

This is a self-archived version of an original article. This version may differ from the original in pagination and typographic details.

Author(s): Schubert, Florian; Rosseel, Yves; Rönkkö, Mikko; Trinchera, Laura; Kline, Rex B.; Henseler, Jörg

Title: Structural Parameters under Partial Least Squares and Covariance-Based Structural Equation Modeling : A Comment on Yuan and Deng (2021)

Year: 2023

Version: Published version

Copyright: © 2022 The Author(s). Published with license by Taylor & Francis Group, LLC

Rights: CC BY 4.0

Rights url: <https://creativecommons.org/licenses/by/4.0/>

Please cite the original version:

Schubert, F., Rosseel, Y., Rönkkö, M., Trinchera, L., Kline, R. B., & Henseler, J. (2023). Structural Parameters under Partial Least Squares and Covariance-Based Structural Equation Modeling : A Comment on Yuan and Deng (2021). *Structural Equation Modeling : A Multidisciplinary Journal*, 30(3), 339-345. <https://doi.org/10.1080/10705511.2022.2134140>



Structural Parameters under Partial Least Squares and Covariance-Based Structural Equation Modeling: A Comment on Yuan and Deng (2021)

Florian Schuberth, Yves Rosseel, Mikko Rönkkö, Laura Trinchera, Rex B. Kline & Jörg Henseler

To cite this article: Florian Schuberth, Yves Rosseel, Mikko Rönkkö, Laura Trinchera, Rex B. Kline & Jörg Henseler (2022): Structural Parameters under Partial Least Squares and Covariance-Based Structural Equation Modeling: A Comment on Yuan and Deng (2021), Structural Equation Modeling: A Multidisciplinary Journal, DOI: [10.1080/10705511.2022.2134140](https://doi.org/10.1080/10705511.2022.2134140)

To link to this article: <https://doi.org/10.1080/10705511.2022.2134140>



© 2022 The Author(s). Published with license by Taylor & Francis Group, LLC.



Published online: 21 Nov 2022.



Submit your article to this journal [↗](#)









View related articles [↗](#)



View Crossmark data [↗](#)

Structural Parameters under Partial Least Squares and Covariance-Based Structural Equation Modeling: A Comment on Yuan and Deng (2021)

Florian Schuberth^a , Yves Rosseel^b , Mikko Rönkkö^c , Laura Trinchera^d , Rex B. Kline^e  and Jörg Henseler^{a,f} 

^aUniversity of Twente; ^bGhent University; ^cUniversity of Jyväskylä; ^dNEOMA Business School; ^eConcordia University; ^fUniversidade Nova de Lisboa

ABSTRACT

In their article, Yuan and Deng argue that a structural parameter under partial least squares structural equation modeling (PLS-SEM) is zero if and only if the same structural parameter is zero under covariance-based structural equation modeling (CB-SEM). Yuan and Deng then conclude that statistical tests on individual structural parameters assessing the null hypothesis of no effect can achieve the same purpose in CB-SEM and PLS-SEM. Our response to their article highlights that the relationship they find between PLS-SEM and CB-SEM structural parameters is not universally valid, and that consequently, tests on individual parameters in CB-SEM and PLS-SEM generally do not fulfill the same purpose.

KEYWORDS

Bias correction; factor score regression; measurement error; partial least squares structural equation modeling (PLS-SEM)

1. Motivation

In a recent article, Yuan and Deng (2021) contribute to a better understanding of different types of factor scores and their relations by showing the relationship between Bartlett factor scores (Bartlett, 1937), regression factor scores (Thomson, 1934; Thurstone, 1935), and scores obtained by partial least squares structural equation modeling (PLS-SEM, Wold, 1982). In doing so, they elaborate on the relationship between regression and Bartlett factor scores and remind us that Bartlett and regression weights are rescaled versions of each other (Bartlett, 1938; Lawley & Maxwell, 1962). Similarly, they elaborate on the findings in the PLS-SEM literature that scores obtained by Mode B weights are (a) asymptotically proportional to regression factor scores if separately computed for each latent variable (see, e.g., Dijkstra, 1985 p. 57 and Dijkstra, 2010 p. 33), and (b) asymptotically univocal, i.e., scores for one factor are asymptotically not contaminated by variance of other factors (Lohmöller, 1989, pp. 100–107; see also Harman, 1976, p. 387). Based on these results, they show that the correlation between regression factor scores and Bartlett factor scores, respectively, equals the canonical correlation between the first pair of canonical variates for cases in which there are only two blocks of indicators. Further, Yuan and Deng (2021) propose a transformation of PLS-SEM Mode A weights into PLS-SEM Mode B weights. Such a transformation can be beneficial since PLS-SEM Mode A weights are known to be more numerically stable than Mode B weights (Dijkstra & Henseler, 2015a), while, as Yuan and Deng (2021) show, factor scores obtained by PLS-SEM Mode B attain maximum reliability asymptotically.

