and all second-order moments are assumed to exist. The measurement errors $\boldsymbol{\varepsilon}_j$ are assumed to satisfy $E(\boldsymbol{\varepsilon}_j|\eta_j) = \boldsymbol{0}$ such that the conditional mean of $\boldsymbol{x}_j$ is given by $\boldsymbol{\lambda}_j \eta_j$. Furthermore, measurement errors are taken as mutually uncorrelated within blocks and between blocks such that the within-block measurement error covariance matrix $\boldsymbol{\Theta}_{jj} \equiv E(\boldsymbol{\varepsilon}_j \boldsymbol{\varepsilon}_j')$ is diagonal and the measurement error covariance matrix across blocks $\boldsymbol{\Theta}_{ij} \equiv E(\boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_j')$ is 0. Based on these assumptions, we have the following covariance matrices:

$$\Sigma_{ij} \equiv E\left(\boldsymbol{x}_i \boldsymbol{x}_j'\right) = \rho_{ij} \boldsymbol{\lambda}_i \boldsymbol{\lambda}_j', \tag{2}$$

and:

$$\Sigma_{jj} \equiv E\left(\boldsymbol{x}_j \boldsymbol{x}_j'\right) = \boldsymbol{\lambda}_j \boldsymbol{\lambda}_j' + \boldsymbol{\Theta}_{jj}, \tag{3}$$

where ρ_{ij} is the correlation between latent variables η_i and η_j. The correlation matrix (ρ_{ij}) will generally be positive definite. It can satisfy rank constraints on sub-matrices as induced by (non-recursive) simultaneous equations for the latent variables (Dijkstra, 1981). In this paper, we work with recursive systems only, so each equation for a latent variable is a regression equation.

2.1 Traditional PLS path modeling

In addition to the setup given above, assume that there are K_j column vectors of standardized indicator observations of length N denoted by $\boldsymbol{x}_{1j}, \boldsymbol{x}_{2j}, \ldots, \boldsymbol{x}_{K_j j}$. For ease of notation, all K_j indicators are stacked in the $(N \times K_j)$ matrix $\boldsymbol{X}_j$. In PLS, proxies for each latent variable are built as the weighted sum of its related indicators. The unknown weight vector $\boldsymbol{w}_j$ is determined in an iterative three-step procedure.

At the outset, initial arbitrary outer weights $\hat{\boldsymbol{w}}_j^{(0)}$ are chosen such that the unit variance condition $\hat{\boldsymbol{w}}_j^{(0)'} \boldsymbol{S}_{jj} \hat{\boldsymbol{w}}_j^{(0)} = 1$ holds, where the $(K_j \times K_j)$ matrix $\boldsymbol{S}_{jj}$ is a consistent estimate of the population correlation matrix $\boldsymbol{\Sigma}_{jj}$[1]. After initialization, the iterative algorithm begins with Step 1, the outer estimation of η_j is as follows:

$$\hat{\boldsymbol{\eta}}_j^{(h)} = \boldsymbol{X}_j \hat{\boldsymbol{w}}_j^{(h)} \quad \text{with} \quad \hat{\boldsymbol{w}}_j^{(h)'} \boldsymbol{S}_{jj} \hat{\boldsymbol{w}}_j^{(h)} = 1 \quad \forall j = 1, \ldots, J, \tag{4}$$

where $\hat{\boldsymbol{\eta}}_j^{(h)}$ is the $(N \times 1)$ vector of outer estimates and $\hat{\boldsymbol{w}}_j^{(h)}$ the $(K_j \times 1)$ estimated weight vector. The superscript indicates the h-th iteration step. Since outer weights are scaled, the outer estimates are scaled as well.

Based on the outer estimates from Step 1, so-called inner estimates of latent variable η_j are computed according to the inner weighting scheme:

$$\tilde{\boldsymbol{\eta}}_j^{(h)} = \sum_{i\ 1}^{J} e_{ji}^{(h)} \hat{\boldsymbol{\eta}}_i^{(h)}, \tag{5}$$

where $e_{ji}^{(h)} = \text{sign}(\hat{\boldsymbol{w}}_j^{(h)'} \boldsymbol{S}_{ji} \hat{\boldsymbol{w}}_i^{(h)})$ is the inner weight with plim $\boldsymbol{S}_{ij} = \boldsymbol{\Sigma}_{ij}$[2]. All inner estimates $\tilde{\boldsymbol{\eta}}_j^{(h)}$ are again scaled such that their variance is 1.

In the third step of each iteration, new outer weights are calculated according to *mode A*. For *mode A*, the new estimated outer weights, also known as correlation weights, $\hat{\boldsymbol{w}}_j^{(h+1)}$ are equal to the coefficients resulting from a sequence of univariate ordinary least squares (OLS) regressions of $\boldsymbol{X}_j$ on $\tilde{\boldsymbol{\eta}}_j^{(h)}$[3]. As a crucial result of *mode A*, the following proportionality relation is obtained:

$$\hat{\boldsymbol{w}}_j^{(h+1)} \propto \sum_{i\ 1}^{J} e_{ij}^{(h)} \boldsymbol{S}_{ij} \hat{\boldsymbol{w}}_i^{(h)} \quad \text{with} \quad \hat{\boldsymbol{w}}_j^{(h+1)'} \boldsymbol{S}_{jj} \hat{\boldsymbol{w}}_j^{(h+1)} = 1. \tag{6}$$

New outer weights $\hat{\boldsymbol{w}}_j^{(h+1)}$ are checked for notable changes compared to the outer weights from the previous iteration step $\hat{\boldsymbol{w}}_j^{(h)}$. If there is a significant change in the weights, the algorithm continues by building new outer proxies based on the newly obtained weights; otherwise, it stops. Assuming that the established model is correct, it can be shown that the PLS algorithms will converge with a probability tending to one as the sample size increases (Dijkstra, 1981). For smaller samples and misspecified models, however, convergence may be an issue (Henseler, 2009). The resulting weights satisfy Equation (6) with all superscripts removed. Moreover, their probability limits satisfy the same equations, with $\boldsymbol{S}_{ij}$ replaced by $\boldsymbol{\Sigma}_{ij}$. Thus, the probability limits of the weights obtained by PLS and PLSc can be obtained by applying them to the population indicator covariance matrix $\boldsymbol{\Sigma}$. Notably, the proof of numerical and probabilistic convergence does not require that the measurement errors within blocks are uncorrelated. To see this, it is crucial to note, that the population weights are unaffected of the precise nature of $\boldsymbol{\Sigma}_{jj}$. Using the final weights $\hat{\boldsymbol{w}}_j$ and taking probability limits on both sides of Equation (6), we have formulated the following:

$$\text{plim } \hat{\boldsymbol{w}}_j \propto \text{plim} \sum_{i\ 1}^{J} e_{ij} \boldsymbol{S}_{ij} \hat{\boldsymbol{w}}_i \to \boldsymbol{w}_j \propto \sum_{i\ 1}^{J} e_{ij} \boldsymbol{\Sigma}_{ij} \boldsymbol{w}_i = \sum_{i\ 1}^{J} e_{ij} \rho_{ij} \boldsymbol{\lambda}_i \boldsymbol{\lambda}_j' \boldsymbol{w}_i, \tag{7}$$

where the last equality crucially assumes uncorrelated measurement errors across blocks, i.e., $\mathbf{\Theta}_{ij} = \mathbf{0}$, but not within blocks of indicators [4].

