Sensitivity Analysis of the Weights of the Composites Under Partial Least-Squares Approach to Structural Equation Modeling

Ke-Hai Yuan[1], Yong Wen[2] and Jiashan Tang[2]

[1]University of Notre Dame

[2]Nanjing University of Posts and Telecommunications

March 5, 2022
Revised, July 1, 2022
Third revision, July 22, 2022

Yuan, K.-H., Wen, Y., & Tang, J. (2023). Sensitivity analysis of the weights of the composites under partial least-squares approach to structural equation modeling. Structural Equation Modeling, 30(1), 53–69. doi: 10.1080/10705511.2022.2106487
IES funding acknowledged: R305D210023.

This work was supported by a grant from the Department of Education (R305D210023), and by a grant from the Natural Science Foundation of China (31971029). However, the contents of the study do not necessarily represent the policy of the funding agencies, and you should not assume endorsement by the Federal Government. Correspondence concerning this article should be addressed to Ke-Hai Yuan (kyuan@nd.edu).

Abstract

Structural equation modeling (SEM) and path analysis using composite-scores are distinct classes of methods for modeling the relationship of theoretical constructs. The two classes of methods are integrated in the partial-least-squares approach to structural equation modeling (PLS-SEM), which systematically generates weighted composites and uses them to conduct path analysis of the structural model via the least-squares method. However, the goodness of PLS-SEM depends on the statistical properties of the composites, which are further determined by the formulations of the weights. This article studies how the formulations of PLS-SEM composites are affected by model specification, with focus on the sensitivity of the weights to common specification errors. Results indicate that the weights under PLS-SEM mode A are not affected by within-block error-covariances but those under mode B are. While between-block error-covariances and cross-loadings only affect the weights of the involved items under both PLS-SEM modes A and B, the weights under mode B are much more sensitive than those under mode A. In contrast, the weights under a recently proposed transformed mode (denoted as $\mathrm{B_A}$) are a compromise between those of modes A and B. The findings not only advance the understanding of the PLS-SEM methodology but also facilitate model diagnostics. Empirical applications of the results are illustrated via the analysis of a real dataset.
Keywords: Structural equation model, path analysis, weighted composite, error-covariance, cross-loading.

1. Introduction

Structural equation modeling (SEM) and path analysis using composite-scores are distinct classes of methods for modeling the relationship of theoretical constructs. The two classes of methods are integrated in the partial-least-squares approach to structural equation modeling (PLS-SEM), which systematically generates weighted composites and uses them to conduct path analysis of the structural model via the least-squares (LS) method. To distinguish the conventional SEM method from PLS-SEM, the former is often called covariance-based SEM (CB-SEM). The advantages of CB-SEM include directly yielding consistent parameter estimates and fit indices for evaluating the goodness of the overall model structure. The strength of PLS-SEM is in prediction and classification of individuals (Deng & Yuan, 2022; Hair et al., 2017; Henseler, 2021; Rigdon, Sarstedt & Ringle, 2017). In particular, path analysis with composite-scores possesses the property of yielding the least mean-square-prediction errors (Fuller, 1987) and more powerful statistical test on the path coefficients (Deng & Yuan, 2022; Yuan & Fang, 2022). However, not all weighted composites are equivalent, and the goodness of path analysis closely depends on the statistical properties of the composites, which are further determined by the formulations of the weights. This article studies how the formulations of PLS-SEM composites are affected by model specification, with focus on the sensitivity of the weights to common specification errors. Our aim is to advance the understanding of PLS-SEM so that the method can be better used in practice.

The formulation of weights is fundamental to understanding the properties of PLS-SEM and the corresponding results. The conventional PLS methodology has two ways of computing weighted composites, termed as modes A and B, respectively (Wold, 1980, 1982). For a correctly specified latent-variable model with reflective indicators, Dijkstra (1983, see also Schneeweiss, 1993) showed that, the weight vector for each block[1] of indicators under PLS-SEM mode A is proportional to the vector of factor loadings of that block, and that under PLS-SEM mode B is proportional to the vector of factor loadings multiplied by the precision matrix in front. These results allow us to analytically study the goodness of PLS-SEM against CB-SEM and other methods of path analysis with weighted composites. In particular, composites under PLS-SEM mode A may not be as reliable as the simple averages

[1]The set of manifest variables measuring a single latent variable is referred to as a block of indicators.

whereas the composites under PLS-SEM mode B (when applying to models with reflective indicators) achieve the maximum reliability (Yuan & Deng, 2021). We would expect that mode B performs better than mode A in practice. However, Dijkstra and Henseler (2015) noted that mode A of PLS-SEM yields numerically more stable results, and a real data example in Yuan and Deng (2021) showed that weights under mode B of PLS-SEM may have opposite signs from the factor loadings estimated under CB-SEM. These results are closely related to the sensitivity and stability of the weights to specification errors. To understand the inconsistency between analytical property and empirical performance of the PLS-SEM methodology, it is necessary to study how weights are affected when different parts of a latent-variable model are misspecified.

While CB-SEM facilitates consistent parameter estimates, the values of the path coefficients are for the relationship among the latent variables that represent the population distribution, and all individuals are equivalent under such a relationship. In practice when observed scores are used for prediction or diagnosis, individuals are no longer equivalent. A person with greater test scores is expected to perform better on the criterion variable, and such a relationship is directly characterized by the regression coefficients under PLS-SEM. But the quality of the prediction under PLS-SEM still depends on how the composite-scores are formulated. In particular, the reliabilities of the composites affect many aspects of path analysis, e.g., bias, R^2, prediction error, etc. (Cochran, 1970; Fuller, 1987). Although results of Dijkstra (1983) and Schneeweiss (1993) provide the guidance for choosing a proper PLS-SEM mode in data analysis, they are based on correctly specified models that are not obtainable in practice. While we expect that the weights under either mode A or B of PLS-SEM to be affected by model misspecification, it is not clear how they are affected or whether a particular misspecification will affect all the weights globally or only the weight local to the location of misspecification is affected. Studies of the sensitivity and/or stability of the weights of composites to model misspecification not only facilitate the proper use of PLS-SEM but also advance our understanding of the particular results in real data analyses.

The purpose of this article is to systematically study how PLS-SEM weights are affected when different parts of a latent-variable model are misspecified. We will have both analytical and numerical results on the sensitivity of weights under both modes A and B as well as of weights under a recently proposed mode by Yuan and Deng (2021). The obtained results will be used to analyze a real dataset for which PLS-SEM mode B yields negative weights for some items. Section 2 contains an overview of PLS-SEM methodology and existing results on weights to set up the context of the study. Analytical results on sensitivity of weights are presented in section 3. Numerical illustrations of the sensitivity of weights are given in section 4. Section 5 contains the analysis of the real dataset. Conclusion and recommendation for path analysis with weighted composites are provided at the end.

2. PLS-SEM and Weights

In this section we give a brief overview of the PLS-SEM methodology to set up the context for the focused study. Readers who are interested in more systematic introductions of the PLS method are referred to Hair et al. (2017), Henseler (2021), and Tenenhaus et al. (2005). Throughout this article we will assume that all the indicators are reflective and contain measurement errors. Readers who are interested in developments for correctly specified models with formative indicators are referred to Dijkstra (2017), and Cho and Choi (2020). Since our interest is the sensitivity of the weights of the composites under PLS-SEM, we will highlight existing results on weights and spell out the issues to address. Because means of the manifest variables or latent variables do not affect the weights under PLS-SEM, all the variables are assumed mean centered in our presentation.

2.1 PLS-SEM

PLS-SEM consists of two stages. Weighted composites or weights are computed in the first stage where the distinction between modes A, B, and $\mathrm{B_A}$ is made. The second stage is the same for the three modes where the structural model is estimated using the composites obtained at stage 1. Both modes A and B are computed by least-squares (LS) regression via the so-called environmental variables, while weights under mode $\mathrm{B_A}$ are obtained by transforming the weights under mode A using a LS method for factor analysis.

Under PLS-SEM, manifest variables are assumed unidimensional[2], where each manifest variable or indicator loads on a single latent variable without correlated errors. Let $\mathbf{x} = (x_1, x_2, \ldots, x_{p_x})'$ be the vector of indicators for a latent variable ξ. Other latent variables in

[2]This terminology aligns with the discussion of Anderson and Gerbing (1988) who wrote "By contrast, measurement models that contain correlated measurement errors or that have indicators that load on more than one estimated construct do not represent unidimensional construct measurement" (p. 415). A reviewer noted that, while an indicator may load on more than one latent variable, the indicators within each cluster can still be regarded as unidimensional because they follow a one-factor model.

the model might be (1) predicted by ξ, (2) predict ξ, (3) just correlate with ξ, (4) neither correlate with ξ nor predict or being predicted by ξ. The latent variables in the first two scenarios are regarded as being directly connected with ξ while correlations or covariances among latent variables are not regarded as direct connections under PLS-SEM. Suppose ξ is directly connected with m other latent variables ξ_1, ξ_2, ..., ξ_m. The *environmental variable* of ξ is formulated by a linear combination $\bar{\xi} = r_1\xi_1 + r_2\xi_2 + \ldots + r_m\xi_m$, where r_j is the sign of the correlation between ξ_j and ξ or the value of the correlation itself (see e.g., Henseler 2021, p. 91). For a model that contains two blocks of indicators, let $\mathbf{x}$ be the vector of indicators for latent variable ξ, $\mathbf{y}$ be the vector of indicators for latent variable η, and $\eta = \gamma\xi + \zeta$ be the structure model with $\gamma > 0$. Then ξ is the environmental variable of η, and η is the environmental variable of ξ.

The weights under both modes A and B are computed iteratively (Tenenhaus et al., 2005; Wold, 1980). The simple average of each block of indicators serves as a good starting value for the corresponding composite, and the starting values for the corresponding environmental variables automatically follow by substituting each latent variable with the corresponding composite. Under mode A, the weight for x_k (an element of $\mathbf{x}$) is updated by the slope parameter of LS regression of x_k on the environmental variable $\bar{\xi}$. That is, $\hat{\mathbf{w}}_a = (\hat{w}_{a1}, \hat{w}_{a2}, \ldots, \hat{w}_{ap_x})'$ with $\hat{w}_{ak} = s_{x_k\bar{\xi}}/s_{\bar{\xi}\bar{\xi}}$, where s is the notation for sample variance-covariance. Under mode B, weights of the indicators in $\mathbf{x}$ are updated by the coefficients of multiple regression of the environmental variable $\bar{\xi}$ on $\mathbf{x}$. That is, $\hat{\mathbf{w}}_b = \mathbf{S}_{xx}^{-1}\mathbf{s}_{x\bar{\xi}}$, where $\mathbf{S}$ is the sample covariance matrix and $\mathbf{s}$ is the vector of sample covariances. The updated weights generate updated composites and the corresponding updated environmental variables, which permit the computation of new weights by LS regression. The iteration alternates across all blocks of indicators until all the weights are stabilized. Because the regression coefficients depend on the scales of the involved variables, it is a common practice to proportionally scale the weight vectors so that each weighted composite has a sample variance of 1.0.

2.2 Weights

PLS-SEM can also be considered at the population level. Suppose we have two blocks of indicators, $\mathbf{x}$ for ξ and $\mathbf{y}$ for η, and they are connected by the structural model $\eta = \gamma\xi + \zeta$. Let $\mathbf{\Sigma}_{xx}$, $\mathbf{\Sigma}_{xy} = \mathbf{\Sigma}_{yx}'$ and $\mathbf{\Sigma}_{yy}$ represent the population covariance matrices of $\mathbf{x}$ with $\mathbf{x}$, $\mathbf{x}$ with $\mathbf{y}$, and $\mathbf{y}$ with $\mathbf{y}$, respectively. Then, regardless of whether the measurement and structural models are correctly specified or not, the PLS-SEM algorithm described in the previous subsection directly implies that the vectors of weights $\mathbf{w}_x$ and $\mathbf{w}_y$ under mode A satisfy the equations (see also Dijkstra, 1983)

$$\mathbf{w}_{ax} = c_x \boldsymbol{\Sigma}_{xy} \mathbf{w}_y \text{ and } \mathbf{w}_{ay} = c_y \boldsymbol{\Sigma}_{yx} \mathbf{w}_x, \tag{1}$$

where c_x and c_y are scalars whose particular values are not material due to standardization. That is, $\mathbf{w}_x' \boldsymbol{\Sigma}_{xx} \mathbf{w}_x = 1$ and $\mathbf{w}_y' \boldsymbol{\Sigma}_{yy} \mathbf{w}_y = 1$ will cancel the effect of c_x and c_y. Under mode B, the PLS-SEM algorithm implies that the weight vectors $\mathbf{w}_x$ and $\mathbf{w}_y$ satisfy

$$\mathbf{w}_{bx} = c_x \boldsymbol{\Sigma}_{xx}^{-1} \boldsymbol{\Sigma}_{xy} \mathbf{w}_y \text{ and } \mathbf{w}_{by} = c_y \boldsymbol{\Sigma}_{yy}^{-1} \boldsymbol{\Sigma}_{yx} \mathbf{w}_x. \tag{2}$$

Wold (1980, 1982) recommended mode A for models with reflective indicators and mode B for models with formative indicators. Such a recommendation is intuitive but is not justified by statistical or psychometric theory (Yuan & Deng, 2021).

