Received: 31 January 2023 | Accepted: 21 March 2023

DOI: 10.1111/bmsp.12304

COMMENTARY

the british psychological society
promoting excellence in psychology

Premature conclusions about the signal-to-noise ratio in structural equation modeling research: A commentary on Yuan and Fang (2023)

Florian Schuberth[1] | **Tamara Schamberger**[1,2] | **Mikko Rönkkö**[3] | **Yide Liu**[4] | **Jörg Henseler**[1,5]

[1]Faculty of Engineering Technology, University of Twente, Enschede, The Netherlands

[2]Faculty of Business Management and Economics, University of Würzburg, Würzburg, Germany

[3]Jyväskylä University School of Business and Economics, University of Jyväskylä, Jyväskylä, Finland

[4]School of Business, Macau University of Science and Technology, Wai Long, Taipa, Macau, China

[5]Nova Information Management School, Universidade Nova de Lisboa, Lisboa, Portugal

Correspondence

Jörg Henseler, Faculty of Engineering Technology, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands.
Email: j.henseler@utwente.nl

Funding information

Information Management Research Center–MagIC/NOVA IMS, Grant/Award Number: UIDB/04152/2020

Abstract

In a recent article published in this journal, Yuan and Fang (*British Journal of Mathematical and Statistical Psychology*, 2023) suggest comparing structural equation modeling (SEM), also known as covariance-based SEM (CB-SEM), estimated by normal-distribution-based maximum likelihood (NML), to regression analysis with (weighted) composites estimated by least squares (LS) in terms of their signal-to-noise ratio (SNR). They summarize their findings in the statement that "[c]ontrary to the common belief that CB-SEM is the preferred method for the analysis of observational data, this article shows that regression analysis via weighted composites yields parameter estimates with much smaller standard errors, and thus corresponds to greater values of the [SNR]." In our commentary, we show that Yuan and Fang have made several incorrect assumptions and claims. Consequently, we recommend that empirical researchers not base their methodological choice regarding CB-SEM and regression analysis with composites on the findings of Yuan and Fang as these findings are premature and require further research.

KEYWORDS

composite model, covariance-based structural equation modeling, effect size, factor score regression, Henseler–Ogasawara specification, partial least squares structural equation modeling, regression analysis with weighted composites, sum scores

1 | MOTIVATION

In a recent article published in this journal, Yuan and Fang (2023) compare structural equation modeling (SEM), also known as covariance-based SEM (CB-SEM), estimated by normal-distribution-based maximum likelihood (NML; see, e.g., Jöreskog, 1970) with various forms of regression analysis with (weighted) composites estimated by least squares (LS), including partial least-squares SEM (PLS-SEM; Hair et al., 2011; Wold, 1975). As the criterion in comparing the two classes of methods, they propose to consider the signal-to-noise ratio (SNR), which they define as the population value of a parameter estimate divided by $\sqrt{N}$ times its standard deviation. As Yuan and Fang (2023, p. 2) emphasize, "the SNR is fundamental in testing the existence of a relationship." Yuan and Fang (2023, abstract) summarize their main finding, stating that "[c]ontrary to the common belief that CB-SEM is the preferred method for the analysis of observational data, this article shows that regression analysis via weighted composites yields parameter estimates with much smaller standard errors, and thus corresponds to greater values of the [SNR]." This implies that by using LS regression analysis with weighted composites instead of CB-SEM, empirical researchers could increase the odds of finding a significant effect.

Considering a broad preference for significant results (the so-called significosis, see Antonakis, 2017), many researchers may find the promise of increasing the likelihood of attaining significant effects appealing. If Yuan and Fang's (2023) statement were correct, we could expect that many empirical researchers would prefer LS regression analysis with composites to CB-SEM estimated by NML in their analyses. But is the statement correct? First doubts appear to be justified, because the NML estimator is known to be consistent and yields asymptotically efficient estimates given that its underlying assumptions are met (Davidson & MacKinnon, 1993). Moreover, Yuan and Fang (2023, p. 13) themselves admit that the results "may seem contrary to expectations." And not least, there is ample evidence that the bias induced by LS regression analysis with (weighted) composites for latent variable models can lead analysts to erroneous conclusions (e.g., Bollen, 1989; Devlieger et al., 2016; Dijkstra, 1985; Dijkstra & Henseler, 2015; Goodhue et al., 2017; McDonald, 1996; Schuberth, 2021; Schuberth et al., 2022; Skrondal & Laake, 2001). Following up on these doubts, we present this commentary.

In our commentary, we discuss three major problems that we believe limit the significance and contribution of Yuan and Fang's article. First, Yuan and Fang (2023) assume that the SNR is method-specific and overlook the fact that the SNR also depends on the model specification. Specifically, if the latent variables in CB-SEM had been scaled differently, the conclusion of Yuan and Fang (2023) would have been different. Second, the computational experiment Yuan and Fang (2023) conducted suffers from aliasing. In particular, their study design does not permit the disentanglement of the estimation procedure's effect, that is, CB-SEM by NML versus regression analysis by LS, from the effect of the representation of the theoretical constructs, that is, latent variables versus (weighted) composites. Consequently, the exact reason for the differences in SNR values remains unclear. Third, Yuan and Fang (2023, p. 2) build on an incorrect assumption that "the population values of the model parameters under SEM are artificial." As we shall explain, this is not the case. Consequently, bias enjoys a substantive interpretation and adjusting the scales of latent variables and composites, respectively, to obtain the same regression coefficients under LS regression analysis and CB-SEM is problematic because the two sets of parameters have different interpretations. In the following sections, we elaborate on these issues.

2 | THE SNR FOR CB-SEM BY NML DEPENDS ON THE SCALING METHOD

The SNR is a measure commonly used in engineering to evaluate the quality of a measurement system (e.g., Kieser et al., 2005; Taguchi et al., 2005). As its name suggests, it is defined as the ratio between a signal and its noise (Taguchi et al., 2005). In the context of regression analysis, there are at least two understandings of the SNR. On the one hand, it can be defined as the ratio of explained variance, that is, R^2, to unexplained variance, that is, $1 - R^2$ (e.g., Cohen, 1988; Czanner et al., 2008). Consequently, the SNR considers the whole regression equation. On the other hand, the SNR can refer to a single

TABLE 1 Estimates of regression coefficients and SNR for different methods.