Once convergence is reached, the resulting stable outer weights $\hat{\boldsymbol{w}}_j$ are used to build the final proxy for the latent variables: $\hat{\boldsymbol{\eta}}_j = \boldsymbol{X}_j \hat{\boldsymbol{w}}_j$. Finally, factor loadings for each block are obtained as the OLS solution of a sequence of regressions of $\boldsymbol{X}_j$ on $\hat{\boldsymbol{\eta}}_j$. Similarly, the path coefficients are the OLS estimates of the equations postulated by the structural model.

2.2 Consistent PLS

The principal idea of PLS is to build proxies as stand-ins for the latent variables and subsequently estimate model parameters based on these proxies. Naturally, it cannot be expected that these stand-ins perfectly reflect the underlying latent variables unless all measurement errors are assumed to be 0 in the population. As a consequence, the probability limit of the estimated correlation between proxies is smaller in absolute value than the population correlation between their corresponding common factors. Hence, path coefficients and factor loadings based on estimated proxy correlations are inconsistent estimates for their population counterpart. PLSc addresses this shortcoming by consistently estimating the composite reliability and subsequently correcting the correlations among the proxies for attenuation (Cohen et al., 2003). Provided that each latent variable is connected to at least two indicators, the population composite reliability of the population proxy $\overline{\eta}_j$ as defined in Dijkstra and Henseler (2015b) is given by:

$$\rho_{A,j} := \left(\boldsymbol{w}_j'\boldsymbol{w}_j\right)^2 \cdot c_j^2, \tag{8}$$

where $c_j := \sqrt{\boldsymbol{\lambda}_j'\Sigma_{jj}\boldsymbol{\lambda}_j}$ is the factor that relates population weights $\boldsymbol{w}_j = \text{plim } \hat{\boldsymbol{w}}_j$ to their corresponding population loadings $\boldsymbol{\lambda}_j$ (Dijkstra, 1981; Dijkstra and Henseler, 2015a):

$$\boldsymbol{w}_j = \frac{\boldsymbol{\lambda}_j}{\sqrt{\boldsymbol{\lambda}_j'\Sigma_{jj}\boldsymbol{\lambda}_j}}. \tag{9}$$

It is crucial to note, that this relationship holds independent of the form of Σ_{jj}. To see this, note that based on Equation (7), the population relation between weights and loadings may simply be written as $\boldsymbol{w}_j = c_j^{-1}\boldsymbol{\lambda}_j$ since $\sum_{i\ 1}^{J} e_{ij}\rho_{ij}\boldsymbol{\lambda}_i'\boldsymbol{w}_i$ is a scalar.

Using the population normalization condition $\boldsymbol{w}_j'\Sigma_{jj}\boldsymbol{w}_j = 1$ now yields the population value c_j:

$$\boldsymbol{w}_j'\Sigma_{jj}\boldsymbol{w}_j = 1, \tag{10}$$

$$\frac{\boldsymbol{\lambda}_j'}{c_j}\Sigma_{jj}\frac{\boldsymbol{\lambda}_j}{c_j} = 1, \tag{11}$$

$$c_j^2 = \boldsymbol{\lambda}_j'\Sigma_{jj}\boldsymbol{\lambda}_j. \tag{12}$$

Consequently, population weights and the proportionality constant c_j clearly vary with Σ_{jj}, however, the fundamental relationship given by Equation (7) is unaffected by Σ_{jj} (and therefore also unaffected by potential within-block error correlation).

To obtain the estimated correction factor $\hat{c}_j$, a variety of approaches are possible (Dijkstra, 2013). Usually, $\hat{c}_j$ is chosen for block j such that the squared Euclidean distance between the off-diagonal elements of the empirical covariance matrix

S_{jj} and the matrix $(c_j \hat{\boldsymbol{w}}_j)(c_j \hat{\boldsymbol{w}}_j)'$ is minimized. In this case, the squared estimated correction factor is given by:

$$\hat{c}_j^2 = \frac{\hat{\boldsymbol{w}}_j'(S_{jj}-\text{diag}(S_{jj}))\hat{\boldsymbol{w}}_j}{\hat{\boldsymbol{w}}_j'\left(\hat{\boldsymbol{w}}_j\hat{\boldsymbol{w}}_j'-\text{diag}\left(\hat{\boldsymbol{w}}_j\hat{\boldsymbol{w}}_j'\right)\right)\hat{\boldsymbol{w}}_j}. \tag{13}$$

Since plim $\hat{\boldsymbol{w}}_j = \boldsymbol{w}_j$ and plim $S_{jj} = \Sigma_{jj}$ and since the functions involved are continuous, the probability limit directly follows:

$$\text{plim } \hat{c}_j^2 = \frac{\boldsymbol{w}_j'(\Sigma_{jj}-\text{diag}(\Sigma_{jj}))\boldsymbol{w}_j}{\boldsymbol{w}_j'\left(\boldsymbol{w}_j\boldsymbol{w}_j'-\text{diag}\left(\boldsymbol{w}_j\boldsymbol{w}_j'\right)\right)\boldsymbol{w}_j}, \tag{14}$$

$$= \lambda_j'\Sigma_{jj}\lambda_j + \lambda_j'\Sigma_{jj}\lambda_j \cdot \frac{\lambda_j'(\Theta_{jj}-\text{diag}(\Theta_{jj}))\lambda_j}{\lambda_j'\left(\lambda_j\lambda_j'-\text{diag}\left(\lambda_j\lambda_j'\right)\right)\lambda_j}. \tag{15}$$

The numerator of the last term in Equation (15) is 0 when all the measurement errors are uncorrelated in the population since, in this case, $\Theta_{jj} = \text{diag}(\Theta_{jj})$. Assuming that Θ_{jj} is indeed a diagonal matrix, the resulting probability limit of the squared estimated correction factor equals the squared correction factor from Equation (12), i.e., the squared distortion of the population weights to population loadings. Hence, consistent factor loading estimates and attenuation-corrected correlations between common factors j and i are readily given by:

$$\hat{\lambda}_j = \hat{c}_j\hat{\boldsymbol{w}}_j \quad \text{and} \quad \widehat{\text{Cor}(\eta_j, \eta_i)} = \frac{\hat{\boldsymbol{w}}_j'S_{ji}\hat{\boldsymbol{w}}_i}{\sqrt{\hat{\rho}_{A,j}\cdot\hat{\rho}_{A,i}}}. \tag{16}$$

Depending on the underlying structural model, consistent path coefficient estimates may be obtained by OLS or two-stage least squares using the estimated disattenuated correlation given above.