Besides these valuable insights, Yuan and Deng (2021) claim that a structural parameter under PLS-SEM equals zero if and only if the corresponding structural parameter is zero under covariance-based structural equation modeling (CB-SEM). Further, they argue that in structural models containing latent variables the population parameters are arbitrary, and then conclude that “for the purpose of modeling the relationship among latent variables, the proxies still permit unbiased test for the significance of each parameter estimate” (p. 558). More specifically, they claim that a “statistical test on individual parameters [...] under PLS-SEM can achieve the same purpose as that under CB-SEM given the same overall-model structure” (p. 562). This is surprising because it is at odds with the various criticisms raised in the literature regarding the use of PLS-SEM for structural models containing latent variables (e.g., Goodhue et al., 2017; Rönkkö et al., 2015, 2016; Rönkkö & Evermann, 2013; Schuberth, 2021) and contradicts recent PLS-SEM guidelines recommending bias correction if PLS-SEM is applied to latent variable models (e.g., Benitez et al., 2020; Evermann & Rönkkö, in press; Schuberth et al., in press).

In this comment on Yuan and Deng’s paper, we shed more light on the relationship between CB-SEM and PLS-SEM parameters and we show it is generally not true that a structural parameter under PLS-SEM is equal to zero if and only if the corresponding structural parameter equals zero under CB-SEM. In doing so, we draw on the literature on error-in-variables models, i.e., the methodological research strand that focuses on the consequences of variables contaminated by error. In addition, we present a scenario analysis which further illustrates the problem. Furthermore, we explain

that statistical tests on individual structural parameters in most cases do not achieve the same purpose under PLS-SEM and CB-SEM. In the last section we give a conclusion.

2. Relationship between PLS-SEM and CB-SEM Structural Parameters

In their article, Yuan and Deng (2021) argue that a structural parameter under PLS-SEM is equal to zero if and only if the corresponding parameter is zero under CB-SEM, assuming that the model is correctly specified (see also Equation (20) in Yuan and Deng, 2021). According to Yuan and Deng, this is because in PLS-SEM each predictor will be correlated with its response variable whenever the latent predictor variable is correlated with the latent response variable and the two correlation coefficients have the same sign.

As we will show in the following argument, their statement is not generally valid and only true under very specific conditions, such as in structural models with a single latent predictor variable or with orthogonal latent predictor variables, which are unlikely to hold in situations where SEM is typically applied. Like Yuan and Deng (2021), we limit our focus to linear recursive structural equation models with uncorrelated structural disturbance terms. Each latent variable η_j is assumed to be standardized and measured reflectively by a set of indicators $x_j = \lambda_j \eta_j + \varepsilon_j$ with $j = 1, \dots, J$. As is common, we assume that the random measurement errors ε have a mean of zero and are mutually uncorrelated. Further, we assume that the random measurement errors are uncorrelated with the latent variables and the structural disturbance terms. Additionally, it is assumed that the structural disturbance terms are uncorrelated with the exogenous latent variables.

To demonstrate that Yuan and Deng's (2021) statement about the relationship between PLS-SEM and CB-SEM structural parameters is not generally true, we draw on the literature on error-in-variables models, i.e., the literature that studies problems associated with variables contaminated by errors.¹ As Yuan and Deng (2021) pointed out, PLS-SEM creates proxies for each latent variable η_j as $\tilde{\eta}_j = \mathbf{w}'_j \mathbf{x}_j = \mathbf{w}'_j \lambda_j \eta_j + \mathbf{w}'_j \varepsilon_j = q_j \eta_j + \delta_j$ to estimate the structural parameters. Obviously, these proxies are representations of the latent variables contaminated by random measurement error. As is common in PLS-SEM, the weights are scaled to ensure that each proxy has unit variance, i.e., $\text{var}(\tilde{\eta}_j) = \mathbf{w}'_j \Sigma_{jj} \mathbf{w}_j = 1$ where Σ_{jj} denotes the variance-covariance matrix of the indicators associated with the latent variable η_j . Since all latent variables are standardized, the variance of the composed random measurement error term δ_j equals $1 - q_j^2$, where q_j^2 is the reliability of proxy $\tilde{\eta}_j$ for latent variable η_j . Following Dijkstra and Schermelleh-Engel (2014, see Equations 10 and 11), in cases where Mode A weights are used to build the proxies, the reliability of proxy $\tilde{\eta}_j$ is given as $q_j^2 = (\mathbf{w}'_j \lambda_j)^2 = (\lambda'_j \lambda_j) / (\lambda'_j \Sigma_{jj} \lambda_j)$ (see Yuan & Deng, 2021, p. 562 and also Dijkstra, 1985, Chapter 2). Similarly, if

Mode B weights are used to create the proxies, their reliability is given as $q_j^2 = \lambda'_j \Sigma_{jj}^{-1} \lambda_j$ (Dijkstra, 2010, Equation 1.24).