(a) Parameter estimates, their standard errors (SEs), and z-statistics for confirmatory two-factor model (T_{ml} = 2.073, df = 4, N = 88)

θ	$\hat{\theta}$	SE	z
λ_{y_1}	12.253	1.843	6.649
λ_{y_2}	10.383	1.379	7.530
λ_{x_1}	9.834	.929	10.588
λ_{x_2}	11.490	1.403	8.192
λ_{x_3}	12.517	1.667	7.508
ϕ_{21}	.818	.073	11.258
ψ_{y_1}	155.632	31.679	4.913
ψ_{y_2}	65.036	18.099	3.593
ψ_{x_1}	16.186	7.261	2.229
ψ_{x_2}	88.352	16.773	5.268
ψ_{x_3}	141.074	24.881	5.670

(b) Estimate of γ_*, its standard deviation, and SNR (conditional on estimated weights, except for $\text{BFS}_{\hat{w}}$) for SEM model with closed-book exam trait (η) being predicted by open-book exam trait (ξ)

	$T_{\text{ml}} = 2.073$			$T_{\text{ml}} = 0$		
Method	$\hat{\gamma}_*$	$\hat{\text{SD}}_*$	$\hat{\tau}$	$\hat{\gamma}_*$	$\hat{\text{SD}}_*$	$\hat{\tau}$
CB^{i}	10.019	16.915	.592	10.019	16.915	.592
CB^{ii}	1.019	1.686	.604	1.019	1.686	.604
CB^{iii}	.818	.681	1.200	.818	.681	1.200
BFS	.732	.847	.865	.732	.847	.865
PLS_a	.584	.821	.711	.609	.802	.759
PLS_b	.663	.757	.876	.658	.761	.865
EWC	.428	.587	.730	.450	.576	.782
$\text{BFS}_{\hat{w}}$	.732	.647	1.132	.732	.647	1.132

(c) Average and empirical standard deviation of $\hat{\gamma}_*$ over 1000 bootstrap replications, and corresponding empirical SNR for SEM model with closed-book exam trait (η) being predicted by open-book exam trait (ξ)

	$T_{\text{ml}} = 2.097$			$T_{\text{ml}} = 0$		
Method	$\bar{\gamma}_*$	$\hat{\text{SD}}_*$	$\hat{\tau}$	$\bar{\gamma}_*$	$\hat{\text{SD}}_*$	$\hat{\tau}$
CB^{i}	9.955	18.845	.528	9.968	19.136	.521
CB^{ii}	1.020	1.740	.586	1.023	1.841	.556
CB^{iii}	.819	.859	.954	.820	.871	.941
$\text{BFS}_{\hat{w}}$	.740	.812	.911	.739	.786	.941
$\text{PLS}_{a\hat{w}}$	.592	.739	.801	.616	.717	.858
$\text{PLS}_{b\hat{w}}$	.678	.632	1.073	.674	.652	1.033
EWC	.429	.629	.681	.450	.609	.739

Note: CB^{i}: Var(ξ) and λ_{y_1} are fixed to 1; CB^{ii}: λ_{y_1} and λ_{x_1} are fixed to 1; CB^{iii}: Var(ξ) is fixed to 1, and the variance of the error term of η is constrained to ensure that Var(η) equals 1.

independent variable of a regression equation. In this case, the signal is understood as the effect that an independent variable has on the dependent variable, and the noise corresponds to the uncertainty of this effect (Wencheko, 2000). Consequently, the SNR is the ratio of the effect and its uncertainty, and it resembles effect size measures known from psychological research (Czanner et al., 2015; Gibson, 2015).

In their article, Yuan and Fang follow the second understanding, highlighting that "[t]his ratio also plays a key role in determining the power of a statistical test for the null hypothesis of the corresponding parameter" (p. 25) as the empirical SNR equals the z-statistic of that parameter divided by $\sqrt{N}$. As is known in the CB-SEM literature, the value of the z-statistic, which is a special case of the Wald test statistic (Greene, 2012, chapter 14), depends on the method used to fix the latent variables' scales (e.g., Gonzalez & Griffin, 2001; Klopp & Klößner, 2021). Consequently, the SNR for CB-SEM using NML also depends on the scaling method. To demonstrate this issue, we rerun Yuan and Fang's empirical example and parts of their Monte Carlo simulation. In doing so, we perform all the estimations and calculations in the statistical programming environment R (R Core Team, 2021, Version 4.2.2). Specifically, we use the `lavaan` package (Rosseel, 2012, Version 0.6-12) for the NML estimation of CB-SEM and the `matrixpls` package (Rönkkö, 2022, Version 1.0.15) to conduct PLS-SEM.[1]

The empirical example is based on the open- and closed-book test dataset presented by Mardia et al. (1979), which consists of 88 observations. The specified model contains one independent latent variable, that is, the trait for open-book tests (ξ), and one dependent latent variable, that is, the trait for closed-book tests (η). While ξ is measured by three indicators, η is measured by two. As Yuan and Fang (2023) did and as is shown in Table 1, we compare the following methods: (1) CB-SEM estimated by NML (CB), (2) Bartlett factor score regression (BFS), (3) regression analysis with equally weighted composites (EWC), (4) PLS-SEM using mode A (PLS_a), and (5) PLS-SEM using mode B (PLS_b). Considering CB-SEM, we used three different scaling methods. First, we used the scaling method employed by Yuan and Fang (2023), that is, the variance of the independent latent variable and the first loading of the dependent latent variable are fixed to 1 (CB^{i}). Second, we fix the first loading of the independent and dependent latent variable to 1, that is, fixed marker scaling (CB^{ii}). Third, we fix the variances of the independent and dependent latent variables to 1 (CB^{iii}). Additionally, the Bartlett factor scores are based on the confirmatory factor analysis (CFA) results as reported in Table 1a, which are identical to the CFA results reported by Yuan and Fang (2023) in Table 3(a). Finally, for the LS regression analysis with weighted composites, a $\hat{w}$ in the index of a method indicates that the sampling error in the estimated weights has been accounted for in calculating the standard error of the parameter estimate; otherwise, the standard error is conditional on the estimated weights. For a more detailed description of the empirical example and the studied situations, we refer the reader to Yuan and Fang (2023).