3. Correlated measurement errors

As suggested by Equation (15), the consistency of original PLSc was established based on the assumptions of the basic design, including measurement errors that are uncorrelated across and within blocks of indicators; i.e., Θ_{jj} is indeed a diagonal matrix. In fact, if measurement errors in the population are correlated within blocks of indicators, then original PLSc using the correction factor from Equation (13) leads to inconsistent parameter estimates for both factor loadings and path coefficients, where the magnitude of the inconsistency is positively related to the strength of the measurement error correlation and negatively affected by the composite reliability. However, taking into account measurement errors are straightforward provided that the correlation is confined to be within the indicator blocks.

Given a presumption on the measurement error correlation structure, define the set of uncorrelated measurement error pairs as $U_j := \{(k, m)|\theta_{km;jj} = 0\}$, where $\theta_{km;jj}$ denotes the population covariance between the k-th and m-th measurement error of block j. An immediate extension to original PLSc is to minimize the squared Euclidean distance between the off-diagonal elements of the empirical covariance matrix S_{jj} and the matrix $(c_j \hat{\boldsymbol{w}}_j)(c_j \hat{\boldsymbol{w}}_j)'$ with respect to c_j, including only those elements contained in the set U_j:

$$\hat{c}_j^2 = \underset{c_j^2}{\arg\min} \sum_{k,m \in U_j} \left[s_{km;jj}-c_j^2\hat{w}_{kj}\hat{w}_{mj}\right]^2, \tag{17}$$

where $\hat{w}_{kj}$ and $\hat{w}_{mj}$ are the k-th and m-th elements of the weight vector $\hat{\boldsymbol{w}}_j$ and $s_{km;\,jj}$ is the empirical covariance between the k-th and m-th indicators of block j [5]. Provided that the set of uncorrelated measurement error pairs is nonempty, minimization yields:

$$\hat{c}_j^{*2} = \frac{\sum_{k,m\,\in\,U_j}\hat{w}_{kj}\hat{w}_{mj}s_{km;jj}}{\sum_{k,m\,\in\,U_j}\hat{w}_{kj}^2\hat{w}_{mj}^2}. \tag{18}$$

Because of the continuity of the functions involved, the consistency of the sample moments, and the fact that the probability limits of the PLS weight vectors, as given in Dijkstra (1981), are effectively independent of the assumed structure within the blocks, the probability limit of the estimated adjusted squared correction factor is again equal to $\boldsymbol{\lambda}_j'\Sigma_{jj}\boldsymbol{\lambda}_j$. Indeed, replacing the terms in Equation (18) by their population counterparts yields:

$$\text{plim } \hat{c}_j^{*2} = \frac{\sum_{k,m\,\in\,U_j}w_{k,\,j}w_{m,\,j}\sigma_{km;\,jj}}{\sum_{k,m\,\in\,U_j}w_{kj}^2w_{mj}^2}, \tag{19}$$

$$= \boldsymbol{\lambda}_j'\Sigma_{jj}\boldsymbol{\lambda}_j \cdot \frac{\sum_{k,m\,\in\,U_j}\lambda_{kj}^2\lambda_{mj}^2 + \sum_{k,m\,\in\,U_j}\lambda_{kj}\lambda_{mj}\theta_{km;jj}}{\sum_{k,m\,\in\,U_j}\lambda_{kj}^2\lambda_{mj}^2}, \tag{20}$$

$$= \boldsymbol{\lambda}_j'\Sigma_{jj}\boldsymbol{\lambda}_j, \tag{21}$$

where the last term in Equation (20) is one since $\theta_{km;jj}$ is 0 by assumption for all elements contained in U_j. As a consequence, consistent estimates for the attenuation-corrected correlations between common factors, loadings and path coefficients may be obtained along the same lines described in the preceding section.

4. Monte Carlo simulation
To assess the efficacy of the modification, a Monte Carlo simulation is conducted.

To this end, six population models are investigated [6]. The baseline population model to be considered is illustrated in Figure 1. The structural population model contains three latent variables:

$$\eta_2 = \gamma_1\eta_1 + \zeta_1, \tag{22}$$

$$\eta_3 = \gamma_2\eta_1 + \beta\eta_2 + \zeta_2, \tag{23}$$

where $\gamma_1 = 0.6$, $\gamma_2 = 0.4$, $\beta = 0$, $\text{Var}(\zeta_1) = 0.64$, $\text{Var}(\zeta_2) = 0.84$, and $\text{Cov}(\eta_1,\zeta_1) = \text{Cov}(\eta_1,\zeta_2) = \text{Cov}(\eta_2,\zeta_2) = \text{Cov}(\zeta_1,\zeta_2) = 0$. The structural model remains identical across all six population models and is similar to structural models typically applied in the literature (e.g. Paxton $et\ al.$, 2001; Hwang $et\ al.$, 2010).

For each population model, the exogenous latent variable η_1 and the two endogenous latent variables η_2 and η_3 are each connected to three indicators, the minimum requirement for our approach to be feasible since the additional indicator ensures that $U_j \neq \emptyset$ if a correlation between any two measurement errors is allowed. Factor loadings for η_2 and η_3 are fixed at $\lambda_{12} = 0.7$, $\lambda_{22} = 0.85$, $\lambda_{32} = 0.8$ and $\lambda_{13} = 0.8$, $\lambda_{23} = 0.75$, $\lambda_{33} = 0.8$, reflecting average indicator reliabilities. Furthermore, the first two loadings of η_1 are set to $\lambda_{11} = 0.65$ and $\lambda_{21} = 0.8$, respectively. To investigate how different composite reliabilities affect parameter estimates, both the number of indicators per block and the size of the loadings may be varied. Here, we chose the latter by varying λ_{31} within a range of 0.5 to 0.9 in steps of 0.2.
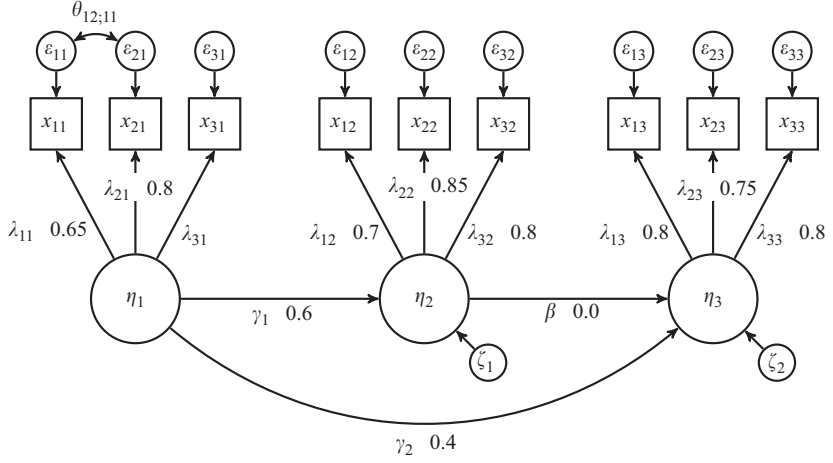
Figure 1.
Baseline population
model