Under PLS-SEM, the structural parameters of each equation of the structural model are estimated by ordinary least squares (OLS). Consequently, its use for structural equation models with latent variables resembles applying the naive estimator, i.e., the OLS estimator, to models with additive random measurement error in all variables, i.e., dependent and independent variables (see, e.g., Carroll et al., 2006, Chapter 2). For this type of model, Buonaccorsi (2010, Equation 5.7) has derived the expected value of the estimates obtained by the naive estimator. Applying Equation 5.7 of Buonaccorsi (2010) to the PLS-SEM context, the probability limit of the PLS-SEM structural parameter estimates is given as follows:²

$$\text{plim } \hat{\gamma}^{PLS} = (\mathbf{Q}\Sigma\mathbf{Q} + \mathbf{I} - \mathbf{Q}^2)^{-1} \mathbf{Q}\Sigma q_{\text{dep}} \gamma, \quad (1)$$

where Σ is the population correlation matrix of the independent latent variables. Moreover, the matrix \mathbf{Q} contains the square root of the reliabilities of the proxies for the independent latent variables on the diagonal and \mathbf{I} is a unit matrix of the same dimension as \mathbf{Q} . Similarly, q_{dep} equals the square root of the reliability of the dependent latent variable's proxy. Note that since in PLS-SEM the proxies are standardized, the random measurement errors comprised in the proxy for the dependent latent variable additionally distort the parameter estimates (e.g., Bollen, 1989, Chapter 5). Further, γ is the vector containing the true standardized structural parameters of that equation, i.e., the probability limit of the standardized structural parameter estimates under CB-SEM. The effects of random measurement errors on PLS-SEM parameter estimates have also been pointed out in prior PLS-SEM literature (see, e.g., Dijkstra, 1985, 2010; Dijkstra & Henseler, 2015a, 2015b).

Given that \mathbf{Q} is a diagonal matrix, Equation (1) reveals that it is generally not true that the probability limit of a structural parameter estimate under PLS-SEM equals zero if and only if the population counterpart equals zero, as Yuan and Deng (2021) have claimed. In fact, this is only the case under one of the following conditions: (i) in simple regression models, i.e., if the model has only one independent latent variable, (ii) if all independent latent variables are uncorrelated, (iii) if all independent latent variables have no effect on the dependent latent

²The original Equation 5.7 of Buonaccorsi (2010) reads as follows:

$$E(\hat{\beta}_{1, \text{naive}}) = \gamma_1 = (\Sigma_{xx} + \Sigma_u)^{-1} \Sigma_{xx} \beta_1 = \kappa \beta_1,$$

where the vector $\hat{\beta}_{1, \text{naive}}$ contains the OLS estimates of the independent variables' parameters in a multiple regression model in which both dependent and independent variables are contaminated by additive random measurement errors. Additionally, Σ_{xx} and Σ_u are the expected variance-covariance matrices of the true independent variables (i.e., the independent variables measured without error) and the random measurement errors comprised in the observed independent variables, respectively. Further, β_1 is the vector containing the expected value of the independent variables' parameters in cases where there is no random measurement error. The matrix κ is also known as the reliability matrix, which determines the degree of bias due to attenuation. Since PLS-SEM relies on estimated weights to form the proxies, we consider the probability limit of the PLS-SEM estimates instead of their expected value. Furthermore, in PLS-SEM all variables, i.e., the latent variables and their proxies, are assumed to be standardized, which is not the case in the formula of Buonaccorsi (2010). Therefore, the reliability of the dependent variables also needs to be taken into account.

¹Error-in-variable models have also been studied in other fields such as econometrics. See, e.g., Wooldridge (2012, pp. 320–323).

Table 1. Scenario 1 in which a structural parameter equals zero under CB-SEM, while the corresponding structural parameter under PLS-SEM is different to zero.

Coefficients	Values
Factor loadings	$\begin{pmatrix} \lambda_{11} \\ \lambda_{12} \\ \lambda_{13} \end{pmatrix} = \begin{pmatrix} 0.5 \\ 0.8 \\ 0.7 \end{pmatrix}; \quad \begin{pmatrix} \lambda_{21} \\ \lambda_{22} \\ \lambda_{23} \end{pmatrix} = \begin{pmatrix} 0.5 \\ 0.8 \\ 0.7 \end{pmatrix}; \quad \begin{pmatrix} \lambda_{31} \\ \lambda_{32} \\ \lambda_{33} \end{pmatrix} = \begin{pmatrix} 0.5 \\ 0.8 \\ 0.7 \end{pmatrix}$
Indicator correlations	$\begin{matrix} \begin{matrix} \underline{x}_{11} & \underline{x}_{12} & \underline{x}_{13} & \underline{x}_{21} & \underline{x}_{22} & \underline{x}_{23} & \underline{x}_{31} & \underline{x}_{32} & \underline{x}_{33} \end{matrix} \\ \begin{pmatrix} 1.00 \\ 0.40 & 1.00 \\ 0.35 & 0.56 & 1.00 \\ 0.15 & 0.24 & 0.21 & 1.00 \\ 0.24 & 0.38 & 0.34 & 0.40 & 1.00 \\ 0.21 & 0.34 & 0.29 & 0.35 & 0.56 & 1.00 \\ 0.11 & 0.17 & 0.15 & 0.17 & 0.28 & 0.25 & 1.00 \\ 0.17 & 0.27 & 0.24 & 0.28 & 0.45 & 0.39 & 0.40 & 1.00 \\ 0.15 & 0.24 & 0.21 & 0.25 & 0.39 & 0.34 & 0.35 & 0.56 & 1.00 \end{pmatrix} \end{matrix}$
Structural parameters under CB-SEM	
Structural parameters under PLS-SEM	

Note: Values are rounded to the second decimal.

variable, or (iv) if the latent variables' proxies show a reliability of zero. Note that in the case of a proxy's reliability being zero, the indicators forming that proxy would be uncorrelated with all other indicators in the model. Consequently, neither CB-SEM nor PLS-SEM estimates are identified. Similarly, we emphasize that in case of isolated latent variables, the PLS-SEM weights used to build the proxies for these latent variables are not defined (Dijkstra, 1985, p. 72).