Suppose the vector $\mathbf{x}$ contains p_x indicators that measure a single factor ξ, and $\mathbf{y}$ contains p_y indicators that measure a single factor η. Then the relationship among $\mathbf{x}$ and $\mathbf{y}$ can be correctly represented by

$$\mathbf{x} = \boldsymbol{\lambda}_x \xi + \boldsymbol{\varepsilon}_x, \quad \mathbf{y} = \boldsymbol{\lambda}_y \eta + \boldsymbol{\varepsilon}_y, \tag{3}$$

and

$$\eta = \gamma \xi + \zeta, \tag{4}$$

where $\boldsymbol{\lambda}_x$ and $\boldsymbol{\lambda}_y$ are vectors of factor loadings, and $\boldsymbol{\Psi}_{xx} = \text{Cov}(\boldsymbol{\varepsilon}_x)$ and $\boldsymbol{\Psi}_{yy} = \text{Cov}(\boldsymbol{\varepsilon}_y)$ are diagonal matrices of error variances. The two equations in (3) are referred to as the measurement model while the one in (4) is referred to as the structural model (Anderson & Gerbing, 1988). Note that both $\mathbf{x}$ and $\mathbf{y}$ in equation (3) are reflective indicators. When applying mode A to the model in equations (3) and (4), the weights for $\mathbf{x}$ and $\mathbf{y}$ are respectively

$$\mathbf{w}_{ax} = c_x \boldsymbol{\lambda}_x \text{ and } \mathbf{w}_{ay} = c_y \boldsymbol{\lambda}_y. \tag{5}$$

When applying mode B to the model in equations (3) and (4), the weights for $\mathbf{x}$ and $\mathbf{y}$ are

$$\mathbf{w}_{bx} = c_x \boldsymbol{\Sigma}_{xx}^{-1} \boldsymbol{\lambda}_x \text{ and } \mathbf{w}_{by} = c_y \boldsymbol{\Sigma}_{yy}^{-1} \boldsymbol{\lambda}_y. \tag{6}$$

The results in (5) and (6) were obtained by Dijkstra (1983) and also presented by Schneeweiss (1993). They can also be directly derived from equations (1), (2), (3) and (4). For correctly specified models (i.e., unidimensional without correlated errors), these results also hold with more than two blocks of indicators.

Using the results in equation (6), Yuan and Deng (2021) showed that, when applying mode B to models with reflective indicators, the resulting composites are equivalent[3] to the Bartlett-factor-scores (under CB-SEM), which are known to enjoy the maximum reliability among all weighted composites. Because mode A is numerically more stable than mode B, Yuan and Deng (2021) proposed a procedure to transform mode A to mode B according to the structure of a one-factor model for each block of indicators. For the transformed mode $\mathrm{B_A}$, the weight vector for each block of indicators is obtained using the formula

$$\mathbf{w}_{b_a} = c_* \mathbf{\Psi}_*^{-1} \mathbf{w}_a, \tag{7}$$

where $\mathbf{w}_a$ is the weight vector (for either the block $\mathbf{x}$ or $\mathbf{y}$) under mode A, and $\mathbf{\Psi}_*$ is a diagonal matrix obtained by fitting the structural model

$$\mathbf{\Sigma}(\boldsymbol{\theta}) = \theta_* \mathbf{w}_a \mathbf{w}_a' + \mathbf{\Psi}_* \tag{8}$$

to the covariance matrix $\mathbf{\Sigma}$ of the block of indicators corresponding to the weight vector $\mathbf{w}_a$. Note that only θ_* and the diagonal elements of $\mathbf{\Psi}_*$ in (8) are subject to estimation whereas the vector $\mathbf{w}_a$ is regarded as given. Yuan and Deng (2021) suggested using the LS method to estimate θ_* and $\mathbf{\Psi}_*$, which enjoy analytical solutions. Also note that, for correctly specified models, $\mathbf{\Psi}_* = \mathbf{\Psi}$ and $\mathbf{w}_{b_a} = \mathbf{w}_b$ in the population.

2.3 Sensitivity analysis

The results in equations (5), (6) and (7) constitute the cornerstone for analytically studying the properties of PLS-SEM methodology. However, they are for correctly specified models, which may not be obtainable in the real world. We are interested in how model misspecifications are going to affect the weight vectors in these equations. Sensitivity analysis of the weights under modes A, B and $\mathrm{B_A}$ will provide the information on why they are empirically different from their expected values. The study will also provide the key for understanding why mode B results in negative item weights. Sensitivity analysis also allows us to better understand the differences between PLS-SEM and path analysis with other types of composite-scores.

[3] Yuan and Deng (2021) showed that Bartlett-factor-scores are also statistically equivalent to regression-factor-scores in conducting path analysis.

In developing a procedure for correcting the bias of parameter estimates under PLS-SEM mode A, Rademaker, Schuberth and Dijkstra (2019) noted that weights under mode A are not affected by within-block error-covariances. They did not study the effect of between-block error-covariances or cross-loadings on weights following PLS-SEM mode A nor the effect of any misspecification on weights under PLS-SEM modes B and $\mathrm{B_A}$. While our main interest is to identify the causes for mode B to yield negative weights, we will also study the effects of different types of model misspecification on modes A and $\mathrm{B_A}$, and compare their sensitivities to model misspecification against that of mode B.

We only consider models with reflective indicators in this article. This is because formative indicators typically do not consider measurement errors and their correlations do not need to be due to sharing a common construct. The results in (5), (6) and (7) are all for models with reflective indicators. The property for composites under PLS-SEM mode B and $\mathrm{B_A}$ to enjoy maximum reliability is also for models with reflective indicators. Models with reflective indicators are also called latent variable models, which typically include a measurement model and a structural model. We will only consider misspecifications in the measurement model, mostly because indicators under PLS-SEM are unidimensional. In particular, under PLS-SEM the indicators within each block share a single common factor and there is no mechanism to account for cross-loadings or correlated errors (see e.g., Tenenhaus et al., 2005). While the structural model under PLS-SEM can be misspecified as well, it also enjoys the freedom of being specified as saturated. There are three types of specification errors in the measurement model. They are (1) within-block error-covariances, (2) between-block error-covariances, and (3) cross-loadings. Because standard PLS-SEM does not allow the inclusion of these parameters, a model misspecification occurs when a parameter in one of the three types is non-zero. Other specification errors are also possible, e.g., nonlinear relationship or lack of additional latent variables. These are beyond the scope of this article.

3. Analytical Results on Sensitivity of Weights

In this section we examine how the weights in equation (5), (6) and (7) are affected by model misspecification, including error-covariances and cross-loadings. For the purpose of clean results, we mainly consider a model with two latent variables. The results also hold with more latent variables as will be shown in a separate subsection.

Insert Figure 1 about here

Consider the model in Figure 1, where the solid arrows represent a typical two-latent-variable model corresponding to the one specified by equations (3) and (4). Our analysis in subsections 3.1 to 3.7 is for this model. A misspecification occurs whenever any of the coefficients for the dashed arrows in Figure 1 is non-zero. When the method is clear from the context, we will use $\mathbf{w}_x$ and $\mathbf{w}_y$ to represent the vectors of weights corresponding to the blocks $\mathbf{x}$ and $\mathbf{y}$, respectively. When the block is clear from the context, we will use $\mathbf{w}_a$, $\mathbf{w}_b$ and $\mathbf{w}_{b_a}$ to represent the vectors of weights corresponding to PLS-SEM modes A, B and $\mathrm{B_A}$, respectively. Note that the scales of the weighted composites are arbitrary, so we can focus on the relative change of the weights using their counterpart under the correctly specified model as the reference. In particular, the words "greater" and "smaller" are also in comparison with the counterparts under the correctly specified model (i.e., the coefficients for all the dashed-arrows in Figure 1 are zero). Throughout this section, we will use the letter "c" to represent a constant whose particular value is not material.

3.1 Weights under mode A by within-block error-covariances

Note that within-block error-covariances only affect the covariance matrix $\mathbf{\Sigma}_{xx}$ or $\mathbf{\Sigma}_{yy}$, not $\mathbf{\Sigma}_{xy}$ nor $\mathbf{\Sigma}_{yx}$. In particular, $\mathbf{\Sigma}_{xy} = \mathbf{\Sigma}'_{yx} = \gamma\boldsymbol{\lambda}_x\boldsymbol{\lambda}'_y$ for the model in Figure 1 regardless of whether within-block error-covariances are zero or not. Also note that, for two vectors $\mathbf{a}$ and $\mathbf{b}$ of the same dimension, regardless of their values, $\mathbf{a}'\mathbf{b}$ is a scalar. It follows from equation (1) that

$$\mathbf{w}_x = c_x\mathbf{\Sigma}_{xy}\mathbf{w}_y = c_x\gamma(\boldsymbol{\lambda}_x\boldsymbol{\lambda}'_y)\mathbf{w}_y = c_x\gamma\boldsymbol{\lambda}_x(\boldsymbol{\lambda}'_y\mathbf{w}_y) = c_{x*}\boldsymbol{\lambda}_x$$

and

$$\mathbf{w}_y = c_y\mathbf{\Sigma}_{yx}\mathbf{w}_x = c_y\gamma(\boldsymbol{\lambda}_y\boldsymbol{\lambda}'_x)\mathbf{w}_x = c_y\gamma\boldsymbol{\lambda}_y(\boldsymbol{\lambda}'_x\mathbf{w}_x) = c_{y*}\boldsymbol{\lambda}_y.$$

Thus, a within-block error-covariance does not affect the weights under mode A, and they are still proportional to the factor loadings of the block (Rademaker et al., 2019). But the size of the absolute values of the elements of the weight vector changes, due to standardization (i.e., $\mathbf{w}'_x\mathbf{\Sigma}_{xx}\mathbf{w}_x = 1$) under PLS-SEM. They become proportionally smaller when positive within-block error-covariances exist, and the other way around when negative within-block error-covariances exist.

3.2 Weights under mode A by between-block error-covariances

Using the notation in equation (3), suppose the jth element of $\boldsymbol{\varepsilon}_x$ in Figure 1 is correlated with the kth element of $\boldsymbol{\varepsilon}_y$, with $\psi_{jk} = \text{Cov}(\varepsilon_{x_j}, \varepsilon_{y_k})$. Then we have

$$\boldsymbol{\Sigma}_{xy} = \boldsymbol{\Sigma}'_{yx} = \gamma \boldsymbol{\lambda}_x \boldsymbol{\lambda}'_y + \psi_{jk} \mathbf{e}_j \mathbf{e}'_k, \tag{9}$$

where $\mathbf{e}_j$ is a vector of length p_x with the jth element being 1 and others being zero, and $\mathbf{e}_k$ is a vector of length p_y with the kth element being 1 and others being zero. It follows from equations (1) and (9) that

$$\mathbf{w}_x = c_x \boldsymbol{\Sigma}_{xy} \mathbf{w}_y = c_x(\gamma \boldsymbol{\lambda}_x \boldsymbol{\lambda}'_y + \psi_{jk} \mathbf{e}_j \mathbf{e}'_k)\mathbf{w}_y = c_{x*} \boldsymbol{\lambda}_x + c_* \mathbf{e}_j.$$

Thus, except for the jth element of $\mathbf{w}_x$, the other elements of $\mathbf{w}_x$ are still proportional to those of $\boldsymbol{\lambda}_x$. That is, only the weights for items that have non-zero between-block error-covariances are affected by the misspecification. The weights for the items that have null between-block error-covariances are still proportional to the factor loadings[4]. But their values will proportionally change due to scaling via standardization.

3.3 Weights under mode A by cross-loadings

Suppose some items in $\mathbf{x}$ also load on the latent variable η in Figure 1. Then the measurement model in equation (3) does not correctly represent the relationship between $\mathbf{x}$ and the latent variables. We need to write the measurement model as

$$\begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\lambda}_x & \boldsymbol{\lambda}_h \\ \mathbf{0} & \boldsymbol{\lambda}_y \end{pmatrix} \begin{pmatrix} \xi \\ \eta \end{pmatrix} + \begin{pmatrix} \boldsymbol{\varepsilon}_x \\ \boldsymbol{\varepsilon}_y \end{pmatrix}, \tag{10}$$

where $\boldsymbol{\lambda}_h$ is the vector of cross-loadings of $\mathbf{x}$ on η. Thus, the covariance matrix of the $(p_x + p_y)$-vector $(\mathbf{x}', \mathbf{y}')'$ is given by

$$\begin{pmatrix} \boldsymbol{\Sigma}_{xx} & \boldsymbol{\Sigma}_{xy} \\ \boldsymbol{\Sigma}_{yx} & \boldsymbol{\Sigma}_{yy} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\lambda}_x & \boldsymbol{\lambda}_h \\ \mathbf{0} & \boldsymbol{\lambda}_y \end{pmatrix} \begin{pmatrix} \phi_{\xi\xi} & \phi_{\xi\eta} \\ \phi_{\eta\xi} & \phi_{\eta\eta} \end{pmatrix} \begin{pmatrix} \boldsymbol{\lambda}'_x & \mathbf{0} \\ \boldsymbol{\lambda}'_h & \boldsymbol{\lambda}'_y \end{pmatrix} + \begin{pmatrix} \boldsymbol{\Psi}_{xx} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Psi}_{yy} \end{pmatrix},$$

where the ϕs are for the variances-covariance of ξ and η. Direct matrix multiplication yields

$$\begin{gathered} \boldsymbol{\Sigma}_{xx} = \boldsymbol{\lambda}_x \phi_{\xi\xi} \boldsymbol{\lambda}'_x + \boldsymbol{\lambda}_h \phi_{\eta\xi} \boldsymbol{\lambda}'_x + \boldsymbol{\lambda}_x \phi_{\xi\eta} \boldsymbol{\lambda}'_h + \boldsymbol{\lambda}_h \phi_{\eta\eta} \boldsymbol{\lambda}'_h + \boldsymbol{\Psi}_{xx}, \\ \boldsymbol{\Sigma}_{xy} = \boldsymbol{\Sigma}'_{yx} = \boldsymbol{\lambda}_x \phi_{\xi\eta} \boldsymbol{\lambda}'_y + \boldsymbol{\lambda}_h \phi_{\eta\eta} \boldsymbol{\lambda}'_y, \text{ and } \boldsymbol{\Sigma}_{yy} = \boldsymbol{\lambda}_y \phi_{\eta\eta} \boldsymbol{\lambda}'_y + \boldsymbol{\Psi}_{yy}. \end{gathered} \tag{11}$$

With the $\boldsymbol{\Sigma}_{xy}$ in equation (11), it follows from equation (1) that

$$\mathbf{w}_x = c_x(\boldsymbol{\lambda}_x \phi_{\xi\eta} \boldsymbol{\lambda}'_y + \boldsymbol{\lambda}_h \phi_{\eta\eta} \boldsymbol{\lambda}'_y)\mathbf{w}_y = c_{x*} \boldsymbol{\lambda}_x + c_* \boldsymbol{\lambda}_h.$$

[4]The factor loadings are defined as the population values when the SEM model is correctly specified by including the non-zero error-covariances as free parameters.