All measurement errors (ε_{kj}) have a mean of 0 and are uncorrelated across and within blocks except for the first and the second measurement errors of the first indicator block: $\theta_{12;11} = \sqrt{0.360 \cdot 0.578} \cdot \rho_{12;11}$, where $\rho_{12;11}$ denotes the correlation between ε_{11} and ε_{21}. To assess how the strength of the correlation affects parameter estimates, we include a case with comparatively low ($\rho_{12;11} = 0.1$) and high correlation ($\rho_{12;11} = 0.6$).

The simulation is conducted in the statistical software environment R (R Core Team, 2017). The data sets for each of the six resulting population models ($= 3$ different loading magnitudes $\times$ 2 different measurement error correlations) are drawn according to the following baseline population indicator correlation matrix using the MASS package (Venables and Ripley, 2002). Samples of size $n = 100$, 200 and 1,000 are drawn from a multivariate normal distribution with the mean of each indicator set to 0 and the covariance matrix displayed in Equation (24):

$$\Sigma = \begin{array}{ccccccccc} x_{11} & x_{21} & x_{31} & x_{12} & x_{22} & x_{32} & x_{13} & x_{23} & x_{33} \\ \begin{pmatrix} 1.00 & 0.52 + \theta_{12;11} & 0.65 \cdot \lambda_{31} & 0.27 & 0.33 & 0.31 & 0.21 & 0.20 & 0.21 \\ & 1.00 & 0.80 \cdot \lambda_{31} & 0.34 & 0.41 & 0.38 & 0.26 & 0.24 & 0.26 \\ & & 1.00 & 0.21 & 0.26 & 0.24 & 0.16 & 0.15 & 0.16 \\ & & & 1.00 & 0.60 & 0.56 & 0.13 & 0.13 & 0.13 \\ & & & & 1.00 & 0.68 & 0.16 & 0.15 & 0.16 \\ & & & & & 1.00 & 0.15 & 0.14 & 0.15 \\ & & & & & & 1.00 & 0.60 & 0.64 \\ & & & & & & & 1.00 & 0.60 \\ & & & & & & & & 1.00 \end{pmatrix} \end{array}. \quad (24)$$

The number of replications per population model is set to 1,000, resulting in a total of 18,000 data sets (6 population models $\times$ 3 sample sizes $\times$ 1,000 replications).

To estimate the underlying population parameters for each data set, two models were specified. The first model M_1 correctly reflects the corresponding underlying population model in terms of the structural and the measurement model but does not explicitly account

for the correlation between the measurement errors ε_{11} and ε_{21}. Here, estimation by traditional PLSc is expected to yield estimates that systematically deviate from their corresponding population values. The second model M_2 is similar to the first model but acknowledges the measurement error correlation as present in the population models. Estimation is performed using our contributed modification. To this end, we use the MoMpoly function provided by the MoMpoly package (Schuberth *et al.*, 2017), which implements the procedure as described in this paper [7]. Here, the enhanced procedure is expected to yield estimates close to the corresponding population parameters. However, this is likely to come at the cost of a loss in precision, as the calculation of the correction factor is based on less information. In addition to the estimations based on the simulated data sets, we retrieve the parameters for each population model using the population covariance matrix as input. This serves to verify Fisher consistency, i.e., whether a given estimator is in fact able to yield population parameters if supplied by the population covariance matrix.

To compare the estimates across the different designs, two common quality measures are considered: the estimated bias and the root mean squared error (RMSE). The bias is estimated as follows:

$$\widehat{\text{Bias}} = \frac{1}{M} \sum_{i}^{M} \left(\widehat{\psi}_i - \psi \right), \tag{25}$$

where ψ denotes a generic population parameter and $\hat{\psi}$ is its corresponding estimate for a given model and sample size. The number of elements M is equal to the number of replications corrected for the number of Heywood cases and outliers [8]. The latter is defined as all estimates larger than the median ± 3 times the interquartile range.

Consistency of our modification is essentially achieved by discarding information. Hence, finite sample comparisons between modified PLSc and original PLSc should take the expected trade-off between bias and variability into account. A well-established measure in this respect is the (estimated) RMSE given by:

$$\widehat{\text{RMSE}} = \sqrt{\frac{1}{M} \sum_{i}^{M} \left(\widehat{\psi}_i - \psi \right)^2}. \tag{26}$$

The population RMSE essentially combines standard deviation and bias. For an unbiased estimator, it equals to the standard deviation.

5. Results
Below, we present the results of the simulation study. We report the results for the path coefficients γ_1, γ_2 and β and the factor loadings λ_{21} and λ_{31} of the indicator block affected by measurement error correlation. In addition, the share of Heywood cases and the share of outliers are given for each setup. Omission of the other loadings is justified because the results for λ_{11} are virtually identical to those of λ_{21} and λ_{31}, while the loadings of those indicator blocks whose measurement errors are assumed to be uncorrelated are by construction unaffected by the correlated measurement errors of other blocks within the structural model.

Tables I and II summarize the results. Each major column contains the results for a given population factor loading λ_{31} (i.e. 0.5, 0.7, 0.9) spread across two minor columns representing the varying population measurement error correlation $\rho_{12;11}$, where $\rho_{12;11} \in \{0.1, 0.6\}$. Each major-minor combination is again split by model (i.e. model M_1 and model M_2) to facilitate the comparison.