To further highlight this point, we consider a latent variable model with two correlated latent predictor variables η_1 and η_2 and one latent response variable η_3 . All latent variables are standardized, i.e., they have a mean zero and a unit variance. The structural model is given as follows:

$$\eta_3 = \gamma_1 \eta_1 + \gamma_2 \eta_2 + \zeta. \quad (2)$$

To ensure that the latent response variable has a unit variance, the variance of the structural disturbance term is chosen as $\text{var}(\zeta) = 1 - \gamma_1^2 - 2\gamma_1\gamma_2\phi_{12} - \gamma_2^2$, where the correlation between the two latent predictor variables is denoted

by ϕ_{12} . Considering Equation (1), the structural parameter estimate under PLS-SEM $\hat{\gamma}_1^{\text{PLS}}$ converges in probability to:

$$\text{plim} \hat{\gamma}_1^{\text{PLS}} = \frac{\gamma_1 q_1 q_3 (1 - \phi_{12}^2 q_2^2) + \gamma_2 \phi_{12} q_1 q_3 (1 - q_2^2)}{1 - \phi_{12}^2 q_1^2 q_2^2}. \quad (3)$$

Assuming that the true structural parameter between the two latent variables η_1 and η_3 is equal to zero, i.e., $\gamma_1 = 0$, Equation (3) simplifies as follows:

$$\text{plim} \hat{\gamma}_1^{\text{PLS}} = \gamma_2 \frac{\phi_{12} q_1 q_3 (1 - q_2^2)}{1 - \phi_{12}^2 q_1^2 q_2^2}. \quad (4)$$

Considering Equation (4), $\hat{\gamma}_1^{\text{PLS}}$ would be consistent, i.e., $\text{plim} \hat{\gamma}_1^{\text{PLS}} = \gamma_1 = 0$, if γ_2 were equal to zero or if the latent variables η_1 and η_2 were uncorrelated, i.e., $\phi_{12} = 0$. Similarly, $\hat{\gamma}_1^{\text{PLS}}$ would be consistent if the reliabilities of the proxies $\tilde{\eta}_1$ or $\tilde{\eta}_3$ were equal to zero. This is arguably not the case in most empirical settings. Moreover, a completely unreliable proxy would require that the indicators making up that

Table 2. Scenario 2 in which a structural parameter is different to zero under CB-SEM, while the corresponding structural parameter under PLS-SEM equals zero.

Coefficients	Values
Factor loadings	$\begin{pmatrix} \lambda_{11} \\ \lambda_{12} \\ \lambda_{13} \end{pmatrix} = \begin{pmatrix} 0.5 \\ 0.8 \\ 0.7 \end{pmatrix}; \quad \begin{pmatrix} \lambda_{21} \\ \lambda_{22} \\ \lambda_{23} \end{pmatrix} = \begin{pmatrix} 0.5 \\ 0.8 \\ 0.7 \end{pmatrix}; \quad \begin{pmatrix} \lambda_{31} \\ \lambda_{32} \\ \lambda_{33} \end{pmatrix} = \begin{pmatrix} 0.5 \\ 0.8 \\ 0.7 \end{pmatrix}$
Indicator correlations	$\begin{matrix} & \underline{x_{11}} & \underline{x_{12}} & \underline{x_{13}} & \underline{x_{21}} & \underline{x_{22}} & \underline{x_{23}} & \underline{x_{31}} & \underline{x_{32}} & \underline{x_{33}} \\ \begin{pmatrix} 1.00 \\ 0.40 & 1.00 \\ 0.35 & 0.56 & 1.00 \\ -0.17 & -0.27 & -0.24 & 1.00 \\ -0.27 & -0.43 & -0.38 & 0.40 & 1.00 \\ -0.24 & -0.38 & -0.33 & 0.35 & 0.56 & 1.00 \\ -0.08 & -0.14 & -0.12 & 0.17 & 0.27 & 0.24 & 1.00 \\ -0.14 & -0.22 & -0.19 & 0.27 & 0.43 & 0.38 & 0.40 & 1.00 \\ -0.12 & -0.19 & -0.17 & 0.24 & 0.38 & 0.33 & 0.35 & 0.56 & 1.00 \end{pmatrix} \end{matrix}$
Structural parameters under CB-SEM	
Structural parameters under PLS-SEM	

Note: Values are rounded to the second decimal.

proxy be uncorrelated with all other indicators in the model. Consequently, neither the ML estimates nor the PLS-SEM estimates would be identified. Therefore, it can be concluded that a structural parameter of zero under CB-SEM does not generally imply a structural parameter of zero under PLS-SEM.