Table 1b presents the estimates of γ_*, the corresponding standard deviation SD_*, and the empirical SNR $\hat{\tau}$. Note that, except for CB^{i}, CB^{ii}, CB^{iii}, and $\text{BFS}_{\hat{w}}$, the SD_* and, thus, the SNR are calculated conditional on the estimated weights. Our results are almost identical to those Yuan and Fang (2023) reported in tab. 3(b). The only difference is observed for CB^{i}, that is, CB-SEM with the scaling method as employed by Yuan and Fang (2023). Although we get the same point estimate, our standard deviation and, thus, the resulting SNR are slightly different. It seems that Yuan and Fang (2023) used in this case $\sqrt{88-1}$ instead of $\sqrt{88}$ times the standard error to obtain the standard deviation.

As can be seen in Table 1b, regardless of whether the original dataset ($T_{\text{ml}} = 2.097$) or the transformed dataset ($T_{\text{ml}} = 0$) is used, among the three scaling methods the one Yuan and Fang (2023) employed, that is, CB^{i}, leads to the smallest SNR for CB-SEM. However, using a different scaling method, that is, CB^{iii}, CB-SEM leads to the largest SNR of the considered methods. Similar outcomes are observed for the SNR calculated based on bootstrap resampling as shown in Table 1c, namely, that CB^{iii} produces the largest SNR of the considered scaling methods for CB-SEM. In addition, CB^{iii} outperforms most of the LS regression analyses with (weighted) composites. Note that, due to resampling, our results are slightly different to those Yuan and Fang (2023) reported in tab. 3(c).

Besides the empirical example, we reran the Monte Carlo simulation presented in section 3.3 of Yuan and Fang (2023), which compares the empirical SNR between CB-SEM by NML and LS regression anal-

[1]The complete R code can be downloaded here: https://osf.io/tcnxy/.

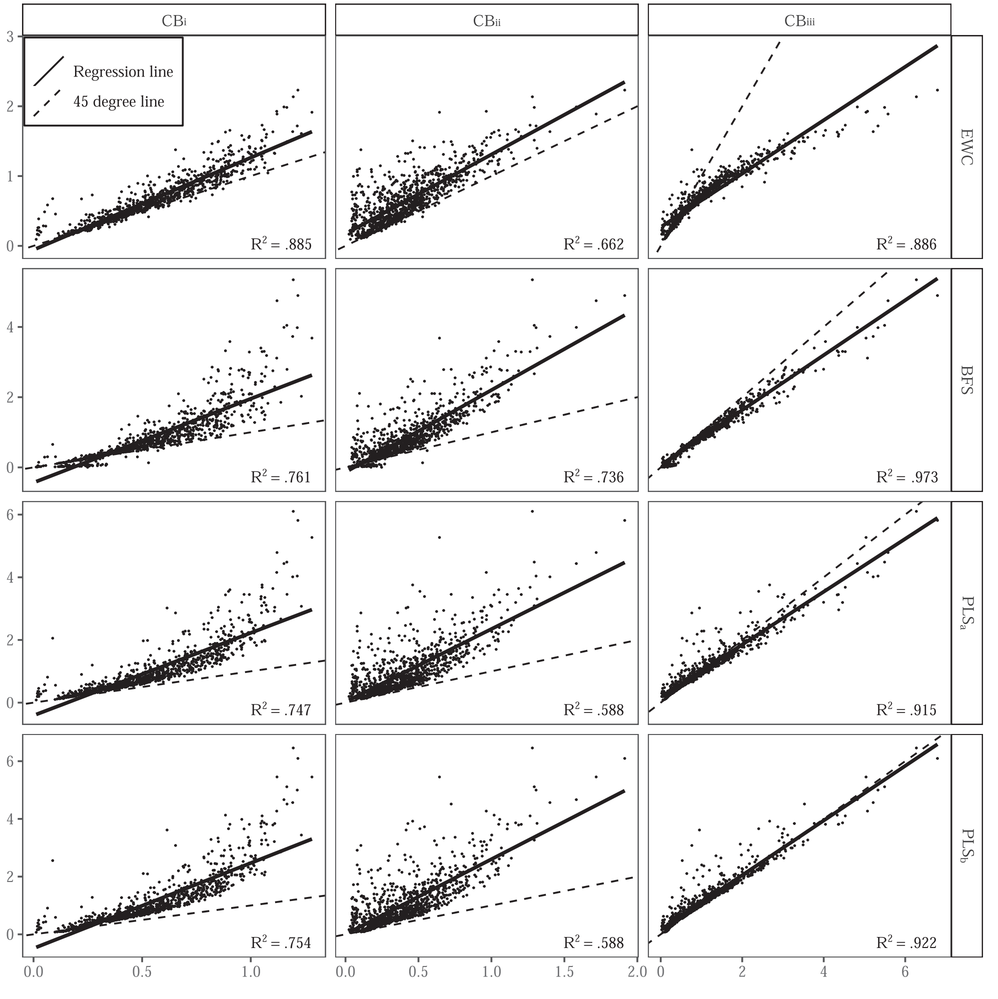

FIGURE 1 Comparison of the empirical signal-to-noise ratio across different methods.

ysis with composites. Specifically, we considered the methods Yuan and Fang (2023) studied, that is, CB^{i}, EWC, BFS, PLS_a, and PLS_b, and in addition also CB-SEM using the two other scaling methods described previously, that is, CB^{ii} and CB^{iii}. Note that for regression analysis with composites the weights are estimated.[2] To generate the datasets, we used the 1000 sets of population parameters Yuan and Fang (2023) reported.[3] For each set of population parameters, 1000 samples with 200 observations were drawn from the multivariate normal distribution with mean zero and the corresponding variance-covariance matrix using the `mvrnorm()` function of the R package `MASS` (Venables & Ripley, 2002, Version 7.3.58.1). Subsequently, we applied each method to the 1000 samples corresponding to a set of population parameters and obtained the SNR as the average parameter estimate divided by the empirical standard deviation of $\sqrt{200}$ times the parameter estimate.[4] For more details on the simulation study, we refer the reader to section 3 of Yuan and Fang (2023). Figure 1 illustrates our results.