| n | Parameter | $\lambda_{31}=0.5$ | | | | $\lambda_{31}=0.7$ | | | | $\lambda_{31}=0.9$ | | | |
| | | $\rho_{12;11}=0.1$ | | $\rho_{12;11}=0.6$ | | $\rho_{12;11}=0.1$ | | $\rho_{12;11}=0.6$ | | $\rho_{12;11}=0.1$ | | $\rho_{12;11}=0.6$ | |
		M_1 ( )	M_2 (+)	M_1 ( )	M_2 (+)	M_1 ( )	M_2 (+)	M_1 ( )	M_2 (+)	M_1 ( )	M_2 (+)	M_1 ( )	M_2 (+)
100	γ_1	0.00	0.01	0.06	0.01	0.00	0.01	0.04	0.00	0.00	0.01	0.01	0.01
100	γ_2	0.02	0.05	0.06	0.04	0.02	0.02	0.04	0.00	0.00	0.01	0.02	0.02
100	β	0.02	0.00	0.08	0.00	0.02	0.02	0.07	0.03	0.04	0.01	0.06	0.02
100	λ_{21}	0.01	0.01	0.08	0.02	0.01	0.00	0.06	0.00	0.01	0.00	0.06	0.01
100	λ_{31}	0.01	0.01	0.06	0.01	0.01	0.00	0.06	0.00	0.01	0.01	0.02	0.01
Heywood cases (%)		28.70	34.00	33.30	36.30	27.00	31.20	30.60	27.60	30.50	27.50	46.10	32.80
Outliers (%)		1.96	3.64	3.75	5.81	4.11	2.47	2.59	1.93	2.45	3.45	2.97	2.53
200	γ_1	0.01	0.01	0.06	0.01	0.01	0.00	0.04	0.01	0.01	0.00	0.03	0.00
200	γ_2	0.01	0.02	0.08	0.02	0.01	0.00	0.05	0.02	0.01	0.00	0.03	0.00
200	β	0.02	0.00	0.08	0.01	0.02	0.01	0.06	0.01	0.01	0.01	0.04	0.01
200	λ_{21}	0.02	0.01	0.09	0.00	0.01	0.00	0.07	0.00	0.01	0.00	0.06	0.00
200	λ_{31}	0.01	0.00	0.06	0.00	0.01	0.00	0.06	0.00	0.00	0.00	0.03	0.00
Heywood cases (%)		9.00	13.70	14.20	13.20	10.40	8.30	10.10	10.70	12.50	8.00	25.90	10.80
Outliers (%)		0.88	0.81	0.93	0.92	0.67	1.31	0.67	0.78	1.14	0.65	0.27	0.56
1,000	γ_1	0.01	0.01	0.07	0.00	0.01	0.00	0.05	0.00	0.01	0.00	0.03	0.00
1,000	γ_2	0.02	0.00	0.08	0.00	0.01	0.00	0.06	0.00	0.01	0.00	0.05	0.00
1,000	β	0.02	0.00	0.07	0.00	0.01	0.00	0.05	0.01	0.01	0.00	0.04	0.00
1,000	λ_{21}	0.02	0.00	0.10	0.00	0.01	0.00	0.07	0.00	0.01	0.00	0.05	0.00
1,000	λ_{31}	0.01	0.00	0.06	0.00	0.01	0.00	0.06	0.00	0.01	0.00	0.05	0.00
Heywood cases (%)		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	3.50	0.10
Outliers (%)		0.30	0.10	0.10	0.30	0.00	0.10	0.40	0.00	0.00	0.00	0.10	0.00
Pop.	γ_1	0.01	0.00	0.07	0.00	0.01	0.00	0.05	0.00	0.01	0.00	0.03	0.00
Pop.	γ_2	0.02	0.00	0.08	0.00	0.01	0.00	0.06	0.00	0.01	0.00	0.05	0.00
Pop.	β	0.02	0.00	0.07	0.00	0.01	0.00	0.05	0.00	0.01	0.00	0.04	0.00
Pop.	λ_{21}	0.02	0.00	0.10	0.00	0.01	0.00	0.07	0.00	0.01	0.00	0.05	0.00
Pop.	λ_{31}	0.01	0.00	0.06	0.00	0.01	0.00	0.06	0.00	0.01	0.00	0.05	0.00

Table I.
Estimated bias

| n | Parameter | $\lambda_{31}=0.5$ | | | | $\lambda_{31}=0.7$ | | | | $\lambda_{31}=0.9$ | | | |
| | | $\rho_{12;11}=0.1$ | | $\rho_{12;11}=0.6$ | | $\rho_{12;11}=0.1$ | | $\rho_{12;11}=0.6$ | | $\rho_{12;11}=0.1$ | | $\rho_{12;11}=0.6$ | |
		M_1 ( )	M_2 (+)	M_1 ( )	M_2 (+)	M_1 ( )	M_2 (+)	M_1 ( )	M_2 (+)	M_1 ( )	M_2 (+)	M_1 ( )	M_2 (+)
100	γ_1	0.09	0.10	0.11	0.11	0.09	0.10	0.10	0.10	0.08	0.08	0.08	0.09
100	γ_2	0.16	0.20	0.16	0.22	0.17	0.17	0.15	0.17	0.15	0.14	0.14	0.15
100	β	0.17	0.19	0.17	0.20	0.17	0.18	0.17	0.17	0.16	0.15	0.16	0.17
100	λ_{21}	0.09	0.10	0.11	0.10	0.08	0.09	0.09	0.09	0.08	0.08	0.08	0.07
100	λ_{31}	0.12	0.09	0.14	0.10	0.10	0.08	0.11	0.09	0.06	0.05	0.06	0.06
200	γ_1	0.07	0.08	0.09	0.08	0.07	0.07	0.08	0.07	0.06	0.06	0.06	0.06
200	γ_2	0.12	0.15	0.13	0.15	0.11	0.11	0.12	0.13	0.11	0.11	0.10	0.11
200	β	0.12	0.14	0.13	0.14	0.11	0.11	0.12	0.12	0.11	0.11	0.11	0.11
200	λ_{21}	0.07	0.08	0.11	0.08	0.06	0.07	0.08	0.06	0.05	0.05	0.07	0.05
200	λ_{31}	0.08	0.07	0.11	0.07	0.07	0.06	0.09	0.06	0.04	0.04	0.05	0.04
1,000	γ_1	0.03	0.04	0.07	0.04	0.03	0.03	0.06	0.03	0.03	0.03	0.04	0.03
1,000	γ_2	0.05	0.06	0.10	0.06	0.05	0.05	0.08	0.05	0.05	0.05	0.06	0.05
1,000	β	0.05	0.06	0.08	0.06	0.05	0.05	0.07	0.05	0.05	0.05	0.06	0.05
1,000	λ_{21}	0.03	0.04	0.10	0.04	0.03	0.03	0.07	0.03	0.02	0.02	0.05	0.02
1,000	λ_{31}	0.04	0.03	0.07	0.03	0.03	0.03	0.07	0.03	0.02	0.02	0.06	0.02

Table II.
Root mean squared
error (RMSE)

Table I displays the simulation results with respect to the estimated bias. Each row displays the average deviation of the estimated parameters from their corresponding population values split by sample size $n = 100, 200, 1000$ (across rows), population and estimated model (across columns). In the presence of unmodeled measurement error correlation within a block of indicators, parameter estimates obtained by PLSc using the traditional correction factor (model M_1) systematically deviate on average from their pre-specified population value, where the deviation per population model and parameter is stable across sample sizes. This finding is in line with the fact that original PLSc is indeed unable to retrieve population parameters when supplied with the corresponding population indicator covariance matrix, as displayed at the bottom of Table I. Comparing results for a given sample size, the magnitude of the deviations varies between virtually no bias (e.g. for $\lambda_{31} = 0.9$ and $\rho_{12;11} = 0.1$) and values of up to 0.1 (e.g. for $\lambda_{31} = 0.5$ and $\rho_{12;11} = 0.6$), depending on the strength of the measurement error correlation $\rho_{12;11}$ and the size of the population loading λ_{31}. In this respect, the effect of the strength of the correlation between measurement errors on the estimated bias is most pronounced with higher error correlations leading to increased deviation.