Suppose only the jth element of $\mathbf{x}$ loads on η, then $\boldsymbol{\lambda}_h$ becomes $\lambda_h \mathbf{e}_j$. Thus, the cross-loading of x_j on η in Figure 1 only affects the value of the jth element of $\mathbf{w}_x$. The other elements of $\mathbf{w}_x$ are still proportional to those of $\boldsymbol{\lambda}_x$. But the size of all the elements of $\mathbf{w}_x$ will be adjusted proportionally after standardization (i.e., $\mathbf{w}_x' \boldsymbol{\Sigma}_{xx} \mathbf{w}_x = 1$).

For the weight vector $\mathbf{w}_y$, we have, according to equations (1) and (11),

$$\mathbf{w}_y = c_y \boldsymbol{\Sigma}_{yx} \mathbf{w}_x = c_y (\boldsymbol{\lambda}_y \phi_{\eta\xi} \boldsymbol{\lambda}_x' + \boldsymbol{\lambda}_y \phi_{\eta\eta} \boldsymbol{\lambda}_h') \mathbf{w}_x = c_{y*} \boldsymbol{\lambda}_y.$$

Thus, cross-loadings of items in $\mathbf{x}$ on η do not affect the weights for the block $\mathbf{y}$ under mode A, and $\mathbf{w}_y$ is still proportional to $\boldsymbol{\lambda}_y$.

In parallel, cross-loadings of items in $\mathbf{y}$ on ξ in Figure 1 do not affect the weights of $\mathbf{w}_x$ under mode A. They are still proportional to $\boldsymbol{\lambda}_x$. The elements of $\mathbf{w}_y$ corresponding to items that do not cross-load on ξ are also proportional to those of $\boldsymbol{\lambda}_y$. But weights for items in $\mathbf{y}$ that have loadings on ξ will be affected, they might become greater or smaller according to the sign and size of the cross-loadings. We will illustrate such a relationship numerically in the next section.

3.4 Weights under mode B by within-block error-covariances

When errors within the block $\mathbf{x} = (x_1, x_2, \ldots, x_{p_x})'$ in Figure 1 are correlated, we can write the covariance matrix of $\mathbf{x}$ as

$$\boldsymbol{\Sigma}_{xx} = \boldsymbol{\lambda}_x \phi_{\xi\xi} \boldsymbol{\lambda}_x' + \boldsymbol{\Psi}_{xx},$$

where the matrix $\boldsymbol{\Psi}_{xx} = \mathrm{Cov}(\boldsymbol{\varepsilon}_x)$ has some non-zero off-diagonal elements corresponding to the error covariances. Let $\phi_{\xi\xi} = \mathrm{Var}(\xi) = 1$ for model identification. Then the inverse of the matrix $\boldsymbol{\Sigma}_{xx}$ is given by

$$\boldsymbol{\Sigma}_{xx}^{-1} = \boldsymbol{\Psi}_{xx}^{-1} - \frac{\boldsymbol{\Psi}_{xx}^{-1} \boldsymbol{\lambda}_x \boldsymbol{\lambda}_x' \boldsymbol{\Psi}_{xx}^{-1}}{1 + \boldsymbol{\lambda}_x' \boldsymbol{\Psi}_{xx}^{-1} \boldsymbol{\lambda}_x}. \tag{12}$$

Thus,

$$\boldsymbol{\Sigma}_{xx}^{-1} \boldsymbol{\lambda}_x = \boldsymbol{\Psi}_{xx}^{-1} \boldsymbol{\lambda}_x - \frac{\boldsymbol{\Psi}_{xx}^{-1} \boldsymbol{\lambda}_x \boldsymbol{\lambda}_x' \boldsymbol{\Psi}_{xx}^{-1} \boldsymbol{\lambda}_x}{1 + \boldsymbol{\lambda}_x' \boldsymbol{\Psi}_{xx}^{-1} \boldsymbol{\lambda}_x} = c \boldsymbol{\Psi}_{xx}^{-1} \boldsymbol{\lambda}_x, \tag{13}$$

where c is a constant. Because within-block error-covariances do not affect the between-block covariances, we continue to have $\boldsymbol{\Sigma}_{xy} = \boldsymbol{\Sigma}_{yx}' = \gamma \boldsymbol{\lambda}_x \boldsymbol{\lambda}_y'$. It follows from equation (2) that

$$\mathbf{w}_x \quad = c_x \boldsymbol{\Sigma}_{xx}^{-1} \boldsymbol{\Sigma}_{xy} \mathbf{w}_y = c_x \gamma \boldsymbol{\Sigma}_{xx}^{-1} (\boldsymbol{\lambda}_x \boldsymbol{\lambda}_y') \mathbf{w}_y = c_x \gamma (\boldsymbol{\lambda}_y' \mathbf{w}_y) \boldsymbol{\Sigma}_{xx}^{-1} \boldsymbol{\lambda}_x = c_{x*} \boldsymbol{\Psi}_{xx}^{-1} \boldsymbol{\lambda}_x,$$

where equation (13) is used for the last equal sign. Without loss of generality, suppose the first two errors in $\boldsymbol{\varepsilon}_x$ are correlated, then

$$\boldsymbol{\Psi}_{xx} = \begin{pmatrix} \boldsymbol{\Psi}_{x11} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Psi}_{x22} \end{pmatrix} \text{ and } \boldsymbol{\Psi}_{xx}^{-1} = \begin{pmatrix} \boldsymbol{\Psi}_{x11}^{-1} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Psi}_{x22}^{-1} \end{pmatrix},$$

where $\boldsymbol{\Psi}_{x11}$ is a 2×2 matrix corresponding to the first two correlated errors and $\boldsymbol{\Psi}_{x22}$ is a diagonal matrix corresponding to the last $(p_x - 2)$ elements of $\boldsymbol{\varepsilon}_x$. Thus, under mode B, weights of items having non-zero within-block error-covariances are affected by the covariances. Weights of items whose errors do not correlate with those of the other items are still proportional to those of the correctly specified model (i.e., λ_j/ψ_{jj}). However, because the values of the weights for items that have non-zero error-covariances are changed, the absolute values of the weights of items that have zero error-covariances will be proportionally changed due to scaling via standardization.

Under model B, when only $\mathbf{x}$ contains within-block error-covariances in Figure 1, the weight vector $\mathbf{w}_y$ satisfies (see equation 2)

$$\mathbf{w}_y = c_y \boldsymbol{\Sigma}_{yy}^{-1} \boldsymbol{\Sigma}_{yx} \mathbf{w}_x = \gamma \boldsymbol{\Sigma}_{yy}^{-1} \boldsymbol{\lambda}_y (\boldsymbol{\lambda}_x' \mathbf{w}_x) = c_{y*} \boldsymbol{\Sigma}_{yy}^{-1} \boldsymbol{\lambda}_y,$$

which is also proportional to $\boldsymbol{\Psi}_{yy}^{-1} \boldsymbol{\lambda}_y$. Thus, weights for the block $\mathbf{y}$ under mode B are not affected by the error-covariances within the block $\mathbf{x}$, and they remain the same as those when the whole model is correctly specified.

3.5 Weights under mode B by between-block error-covariances

Suppose the jth item in $\mathbf{x}$ and the kth item in $\mathbf{y}$ of Figure 1 have a non-zero error-covariance, it follows from equations (2) and (9) that the weight vector $\mathbf{w}_x$ under mode B is given by

$$\mathbf{w}_x = c \boldsymbol{\Sigma}_{xx}^{-1} \boldsymbol{\Sigma}_{xy} \mathbf{w}_y = c_x \boldsymbol{\Sigma}_{xx}^{-1} (\gamma \boldsymbol{\lambda}_x \boldsymbol{\lambda}_y' + \psi_{jk} \mathbf{e}_j \mathbf{e}_k') \mathbf{w}_y = c_{x1} \boldsymbol{\Sigma}_{xx}^{-1} \boldsymbol{\lambda}_x + c_{x2} \boldsymbol{\Sigma}_{xx}^{-1} \mathbf{e}_j.$$

Note that between-block error-covariances do not affect the within-block covariance matrix of the observed variables. Replacing the $\boldsymbol{\Sigma}_{xx}^{-1}$ in the above equation by the expression in (12) yields

$$\mathbf{w}_x = c_{x*} \boldsymbol{\Psi}_{xx}^{-1} \boldsymbol{\lambda}_x + c_* \boldsymbol{\Psi}_{xx}^{-1} \mathbf{e}_j = c_{x*} \boldsymbol{\Psi}_{xx}^{-1} \boldsymbol{\lambda}_x + c_{1*} \mathbf{e}_j,$$

where the 2nd equal sign used the assumption that the errors within the block $\mathbf{x}$ are still uncorrelated so that $\boldsymbol{\Psi}_{xx}$ is a diagonal matrix. Thus, under mode B, only the weights of items that have non-zero between-block error-covariances are affected. Weights for items that do not have between-block error-covariances are still proportional to those of the correctly specified model, i.e, $w_j = c\lambda_j/\psi_{jj}$.

3.6 Weights under mode B by cross-loadings

Suppose $\mathbf{x}$ has cross-loadings on η according to equation (10). It follows from equations (2) and (11) that

$$\mathbf{w}_y = c_y\boldsymbol{\Sigma}_{yy}^{-1}\boldsymbol{\Sigma}_{yx}\mathbf{w}_x = c_y\boldsymbol{\Sigma}_{yy}^{-1}(\boldsymbol{\lambda}_y\phi_{yx}\boldsymbol{\lambda}_x' + \boldsymbol{\lambda}_y\phi_{yy}\boldsymbol{\lambda}_h')\mathbf{w}_x = c_{y*}\boldsymbol{\Sigma}_{yy}^{-1}\boldsymbol{\lambda}_y.$$

Consequently, under mode B cross-loadings of $\mathbf{x}$ on η do not affect the weight vector $\mathbf{w}_y$. They are still proportional to $\boldsymbol{\Sigma}_{yy}^{-1}\boldsymbol{\lambda}_y = c\boldsymbol{\Psi}_{yy}^{-1}\boldsymbol{\lambda}_y$.

To find out how the weight vector $\mathbf{w}_x$ is affected when elements of $\mathbf{x}$ load on η in Figure 1, we will directly solve for $\mathbf{w}_x$ according to equations (2) and (11). Multiplying both sides of the first equation in (2) by $\boldsymbol{\Sigma}_{xx}$ in front yields

$$\boldsymbol{\Sigma}_{xx}\mathbf{w}_x = c_x\boldsymbol{\Sigma}_{xy}\mathbf{w}_y.$$

Using equation (11), we can rewrite the above equation as

$$(\boldsymbol{\lambda}_x\phi_{\xi\xi}\boldsymbol{\lambda}_x' + \boldsymbol{\lambda}_h\phi_{\eta\xi}\boldsymbol{\lambda}_x' + \boldsymbol{\lambda}_x\phi_{\xi\eta}\boldsymbol{\lambda}_h' + \boldsymbol{\lambda}_h\phi_{\eta\eta}\boldsymbol{\lambda}_h' + \boldsymbol{\Psi}_{xx})\mathbf{w}_x = c_x(\boldsymbol{\lambda}_x\phi_{\xi\eta}\boldsymbol{\lambda}_y' + \boldsymbol{\lambda}_h\phi_{\eta\eta}\boldsymbol{\lambda}_y')\mathbf{w}_y.$$

Since $(\mathbf{a}'\mathbf{b})$ is a scalar regardless of the values of the elements of $\mathbf{a}$ and $\mathbf{b}$, except for the term $\boldsymbol{\Psi}_{xx}\mathbf{w}_x$, all the other terms in the above equation (after the operation of the multiplication on both sides) are either a scalar times $\boldsymbol{\lambda}_x$ or a scalar times $\boldsymbol{\lambda}_h$. Move the first four terms on the left of the equal sign to the right side in the above equation, we have

$$\boldsymbol{\Psi}_{xx}\mathbf{w}_x = c_{x1}\boldsymbol{\lambda}_x + c_{x2}\boldsymbol{\lambda}_h.$$

Consequently,

$$\mathbf{w}_x = c_{x*}\boldsymbol{\Psi}_{xx}^{-1}\boldsymbol{\lambda}_x + c_*\boldsymbol{\Psi}_{xx}^{-1}\boldsymbol{\lambda}_h.$$

Thus, under mode B the weights of items in $\mathbf{x}$ that do not cross-load on η are still proportional to λ_j/ψ_{jj}. Only the weights of items that have non-zero cross-loadings on η are affected.

In parallel, cross-loadings by items of $\mathbf{y}$ on ξ in Figure 1 do not affect $\mathbf{w}_x$ under mode B, which is still proportional to $\boldsymbol{\Psi}_{xx}^{-1}\boldsymbol{\lambda}_x$. The elements of $\mathbf{w}_y$ corresponding to items that do not load on ξ are also proportional to λ_j/ψ_{jj}. But weights for items in $\mathbf{y}$ that have non-zero loadings on ξ will be affected, they might become greater or smaller according to the sign and size of the cross-loadings. We will illustrate such a pattern numerically in the next section.

3.7 Weights under PLS-SEM mode $\mathrm{B_A}$

We have noted that the weight vector under PLS-SEM mode $\mathrm{B_A}$ is given by $\mathbf{w}_{b_a} = c_*\mathbf{\Psi}_*^{-1}\mathbf{w}_a$, where $\mathbf{\Psi}_*^{-1}$ is a diagonal matrix obtained by fitting a one-factor model to each block of indicators. Thus, the sensitivity of $\mathbf{w}_{b_a}$ to model misspecification closely depends on the sensitivity of $\mathbf{w}_a$. The following properties of $\mathbf{w}_{b_a}$ are directly derived from the analytical results on $\mathbf{w}_a$ presented in subsections 3.1-3.3.

Within-block error-covariances

When the $\mathbf{x}$ in Figure 1 contains within-block error-covariances, then the one-factor model structure in (8) is misspecified. The value of θ_* will deviate from the population value corresponding to a correctly specified model, which will cause all the diagonal values of $\mathbf{\Psi}_*$ deviate from the population value corresponding to a correctly specified model. Consequently, all the elements of $\mathbf{w}_{b_a}$ are affected even when only a single pair of errors within the block $\mathbf{x}$ are correlated. However, the weight vector $\mathbf{w}_{b_a}$ for the block $\mathbf{y}$ will be identical to those corresponding to the correctly specified model, that is, $\mathbf{w}_{b_a} = \mathbf{w}_b = c\mathbf{\Psi}_{yy}^{-1}\boldsymbol{\lambda}_y$.