[2]Yuan and Fang (2023) additionally considered the performance of BFS and PLS_a using the population weights.
[3]https://www3.nd.edu/~kyuan/PLS-SEM/SNR/.
[4]Estimations that did not converge were removed before calculating the SNR. Very similar to Yuan and Fang (2023, tab. 1) we obtained the following convergence rates per method: CB^{i}: 99.16%; CB^{ii}: 99.78%; CB^{iii}: 100%; PLS_a: 99.99%; PLS_b: 99.71%, and BFS: 99.93%.

The first column of Figure 1 compares the empirical SNR for the four methods of regression analysis with (weighted) composites to CB^{i}, that is, CB-SEM using the scaling method Yuan and Fang (2023) proposed. This represents the situation Yuan and Fang (2023) studied, and, as expected, our results are very similar to those reported in fig. 2 of Yuan and Fang (2023). Specifically, in most of the 1000 cases, regression analysis with (weighted) composites outperforms CB-SEM in terms of the empirical SNR ($\hat{\tau}$): $\#\hat{\tau}_{\text{CB}^i} > \hat{\tau}_{\text{EWC}}$: 148; $\#\hat{\tau}_{\text{CB}^i} > \hat{\tau}_{\text{BFS}}$: 102; $\#\hat{\tau}_{\text{CB}^i} > \hat{\tau}_{\text{PLS}_b}$: 40; and $\#\hat{\tau}_{\text{CB}^i} > \hat{\tau}_{\text{PLS}_a}$: 2; for Yuan and Fang's (2023) results, see Table 2 in their article. In contrast, the third column of Figure 1 shows the results for CB^{iii}, that is, CB-SEM where the variance of the dependent and independent latent variables is fixed to 1. As can be seen, for the majority of population models CB^{iii} produces a larger empirical SNR value than EWC ($\#\hat{\tau}_{\text{CB}^{iii}} > \hat{\tau}_{\text{EWC}}$: 891) and BFS ($\#\hat{\tau}_{\text{CB}^{iii}} > \hat{\tau}_{\text{BFS}}$: 941). Considering the comparison between CB^{iii} and PLS-SEM, the SNR values are closer to the 45 degree line than in the comparison between CB^{i} and PLS-SEM. Hence, CB-SEM by NML performs more like PLS-SEM in terms of the empirical SNR using this scaling method ($\#\hat{\tau}_{\text{CB}^{iii}} > \hat{\tau}_{\text{PLS}_b}$: 218; $\#\hat{\tau}_{\text{CB}^{iii}} > \hat{\tau}_{\text{PLS}_a}$: 403).

Conclusion: The empirical SNR for CB-SEM estimated by NML depends on the scaling method. Consequently, Yuan and Fang's findings are limited and only apply to the scaling method they used.

3 | COMPARING CB-SEM WITH LATENT VARIABLES TO REGRESSION ANALYSIS WITH COMPOSITES IS LIKE COMPARING APPLES AND ORANGES

The focus of Yuan and Fang (2023, p. 24) "is to compare CB-SEM against regression analysis with weighted composites." Between the two methods, they would like to determine "[w]hich method delivers greater [SNR]?" Later in their article, Yuan and Fang (2023, p. 24) mention that the focus of their "paper is on the SNRs of the estimates of the regression coefficients when the underlying populations are generated by CB-SEM models, and the constructs are represented by either the latent variables or weighted composites in operation."

These two different perspectives on the same comparison reveal that what at first glance looks like two methods is actually a conflation of two methodological choices: on the one hand, the issue is an estimation procedure – CB-SEM estimated by NML versus regression analysis estimated by LS – and on the other hand, the issue is how to represent the theoretical constructs – as latent variables or as (weighted) composites.[5] As long as the two methodological choices are not disentangled, it is impossible to know whether the different levels of SNR are attributable to the estimation procedure or to how the theoretical constructs are represented. Hence, we see a clear case of unintended aliasing.

Figure 2 visualizes three research designs: the alleged research design of Yuan and Fang (2023), their actually employed research design, and an option for an appropriate (full-factorial) research design.

Yuan and Fang (2023) focus on comparing two methods: CB-SEM with latent variables estimated by NML and regression analysis with (weighted) composites estimated by LS. Both the title of their article ("Which method…") and later elaborations ("CB-SEM and the other four methods") indicate their interest in the differences between methods. This coincides with a research design as depicted in Figure 2a. In fact, Yuan and Fang's (2023) research design consists of two factors with two levels each: the factor *estimation procedure* with the levels (1) CB-SEM estimated by NML and (2) regression analysis estimated by LS, and the factor *representation of theoretical constructs* with the levels (1) latent variable and (2) composite. As Figure 2b shows, this research design contains two blank cells. Since one factor is never varied in isolation but only in combination with the other factor, the research design suffers from aliasing. This means that one cannot say which of the two factors is responsible for the variation in the SNR. Hence, Yuan and Fang's (2023) study

[5]One could even argue that such a comparison involves three methodological choices. First, the choice concerns how the model is estimated, that is, in one step or two. For instance, in regression analysis with weighted composites, the composites are formed in a first step, and in a second step the structural model is estimated; in contrast, in CB-SEM with latent variables, the model is estimated in one step. Second, the choice concerns which estimator to use, that is, NML or LS.

For instance, a regression analysis can also be estimated by NML. Third, the choice concerns how to represent the theoretical constructs.