Looking across columns for a given measurement error correlation, deviations vary only marginally, although an increasing reliability – as induced by the higher loadings – slightly decreases bias overall. These findings are again supported by the parameters obtained based on the corresponding population covariance matrix shown in the last four rows of Table I: deviations for all parameters are lowest for estimates based on population models with a higher composite reliability, i.e., $\lambda_{31} = 0.9$.

In contrast to Model M_1, population model parameters are retrieved when errors are taken into account along the lines described in Section 3 (model M_2). The finite sample results for model M_2 are largely in line with these findings, although small deviations are found; e.g., with values of 0.04 and 0.05, the estimated bias for path coefficient γ_2 is comparatively high.

For a given parameter, the sign of the deviations is relatively stable across sample sizes, population model and estimated model. The results show a small but almost consistently negative deviation for γ_1 and γ_2, while β, the path coefficient connecting the two endogenous latent variables η_2 and η_3, as well as the loadings λ_{21} and λ_{31} are uniformly overestimated.

Overall, the difference between M_1 and M_2 is most pronounced for the estimated loadings, while deviations for the path coefficients are generally small, with modified PLSc outperforming original PLSc for large samples sizes and strong measurement error correlation only.

Table II reports the results for the RMSE. Here, the picture is mixed. For medium ($\lambda_{31} = 0.7$) and high ($\lambda_{31} = 0.9$) composite reliability, the RMSE for both loading and path coefficient estimates is virtually identical for M_1 and M_2. In contrast to the results in Table I, the RMSE does not differ systematically with the magnitude of the error correlation. For $\lambda_{31} = 0.5$, however, original PLSc is superior to the modified approach in small samples ($n = 100, 200$). Only for a large sample size and a high composite reliability does M_2 produce strictly smaller RMSEs compared to the values produced by M_1.

Regarding Heywood cases and outliers, no significant difference between M_1 and M_2 is visible. While the number of Heywood cases is close to 0 or is 0 for large samples, roughly 300 of the 1,000 replications were discarded for a sample size of $n = 100$. In each instance, Heywood cases occur because of the loading estimates that are larger than one in absolute value.

6. Discussion and future research

Correlated measurement errors are a common feature in SEM. However, research regarding issues and potential remedies related to measurement error correlations in the context of VB estimation is scarce. While prior research papers (e.g. Charles, 2005; Zimmerman, 2007; Padilla and Veprinsky, 2012; Raykov *et al.*, 2014) have discussed and addressed the issue of correlated measurement errors in the common factor framework, none of these are

based on a VB approach like PLS. Against this background, we contribute to the ongoing development and assessment of VB estimation approaches by filling two gaps in the literature.

First, this study enhanced PLSc to yield consistent parameter estimates for population models whose indicator blocks contain a subset of correlated measurement errors – provided that all correlated errors are accounted for in the estimated model. Since PLS and PLSc are viable options for estimating interactions and other non-linear relationships between constructs (e.g. Dijkstra and Henseler, 2011; Dijkstra and Schermelleh-Engel, 2014), our findings may help in advancing current approaches in this field. Notable examples of this kind would be the product-indicator approach (Chin *et al.*, 2003) and the orthogonalizing approach (Henseler and Chin, 2010) – both of which rely on indicators whose errors can safely be assumed to be correlated for technical reasons. The proposed correction can help to make these two approaches consistent.

Second, initial evidence on the implications of neglecting measurement error correlation in PLSc was provided. To this end, a Monte Carlo simulation was conducted to investigate the average difference between estimated parameters and their respective population counterpart as well as the RMSE across a range of pre-specified population models for original and modified PLSc.

For original PLSc, the simulation results showed a generally small yet persistent average deviation between the estimated parameters and their corresponding population value (estimated bias) across all population models if measurement error correlation was neglected in the estimated model (model M_1). For our proposed approach (model M_2), the average deviation between the estimated parameters and their corresponding population value was virtually 0 across all samples sizes, indicating that the procedure works well in finite samples. These findings were in line with theoretical considerations regarding the inconsistency of original PLSc when measurement errors within indicator blocks are ignored. Overall, however, differences were generally rather small. In particular, when efficiency is considered with respect to the RMSE, M_1 and M_2 produce virtually identical results unless both the sample size and the population error correlation are high and the population composite reliability is low.

Regarding the magnitude of the estimated bias, we found a positive relation with the strength of the measurement error correlation, while higher composite reliability can be seen as a catalyst that essentially mitigates the effect of a given neglected measurement error correlation. The latter is intuitively appealing since an increase in composite reliability implies a decrease in attenuation of the latent variable correlation. Hence, correction for attenuation and, by the same token, any inconsistency caused by unmodeled measurement error correlation becomes less and less influential. Regarding the RMSE, the relation is less clear, although the RMSE for both the modified approach and original PLSc is higher when the population measurement error is comparatively high.

These findings are regarded as initial evidence that – although our approach is theoretically superior – original PLSc is comparatively robust with respect to misspecification of the structure of the measurement error correlations within blocks of indicators. Indeed, some preliminary simulation results by the authors confirm that PLSc outperforms common CB estimators (including maximum likelihood) in terms of bias if measurement error correlation within blocks of indicators is neglected. However, a generalization of these findings requires separate attention.

The observed tendency of PLSc to produce Heywood cases (loadings larger than one in absolute value), or incorrect signs of regression coefficients in PLS, should be addressed. We chose the simplest method to demonstrate our modification, but more robust approaches for estimating the correction factor may be applied. In fact, initial Monte Carlo evidence confirms that using, e.g., Equation (11) of Dijkstra (2013), does indeed improve the share of

admissible results by roughly 10 percentage points without affecting any of the results described above. Whether these findings hold in general, however, is an open question. Furthermore, we have developed a simple approach – essentially empirically Bayes – where we use a posterior mean, median or mode that does lie in the appropriate range to address these issues. The merits of this approach, however, are not yet fully investigated (Dijkstra, 2018).

This study provided initial evidence on the implications of neglecting measurement error correlation in terms of parameter accuracy. Clearly, this is of limited scope. Future research should investigate the consequences of our modified approach for model fit. Critics have repeatedly cautioned against pre-specifying measurement error correlations, claiming that these correlations often lack a substantive meaning, which would in turn only obfuscate a meaningful interpretation of the specified model. In fact, for CB estimators such as maximum likelihood freeing, one or more measurement error correlations naturally leads to an increase in model fit, as the estimated model-implied covariance matrix is closer to its empirical counterpart. Similarly, common fit indices based on the distance between the estimated model-implied and empirical covariance matrix – such as the standardized root mean squared residual or the geodesic distance – generally indicate a better fit.