Between-block error-covariances

When between-block error-covariances exist, then the one-factor model in equation (8) is misspecified for each of the involved blocks. This is because the vector $\mathbf{w}_a$ is fixed at the values obtained under PLS-SEM mode A, although the covariance matrices $\mathbf{\Sigma}_{xx}$ and $\mathbf{\Sigma}_{yy}$ are not affected by the between-block error-covariances. The diagonal elements of the $\mathbf{\Psi}_*$ in equation (8) need to adjust their values due to the misspecification. Consequently, the weight vectors $\mathbf{w}_{b_a}$ for both $\mathbf{x}$ and $\mathbf{y}$ in Figure 1 will deviate from those corresponding to the correctly specified model. In particular, even the weights of items in $\mathbf{x}$ and $\mathbf{y}$ that have null between-block error-covariances will be affected.

Cross-loadings

When items of $\mathbf{x}$ load on η in Figure 1, the weight vector $\mathbf{w}_{b_a}$ for the block $\mathbf{x}$ will deviate from that corresponding to the correctly specified model, due to the model in (8) being misspecified. But for the block $\mathbf{y}$, the weight vector $\mathbf{w}_{b_a}$ remains proportional to $\mathbf{\Psi}_{yy}^{-1}\boldsymbol{\lambda}_y$.

3.8 Weights for models with more than two blocks of indicators

With more than two blocks of indicators, the weights under PLS-SEM are computed via the environmental variables. Consider the model in Figure 2 that contains three latent variables ξ_1, ξ_2 and η. The environmental variable of η is given by $\bar{\eta} = c(r_1\xi_1 + r_2\xi_2)$, where[5] $r_j = \text{sgn}(\rho_{\eta\xi_j})$ or $\rho_{\eta\xi_j}$ and c is a constant so that $\text{Var}(\bar{\eta}) = 1$. The environmental variables of both ξ_1 and ξ_2 are η because they each has a single direct connection. That is, $\bar{\xi}_1 = \eta$ and $\bar{\xi}_2 = \eta$. With more than two blocks of indicators, the weight vectors under modes A and B in equations (1) and (2) are replaced by

$$\mathbf{w}_a = c_a\boldsymbol{\sigma}_{zt} \text{ and } \mathbf{w}_b = c_b\boldsymbol{\Sigma}_{zz}^{-1}\boldsymbol{\sigma}_{zt}, \tag{14}$$

respectively, where $\mathbf{z}$ is the block of indicators for latent variable ξ_1, ξ_2 or η in Figure 2, and $\boldsymbol{\sigma}_{zt}$ is the vector of covariances between $\mathbf{z}$ ($= \mathbf{x}_1$, $\mathbf{x}_2$, or $\mathbf{y}$) and the environmental variable t ($= \eta$ or $\bar{\eta}$). The results stated below can be obtained analytically via equation (14), parallel to the analysis presented in the previous subsections.

Insert Figure 2 about here

Within-block error-covariances

Within-block error-covariances do not affect the weights of any blocks under PLS-SEM mode A. But they will affect the weights of the involved items under PLS-SEM mode B. All the elements of the weight vector in the same block under PLS-SEM mode B_A are affected by within-block error-covariances. For each of the three modes, within-block error-covariances do not affect the weights for the other blocks that have null within-block error-covariances.

Between-block error-covariances

Between-block error-covariances only affect the weights of the involved items under PLS-SEM modes A and B. They affect the weights of all the items of the involved blocks under PLS-SEM mode B_A. For the blocks whose items do not have between-block error-covariances, the weight vectors are not affected under any of the three modes.

If two constructs are just correlated, as the ξ_1 and ξ_2 in Figure 2, error-covariances between the two blocks do not affect the weights of any of the blocks under PLS-SEM modes A, B and B_A, because neither $\boldsymbol{\sigma}_{zt}$ nor $\boldsymbol{\Sigma}_{zz}$ in equation (14) is affected by such error-covariances. For the same reason, weights under the three modes are not affected by error-covariances between the blocks that are not directly connected.

[5] The choice of the coefficient r_j does not change the conclusion stated in this subsection.

Cross-loadings

A cross-loading of an item z_j only affects the weight of this individual item under PLS-SEM modes A and B. A cross-loading of z_j will affect the weights of all the items that share the same block with z_j for PLS-SEM mode $\mathrm{B_A}$. But weights of the items that do not share the block with z_j are not affected under any of the three modes.

4. Numerical Illustration

This section provides numerical illustrations of the analytical results obtained in the previous section. The numerical results also allow us to see the size of the weight changes relative to their counterparts under a correctly specified model. Two models will be considered, one has two latent variables and the other has four latent variables. To save space, we only present the results for the 2-latent-variable model. For the model with four latent variables, we will point out the notable results while leave the numerical tables to the supplementary material online.

Consider the model as represented by the solid arrows in Figure 1, where the vectors of factor loadings for ξ and η are given by $\boldsymbol{\lambda}_{x0} = (.80, 1.0, 1.2)'$ and $\boldsymbol{\lambda}_{y0} = (1.0, 1.2, .80, 1.5)'$, respectively; and the matrices of error variances in the measurement model are given by $\boldsymbol{\Psi}_{xx0} = \mathrm{diag}(.55, .60, .40)$ and $\boldsymbol{\Psi}_{yy0} = \mathrm{diag}(.58, .65, .60, .40)$. Let $\phi_0 = \mathrm{Var}(\xi) = 1.0$, the regression coefficient for the structural model $\eta = \gamma\xi + \zeta$ is set at $\gamma_0 = .70$, and the prediction-error variance is set at $\sigma^2_{\zeta 0} = .40$. These values are chosen because they are also the population values of the model parameters when the CB-SEM model is identified by fixing the variance of ξ at 1.0 and the first factor loading of $\mathbf{y}$ at 1.0. Let $\boldsymbol{\theta}_0$ denote the vector of parameters corresponding to these population values. Alternative choices for $\boldsymbol{\theta}_0$ can also be used for the purpose of illustration since the changes in weights of composites under PLS-SEM are relative to the size of model misspecification, which will be specified according to the value of the fit index RMSEA under CB-SEM (Steiger & Lind, 1980). Let $\mathbf{w}_{a0}$, $\mathbf{w}_{b_a0}$ and $\mathbf{w}_{b0}$ be the vectors of weights for PLS-SEM modes A, $\mathrm{B_A}$ and B corresponding to the population values in $\boldsymbol{\theta}_0$, respectively. Because the model is correctly specified, the weight vectors $\mathbf{w}_{a0}$ for $\mathbf{x}$ and $\mathbf{y}$ are proportional to $\boldsymbol{\lambda}_{x0}$ and $\boldsymbol{\lambda}_{y0}$, respectively; and the weight vectors $\mathbf{w}_{b0} = \mathbf{w}_{b_a0}$ for $\mathbf{x}$ and $\mathbf{y}$ are proportional to $\boldsymbol{\Psi}^{-1}_{xx0}\boldsymbol{\lambda}_{x0}$ and $\boldsymbol{\Psi}^{-1}_{yy0}\boldsymbol{\lambda}_{y0}$, respectively.

To examine the sensitivity of the weights to model misspecification, corresponding to the analytical results presented in the previous section, we will consider conditions when error-covariances or cross-loadings exist in the population but are ignored in the model. Three conditions of error-covariance are picked. They are: 1) e_{x_2} and e_{x_3} are correlated, 2) e_{y_2} and e_{y_3} are correlated, 3) e_{x_2} and e_{y_2} are correlated. They respectively represent correlated errors within the block $\mathbf{x}$, within the block $\mathbf{y}$, and between the blocks $\mathbf{x}$ and $\mathbf{y}$. Two conditions of cross-loading are picked, and they are: 1) x_2 has a non-zero loading on η, and 2) y_2 has a non-zero loading on ξ. Let the values of the error-covariance and cross-loading be denoted as ψ_h and λ_h, respectively. For each of the 3 conditions of error-covariance, we let $\psi_h = 0.2$ and -0.2, respectively; and for each of the 2 conditions of cross-loadings, we let $\lambda_h = 0.3$ and -0.3, respectively. So we have a total of 10 conditions of population. The values of the RMSEA corresponding to these 10 conditions range from .034 to .073, with a mean of .054. They represent minor to moderate model misspecification according to established cutoffs (e.g., Hu & Bentler, 1999; MacCallum, Browne & Sugawara, 1996) and are reported in Tables 1 and 2.

Let $\mathbf{\Sigma}_*$ be the population covariance matrix of $(\mathbf{x}', \mathbf{y}')'$ corresponding to a condition when $\psi_h \neq 0$ or when $\lambda_h \neq 0$. We then fit the structural model as represented by the solid arrows in Figure 1 to $\mathbf{\Sigma}_*$ (under CB-SEM) using normal-distribution-based maximum likelihood (NML). Factor loadings $\boldsymbol{\lambda}_*$ and error variances $\mathbf{\Psi}_*$ are obtained, and so are the ratios λ_*/λ_0 for both the blocks $\mathbf{x}$ and $\mathbf{y}$. Similarly, PLS-SEM modes A and B are applied to $\mathbf{\Sigma}_*$ via equations (1) and (2), respectively. We denote the obtained weights as $\mathbf{w}_{a*}$ and $\mathbf{w}_{b*}$. The weights $\mathbf{w}_{a*}$ is further transformed to $\mathbf{w}_{b_a*}$ by fitting a one-factor model to the submatrix of $\mathbf{\Sigma}_*$ for each block of indicators, according to equation (8). We will compare the values of $\mathbf{w}_*$ against those of $\mathbf{w}_0$ using the ratio $r_j = w_{j*}/w_{j0}$. Because the values of the weights in $\mathbf{w}_*$ are proportionally changed due to standardization, we further scaled the ratio r_j via

$$r_j^{(s)} = \frac{w_{j*}/w_{1*}}{w_{j0}/w_{10}} \tag{15}$$

so that the value of $r_j^{(s)}$ for the first item in each block is 1.0 (i.e., $r_1^{(s)} = 1.0$). Such a scaling makes it easier to compare the relative sensitivity of different methods.

Insert Table 1 about here

The results of the weight ratio $r_j^{(s)}$ under the 6 conditions of correlated errors are given in Table 1, where the ratio λ_*/λ_0 of factor loadings under CB-SEM is also included for comparison purpose. The values of $r_j^{(s)}$ for the condition $\psi_{x_2x_3} = .2$ are given on the left side of the top panel of Table 1, where all the three factor loadings within the block $\mathbf{x}$ under CB-SEM are affected. While weights under mode A are not affected by the non-zero error-covariance $\psi_{x_2x_3}$, all the elements of $\mathbf{w}_x$ under mode $\mathrm{B_A}$ are affected. However, because of the rescaling via equation (15), the first value of $r_j^{(s)}$ under PLS-SEM mode $\mathrm{B_A}$ for the block $\mathbf{x}$ is still 1.0. In contrast, only the 2nd and 3rd elements of $\mathbf{w}_x$ under PLS-SEM mode B are affected, as indicated by the corresponding values of $r_j^{(s)}$. An interesting phenomenon is that a positive $\psi_{x_2x_3}$ makes the weights of x_2 and x_3 under mode B smaller, while under mode $\mathrm{B_A}$ these weights become greater. When $\psi_{x_2x_3} = -.2$, the results for the relative changes in weights are given on the right side of the top panel of Table 1. While the weights under mode A remain intact, the weight changes for the other two modes are in the opposite directions of those when $\psi_{x_2x_3} = .2$.

The middle panel of Table 1 contains the results of $r_j^{(s)}$ when the error terms of y_2 and y_3 are correlated. Again, all the elements of $\mathbf{w}_y$ under PLS-SEM mode $\mathrm{B_A}$ are affected, but only those corresponding to y_2 and y_3 are affected under mode B. The pattern of the weight changes for the condition $\psi_{y_2y_3} = -.2$ is similar to that for the condition $\psi_{y_2y_3} = .2$ but in the opposite directions.

When the error terms of x_2 and y_2 are correlated, the results of $r_j^{(s)}$ are given in the bottom panel of Table 1. As characterized by the analytical results in the previous section, for PLS-SEM modes A and B, only the weights for items x_2 and y_2 are affected. All the weights under PLS-SEM mode $\mathrm{B_A}$ are affected. The changes in weights that are affected by a positive $\psi_{x_2y_2}$ are in the opposite directions from those by a negative $\psi_{x_2y_2}$. Note that the ratio λ_*/λ_0 for y_3 in the bottom panel of Table 1 has also changed, although not reflected in the first 3 decimal places.

While the same sets of elements of $\mathbf{w}_a$ and $\mathbf{w}_b$ are affected by a non-zero between-block error-covariance, the changes in elements of $\mathbf{w}_a$ are much smaller than those of $\mathbf{w}_b$, indicating more stability of mode A. Also, while all the elements of $\mathbf{w}_{b_a}$ are affected by within-block and between-block error-covariances, the changes of the individual elements of $\mathbf{w}_{b_a}$ are much smaller than those of the affected elements of $\mathbf{w}_b$.

Insert Table 2 about here

The results on relative weight-change ($r_j^{(s)}$) due to cross-loadings ($\lambda_{x_2\eta} = \pm.3$, $\lambda_{y_2\xi} = \pm.3$)

are in Table 2. For both PLS-SEM modes A and B, only the weights of the items that have a cross-loading are affected. For PLS-SEM mode B_A, the weights of all the items in the block are affected. For the affected weights, the changes under mode A are much smaller than those under mode B. For the item that has a cross-loading, the value of $r_j^{(s)}$ under mode B_A is always between those under modes A and B, indicating that mode B_A is a compromise of mode A and mode B.