(a) Alleged research design

Experimental Factor: Method	
Level 1: CB-SEM with Latent Variables by NML	Level 2: Regression Analysis with (Weighted) Composites by LS
SNR of CB-SEM with latent variables by NML	SNR of regression analysis with (weighted) composites by LS

(b) Actual research design

		Experimental Factor 1: Estimation Procedure	
		Level 1: CB-SEM by NML	Level 2: Regression Analysis by LS
Experimental Factor 2: Representation of the Theoretical Constructs	Level 1: Latent Variable	SNR of CB-SEM with latent variables by NML	
	Level 2: Composite		SNR of regression analysis with (weighted) composites by LS

(c) Appropriate research design

		Experimental Factor 1: Estimation Procedure	
		Level 1: CB-SEM by NML	Level 2: Regression Analysis by LS
Experimental Factor 2: Representation of the Theoretical Constructs	Level 1: Latent Variable	SNR of CB-SEM with latent variables by NML	SNR of regression analysis with weighted composites by LS and correction for attenuation
	Level 2: Composite	SNR of CB-SEM with (weighted) composites by NML	SNR of regression analysis with (weighted) composites by LS

FIGURE 2 Alleged, actual, and appropriate research design.

lacks internal validity. Their study is certainly not the first to suffer from aliasing in comparing CB-SEM with latent variables estimated by NML to regression analysis with (weighted) composites estimated by LS. The same things happened in many other studies, including, for example, Astrachan et al. (2014), Barroso et al. (2010), Chin and Newsted (1999), Dash and Paul (2021), Goodhue et al. (2011); (2012), Lu et al. (2011), Hwang et al. (2010), Reinartz et al. (2009), Sharma and Kim (2013), and Vilares et al. (2010). However, at the time when these studies were conducted, it was not yet clear how to fill the blank cells. In the meantime, methodological research has developed to the point where the literature now offers numerous possibilities for filling the blank cells. Consequently, there is no longer any reason to accept unintended aliasing. Figure 2c provides a suggestion for how to fill the blank cells and, hence, dissolve the aliasing in the research design: LS regression analysis combined with latent variables as representations of theoretical constructs (top right cell) and CB-SEM by NML relying on composites as representations of theoretical constructs (bottom left cell).

Several methods exist for combining LS regression analysis with latent variables as representations of theoretical constructs (e.g., Takane & Hwang, 2018). One way to obtain estimates for a latent variable model on the basis of weighted composites is to apply a correction for attenuation. Such a correction

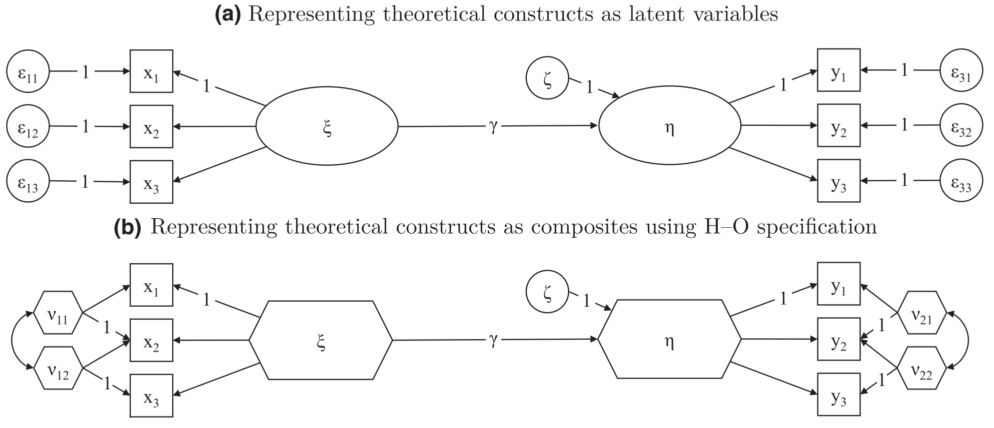

FIGURE 3 Different ways of representing theoretical concepts.

requires reliability estimates of the weighted composites. As reliability estimates, analysts can employ McDonald's ω (McDonald, 1999) for equally weighted composites, Dijkstra–Henseler's ρ_A (Dijkstra & Henseler, 2015) for composites obtained with PLS-SEM using Mode A,[6] or Raykov's r^2 (Raykov, 1997) for weighted composites in general. Use of this new cell in the research design would make it possible to investigate how the estimation procedure (CB-SEM by NML vs. regression analysis by LS) affects the SNR if theoretical constructs are represented by latent variables.

Although CB-SEM has mainly been used for latent variables, researchers have also frequently investigated its use for modeling composites (e.g., Fan, 1997; Grace & Bollen, 2008; Hancock et al., 2013; Rose et al., 2019). The recently introduced Henseler–Ogasawara (H–O) specification (Henseler, 2021; Schuberth, 2023; Yu et al., 2023) offers a way of modeling composites in CB-SEM with the same ease as modeling latent variables. An example H–O specification is shown in Figure 3b, where ξ and η depict the composites of interest. Note that the H–O specification models a composite in such a way that it accounts for the covariances between its components and other variables of the model. An alternative that specifies composites without that characteristic is the pseudo-indicator model (Rose et al., 2019).

As Figure 3 shows, the two contrasting ways of representing theoretical constructs are simply two different models. In light of the two different models depicted in Figure 3, a question that naturally arises is this: If the model in Figure 3a is the true model (as in Yuan & Fang, 2023) and a researcher would like to make a statistical inference about the coefficient γ in this model, why should he or she use the coefficient γ of the model displayed in Figure 3b for this purpose, while the latter coefficient typically deviates from the original coefficient in terms of meaning and magnitude?

Conclusion: The study by Yuan and Fang (2023) suffers from aliasing because the effect of the representation of theoretical constructs (latent variables vs. composites) cannot be disentangled from the effect of the estimation procedure (CB-SEM by NML vs. regression analysis by LS).

4 | PARAMETER VALUES IN SEM ARE NOT ARTIFICIAL

In their article, Yuan and Fang compare the SNR for an individual independent variable across CB-SEM estimated by NML and various forms of LS regression analysis with composites. To justify their comparison, they explain that the parameter values in CB-SEM depend on an arbitrarily chosen method to fix the scale of the latent variables. The same holds for the parameter values in LS regression analysis with (weighted) composites. This leads Yuan and Fang (2023, p. 2) to conclude that "the population values of

[6]Something similar can be done for Mode B; see Dijkstra (1985, chapter 2).

the model parameters under SEM are artificial" and that "bias in parameter estimates with latent variable models does not enjoy a clear substantive interpretation."