The focus of this paper was on within indicator block measurement error correlation only. In the presence of unmodeled population measurement errors across blocks, the modification does not yield consistent estimates because the proportionality between weights and loadings as used to derive the correction factor no longer holds. As a consequence, loadings, reliabilities and path coefficients pertaining to the blocks affected by measurement error correlation are generally inconsistent. Strategies to address unmodeled population measurement errors across blocks within the PLS/PLSc framework are thus needed.

Notes

1. Throughout the iteration, the unit variance condition is maintained by using the scaling factor $(\hat{\boldsymbol{w}}_j^{(h)\prime}\boldsymbol{S}_{jj}\hat{\boldsymbol{w}}_j^{(h)})^{-(1/2)}$ for the outer weights $\hat{\boldsymbol{w}}_j^{(h)}$ in each iteration step h.

2. The inner weight e_{ji} defines how the inner estimates are built. Three inner weighting schemes are common: the centroid, the factorial and the path weighting scheme. For linear structural models, however, all schemes yield essentially the same results (Noonan and Wold, 1982) and therefore do not affect our proposed approach. For the purpose of our simulation, we employed the centroid scheme. For more details on the schemes, see, e.g., Tenenhaus *et al.* (2005).

3. Only correlation weights are considered, as these were originally used by Dijkstra and Henseler (2015a) to obtain consistent parameter estimates. However, consistent parameter estimates can be also obtained from the weights calculated by *mode B* or *mode C* (Dijkstra, 1981, Chap. 2 par. 5.2). Moreover, weights obtained by *mode A* are generally more stable, since those from *mode B* (regression weights) tend to suffer from multicollinearity. For an overview of outer weighting schemes and their properties, see Dijkstra (1981).

4. In fact, Equation (7) is not tied to using "converged" weights such as those obtained by PLS. Dijkstra and Schermelleh Engel (2014), for example, discuss what they call "one step" weight (essentially weight obtained after one iteration). In theory, any weight vector obtained after an arbitrary number of iterations (converged or not) will satisfy Equation (7).

5. The extension suggested here is not necessarily tied to using the squared Euclidean distance. As pointed out by Dijkstra (2013), weights could be introduced in Equation (17) to potentially reap efficiency gains. More generally, functions of ratios may be minimized; however, the solution will require iterative procedures. In this paper, the simplest approach was chosen to keep the main focus on our enhancement.

6. To draw a comprehensive picture of each modeling decision's influence on the results, we examined numerous alternative setups where we varied, for instance, the number of indicators, the number of observations, the indicator block whose errors where correlated and the magnitude of

different loadings. Additionally, as a robustness check, we conducted the simulation using non normally distributed data as in Dijkstra and Henseler (2015a) and applied all of the alternative approaches to obtain the correction factor described in Dijkstra (2013). Here, we describe only those setups that we deem most informative and most general, but note that none of the results of any other specifications were contrary to the central findings of the paper at hand. The results for the alternative specifications or the necessary R files to reproduce these can be obtained from the authors upon request.

7. The MoMpoly package is currently not on the Comprehensive R Archive Network. To replicate the results, a development version is available upon request.

8. Heywood cases in PLSc may occur for three reasons: the attenuation corrected or uncorrected estimated covariance matrix between proxies is not semi positive definite; standardized absolute loading estimates are larger than one; and the PLS algorithm has not converged.

References

Brown, T.A. (2015), *Confirmatory Factor Analysis for Applied Research*, Guilford Publications, New York, NY.

Charles, E.P. (2005), "The correction for attenuation due to measurement error: clarifying concepts and creating confidence sets", *Psychological Methods*, Vol. 10 No. 2, pp. 206 226.

Chiang, H. S. and Hsiao, K. L. (2015), "YouTube stickiness: the needs, personal, and environmental perspective", *Internet Research*, Vol. 25 No. 1, pp. 85 106.

Chin, W.W., Marcolin, B.L. and Newsted, P.R. (2003), "A partial least squares latent variable modeling approach for measuring interaction effects: results from a Monte Carlo simulation study and an electronic mail emotion/adoption study", *Information Systems Research*, Vol. 14 No. 2, pp. 189 217.

Cohen, J., Cohen, P., West, S.G. and Aiken, L.S. (2003), *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*, Lawrence Erlbaum Associates, Mahwah, NJ.

Dijkstra, T.K. (1981), "Latent variables in linear stochastic models", PhD thesis, Rijksuniversiteit te Groningen, Groningen.

Dijkstra, T.K. (2011), "Consistent partial least squares estimators for linear and polynomial factor models", available at: www.researchgate.net/publication/249998080 Consistent Partial Least Squares estimators for linear and polynomial factor models?channel=doi&linkId=00 b7d51e8d58fa0c3e000000&showFulltext=true (accessed September 6, 2018).

Dijkstra, T.K. (2013), "A note on how to make partial least squares consistent", available at: www. researchgate.net/publication/256089237 A note on how to make Partial Least Squares consistent (accessed September 6, 2018).

Dijkstra, T.K. (2018), "A suggested quasi empirical Bayes approach for handling 'Heywood' cases, very preliminary", available at: www.researchgate.net/publication/322331700 A suggested quasi empirical Bayes approach for handling %27Heywood%27 cases very preliminary?channel= doi&linkId=5a54cd6ca6fdcc51a61a5b22&showFulltext=true (accessed September 6, 2018).

Dijkstra, T.K. and Henseler, J. (2011), "Linear indices in nonlinear structural equation models: best fitting proper indices and other composites", *Quality & Quantity*, Vol. 45 No. 6, pp. 1505 1518.

Dijkstra, T.K. and Henseler, J. (2015a), "Consistent and asymptotically normal PLS estimators for linear structural equations", *Computational Statistics & Data Analysis*, Vol. 81, pp. 10 23, available at: www.sciencedirect.com/science/article/pii/S0167947314002126

Dijkstra, T.K. and Henseler, J. (2015b), "Consistent partial least squares path modeling", *MIS Quarterly*, Vol. 39 No. 2, pp. 297 316.

Dijkstra, T.K. and Schermelleh Engel, K. (2014), "Consistent partial least squares for nonlinear structural equation models", *Psychometrika*, Vol. 79 No. 4, pp. 585 604.

Gerbing, D.W. and Anderson, J.C. (1984), "On the meaning of within factor correlated measurement errors", *Journal of Consumer Research*, Vol. 11 No. 1, pp. 572 580.

Gu, H., Wen, Z. and Fan, X. (2017), "Examining and controlling for wording effect in a self report measure: a Monte Carlo simulation study", *Structural Equation Modeling: A Multidisciplinary Journal*, Vol. 24 No. 4, pp. 545 555.