With a cross-loading $\lambda_h = -.30$, the weight of the corresponding item under mode B becomes negative. A negative weight is not consistent with the expectation for the role of the item ($\lambda_{x_2\xi} = 1.0$, $\lambda_{x_2\eta} = -.3$; $\lambda_{y_2\eta} = 1.2$, $\lambda_{y_2\xi} = -.3$), and is also very undesirable in practice. Such a phenomenon does not occur to PLS-SEM modes A or B_A, nor to the factor loadings λ_* under CB-SEM. A rationale for PLS-SEM mode B to generate negative weights is that the cross-loading is a direct effect between the involved item and the latent variable of the other block, while the main loading plays the role of an indirect effect between the item and the other latent variable. Also, the cross-loading of a single item has changed the covariances among all the observed variables in the block. Mode A does not use these covariances but mode B does (see equations 1, 2 & 11). The negative weights or the relative weight-changes with $r_j^{(s)} = 3.077$ or 3.744 in Table 2 reflect the hyper-sensitivity of PLS-SEM mode B to specification errors.

Note that the pattern of weight-changes due to either error-covariances or cross-loadings also depends on the value and sign of the regression coefficient γ of η on ξ in Figure 1. For example, if we change the population value of γ from .70 to -.70, then a cross-loading of x_2 on η at .30 in Figure 1 will result in a negative weight of x_2 under mode B.

The results in Tables 1 and 2 indicate that, with either between-block error-covariances or cross-loadings, the changes in weights of the involved items under both PLS-SEM modes A and B are always in the same direction. But the relative change of w_b is a lot greater than that of w_a. A mechanism for this phenomenon is that w_b is obtained via multiple regression. Due to the indicators within the same block being correlated, each regression coefficient or weight under mode B has more freedom to account for the extra needs of the corresponding item. In contrast, the regression coefficient or weight under mode A has to account for the combined interest of the main-loading and cross-loading by itself, and the main-loading will have a dominant role unless its size is comparable with that of the cross-loading.

Insert Figure 3 about here

Figure 3 presents a path diagram for a model with 4 latent constructs and 13 indicators. The solid arrows in the figure represent a unidimensional model that PLS-SEM methodology aims to estimate, while the dashed arrows represent additional parameters/relationships that may exist in the population. Note that ξ_1 does not directly predict η_2 in the model. The two-way arrow between ξ_1 and ξ_2 is not considered as a direct connection under PLS-SEM.

In Figure 3, there are a total of 22 dashed arrows: four two-way arrows representing within-block error covariances, six two-way arrows representing between-block error covariances, and twelve one-way arrows representing cross-loadings. The population values of the parameters corresponding to the solid arrows as well as the dashed ones are described in the online supplementary material (`https://www3.nd.edu/~kyuan/PLS-SEM/sensitivity/Supplementary.pdf`). The numerical results of relative changes of w_a, w_{b_a} and w_b for the 13 indicators are in Tables A1 to A3 of the supplementary material. They all agree with the analytical results obtained in the previous section. Because the model structure in Figure 3 is different from that in Figure 1, results in Tables A1 to A3 also have some notable features. They include: (1) PLS-SEM mode B can have negative weights when there exist within-block error-covariances in the population; (2) weights under PLS-SEM modes A, $\mathrm{B_A}$ and B are not affected by error-covariances between the blocks of ξ_1 and ξ_2; (3) weights under the three modes are also not affected by error-covariances between the blocks of ξ_1 and η_2; (4) mode A yields a negative weight when x_2 has a negative loading on η_1, and consequently mode $\mathrm{B_A}$ also has a negative weight under this condition; (5) PLS-SEM mode B may yield negative weights under the condition of positive cross-loadings; and (6) there are multiple conditions of Heywood case (negative error variance) when the model in equation (8) is estimated by the LS method. Negative elements of $\mathbf{\Psi}_*$ are always associated with large elements of $\mathbf{w}_a$, and the negative error variance is changed to .05 in computing the $\mathbf{w}_{b_a}$ via equation (7). Readers interested in the details of the numerical results are referred to the supplementary material (`https://www3.nd.edu/~kyuan/PLS-SEM/sensitivity/Supplementary.pdf`).

5. Analysis with a Real Dataset

Yuan and Deng (2021) introduced a dataset for a study of health and stress. The dataset consists of 264 cases and 24 variables, which are indicators of four constructs. The three exogenous constructs are respectively: *emotional exhaustion* (ξ_1, 5 indicators), *cynicism* (ξ_2, 4 indicators), and *professional efficacy* (ξ_3, 6 indicators). The single endogenous construct *depression* (η) has 9 indicators. The path diagram for the original model is represented by the solid arrows in Figure 4, which has three path coefficients (γ_1, γ_2, γ_3) in the structural model. We will refer to the indicators for ξ_1, ξ_2, ξ_3 and η as block 1, block 2, block 3, and block 4 (the block of outcome indicators), respectively. Note that the three exogenous constructs are correlated but they are not considered as directly connected under PLS-SEM, while each of them is directly connected with depression. The raw dataset of the 24 variables has a standardized multivariate kurtosis $M_s = 23.292$ (Mardia, 1970), which is highly significant when referred to the distribution $N(0, 1)$. Yuan and Deng (2021) applied a robust transformation to the dataset, and the standardized multivariate kurtosis for the transformed sample is $M_s = -.003$. Our analysis and results presented below are based on fitting the model in Figure 4 to this robustly transformed sample.

Insert Figure 4 about here

5.1 Three models

When the model represented by the solid arrows in Figure 4 is estimated by normal-distribution-based maximum likelihood (NML) under CB-SEM, Table 3 contains the parameter estimates, their standard errors (SEs) and the corresponding z-statistics. All the loading estimates are positive and statistically significant, and so are all the error-variance estimates. While the estimates of γ_1 and γ_3 are not statistically significant at the level of .05, it does not imply that we can regard them as zero. Note that the estimates of γ_3, ϕ_{31} and ϕ_{32} are negative. The likelihood ratio statistic for the overall model structure is $T_{ml} = 469.579$, highly significant when referred to χ^2_{246}. With fit indices CFI=.948 (Bentler, 1990) and RMSEA=.059 (Steiger & Lind, 1980), the model can be regarded as being acceptable in practice (Hu & Bentler, 1999). However, the model might be far from being correctly specified. We will use the Lagrange multiplier (LM) test (Bentler, 2006; Silvey, 1959) to identify additional parameters to improve the goodness of model-fit. One modified model is to include cross-loadings identified by the LM test, and another model is to include error-covariances identified by the LM test. We will show how the weights of the composites under PLS-SEM are affected by excluding these parameters.

Insert Tables 3 and 4 about here

For the model represented by the solid arrows in Figure 4, the estimated weights under PLS-SEM modes A, B_A, and B are given in Table 4. For comparison purposes, these weights are scaled so that the first element of each block is 1.0, and a superscript s is used to refer to this scaling. Table 4 also includes a scaled version of the factor loadings as reported in Table 3, under the notation $\lambda^{(s)}$. Because we do not have the population values of the parameters for this real-data example, the scaled values are obtained by dividing the vectors of weights and factor loadings for each block by its first element. We put the first element of $w_b^{(s)}$ for the 3rd block (ξ_3) at -1.0 so that the estimated path coefficients and correlations among the constructs by PLS-SEM mode B can have the same signs with those by the other methods. The estimates of the path coefficients and correlations of the constructs are also included for reference. The estimated loadings under $\lambda_{cl}^{(s)}$ and $\lambda_{ec}^{(s)}$ in Table 4 are for the modified models that include cross-loadings (cl) and error-covariances (ec), respectively, which will be further described below.

If the model represented by the solid arrows in Figure 4 were correctly specified, we would have $w_a^{(s)} = \lambda^{(s)}$ and $w_b^{(s)} = w_{b_a}^{(s)} = c\psi^{-1}\lambda^{(s)}$ in the population. While the three columns of factor loadings in Table 4 are close, the estimated weights are quite different from what would be expected for a correctly specified model. In particular, multiple elements under $w_b^{(s)}$ are negative and occurred in three of the four blocks. Multiple elements under $w_{b_a}^{(s)}$ are several times of its counterparts under $w_b^{(s)}$; and a few elements under $w_b^{(s)}$ are also much greater than their counterparts under $w_{b_a}^{(s)}$. In contrast, the elements under $w_a^{(s)}$ are much closer to their counterparts of factor loadings, although differences exist. A Heywood case also occurred in estimating the one-factor model via equation (8) for the 1st block, and the negative error variance (-.071) was changed to .05 in computing the weights under mode B_A.

Although we do not know the specific locations of misspecification with real data, statistical techniques such as model modification index (Sörbom, 1989) and LM test (Silvey, 1959) allow us to empirically identify parameters that can effectively improve the goodness of the model-fit. We will use the LM test in the software EQS (Bentler, 2006) to identify these parameters. In particular, the default option of the LM test in EQS searches for all cross-loadings and ranks those that potentially improve the goodness of fit; and it also searches for error-covariances when specified. For the health-stress dataset and the model represented by the solid arrows in Figure 4, the default LM test identifies a total of eight cross-loadings whose inclusions are expected to significantly improve the model-fit at the level of .05. They are reported in Table 5 in the order of their expected contributions to the reduction of T_{ml}. The model is re-estimated by NML after these 8 cross-loadings are included. The resulting statistic $T_{ml} = 399.288$ (CFI=0.962, RMSEA=.051) drops by 70 from that of the unidimensional model although the corresponding p-value is still essentially 0 when the T_{ml} is referred to χ^2_{238}. The estimates of the cross-loadings and their corresponding z-statistics are reported in Table 5. Except for the parameter $\lambda_{x_{15}\xi_1}$, the order of the size of the z-statistics is consistent with the order of reduction in T_{ml} as predicted by the LM test. The estimates of the rescaled loadings and other parameters corresponding to the original solid arrows in Figure 4 for this second model are reported in Table 4, under $\lambda^{(s)}_{cl}$.

Insert Table 5 about here

With the model represented by the solid arrows in Figure 4, the search for error-covariances by the multivariate version of the LM test identifies 25 parameters that might significantly improve the model-fit at the level of .05. The top 10 are reported on the right side of Table 5 according to the order of their predicted contributions to the drop of T_{ml}. After including the 10 error-covariances, their estimates and the corresponding z-statistics are also reported in Table 5. Except for the error-covariance $\psi_{x_{12}x_{10}}$, the squared values of the z-statistics for the other 9 parameters are all greater than 9.0, although the order of the size of the z-statistics differs from that predicted by the LM test. Inclusion of the 10 error-covariances results in $T_{ml} = 313.640$ ($df = 236$), corresponding to a p-value=0.00053 , CFI=.982, RMSEA=.035. The estimates of the rescaled loadings and other parameters corresponding to the original solid arrows in Figure 4 for this third model are reported in Table 4, under $\lambda^{(s)}_{ec}$.

All the parameters reported in Table 5 are marked in Figure 4 using dashed arrows, one-way arrows for factor loadings and two-way arrows for error-covariances. The signs of the estimated values of the cross-loadings are also included in the figure. With the above information, we next conduct a post-hoc analysis for the results of weights in Table 4, by treating the estimated values in Table 5 as representatives for the population values.

5.2 Effects of cross-loadings

For the effect of cross-loadings, we will first discuss the weights under mode A before turning to those under modes B and B_A. Note that we have three sets of factor loading estimates in Table 4. The differences between $w_a^{(s)}$ and $\lambda^{(s)}$ are due to different methods applying to the same misspecified model. The differences between $w_a^{(s)}$ and $\lambda_{cl}^{(s)}$ or those between $w_a^{(s)}$ and $\lambda_{ec}^{(s)}$ are due to different methods as well as different models. Because the effect of model misspecification on factor loadings has been discussed in Yuan et al. (2003, 2008), we will not further discuss the differences among $\lambda^{(s)}$, $\lambda_{cl}^{(s)}$, and $\lambda_{ce}^{(s)}$ here. Actually, the three versions of factor loadings are much closer to each other than to the weights of PLS-SEM, simply because the factor loadings are subject to many common constraints in their estimation that the weights under PLS-SEM are not subject to.

Mode A

According to the path diagram in Figure 4, three indicators in the first block (ξ_1, emotional exhaustion) have cross-loadings. They are respectively $x_2 \leftarrow\!\!-\!- \xi_3$, $x_3 \leftarrow\!\!-\!- \xi_2$ and $x_5 \leftarrow\!\!-\!- \eta$, and the loadings are all positive. For the unidimensional model represented by the solid arrows in Figure 4, there are two routes for ξ_1 and ξ_3 to be correlated ($\xi_1 \leftrightarrow \xi_3$, $\xi_1 \leftrightarrow \xi_2 \leftrightarrow \xi_3$) and they both deliver negative association. The positive cross-loading $x_2 \leftarrow\!\!-\!- \xi_3$ implies that x_2 has an extra positive association with ξ_3. Under the restriction of the unidimensional model, a smaller w_{x_2} is necessary to satisfy this need of x_2 and ξ_3. In parallel, because the overall association between ξ_1 and ξ_2 ($\xi_1 \leftrightarrow \xi_2$, $\xi_1 \leftrightarrow \xi_3 \leftrightarrow \xi_2$) is positive, a greater w_{x_3} is needed to account for the extra association between x_3 and ξ_2, due to the positive cross-loading. Similarly, because the association between ξ_1 and η is positive, a greater w_{x_5} is needed to account for the extra positive association between x_5 and η. The comparison between $\lambda_{cl}^{(s)}$ and $w_a^{(s)}$ in Table 4 agrees with our analysis.

One indicator (x_9) in the 2nd block (ξ_2, cynicism) has a negative cross-loading on ξ_1. Because the overall association between ξ_1 and ξ_2 is negative, a greater w_{x_9} will account for the extra negative association between x_9 and ξ_1. The results of $\lambda_{cl}^{(s)}$ and $w_a^{(s)}$ in Table 4 also agree with the conclusion.