That the choice of the method to set the scale of a latent variable is arbitrary is true. However, the same is true of any other measured variable, such as length, whose scale is also arbitrary. As Markus and Borsboom (2013, p. 23) explain, "[n]ormally, the estimation of ratios happens by choosing an arbitrary level of the quantitative attribute as a unit (for instance, the meter), and estimating the ratio in which other levels of the attribute stand to that unit." Consequently, following Yuan and Fang (2023), the population parameter values among variables measured on a ratio scale would also be artificial because a researcher's choice to measure length, for example, in meters is arbitrary; thus, the population parameter values depend on arbitrarily chosen metrics of the measured variables.

In contrast, we argue that once the scale of a latent variable has been fixed, the arbitrariness of its scale disappears and the involved parameter values can be interpreted in terms of the latent variable's scale. To fix the scale of a latent variable, various approaches have been proposed, such as fixed marker scaling or effects coding scaling (e.g., Little et al., 2006). As Klopp and Klößner (2021, p. 192) explain, given that the model is correctly specified, "population and estimated parameter vectors are related by a change of scale, entailing that specific estimated parameters such as loadings and regression coefficients constitute algebraic transformations of certain population parameters"; see also tab. 9 in Klößner and Klopp (2018). Hence, although the values of the parameter estimates typically differ across scaling methods, the population parameters under SEM are not artificial. It is just that the interpretation of the parameter estimates and their relationship with the population parameters depends on the method used to fix the scale of the latent variables (see, e.g., tab. 4 in Klopp and Klößner (2021) for the interpretation of the limit estimates in CB-SEM given that the model is correctly specified). The same applies to variables measured on a ratio scale and their associated parameters. Once the metric of a measured variable has been chosen, for example, meter for length, its involved parameters can be interpreted in terms of that metric. If a different metric is chosen, the values of involved parameters will change, as the interpretation of the parameters has changed. Consequently, it is problematic to compare parameters between CB-SEM and regression analysis with composites, even if they are identical in magnitude, as they are differently interpreted.

Conclusion: The parameter values in SEM are no more artificial than the parameter values in regression analysis with measured variables. Once the latent variables have been scaled, the involved parameters can be interpreted substantively and bias in parameter estimates enjoys a substantial interpretation.

5 | DISCUSSION

According to their title, Yuan and Fang (2023) pose the question "[w]hich method delivers greater SNR: Structural equation modelling or regression analysis with weighted composites?" In the abstract, they seem to give the following answer: "Contrary to the common belief that CB-SEM is the preferred method for the analysis of observational data, this article shows that regression analysis via weighted composites yields parameter estimates with much smaller standard errors, and thus corresponds to greater values of the SNR." Based on the information they provide, those readers who only get to see the parts of the article that are not behind a paywall or who only read Yuan and Fang's (2023) introduction would conclude that regression analysis with (weighted) composites is preferable to CB-SEM. Only in the last section of their paper do Yuan and Fang emphasize the limitation of their study, namely, that they "only studied the regression relationship for a model with two constructs" (p. 25), that is, a simple regression model. In fact, this is a very strong limitation, because findings of simple regression models can often not be transferred to a multiple regression context (e.g., Goodhue et al., 2017). Given that models with only two constructs are very rare in empirical research (e.g., Henseler et al., 2014, found that in some disciplines <0.2% of the models consist of only two constructs), readers should realize that Yuan and Fang's findings are applicable only to a very few research situations – if at all.

Yuan and Fang (2023) proposed a comparison of CB-SEM with latent variables estimated by NML and regression analysis with (weighted) composites estimated by LS in terms of the SNR. They present empirical results that support the notion that LS regression analysis with (weighted) composites performs as well as or better than CB-SEM estimated by NML. In conducting such a comparison, Yuan and Fang make two assumptions. First, they claim that parameter values under CB-SEM (and LS regression with (weighted) composites) are artificial and that they do not enjoy a substantive interpretation. Consequently, bias in the parameter estimates would not have substantive meaning, either. Second, Yuan and Fang (2023, p. 2) claim that "one can always adjust the scales of either the composites or those of the latent variables to make the two sets of regression coefficients mathematically equal."

In our commentary, we show that Yuan and Fang (2023) are mistaken on several points, which limits the contribution of their findings. First, we have shown that the empirical SNR of CB-SEM estimated by NML depends on which method is used to fix the latent variables' scales, that is, the scaling method. Hence, Yuan and Fang (2023, p. 25) are wrong in that the "SNR does not depend on the scale of the variables being chosen." Second, in their design, Yuan and Fang (2023) vary two factors simultaneously: how theoretical constructs are represented, that is, as latent variables or as composites, and the estimation procedure, that is, CB-SEM by NML or regression analysis by LS. Therefore, it is not clear whether the difference in the SNR between CB-SEM with latent variables estimated by NML and regression analysis with (weighted) composites estimated by LS is due to a difference in the estimation procedure or to the construct representation. Third, we explain that Yuan and Fang's assumption about the parameter values in CB-SEM, which is necessary to properly compare CB-SEM and regression analysis with composites in terms of the SNR, is incorrect. Specifically, Yuan and Fang claim that parameter values under CB-SEM [and regression analysis with (weighted) composites] are artificial, and consequently parameter estimates and potential bias do not enjoy a clear substantive meaning. As we explained, once the scales of the latent variables are fixed, the involved parameters enjoy a substantive interpretation and are not artificial. However, the exact interpretation of the parameters depends on which scaling method is chosen.