In a similar way, Equation (4) shows that we cannot conclude that a structural parameter other than zero under PLS-SEM implies a structural parameter other than zero under CB-SEM. Further, one can directly show that a structural parameter of zero under PLS-SEM does not necessarily imply a structural parameter of zero under CB-SEM. Considering Equation (3), this is illustrated if $\gamma_1 = -\gamma_2 \phi_{12} (1 - q_2^2) / (1 - \phi_{12}^2 q_2^2)$, where it is assumed that $\gamma_2 \neq 0$. In this situation, the structural parameter estimate under PLS-SEM $\hat{\gamma}_1^{\text{PLS}}$ converges in probability to zero, while the CB-SEM counterpart converges in probability to a value other than zero. Note how this also shows that it is not true that a structural parameter other than zero under CB-SEM implies a structural parameter other than zero under PLS-SEM.

To show this issue numerically, we focus on two scenarios illustrated in Tables 1 and 2. Scenario 1 assumes that the

structural parameters of η_1 and η_2 on η_3 are 0 and 0.7, respectively. The correlation between the two latent variables η_1 and η_2 is assumed to be 0.6. Further, we measure each latent variable η_j by three indicators with the factor loadings 0.5, 0.8, and 0.7. For simplicity, it is assumed that the indicators are standardized. To form the PLS-SEM proxies, we use Mode A weights; consequently, the reliability of the three proxies is given as $q_j^2 = (\lambda_j' \lambda_j)^2 / (\lambda_j' \Sigma_{jj} \lambda_j) \approx 0.74$. As the following equation shows, the structural parameter estimate under PLS-SEM $\hat{\gamma}_1^{\text{PLS}}$ does not converge in probability to zero, i.e., the true value of the structural parameter (which equals the probability limit of the structural parameter estimate under CB-SEM):

$$\text{plim} \hat{\gamma}^{\text{PLS}} = \begin{pmatrix} 1 & 0.6\sqrt{0.74 \cdot 0.74} \\ 0.6\sqrt{0.74 \cdot 0.74} & 1 \end{pmatrix}^{-1} \begin{pmatrix} 0.62\sqrt{0.74 \cdot 0.74} \\ 0.7\sqrt{0.74 \cdot 0.74} \end{pmatrix} = \begin{pmatrix} 0.10 \\ 0.47 \end{pmatrix}. \quad (5)$$

Scenario 2 is very similar to Scenario 1, but here the structural parameters for the latent variables are set to $\gamma_1 = 0.22$ and $\gamma_2 = 0.82$. Additionally, the correlation between

the two exogenous latent variables is set to $\phi_{12} = -0.68$. As is clear in Table 2, the structural parameter estimate of γ_1 under CB-SEM converges in probability to 0.22, while the PLS-SEM counterpart converges to 0.

The situation outlined above also shows that Yuan and Deng (2021, p. 562) are mistaken in stating that “statistical test[s] on individual parameters (i.e., $H_0 : \gamma_j = 0$) under PLS-SEM can achieve the same purpose as that under CB-SEM given the same overall-model structure.” According to the discussion in the literature on error-in-variables models, it is well known that the validity of statistical tests regarding single regression coefficients is very limited in the presence of measurement error when the tests are based on OLS estimates (e.g., Carroll et al., 2006, Chapter 10). For instance, Brunner and Austin (2009) studied the behavior of the t test to test the null hypothesis that a single regression coefficient equals zero for a regression model with two independent variables that are both contaminated by random measurement error. In their study, they conclude that researchers relying on the OLS estimator and the t test to test $H_0 : \gamma_j = 0$ will almost certainly make a type I error when there are additive random measurement errors in the independent variables. Westfall and Yarkoni (2016) reached the same conclusion when they showed that tests on individual parameters are more reliable in CB-SEM than in OLS regression if random measurement errors are present.

The same holds for PLS-SEM when it is used to estimate structural models containing latent variables. This has also been observed in the PLS-SEM literature where, for instance, Goodhue et al. (2017, pp. 678) conclude that “[s]pecifically, in both regression and PLS[-SEM], excessive false positives are possible, and the incidence increases with measurement error, with the size of the correlation between predictor constructs and with sample size.” Hence, we can conclude that statistical tests on individual parameters do in general not achieve the same purpose under PLS-SEM and CB-SEM.

3. Conclusion

The study of Yuan and Deng (2021) contains various contributions including PLS-SEM Mode A weights transformed into Mode B weights. These transformed weights will enjoy the numerical stability of classical Mode A weights and preserve the asymptotic maximum reliability known from Mode B weights. Moreover, their study contributes to a better understanding of different types of factor scores, including those obtained by PLS-SEM and their relationships. Although it is not new in the PLS-SEM literature that Mode B weights asymptotically attain maximum reliability and asymptotically produce scores that are univocal (Dijkstra, 1985; Lohmöller, 1989), Yuan and Deng further formalize these properties and link them to other more traditional ways of obtaining factor scores. However, both reliability and univocality of PLS-SEM factor scores have been studied before by means of simulation studies and the published results show that their finite sample behavior is not always as good as the asymptotic results suggest (Rönkkö, 2014;

Rönkkö & Ylitalo, 2010). Particularly, in PLS-SEM, the weights depend strongly on the inter-block correlations, i.e., on correlations between indicators that belong to different latent variables. Consequently, in finite samples, particularly in small samples, PLS-SEM factor scores are contaminated by variance of other factors (i.e., univocality is sacrificed; Rönkkö & Ylitalo, 2010, refers to this as *bias*). Hence, the finite sample behavior of PLS-SEM factor scores more likely resembles the behavior of regression factor scores than that of Bartlett factor scores. Similarly, and as Yuan and Deng (2021) noted, only anecdotal evidence is available regarding the performance of PLS-SEM under misspecified models (Rönkkö et al., *in press*). Consequently, we need future research to shed more light on these issues.