Three indicators in the 3rd block (ξ_3, professional efficacy) in Figure 4 have cross-loadings. The loading over $x_{10} \leftarrow\!\!-\!- \xi_2$ is positive, and those over $x_{14} \leftarrow\!\!-\!- \xi_2$ and $x_{15} \leftarrow\!\!-\!- \xi_1$ are negative. Because ξ_2 and ξ_3 are negatively associated, $w_{x_{10}}$ needs to be smaller to react to the need of the extra positive association between x_{10} and ξ_2. In contrast, $w_{x_{14}}$ needs to be greater to account for the extra association due to the negative cross-loading over $x_{14} \leftarrow\!\!-\!- \xi_2$. With essentially the same mechanism, because the overall association between ξ_3 and ξ_1 is negative, a greater $w_{x_{15}}$ is needed to account for the extra negative association over the cross-loading $x_{15} \dashleftarrow \xi_1$. The results of $\lambda_{cl}^{(s)}$ and $w_a^{(s)}$ in Table 4 also agree with our analysis. Because the numbers under $w_a^{(s)}$ for the 3rd block are divided by $w_{x_{10}}$ (a smaller value) for rescaling purpose, the resulting weights for x_{11} to x_{15} all become greater than their counterpart under $\lambda_{cl}^{(s)}$.

One indicator in the 4th block (η, depression) has a positive cross-loading ($y_6 \dashleftarrow \xi_3$). There are three routes for ξ_3 to be associated with y_6 under the unidimensional model ($\xi_3 \rightarrow \eta \rightarrow y_6$, $\xi_3 \leftrightarrow \xi_2 \rightarrow \eta \rightarrow y_6$, $\xi_3 \leftrightarrow \xi_1 \rightarrow \eta \rightarrow y_6$), and they are all negative according to the marked paths in Figure 4. Consequently, a smaller w_{y_6} is needed to account for the extra positive association. The analytical conclusion agrees with the results in Table 4.

Mode B

As showed in section 3 and further clarified in section 4, with cross-loadings, the mechanism for weight changes under mode B is essentially the same as that for weight changes under mode A. However, due to indicators within each block being correlated and weights under mode B are computed by multiple regression, each element of $\mathbf{w}_b$ has more freedom in accounting for the additional needs of its corresponding indicator, by leaving the common interest of the block to other indicators.

According to Figure 4, three indicators in block 1 (ξ_1, emotional exhaustion) have cross-loadings, they are respectively x_2 on ξ_3, x_3 on ξ_2, and x_5 on η. For the unidimensional model represented by the solid arrows in Figure 4, the association between ξ_1 and ξ_3 is negative. A smaller w_{x_2} is needed to account for the extra positive association between x_2 and ξ_3. The weight can even be negative if this relationship is strong enough, as is the case under PLS-SEM mode B in Table 4. In contrast, the cross-loading of x_3 on ξ_2 is positive, implying an additional positive association between x_3 and ξ_2. Under the unidimensional model represented by the solid arrows in Figure 4, the association between x_3 and ξ_2 is positive. The extra positive association between x_3 and ξ_2 can be picked up by increasing the value of the weight, not a sign change. In parallel, the positive cross-loading of x_5 on η implies that they have a stronger association than implied by the unidimensional model. Since the association over the path $x_5 \leftarrow \xi_1 \rightarrow \eta$ is positive, the extra association between x_5 and η can be realized by a greater weight of w_{x_5} instead of a sign change. However, we cannot explain why x_4 also has a negative weight, which might be caused by model misspecification not detected by the LM test, as to be further discussed in the concluding section.

The 2nd block (ξ_2, cynicism) has one cross-loading ($x_9 \dashleftarrow \xi_1$) that is negative. While the negative cross-loading of x_9 on ξ_1 may need a smaller or negative w_{x_9} to account for the needed association, the variable x_9 is also subject to the constraint of a negative within-block error-covariance $\psi_{x_6x_9}$, which works against such a move. Consequently, all the weights within the 2nd block are positive.

Three indicators in the 3rd block (ξ_3, professional efficacy) have cross-loadings ($x_{10} \dashleftarrow \xi_2$, $x_{14} \dashleftarrow \xi_2$, $x_{15} \dashleftarrow \xi_1$). The positive loading of x_{10} on ξ_2 implies that they have an extra positive association beyond what the unidimensional model can deliver. Because both the routes $\xi_3 \leftrightarrow \xi_2$ and $\xi_3 \leftrightarrow \xi_1 \leftrightarrow \xi_2$ are negative, the weight for x_9 needs to be negative in order to account for the extra positive association. In contrast, the association between ξ_3 and ξ_2 (via the paths $\xi_3 \leftrightarrow \xi_2$ and $\xi_3 \leftrightarrow \xi_1 \leftrightarrow \xi_2$) is negative, the extra association due to the negative loading of x_{14} on ξ_2 can be accounted by a greater positive value of $w_{x_{14}}$. With the same mechanism, the extra association due to the negative loading of x_{15} on ξ_1 can be accounted for by a greater positive value of $w_{x_{15}}$. Such a pattern is clearly reflected by the results in Table 4.

For the block of the outcome indicators (η, depression), the cross-loading of y_6 on ξ_3 suggests that the two have an extra positive association. However, under the unidimensional model, all the paths between ξ_3 and η ($\xi_3 \rightarrow \eta$, $\xi_3 \leftrightarrow \xi_2 \rightarrow \eta$, $\xi_3 \leftrightarrow \xi_2 \leftrightarrow \xi_1 \rightarrow \eta$) deliver negative association. The negative weight of y_6 under mode B is to account for the needed positive association between ξ_3 and y_6. Again, the negative weights of y_3 under PLS-SEM mode B can be because of correlated errors within the block or misspecification not identified by the LM test.

The results in Table 4 under PLS-SEM mode B are mostly consistent with our analysis, although variations exist due to misspecification not identified by the LM test or by interactions from the misspecification within each block.

Mode B$_\text{A}$

The values of $\mathbf{w}_{b_a}$ are closely related to those of $\mathbf{w}_a$ and the variances of the observed variables. This is because the solution to equation (8) satisfies $\psi_{jj*} = \sigma_{jj} - \theta_* w_{aj}^2$. If the value-change of w_{aj} is to such a degree that ψ_{jj*} is close to zero, then the corresponding value of w_{b_a} will be rather large, as for x_5 and x_{15} in Table 4. Heywood case can even occur if the value of w_{aj} is large enough, as for x_3 in Table 4 as well as cases reported in the supplementary material. Because the negative error variance is replaced by .05 in the computation, the weights under mode B_A always have the same sign as those of mode A.

5.3 Effects of error-covariances

Mode A

According to the results in section 3, within-block correlated errors do not affect the weights under PLS-SEM mode A. For the 10 error-covariances in Table 5, nine are within-block covariances. Only the last one ($\psi_{x_7y_7}$) is between-block error-covariance, and its estimated value (.086) is small. This explains why the weights under mode A are rather close to the factor loadings ($\lambda_{ec}^{(s)}$) under CB-SEM, and their fine differences can be due to misspecification not on the top ten list of the LM test, which identifies 25 error-covariances. The observed differences between $w_a^{(s)}$ and $\lambda_{ec}^{(s)}$ or $\lambda_{cl}^{(s)}$ are due to cross-loadings or other model misspecification not detected by the LM-test.

Mode B

Within-block error-covariances do affect the weights under PLS-SEM mode B. Cases of negative w_b caused by within-block error-covariances were reported in the supplementary material for the analysis of the model in Figure 3. The accumulation of the effects by positive and negative error-covariances makes weights to change in different directions. For the four blocks of indicators in Figure 4, the 3rd block (ξ_3, professional efficacy) has most correlated errors, the weight w_b in this block also varies most, from -1.0 to 1.998. Actually, the first error of the block is correlated with three other errors within the same block, this also explains why we need to set the first w_b of this block as negative in order to keep the correlations and path coefficients among the constructs under mode B consistent with the other methods.

The block that does not have a negative weights under mode B is the 2nd block (ξ_2, cynicism). This block has one-cross loading, one within-block error-covariance, and one between-block error-covariance. While the other blocks have either more correlated errors or more cross-loadings.

The results in sections 3 and 4 imply that weight-changes for composites under PLS-SEM depend on the size and signs of the cross-loadings and error-covariances as well as those of the path coefficients in the structural model. With real data, sampling errors also contribute to the fine differences among the estimated weights in Table 4.

6. Discussion and Recommendation

Many aspects of path analysis depend on the properties of the composites, which are further determined by how the composites are formulated. For a given set of indicators, the properties of a weighted composite are totally determined by the corresponding weights. Equally-weighted composites are easy to formulate but do not use any psychometric properties of the indicators. In contrast, composites under PLS-SEM mode B possess the property of maximum reliability among all weighted composites. But such a property is obtained under the assumption of a correctly specified measurement model. Because model misspecification cannot be avoided in practice, sensitivity analysis of the weights of composites is necessary to better understand the results of different approaches to path analysis. The results of our study indicate that weights of composites under PLS-SEM mode A are not affected by within-block error-covariances; and between-block error-covariances and cross-loadings only affect the weights of the involved individual items. The weights under PLS-SEM mode B are affected by all three types of model misspecification. While only the weights of the involved items under mode B are affected, they can have substantial value changes. In contrast, a single specification error has a block-wise effect on the weights under mode B_A. But w_{b_a} of the involved items are not as sensitive as the corresponding w_b to each type of specification errors. Unless Heywood case occurs, the sign of w_{b_a} is always the same as that of w_a. However, the values of w_{b_a} can be rather large when the corresponding error variances in equation (8) are close to zero. Studies on Heywood cases indicated that they are related to sampling error, specification error and/or small values of error variances (Kano, 1998; van Driel, 1978). We suspect that the same causes apply to the model in equation (8) while additional studies are needed.

Both "sensitivity" and "stability" can be used to describe the properties of weights of composites under PLS-SEM. Sensitivity might be a good property if our purpose is to detect model misspecification or to maximize the power in detecting the relationship among the involved constructs. However, a weight that can change from positive to negative might be too sensitive to model specification. This is the case of PLS-SEM mode B. In contrast, stability to minor or moderate model misspecification can be regarded as a good property. However, a rather stable method may not be able to reflect key properties of the data and model. This is the case for PLS-SEM mode A. Actually, mode A ignores the size of measurement error and the resulting composite can be less reliable than the equally-weighted composite or a single item of its block. The weights under PLS-SEM mode B_A are a compromise between those under modes A and B. The resulting composites of mode B_A keep the stability of mode A and also the sensitivity of mode B. In particular, composites under mode B_A account for measurement errors as the Bartlett-factor-scores do and possess the property of maximum reliability. Consequently, we recommend the use of mode B_A for models with reflective indicators while additional studies are needed regarding the sensitivity of w_{b_a} to sampling and specification errors when the true error variance is small.

The results of the article show that the patterns of the different weights may contain valuable information about model specification. If the scaled versions of w_a are close to the factor loadings for a block of indicators, then the block may not have significant cross-loadings or between-block error-covariances. But significant within-block error-covariances may still exist. If w_b is close to w_{b_a} for a block of indicators, then the block may not have significant (between- & within-block) error-covariances nor cross-loadings. However, indicators in the other blocks may still have cross-loadings on the factor of this block. Consequently, approximate equality of $\mathbf{w}_b$ and $\mathbf{w}_{b_a}$ for all the blocks implies that the measurement model is well-specified and the corresponding composites approximately possess the maximum reliability. If the three weights ($\mathbf{w}_a$, $\mathbf{w}_{b_a}$, $\mathbf{w}_b$) are approximately equal across all the blocks, then the items of the observed data may be parallel or the error variances of each block are approximately homogeneous. However, additional studies are needed to make the above observations statically sound or to develop proper statistical tests to compare the different weights.

Our analysis of the real data is based on the results of the LM test, which is an important tool for post-hoc model modification. In particular, given the formulation of the model, the LM test can identify the parameters whose inclusion will most effectively improve the model-fit under CB-SEM. However, the LM test may not be able to identify the correct paths when the model is far from being correct or when multiple locations of the model contain specification errors (Yuan & Bentler, 2004). The parameters identified by the LM test may be affected by chance errors, especially when the sample size is small. Nevertheless, when the sample size is sufficiently large, the test can provide valuable information on lack of fit locally when the base model has the correct number of latent variables and each indicator has a major loading on the construct of its assigned block (MacCallum, Roznowski & Necowitz, 1992).

When the model is correctly specified or when within-block error-covariances exist, consistent parameter estimates can be obtained by a procedure called PLSc or its modification (Dijkstra & Henseler, 2015; Rademaker et al., 2019), based on estimated weights by PLS-SEM mode A. Although CB-SEM automatically deliver consistent parameter estimates when error-covariances or cross-loadings are known *a priori*, it would be an advance of the PLS-SEM methodology if procedures can be developed to yield consistent parameter estimates by accounting for cross-loadings or between-block error-covariances. While additional studies are needed in this direction, the analytical results in this article will provide the needed guidance for such a development.

While the focus of this article is on models with reflective indicators, PLS-SEM mode B was conventionally recommended for models with formative indicators, which do not need to contain measurement errors nor do they share a common latent trait (Treiblmaier, Bentler & Mair, 2011). Note that the algorithm for computing the weights under mode B does not know whether the indicators are truly formative or reflective. When prediction accuracy is a primary concern, model specification or the signs of the weights might be only a secondary concern (Rigdon et al., 2017). Then the hyper-sensitivity of the weights under mode B can be a strength. This sensitivity allows the method to fully pick up the extra associations that are neglected by the specified path model. Such a strength has been proven with two blocks of indicators, where mode B yields the greatest possible correlation between the two composites (Areskoug, 1982). With error-free formative indicators, we may call the covariances of items not fully explained by the path model *excess item-covariances*. The results of this article shed a different perspective on the strength of PLS-SEM mode B in picking up the excess item-covariances. Because excess item-covariances exert a more direct effect on the relationship of the composites than covariances transferred via the path coefficients, mode B has to adjust the individual weights in order to maximize the predictive power.

There are also several developments for composite models where formative indicators are assumed to be error-free (see e.g., Cho & Choi, 2020; Cho, Sarstedt & Hwang, 2022; Dijkstra, 2017; Henseler, 2021; Hwang et al., 2020; 2021; Schuberth, 2021). In particular, Dijkstra

(2017) showed that, under a composite model with some regularity conditions, PLS-SEM mode B yields consistent estimates of the population weights for all the formative indicators. The regularity conditions include that all the covariance matrices between different blocks of indicators are rank 1. When cross-loadings or excess item-covariances exist, the rank 1 condition will be violated. Additional studies are needed to further examine how weights under mode B are affected under misspecified composite models.