Yuan and Fang's second assumption, made to compare CB-SEM and regression analysis with (weighted) composites, that is, "one can always adjust the scales of either the composites or those of the latent variables to make the two sets of regression coefficients mathematically equal" referring to Skrondal and Laake (2001), Yuan and Deng (2021), and Devlieger et al. (2016), should be considered with extreme caution. Although it is true that for models with one independent and one dependent latent variable, the scales of the latent variables can be adjusted to obtain the same coefficient as in regression analysis with composites, this is not necessarily the case for models involving more than one independent latent variable. Yuan and Fang (2023, p. 25) also implicitly admit this in stating that "[f]or a model with more latent variables, a single parameter [footnote omitted] $\gamma_j = 0$ under CB-SEM may not imply $\gamma_{wj} = 0$ under regression analysis with weighted composites, due to correlations among the latent constructs." If a single parameter is zero under CB-SEM but different to zero under regression analysis with (weighted) composites, the scales of the latent variables cannot be adjusted to obtain a value different to zero under CB-SEM. In the PLS-SEM context, Schuberth et al. (2022) show that it is only possible under very special circumstances, for example, if the independent latent variables are uncorrelated, to obtain a path coefficient of zero under PLS-SEM if the corresponding coefficient is equal to zero under CB-SEM and vice versa. Hence, in situations in which a path coefficient equals zero under CB-SEM while it is different to zero under PLS-SEM, neither the scale of the composites nor the scale of the latent variables can be adjusted to obtain a value of zero for this path coefficient under PLS-SEM or a value different to zero under CB-SEM, respectively. Similarly, the same holds for regression analysis with other types of factor-wise composites, that is, only the indicators belonging to a latent variable are used for calculating the corresponding composite (Skrondal & Laake, 2001, p. 572). We find an exception where composites are created block-wise, that is, all the indicators belonging to the independent latent variables and dependent latent variables, respectively, are used to calculate their composites. In particular, this happens if block-wise Bartlett factor scores are used for the dependent latent variables and regression factor scores are used for the independent latent variables (Skrondal & Laake, 2001). In this case, regression analysis with weighted composites can lead asymptotically to the same results as CB-SEM if the model is correctly specified. This is also acknowledged by Devlieger et al. (2016), who refer to this approach as a *bias avoiding method* and highlight its limitations.

Against this background, researchers should not rely on the SNR only to make a grounded choice between CB-SEM by NML and LS regression analysis with (weighted) composites. Even if they did so and accepted the drawbacks of regression analysis with composites such as inconsistency and the risk of incorrect inference, the results provided by Yuan and Fang (2023) do not allow the conclusion that LS regression analysis with (weighted) composites is preferable to CB-SEM with latent variables estimated by NML in terms of the SNR. Consequently, empirical researchers should not let their methodological choices be guided by Yuan and Fang's (2023) findings. Additionally, if analysts are interested in obtaining a high SNR value, they should consider methods that are specially designed for that purpose, such as ridge regression (Hoerl & Kennard, 1970) or the lasso (Tibshirani, 1996), which trade variance for bias. However, such an investigation has not been done and therefore awaits future research.

AUTHOR CONTRIBUTIONS

Florian Schuberth: investigation; validation; writing – review and editing. **Tamara Schamberger:** formal analysis; software; writing – original draft. **Mikko Rönkkö:** conceptualization; writing – review and editing. **Yide Liu:** conceptualization; writing – original draft. **Jörg Henseler:** conceptualization; supervision; writing – original draft.

ACKNOWLEDGEMENTS

Jörg Henseler served as a reviewer for Yuan and Fang's (2023) manuscript. He 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). We thank Hao Wu, Associate Editor of the *British Journal of Mathematical and Statistical Psychology*, for giving us the opportunity to write this commentary. Moreover, we thank Alexandra Elbakyan for her efforts in making science accessible. Finally, we thank Yves Rosseel for his support in replicating Yuan and Fang's results in lavaan.

CONFLICT OF INTEREST STATEMENT

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

DATA AVAILABILITY STATEMENT

Scripts and model specifications to replicate the analyses presented in this study are available at https://osf.io/tcnxy/.

ORCID

Florian Schuberth https://orcid.org/0000-0002-2110-9086
Tamara Schamberger https://orcid.org/0000-0002-7845-784X
Mikko Rönkkö https://orcid.org/0000-0001-7988-7609
Yide Liu https://orcid.org/0000-0001-9268-3555
Jörg Henseler https://orcid.org/0000-0002-9736-3048

REFERENCES

Antonakis, J. (2017). On doing better science: From thrill of discovery to policy implications. *The Leadership Quarterly*, *28*(1), 5–21. https://doi.org/10.1016/j.leaqua.2017.01.006

Astrachan, C. B., Patel, V. K., & Wanzenried, G. (2014). A comparative study of CB-SEM and PLS-SEM for theory development in family firm research. *Journal of Family Business Strategy*, *5*(1), 116–128. https://doi.org/10.1016/j.jfbs.2013.12.002

Barroso, C., Cepeda Carrión, G., & Roldán, J. L. (2010). Applying maximum likelihood and PLS on different sample sizes: Studies on SERVQUAL model and employee behavior model. In V. E. Vinzi, W. W. Chin, J. Henseler, & H. Wang (Eds.), *Handbook of partial least squares: Concepts, methods and applications* (pp. 427–447). Springer.

Bollen, K. A. (1989). *Structural equations with latent variables*. Wiley.

Chin, W. W., & Newsted, P. R. (1999). Structural equation modeling analysis with small samples using partial least squares. In R. H. Hoyle (Ed.), *Statistical strategies for small sample research* (pp. 307–341). Sage.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Lawrence Erlbaum Associates.

Czanner, G., Sarma, S. V., Ba, D., Eden, U. T., Wu, W., Eskandar, E., & Brown, E. N. (2015). Measuring the signal-to-noise ratio of a neuron. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(23), 7141–7146. https://doi.org/10.1073/pnas.1505545112

Czanner, G., Sarma, S. V., Eden, U. T., & Brown, E. N. (2008). A signal-to-noise ratio estimator for generalized linear model systems. *Proceedings of the World Congress on Engineering*, *2*, 1063–1069.

Dash, G., & Paul, J. (2021). CB-SEM vs PLS-SEM methods for research in social sciences and technology forecasting. *Technological Forecasting and Social Change*, *173*, 121092. https://doi.org/10.1016/j.techfore.2021.121092

Davidson, R., & MacKinnon, J. G. (1993). *Estimation and inference in econometrics*. Oxford University Press.