Besides providing insight on the different types of factor scores, Yuan and Deng (2021) draw a connection between structural parameters under PLS-SEM and CB-SEM. As we have shown in our comment, the connection they make is only true under very special conditions such as structural equations with one independent latent variable or a situation in which all independent latent variables are uncorrelated. Therefore, it is generally not possible to draw conclusions from a CB-SEM structural parameter about the PLS-SEM counterpart and vice versa. Consequently, effect size measures and statistical tests on individual structural parameters can generally not achieve the same purpose under CB-SEM and PLS-SEM. This highlights that bias caused by random measurement errors, i.e., attenuation bias (Cohen et al., 1990), should be taken seriously even if one is only interested in whether a parameter equals zero in the population.

Notably, testing only for the existence of effects is not a recommended research practice. Researchers should additionally pay attention to the magnitude of the estimates (Wasserstein et al., 2019; Wasserstein & Lazar, 2016). Following Yuan and Deng (2021), this would be a fruitless endeavor in the context of latent variable models because the scales of latent variables, and thus the involved parameters, are arbitrary. However, the scale of a latent variable is no more arbitrary than the scale of any other measured variable, such as height. The choice to measure height in centimeters is arbitrary and the parameters of an analysis would be different if another metric (e.g., millimeters, inches, feet) were to be used. Once the metric of a measured variable has been set and fixed, the measured variable and the involved parameters are no longer arbitrary. It is true that any parameter involving a latent variable can be thought of as an algebraic transformation of an unobserved population quantity (Klopp & Klößner, 2021); however, the same applies to any parameter involving a measured variable because of the measurement scales' arbitrariness (Markus & Borsboom, 2013, Chapter 2). Fortunately, there are ways to deal with the arbitrariness of scales. In the case of latent variables studied in the SEM context various methods have been proposed to choose the metric, such as the fixed marker scaling, the effects coding scaling, and the fixed factor scaling. For an explanation, see, e.g., Little et al. (2006). Similar to choosing the metric of height (e.g., centimeters, inches, feet), choosing a particular scaling method is

arbitrary. Once the scaling method has been chosen, the scale of a latent variable is fixed and the arbitrariness of its scale disappears. Consequently, the parameters associated with the latent variables can be interpreted (see Table 4 of Klopp & Klößner, 2021) and estimation error, bias, and consistency can be assessed in a meaningful way.³

To correct the relationship Yuan and Deng (2021) posit between a structural parameter under PLS-SEM and CB-SEM, PLS-SEM's inconsistency could be addressed by correcting for attenuation as has been done in consistent partial least squares (Dijkstra & Henseler, 2015a). A similar proposal has been made in the context of factor score regression to address the problem of random measurement error comprised in the factor scores (e.g., Devlieger et al., 2016). However, correcting for bias often increases an estimator's variability. To quantify the trade-off between an estimator's bias and its variability, criteria such as mean squared error can be used (Casella & Berger, 2001, Chapter 7). The error-in-variables model literature has shown that it depends on the sample size and the reliability of the independent variables whether the estimator corrected for attenuation bias outperforms the OLS estimator in terms of mean squared error (see, e.g., Fuller, 1987, Table 1.1.1). In the CB-SEM and PLS-SEM context, recent simulation studies suggest that the mean squared error of structural parameters estimated by CB-SEM (or other bias-corrected methods) is usually smaller than the mean squared error of these parameters estimated by methods such as uncorrected factor score regression or PLS-SEM that ignore the presence of measurement error in the scores (see, e.g., Devlieger et al., 2016; Yuan et al., 2020). In other words, the observed bias in methods that ignore random measurement error is often much larger than the added variability that is inherent to methods used in correcting for this bias. Possible exceptions would be in situations where researchers work with a very small sample size.

In the past decade, PLS-SEM has been heavily criticized because it produces inconsistent estimates for structural models that contain latent variables due to attenuation bias (e.g., Cadogan & Lee, *in press*; Henseler & Schubert, *in press*; Rönkkö & Evermann, 2013; Rönkkö et al., 2015, 2016, *in press*; Schubert, 2021). However, these criticisms have rarely been echoed in PLS-SEM guidelines which continue to recommend using PLS-SEM for structural models that contain latent variables including assessment criteria developed under CB-SEM, such as average variance extracted, indicator reliability, and composite reliability (e.g., Hair et al., 2019, 2020, 2021). Against this background, our comment is a further call for future research on PLS-SEM to pay serious attention to random measurement errors.

³This can be understood via an example of physical measurement. Assume that we are interested in understanding whether a scale provides biased measures of mass in grams. It is important to note that the question of interest is not about correct measurement of mass in any absolute sense, but about the correct quantification of mass in grams. As such, bias should be understood as a difference between quantification applied to sample(s) and the same quantification applied to the population. No measure can be understood without understanding the unit, i.e., without scaling (see also Klopp & Klößner, 2021).