References

Anderson, J.C., & Gerbing, D.W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin, 103*(3), 411–423. https://doi.org/10.1037/0033-2909.103.3.411

Areskoug, B. (1982). The first canonical correlation: Theoretical PLS analysis and simulation experiments. In K. G. Jöreskog & H. Wold (Eds.), *Systems under indirect observations: Causality, structure, prediction* (Part 2, pp. 95–118). Amsterdam, Netherlands: North-Holland.

Bentler, P.M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*, 238–246. doi: 10.1037/0033-2909.107.2.238

Bentler, P.M. (2006). *EQS 6 structural equations program manual.* Encino, CA: Multivariate Software.

Cho, G., & Choi, J. Y. (2020). An empirical comparison of generalized structured component analysis and partial least squares path modeling under variance-based structural equation models. *Behaviormetrika, 47*(1), 243–272. doi: 10.1007/s41237-019-00098-0

Cho, G., Sarstedt, M., & Hwang, H. (2022). A comparison of covariance structure analysis, partial least squares path modeling and generalized structured component analysis in factor- and composite models. *British Journal of Mathematical and Statistical Psychology.* Advance online publication.

Cochran, W.G. (1970). Some effects of errors of measurement on multiple correlation. *Journal of the American Statistical Association, 65*(329), 22–34. doi: 10.1080/01621459.1970.10481059

Deng, L., & Yuan, K.-H. (2022). Which method is more powerful in testing the relationship of theoretical constructs: A meta comparison of structural equation modeling and path analysis with weighted-composites. *Behavior Research Methods.* doi: 10.3758/s13428-022-01838-z.

Dijkstra, T. (1983). Some comments on maximum likelihood and partial least squares estimates. *Journal of Econometrics, 22*, 67–90. doi: 10.1016/0304-4076(83)90094-5

Dijkstra, T.K. (2017). A perfect match between a model and a mode. In H. Latan & R. Noonan (Eds.), *Partial least squares path modeling: Basic concepts, methodological issues and applications* (pp. 55–80). Cham, Switzerland: Springer. doi: /10.1007/978-3-319-64069-3_4

Dijkstra, T.K., & Henseler, J. (2015). Consistent and asymptotically normal PLS estimators for linear structural equations. *Computational Statistics and Data Analysis, 81*, 10–23. doi: 10.1016/j.csda.2014.07.008

Fuller, W. A. (1987). *Measurement error models*. New York: Wiley.

Hair, J.F., Hult, G.T.M., Ringle, C.M., & Sarstedt, M. (2017). *A primer on partial least squares structural equation modeling (PLS-SEM)* (2nd ed). Thousand Oaks, CA: Sage.

Henseler, J. (2021). *Composite-based structural equation modeling*. New York: Guilford.

Hu, L.-T., & Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1–55. doi: 10.1080/10705519909540118

Hwang, H., Cho, G., Jung, K., Falk, C., Flake, J., & Jin, M. (2021). An approach to structural equation modeling with both factors and components: Integrated generalized structured component analysis. *Psychological Methods, 26*(3), 273–294. doi:10.1037/met0000336

Hwang, H., Sarstedt, M., Cheah, J.H., & Ringle, C.M. (2020). A concept analysis of methodological research on composite-based structural equation modeling: Bridging PLSPM and GSCA. *Behaviormetrika, 47*, 219–241. doi: 10.1007/s41237-019-00085-5

Kano, Y. (1998). Improper solutions in exploratory factor analysis: Causes and treatments. In A. Rizzi, M. Vichi and H. Bock (Eds.), *Advances in data sciences and classification* (pp. 375–382). Berlin: Springer.

MacCallum, R.C., Browne, M.W., & Sugawara, H.M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods, 1*(2), 130–149. doi: 10.1037/1082-989X.1.2.130

MacCallum, R.C., Roznowski, M., & Necowitz, L.B. (1992). Model modification in covariance structure analysis: The problem of capitalization on chance. *Psychological Bulletin, 111*, 490–504. doi: 10.1037/0033-2909.111.3.490

Mardia, K.V. (1970). Measure of multivariate skewness and kurtosis with applications. *Biometrika, 57*, 519–530. doi:10.1093/biomet/57.3.519

Rademaker, M.E., Schuberth, F., & Dijkstra, T.K. (2019). Measurement error correlation within blocks of indicators in consistent partial least squares: Issues and remedies. *Internet Research, 29*(3), 448–463. doi: 10.1108/IntR-12-2017-0525

Rigdon, E.E., Sarstedt, M., & Ringle, C.M. (2017). On comparing results from CB-SEM and PLS-SEM: Five perspectives and five recommendations. *Marketing ZFP, 39*(3), 4–16. doi: 10.15358/0344-1369-2017-3-4

Schneeweiss, H. (1993). Consistency at large in models with latent variables. In K. Haagen, D. J. Bartholomew, & M. Deistler (Eds.), *Statistical modelling and latent variables* (pp. 299–320). Elsevier Science Publishers.

Schuberth, F. (2021). The Henseler-Ogasawara specification of composites in structural equation modeling: A tutorial. *Psychological Methods*. doi: 10.1037/met0000432

Silvey, S.D. (1959). The Lagrangian multiplier test. *Annals of Mathematical Statistics, 30*, 389–407. doi: 10.1214/aoms/1177706259

Sörbom, D. (1989). Model modification. *Psychometrika, 54*, 371–384. doi: 10.1007/BF02294623

Steiger, J.H., & Lind, J.M. (1980, May). *Statistically based tests for the number of common factors*. Paper presented at the Annual Meeting of the Psychometric Society. Iowa City, IA.

Tenenhaus, M., Vinzi, V.E., Chatelin, Y.-M., & Lauro, C. (2005). PLS path modeling. *Computational Statistics & Data Analysis, 48*, 159–205. doi: 10.1016/j.csda.2004.03.005

Treiblmaier, H., Bentler, P.M., & Mair, P. (2010). Formative constructs implemented via common factors. *Structural Equation Modeling, 18*, 1–17. doi: 10.1080/10705511.2011.532693

van Driel, O.P. (1978). On various causes of improper solutions of maximum likelihood factor analysis. *Psychometrika, 43*, 225–243. doi: 10.1007/BF02293865

Wold, H. (1980). Model construction and evaluation when theoretical knowledge is scarce. In J. Kmenta & J.B. Ramsey (Eds.), *Evaluation of econometric models* (pp. 47–74). New York: Academic Press.

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

Yuan, K.-H., & Bentler, P.M. (2004). On chi-square difference and z tests in mean and covariance structure analysis when the base model is misspecified. *Educational and Psychological Measurement, 64*, 737–757. doi: 10.1177/0013164404264853

Yuan, K.-H., & Deng, L. (2021). Equivalence of partial-least-squares SEM and the methods of factor-score regression. *Structural Equation Modeling, 28*(4), 557–571. doi: 10.1080/10705511.2021.1894940

Yuan, K.-H., & Fang, Y. (2022). Which method delivers greater signal-to-noise ratio: Structural equation modeling or regression analysis with composite scores? Under review.

Yuan, K.-H., Kouros, C.D., & Kelley, K. (2008). Diagnosis for covariance structure models by analyzing the path. *Structural Equation Modeling, 15*, 564–602. doi: 10.1080/10705510802338991

Yuan, K.-H., Marshall, L.L., & Bentler, P.M. (2003). Assessing the effect of model misspecifications on parameter estimates in structural equation models. *Sociological Methodology, 33*, 241–265. doi: 10.1111/j.0081-1750.2003.00132.x

Figure 1. A model with two latent variables and seven indicators, dashed arrows represent parameters whose non-zero values make the unidimensional model misspecified.

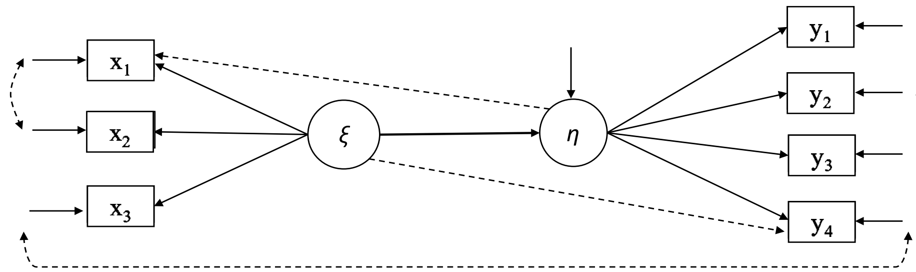

Figure 2. A model with three latent variables and ten indicators, dashed arrows represent parameters whose non-zero values make the unidimensional model misspecified.

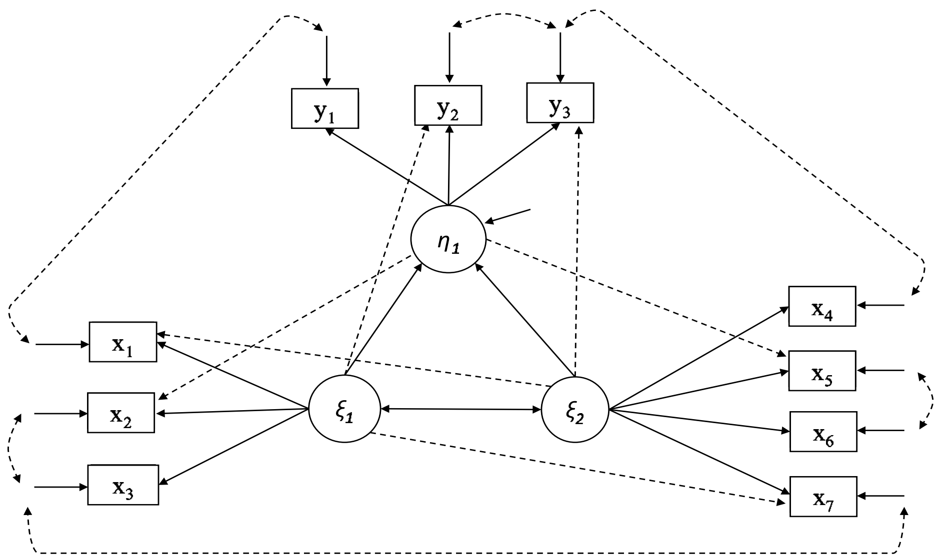

Figure 3. A model with four latent variables and thirteen indicators, dashed arrows represent parameters whose non-zero values make the unidimensional model misspecified.

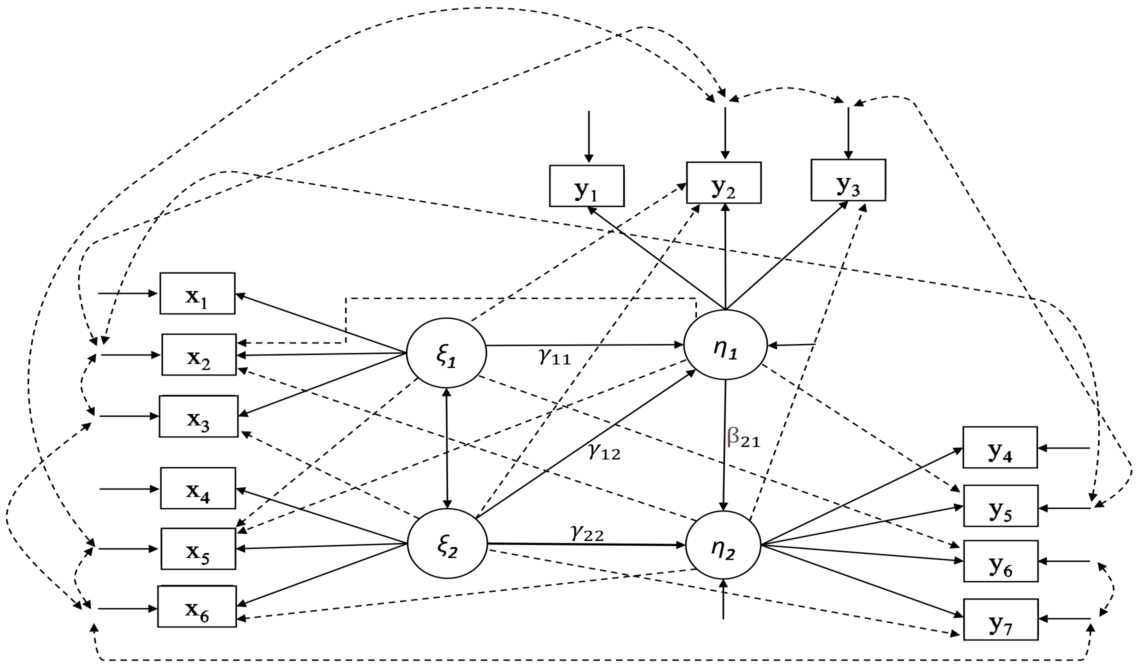

Figure 4. A latent-variable model of health and stress (Yuan & Deng, 2021, $N = 264$, the dashed arrows are model modifications identified by the Lagrange Multiplier test but are not part of the original model).