Hair, J.F., Sarstedt, M., Pieper, T.M. and Ringle, C.M. (2012), "The use of partial least squares structural equation modeling in strategic management research: a review of past practices and recommendations for future applications", *Long Range Planning*, Vol. 45 Nos 5 6, pp. 320 340.

Hair, J.F., Sarstedt, M., Ringle, C.M. and Mena, J.A. (2012), "An assessment of the use of partial least squares structural equation modeling in marketing research", *Journal of the Academy of Marketing Science*, Vol. 40 No. 3, pp. 414 433.

Henseler, J. (2009), "On the convergence of the partial least squares path modeling algorithm", *Computational Statistics*, Vol. 25 No. 1, pp. 107 120.

Henseler, J. and Chin, W.W. (2010), "A comparison of approaches for the analysis of interaction effects between latent variables using partial least squares path modeling", *Structural Equation Modeling: A Multidisciplinary Journal*, Vol. 17 No. 1, pp. 82 109.

Hermida, R. (2015), "The problem of allowing correlated errors in structural equation modeling: concerns and considerations", *Computational Methods in Social Sciences*, Vol. 3 No. 1, pp. 5 17.

Hwang, H., Malhotra, N.K., Kim, Y., Tomiuk, M.A. and Hong, S. (2010), "A comparative study on parameter recovery of three approaches to structural equation modeling", *Journal of Marketing Research*, Vol. 47 No. 4, pp. 699 712.

Jöreskog, K.G. (1978), "Structural analysis of covariance and correlation matrices", *Psychometrika*, Vol. 43 No. 4, pp. 443 477.

Kettenring, J.R. (1971), "Canonical analysis of several sets of variables", *Biometrika*, Vol. 58 No. 3, pp. 433 451.

Khan, G.F., Sarstedt, M., Shiau, W. L., Hair, J.F., Ringle, C.M. and Fritze, M. (forthcoming), "Methodological research on partial least squares structural equation modeling (PLS SEM): an analysis based on social network approaches", *Internet Research*.

Landis, R.S., Edwards, B.D. and Cortina, J.M. (2009), "On the practice of allowing correlated residuals among indicators in structural equation models", in Lance, C.E. and Vandenberg, R.J. (Eds), *On the Practice of Allowing Correlated Residuals Among Indicators in Structural Equation Models*, Routledge and Taylor & Francis Group, New York, NY, pp. 193 214.

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

McNeish, D. (2018), "Thanks coefficient alpha, we'll take it from here", *Psychological Methods*, Vol. 23 No. 3, pp. 412 433.

MacKenzie, S.B. and Podsakoff, P.M. (2012), "Common method bias in marketing: causes, mechanisms, and procedural remedies", *Journal of Retailing*, Vol. 88 No. 4, pp. 542 555.

Müller, T., Schuberth, F. and Henseler, J. (2018), "PLS path modeling   a confirmatory approach to study tourism technology and tourist behaviour", *Journal of Hospitality and Tourism Technology*, Vol. 9 No. 3, pp. 249 266.

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), *PLS Path Modeling with Indirectly Observed Variables: A Comparison of Alternative Estimates for the Latent Variable*, North Holland, Amsterdam, pp. 75 94.

Padilla, M.A. and Veprinsky, A. (2012), "Correlation attenuation due to measurement error", *Educational and Psychological Measurement*, Vol. 72 No. 5, pp. 827 846.

Paxton, P., Curran, P.J., Bollen, K.A., Kirby, J. and Chen, F. (2001), "Monte Carlo experiments: design and implementation", *Structural Equation Modeling: A Multidisciplinary Journal*, Vol. 8 No. 2, pp. 287 312.

Podsakoff, P.M., MacKenzie, S.B. and Podsakoff, N.P. (2012), "Sources of method bias in social science research and recommendations on how to control it", *Annual Review of Psychology*, Vol. 63, pp. 539 569.

Raykov, T., Marcoulides, G.A. and Patelis, T. (2014), "The importance of the assumption of uncorrelated errors in psychometric theory", *Educational and Psychological Measurement*, Vol. 75 No. 4, pp. 634 647.

R Core Team (2017), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna.

Rigdon, E.E., Sarstedt, M. and Ringle, C.M. (2017), "On comparing results from CB SEM and PLS SEM: five perspectives and five recommendations", *Marketing ZFP*, Vol. 39 No. 3, pp. 4 16.

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 No. 1, pp. iii xiv.

Rubio, D.M. and Gillespie, D.F. (1995), "Problems with error in structural equation models", *Structural Equation Modeling: A Multidisciplinary Journal*, Vol. 2 No. 4, pp. 367 378.

Saris, W.E. and Aalberts, C. (2003), "Different explanations for correlated disturbance terms in MTMM studies", *Structural Equation Modeling: A Multidisciplinary Journal*, Vol. 10 No. 2, pp. 193 213.

Sarstedt, M., Ringle, C.M., Henseler, J. and Hair, J.F. (2014), "On the emancipation of PLS SEM: a commentary on Rigdon (2012)", *Long Range Planning*, Vol. 47 No. 3, pp. 154 160.

Schneeweiss, H. (1993), "Consistency at large in models with latent variables", in Haagen, K., Bartholomew, D.J. and Deistler, M. (Eds), *Consistency at Large in Models with Latent Variables*, North Holland, Amsterdam, pp. 299 322.

Schuberth, F., Schamberger, T. and Dijkstra, T.K. (2017), "MoMpoly: non iterative method of moments for polynomial factor models", R package version 0.1.4.

Tenenhaus, M., Vinzi, V.E., Chatelin, Y. M. and Lauro, C. (2005), "PLS path modeling", *Computational Statistics & Data Analysis*, Vol. 48 No. 1, pp. 159 205.

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

Westfall, P.H., Henning, K.S. and Howell, R.D. (2012), "The effect of error correlation on interfactor correlation in psychometric measurement", *Structural Equation Modeling: A Multidisciplinary Journal*, Vol. 19 No. 1, pp. 99 117.

Wold, H. (1975), *Path Models with Latent Variables: The NIPALS Approach*, Academic Press, New York, NY.

Wu, Y. L. and Li, E.Y. (2018), "Marketing mix, customer value, and customer loyalty in social commerce", *Internet Research*, Vol. 28 No. 1, pp. 74 104.

Yan, Y., Zhang, X., Zha, X., Jiang, T., Qin, L. and Li, Z. (2017), "Decision quality and satisfaction: the effects of online information sources and self efficacy", *Internet Research*, Vol. 27 No. 4, pp. 885 904.

Zimmerman, D.W. (2007), "Correction for attenuation with biased reliability estimates and correlated errors in populations and samples", *Educational and Psychological Measurement*, Vol. 67 No. 6, pp. 920 939.

Corresponding author
Florian Schuberth can be contacted at: f.schuberth@utwente.nl