Acknowledgements

Jörg Henseler acknowledges a financial interest in the composite-based SEM software ADANCO and its distributor, Composite Modeling.

Funding

Jörg Henseler gratefully acknowledges financial support from FCT Fundação para a Ciência e a Tecnologia (Portugal), national funding through a research grant from the Information Management Research Center—MagIC/NOVA IMS (UIDB/04152/2020).

ORCID

Florian Schubert  <http://orcid.org/0000-0002-2110-9086>

Yves Rosseel  <http://orcid.org/0000-0002-4129-4477>

Mikko Rönkkö  <http://orcid.org/0000-0001-7988-7609>

Laura Trinchera  <http://orcid.org/0000-0001-9679-0956>

Rex B. Kline  <http://orcid.org/0000-0001-6581-592X>

Jörg Henseler  <http://orcid.org/0000-0002-9736-3048>

References

- Bartlett, M. S. (1937). The statistical conception of mental factors. *British Journal of Psychology*, 28, 97–104. <https://doi.org/10.1111/j.2044-8295.1937.tb00863.x>
- Bartlett, M. S. (1938). Methods of estimating mental factors. *Nature*, 141, 609–610.
- Benitez, J., Henseler, J., Castillo, A., & Schubert, F. (2020). How to perform and report an impactful analysis using partial least squares: Guidelines for confirmatory and explanatory IS research. *Information & Management*, 2, 103168. <https://doi.org/10.1016/j.im.2019.05.003>
- Bollen, K. A. (1989). *Structural equations with latent variables*. John Wiley & Sons.
- Brunner, J., & Austin, P. C. (2009). Inflation of type I error rate in multiple regression when independent variables are measured with error. *Canadian Journal of Statistics*, 37, 33–46. <https://doi.org/10.1002/cjs.10004>
- Buonaccorsi, J. (2010). *Measurement error: Models, methods, and applications*. CRC Press.
- Cadogan, J. W., & Lee, N. (*in press*). A miracle of measurement or accidental constructivism? How PLS subverts the realist search for truth. *European Journal of Marketing*. <https://doi.org/10.1108/EJM-08-2020-0637>
- Carroll, R. J., Ruppert, D., Stefanski, L. A., & Crainiceanu, C. M. (2006). *Measurement error in nonlinear models: A modern perspective*. Chapman & Hall/CRC.
- Casella, G., & Berger, R. (2001). *Statistical inference* (2nd ed.). Duxbury.
- Cohen, P., Cohen, J., Teresi, J., Marchi, M., & Velez, C. N. (1990). Problems in the measurement of latent variables in structural equations causal models. *Applied Psychological Measurement*, 14, 183–196. <https://doi.org/10.1177/014662169001400207>
- Devlieger, I., Mayer, A., & Rosseel, Y. (2016). Hypothesis testing using factor score regression: A comparison of four methods. *Educational and Psychological Measurement*, 76, 741–770. <https://doi.org/10.1177/0013164415607618>
- Dijkstra, T. K. (1985). *Latent variables in linear stochastic models: Reflections on "maximum likelihood" and "partial least squares" methods* (Vol. 1). Sociometric Research Foundation.
- Dijkstra, T. K. (2010). Latent variables and indices: Herman Wold's basic design and partial least squares. In V. Esposito Vinzi, W. W. Chin, J. Henseler, & H. Wang (Eds.), *Handbook of partial least squares* (pp. 23–46). Springer.