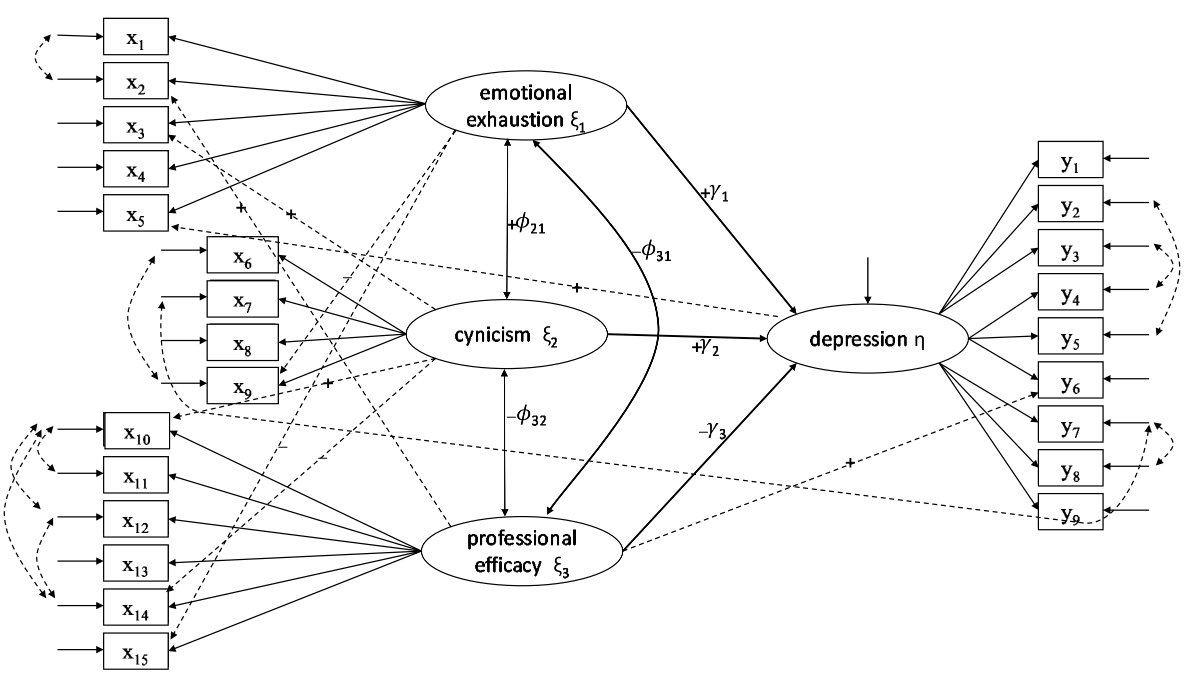

Table 1. Relative population weight changes ($r_j^{(s)}$ in equation 15) for the three modes of PLS-SEM when the model in Figure 1 (2 latent variables and 7 indicators) is misspecified by omitting error covariances of size $\psi_h = \pm 0.2$.

	$\psi_{x_2x_3} = \psi_{x_3x_2} = 0.2$, RMSEA=0.034				$\psi_{x_2x_3} = \psi_{x_3x_2} = -0.2$, RMSEA=0.042			
variable	λ_*/λ_0	PLS-A	PLS-B_A	PLS-B	λ_*/λ_0	PLS-A	PLS-B_A	PLS-B
x_1	0.939	1.000	1.000	1.000	1.042	1.000	1.000	1.000
x_2	1.085	1.000	1.046	0.480	0.925	1.000	0.965	1.920
x_3	1.071	1.000	1.274	0.867	0.943	1.000	0.849	1.533
y_1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
y_2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
y_3	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
y_4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	$\psi_{y_2y_3} = \psi_{y_3y_2} = 0.2$, RMSEA=0.073				$\psi_{y_2y_3} = \psi_{y_3y_2} = -0.2$, RMSEA=0.070			
variable	λ_*/λ_0	PLS-A	PLS-B_A	PLS-B	λ_*/λ_0	PLS-A	PLS-B_A	PLS-B
x_1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
x_2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
x_3	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
y_1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
y_2	1.053	1.000	1.010	0.867	0.975	1.000	0.991	1.362
y_3	1.079	1.000	0.987	0.600	0.959	1.000	1.012	1.629
y_4	0.982	1.000	1.085	1.000	1.023	1.000	0.932	1.000
	$\psi_{x_2y_2} = \psi_{y_2x_2} = 0.2$, RMSEA=0.073				$\psi_{x_2y_2} = \psi_{y_2x_2} = -0.2$, RMSEA=0.072			
variable	λ_*/λ_0	PLS-A	PLS-B_A	PLS-B	λ_*/λ_0	PLS-A	PLS-B_A	PLS-B
x_1	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
x_2	1.013	1.068	1.319	1.746	0.990	0.939	0.787	0.802
x_3	0.990	1.000	0.904	1.000	1.008	1.000	1.124	1.000
y_1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
y_2	1.010	1.081	1.496	2.149	0.993	0.926	0.716	0.505
y_3	1.000	1.000	1.025	1.000	1.000	1.000	0.974	1.000
y_4	0.995	1.000	0.873	1.000	1.005	1.000	1.185	1.000

Table 2. Relative population weight changes ($r_j^{(s)}$ in equation 15) for the three modes of PLS-SEM when the model in Figure 1 (2 latent variables and 7 indicators) is misspecified by omitting cross-loadings of size $\lambda_h = \pm 0.3$.

	$\lambda_{x_2\eta} = 0.3$, RMSEA=0.037				$\lambda_{x_2\eta} = -0.3$, RMSEA=0.044			
variable	λ_*/λ_0	PLS-A	PLS-B$_\mathrm{A}$	PLS-B	λ_*/λ_0	PLS-A	PLS-B$_\mathrm{A}$	PLS-B
x_1	0.986	1.000	1.000	1.000	0.989	1.000	1.000	1.000
x_2	1.259	1.381	2.412	3.077	0.757	0.619	0.381	-0.154
x_3	0.968	1.000	0.820	1.000	1.031	1.000	1.582	1.000
y_1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
y_2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
y_3	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
y_4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	$\lambda_{y_2\xi} = 0.3$, RMSEA=0.045				$\lambda_{y_2\xi} = -0.3$, RMSEA=0.049			
variable	λ_*/λ_0	PLS-A	PLS-B$_\mathrm{A}$	PLS-B	λ_*/λ_0	PLS-A	PLS-B$_\mathrm{A}$	PLS-B
x_1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
x_2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
x_3	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
y_1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
y_2	1.229	1.357	2.750	3.744	0.787	0.643	0.384	-0.425
y_3	1.001	1.000	1.045	1.000	1.000	1.000	0.945	1.000
y_4	0.989	1.000	0.796	1.000	1.015	1.000	1.535	1.000

Table 3. Estimates of the parameters for the model represented by the solid arrows in Figure 4 under CB-SEM.

para	est	SE	z	para	est	SE	z
$\lambda_{x_1\xi_1}$	1.006	0.060	16.820	$\psi_{x_1x_1}$	0.390	0.043	9.084
$\lambda_{x_2\xi_1}$	1.057	0.066	15.984	$\psi_{x_2x_2}$	0.535	0.056	9.558
$\lambda_{x_3\xi_1}$	1.202	0.072	16.621	$\psi_{x_3x_3}$	0.587	0.064	9.210
$\lambda_{x_4\xi_1}$	1.085	0.062	17.460	$\psi_{x_4x_4}$	0.379	0.044	8.612
$\lambda_{x_5\xi_1}$	1.110	0.068	16.365	$\psi_{x_5x_5}$	0.536	0.057	9.359
$\lambda_{x_6\xi_2}$	1.193	0.073	16.442	$\psi_{x_6x_6}$	0.588	0.067	8.814
$\lambda_{x_7\xi_2}$	1.288	0.071	18.155	$\psi_{x_7x_7}$	0.409	0.058	7.014
$\lambda_{x_8\xi_2}$	1.335	0.078	17.107	$\psi_{x_8x_8}$	0.612	0.074	8.235
$\lambda_{x_9\xi_2}$	1.071	0.081	13.157	$\psi_{x_9x_9}$	1.054	0.102	10.311
$\lambda_{x_{10}\xi_3}$	1.221	0.079	15.449	$\psi_{x_{10}x_{10}}$	0.823	0.083	9.935
$\lambda_{x_{11}\xi_3}$	1.399	0.080	17.445	$\psi_{x_{11}x_{11}}$	0.642	0.072	8.875
$\lambda_{x_{12}\xi_3}$	1.340	0.081	16.586	$\psi_{x_{12}x_{12}}$	0.744	0.079	9.421
$\lambda_{x_{13}\xi_3}$	1.234	0.084	14.738	$\psi_{x_{13}x_{13}}$	0.995	0.098	10.177
$\lambda_{x_{14}\xi_3}$	1.380	0.081	16.955	$\psi_{x_{14}x_{14}}$	0.715	0.078	9.207
$\lambda_{x_{15}\xi_3}$	1.304	0.080	16.337	$\psi_{x_{15}x_{15}}$	0.752	0.079	9.550
$\lambda_{y_1\eta}$	1.000			$\psi_{y_1y_1}$	0.174	0.018	9.667
$\lambda_{y_2\eta}$	1.034	0.072	14.441	$\psi_{y_2y_2}$	0.173	0.018	9.529
$\lambda_{y_3\eta}$	0.916	0.076	11.977	$\psi_{y_3y_3}$	0.270	0.026	10.497
$\lambda_{y_4\eta}$	0.918	0.065	14.032	$\psi_{y_4y_4}$	0.154	0.016	9.753
$\lambda_{y_5\eta}$	0.786	0.076	10.359	$\psi_{y_5y_5}$	0.305	0.028	10.835
$\lambda_{y_6\eta}$	0.999	0.076	13.200	$\psi_{y_6y_6}$	0.231	0.023	10.115
$\lambda_{y_7\eta}$	0.947	0.077	12.232	$\psi_{y_7y_7}$	0.270	0.026	10.428
$\lambda_{y_8\eta}$	0.921	0.074	12.440	$\psi_{y_8y_8}$	0.242	0.023	10.369
$\lambda_{y_9\eta}$	0.742	0.075	9.863	$\psi_{y_9y_9}$	0.310	0.028	10.914
γ_1	0.066	0.036	1.831	ϕ_{21}	0.451	0.055	8.260
γ_2	0.269	0.040	6.783	ϕ_{31}	-0.182	0.064	-2.832
γ_3	-0.044	0.032	-1.365	ϕ_{32}	-0.187	0.065	-2.887
σ_ζ^2	0.202	0.028	7.279				

Table 4. Estimates of weights for composites under PLS-SEM mode A (w_a), mode B (w_b), and mode B_A (w_{b_a}) based on the model represented by the solid arrows in Figure 4. Three sets of estimates of factor loadings and structural parameters are also included for reference (λ, λ_{cl}, λ_{ec} for the models represented by the solid arrows in Figure 4, with 8 cross-loadings, and with 10 error-covariances, respectively). The super script s indicates that the weight vectors as well as the factor loadings for each block of indicators are scaled so that its first element is 1.0 while the first element of $\mathbf{w}_b$ for the block of ξ_3 under mode B is kept at -1.0.

variable	$\lambda^{(s)}$	$\lambda_{cl}^{(s)}$	$\lambda_{ec}^{(s)}$	$w_a^{(s)}$	$w_{b_a}^{(s)}$	$w_b^{(s)}$
$x_{1\xi_1}$	1.000	1.000	1.000	1.000	1.000	1.000
$x_{2\xi_1}$	1.050	1.079	1.046	0.706	0.272	-1.248
$x_{3\xi_1}$	1.195	1.057	1.241	1.490	*13.494	1.759
$x_{4\xi_1}$	1.078	1.079	1.115	0.869	0.470	-0.809
$x_{5\xi_1}$	1.103	1.040	1.153	1.337	8.359	1.336
$x_{6\xi_2}$	1.000	1.000	1.000	1.000	1.000	1.000
$x_{7\xi_2}$	1.080	1.079	1.057	1.041	1.117	1.107
$x_{8\xi_2}$	1.119	1.123	1.093	1.240	2.164	2.533
$x_{9\xi_2}$	0.898	0.969	0.921	1.018	0.853	1.357
$x_{10\xi_3}$	1.000	1.000	1.000	1.000	1.000	-1.000
$x_{11\xi_3}$	1.146	1.100	1.139	1.507	2.838	1.365
$x_{12\xi_3}$	1.097	1.051	1.162	1.309	1.678	-0.279
$x_{13\xi_3}$	1.010	0.965	1.054	1.598	4.811	1.121
$x_{14\xi_3}$	1.130	1.054	1.257	1.628	4.662	0.025
$x_{15\xi_3}$	1.068	0.998	1.097	1.690	12.658	1.998
$y_{1\eta}$	1.000	1.000	1.000	1.000	1.000	1.000
$y_{2\eta}$	1.034	1.036	1.067	1.002	0.864	0.659
$y_{3\eta}$	0.916	0.913	0.881	0.761	0.279	-0.302
$y_{4\eta}$	0.918	0.916	0.895	0.899	0.885	0.947
$y_{5\eta}$	0.786	0.784	0.809	0.650	0.222	0.067
$y_{6\eta}$	0.999	1.033	1.001	0.799	0.305	-0.218
$y_{7\eta}$	0.947	0.946	0.929	0.907	0.430	0.628
$y_{8\eta}$	0.921	0.922	0.889	0.831	0.386	0.104
$y_{9\eta}$	0.742	0.745	0.746	0.666	0.244	0.325
γ_1	0.066	0.046	0.072	0.148	0.161	0.198
γ_2	0.269	0.280	0.266	0.436	0.435	0.408
γ_3	-0.044	-0.054	-0.034	-0.079	-0.084	-0.122
ϕ_{21}	0.451	0.420	0.452	0.424	0.463	0.447
ϕ_{31}	-0.182	-0.174	-0.194	-0.181	-0.224	-0.229
ϕ_{32}	-0.187	-0.187	-0.202	-0.176	-0.194	-0.207

Note*: A negative error-variance (-0.071), called Heywood case, occurred for $x_{3\xi_1}$ in the procedure of transforming w_a to w_{b_a}, and the negative value is changed to .05.

Table 5. Model modification by including cross-loadings or error-covariances

	cross-loadings			error-covariances		
order	loading	est	z	error-cov	est	z
1	$\lambda_{x_3\xi_2}$	.283	4.664	$\psi_{x_{10}x_{11}}$	.289	3.907
2	$\lambda_{x_{10}\xi_2}$	.218	3.407	$\psi_{y_3y_4}$	.067	4.127
3	$\lambda_{x_5\eta}$	.279	2.741	$\psi_{x_{10}x_{12}}$	.122	1.853
4	$\lambda_{y_6\xi_3}$	.082	2.463	$\psi_{y_7y_8}$	.084	4.366
5	$\lambda_{x_2\xi_3}$	.120	2.340	$\psi_{x_1x_2}$	.144	3.351
6	$\lambda_{x_{15}\xi_1}$	-.156	-2.514	$\psi_{x_{12}x_{14}}$	-.284	-4.860
7	$\lambda_{x_9\xi_1}$	-.180	-2.308	$\psi_{y_2y_5}$	-.054	-3.473
8	$\lambda_{x_{14}\xi_2}$	-.145	-2.292	$\psi_{x_{10}x_{14}}$	-.228	-3.854
9				$\psi_{x_6x_9}$	-.191	-3.271
10				$\psi_{x_7y_7}$	.086	3.339