- Dijkstra, T. K., & Henseler, J. (2015a). Consistent and asymptotically normal PLS estimators for linear structural equations. *Computational Statistics & Data Analysis*, *81*, 10–23. <https://doi.org/10.1016/j.csda.2014.07.008>
- Dijkstra, T. K., & Henseler, J. (2015b). Consistent partial least squares path modeling. *MIS Quarterly*, *39*, 297–316. <https://doi.org/10.25300/MISQ/2015/39.2.02>
- Dijkstra, T. K., & Schermelleh-Engel, K. (2014). Consistent partial least squares for nonlinear structural equation models. *Psychometrika*, *79*, 585–604. <https://doi.org/10.1007/s11336-013-9370-0>
- Evermann, J., & Rönkkö, M. (in press). Recent developments in PLS. *Communications of the Association for Information Systems*, *44*.
- Fuller, W. A. (1987). *Measurement error models*. Wiley.
- Goodhue, D. L., Lewis, W., & Thompson, R. (2017). A multicollinearity and measurement error statistical blind spot: Correcting for excessive false positives in regression and PLS. *MIS Quarterly*, *41*, 667–684. <https://doi.org/10.25300/MISQ/2017/41.3.01>
- Hair, J. F., Hult, G. T. M., Ringle, C. M., Sarstedt, M., Danks, N. P., & Ray, S. (2021). *Partial least squares structural equation modeling (PLS-SEM) using R: A workbook*. Springer.
- Hair, J. F., Howard, M. C., & Nitzl, C. (2020). Assessing measurement model quality in PLS-SEM using confirmatory composite analysis. *Journal of Business Research*, *109*, 101–110. <https://doi.org/10.1016/j.jbusres.2019.11.069>
- Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review*, *31*, 2–24. <https://doi.org/10.1108/EBR-11-2018-0203>
- Harman, H. H. (1976). *Modern factor analysis*. University of Chicago Press.
- Henseler, J., & Schuberth, F. (in press). Partial least squares as a tool for scientific inquiry: Comments on Cadogan and Lee. *European Journal of Marketing*. <https://doi.org/10.1108/EJM-06-2021-0416>
- Klopp, E., & Klößner, S. (2021). The impact of scaling methods on the properties and interpretation of parameter estimates in structural equation models with latent variables. *Structural Equation Modeling*, *28*, 182–206. <https://doi.org/10.1080/10705511.2020.1796673>
- Lawley, D. N., & Maxwell, A. E. (1962). Factor analysis as a statistical method. *Journal of the Royal Statistical Society*, *12*, 209–229. <https://doi.org/10.2307/2986915>
- Little, T. D., Slegers, D. W., & Card, N. A. (2006). A non-arbitrary method of identifying and scaling latent variables in SEM and MACS models. *Structural Equation Modeling*, *13*, 59–72. https://doi.org/10.1207/s15328007sem1301_3
- Lohmöller, J.-B. (1989). *Latent variable path modeling with partial least squares*. Physica.
- Markus, K. A., & Borsboom, D. (2013). *Frontiers of test validity theory: Measurement, causation, and meaning*. Routledge.
- Rönkkö, M. (2014). The effects of chance correlations on partial least squares path modeling. *Organizational Research Methods*, *17*, 164–181. <https://doi.org/10.1177/1094428114525667>
- Rönkkö, M., & Evermann, J. (2013). A critical examination of common beliefs about partial least squares path modeling. *Organizational Research Methods*, *16*, 425–448. <https://doi.org/10.1177/1094428112474693>
- Rönkkö, M., Lee, N., Evermann, J., McIntosh, C., & Antonakis, J. (in press). Marketing or methodology? Exposing the fallacies of PLS with simple demonstrations. *European Journal of Marketing*.
- Rönkkö, M., McIntosh, C. N., & Antonakis, J. (2015). On the adoption of partial least squares in psychological research: Caveat emptor. *Personality and Individual Differences*, *87*, 76–84. <https://doi.org/10.1016/j.paid.2015.07.019>
- Rönkkö, M., McIntosh, C. N., Antonakis, J., & Edwards, J. R. (2016). Partial least squares path modeling: Time for some serious second thoughts. *Journal of Operations Management*, *47–48*, 9–27. <https://doi.org/10.1016/j.jom.2016.05.002>
- Rönkkö, M., & Ylitalo, J. (2010). Construct validity in partial least squares path modeling. In *ICIS 2010 Proceedings*. Retrieved from http://aisel.aisnet.org/icis2010_submissions/155
- Schuberth, F. (2021). Confirmatory composite analysis using partial least squares: Setting the record straight. *Review of Managerial Science*, *15*, 1311–1345. <https://doi.org/10.1007/s11846-020-00405-0>
- Schuberth, F., Zaza, S., & Henseler, J. (in press). Partial least squares is an estimator for structural equation models: A comment on Evermann and Rönkkö (2021). *Communications of the Association for Information Systems*.
- Thomson, G. H. (1934). The meaning of ‘i’ in the estimate of ‘g’. *British Journal of Psychology*, *25*, 92–99. <https://doi.org/10.1111/j.2044-8295.1934.tb00728.x>
- Thurstone, L. L. (1935). *The vectors of mind*. University of Chicago Press.
- Wasserstein, R. L., & Lazar, N. A. (2016). The ASA statement on *p*-values: Context, process, and purpose. *The American Statistician*, *70*, 129–133. <https://doi.org/10.1080/00031305.2016.1154108>
- Wasserstein, R. L., Schirm, A. L., & Lazar, N. A. (2019). Moving to a world beyond $p < 0.05$. *The American Statistician*, *73*, 1–19. <https://doi.org/10.1080/00031305.2019.1583913>
- Westfall, J., & Yarkoni, T. (2016). Statistically controlling for confounding constructs is harder than you think. *PLoS One*, *11*, e0152719. <https://doi.org/10.1371/journal.pone.0152719>
- Wold, H. (1982). Soft modeling: The basic design and some extensions. In K. G. Jöreskog & H. Wold (Eds.), *Systems under indirect observation: Causality, structure, prediction part II* (pp. 1–54). North-Holland.
- Wooldridge, J. M. (2012). *Introductory econometrics: A modern approach* (5th ed.). Cengage Learning.
- Yuan, K.-H., & Deng, L. (2021). Equivalence of partial-least-squares SEM and the methods of factor-score regression. *Structural Equation Modeling*, *28*, 557–571. <https://doi.org/10.1080/10705511.2021.1894940>
- Yuan, K.-H., Wen, Y., & Tang, J. (2020). Regression analysis with latent variables by partial least squares and four other composite scores: Consistency, bias and correction. *Structural Equation Modeling*, *27*, 333–350. <https://doi.org/10.1080/10705511.2019.1647107>