Devlieger, I., Mayer, A., & Rosseel, Y. (2016). Hypothesis testing using factor score regression: A comparison of four methods. *Educational and Psychological Measurement*, *76*(5), 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., & Henseler, J. (2015). 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

Fan, X. (1997). Canonical correlation analysis and structural equation modeling: What do they have in common? *Structural Equation Modeling: A Multidisciplinary Journal*, *4*(1), 65–79. https://doi.org/10.1080/10705519709540060

Gibson, D. B. (2015). Effect size as the essential statistic in developing methods for mTBI diagnosis. *Frontiers in Neurology*, *6*, 126. https://doi.org/10.3389/fneur.2015.00126

Gonzalez, R., & Griffin, D. (2001). Testing parameters in structural equation modeling: Every "one" matters. *Psychological Methods*, *6*(3), 258–269. https://doi.org/10.1037/1082-989x.6.3.258

Goodhue, D. L., Lewis, W., & Thompson, R. (2011). Measurement error in PLS, regression and CB-SEM. In *MCIS 2011 Proceedings*. Association for Information Systems. http://aisel.aisnet.org/mcis2011/4

Goodhue, D. L., Lewis, W., & Thompson, R. (2012). Does PLS have advantages for small sample size or non-normal data? *MIS Quarterly*, *36*(3), 981–1001. https://doi.org/10.2307/41703490

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*(3), 667–684. https://doi.org/10.25300/misq/2017/41.3.01

Grace, J. B., & Bollen, K. A. (2008). Representing general theoretical concepts in structural equation models: The role of composite variables. *Environmental and Ecological Statistics*, *15*(2), 191–213. https://doi.org/10.1007/s10651-007-0047-7

Greene, W. H. (2012). *Econometric analysis* (7th ed.). Pearson.

Hair, J. F., Ringle, C. M., & Sarstedt, M. (2011). PLS-SEM: Indeed a silver bullet. *Journal of Marketing Theory and Practice*, *19*(2), 139–152. https://doi.org/10.2753/MTP1069-6679190202

Hancock, G. R., Mao, X., & Kher, H. (2013). On latent growth models for composites and their constituents. *Multivariate Behavioral Research*, *48*(5), 619–638. https://doi.org/10.1080/00273171.2013.815579

Henseler, J. (2021). *Composite-based structural equation modeling: Analyzing latent and emergent variables*. Guilford Press.

Henseler, J., Dijkstra, T. K., Sarstedt, M., Ringle, C. M., Diamantopoulos, A., Straub, D. W., Ketchen, D. J., Jr., Hair, J. F., Hult, G. T. M., & Calantone, R. J. (2014). Common beliefs and reality about PLS: Comments on Rönkkö & Evermann (2013). *Organizational Research Methods*, *17*(2), 182–209. https://doi.org/10.1177/1094428114526928

Hoerl, A. E., & Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, *12*(1), 55–67. https://doi.org/10.1080/00401706.1970.10488634

Hwang, H., Malhotra, N. K., Kim, Y., Tomiuk, M. A., & Hong, S. (2010). A comparative study on parameter recovery of three approaches to structural equation modeling. *Journal of Marketing Research*, *47*(4), 699–712. https://doi.org/10.1509/jmkr.47.4.699

Jöreskog, K. G. (1970). A general method for estimating a linear structural equation system. *ETS Research Bulletin Series*, *1970*(2), i–41. https://doi.org/10.1002/j.2333-8504.1970.tb00783.x

Kieser, R., Reynisson, P., & Mulligan, T. J. (2005). Definition of signal-to-noise ratio and its critical role in split-beam measurements. *ICES Journal of Marine Science*, *62*(1), 123–130. https://doi.org/10.1016/j.icesjms.2004.09.006

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: A Multidisciplinary Journal*, *28*(2), 182–206. https://doi.org/10.1080/10705511.2020.1796673

Klößner, S., & Klopp, E. (2018). Explaining constraint interaction: How to interpret estimated model parameters under alternative scaling methods. *Structural Equation Modeling: A Multidisciplinary Journal*, *26*(1), 143–155. https://doi.org/10.1080/10705511.2018.1517356

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: A Multidisciplinary Journal*, *13*(1), 59–72. https://doi.org/10.1207/s15328007sem13013

Lu, I. R., Kwan, E., Thomas, D. R., & Cedzynski, M. (2011). Two new methods for estimating structural equation models: An illustration and a comparison with two established methods. *International Journal of Research in Marketing*, *28*(3), 258268. https://doi.org/10.1016/j.ijresmar.2011.03.006

Mardia, K. V., Kent, J. T., & Bibby, J. M. (1979). *Multivariate analysis*. Academic Press.

Markus, K. A., & Borsboom, D. (2013). *Frontiers of test validity theory: Measurement, causation, and meaning*. Routledge.
McDonald, R. P. (1996). Path analysis with composite variables. *Multivariate Behavioral Research*, *31*(2), 239–270. https://doi.org/10.1207/s15327906mbr3102_5
McDonald, R. P. (1999). *Test theory: A unified treatment*. Psychology Press.
R Core Team. (2021). *R: A language and environment for statistical computing [computer software]*. R Foundation for Statistical Computing. https://www.R-project.org/
Raykov, T. (1997). Estimation of composite reliability for congeneric measures. *Applied Psychological Measurement*, *21*(2), 173–184. https://doi.org/10.1177/01466216970212006
Reinartz, W., Haenlein, M., & Henseler, J. (2009). An empirical comparison of the efficacy of covariance-based and variance-based SEM. *International Journal of Research in Marketing*, *26*(4), 332–344. https://doi.org/10.1016/j.ijresmar.2009.08.001
Rönkkö, M. (2022). *matrixpls: Matrix-based partial least squares estimation [computer software]*. (Version 1.0.15). https://github.com/mronkko/matrixpls
Rose, N., Wagner, W., Mayer, A., & Nagengast, B. (2019). Model-based manifest and latent composite scores in structural equation models. *Collabra Psychology*, *5*(1), 9. https://doi.org/10.1525/collabra.143
Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1–36. https://doi.org/10.18637/jss.v048.i02
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. (2023). The Henseler–Ogasawara specification of composites in structural equation modeling: A tutorial. *Psychological Methods*, in print. https://doi.org/10.1037/met0000432
Schuberth, F., Rosseel, Y., Rönkkö, M., Trinchera, L., Kline, R. B., & Henseler, J. (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*, 1–7. https://doi.org/10.1080/10705511.2022.2134140
Sharma, P. N., & Kim, K. H. (2013). A comparison of PLS and ML bootstrapping techniques in SEM: A Monte Carlo study. In H. Abdi, W. W. Chin, V. E. Vinzi, G. Russolillo, & L. Trinchera (Eds.), *New perspectives in partial least squares and related methods* (pp. 201–208). Springer.
Skrondal, A., & Laake, P. (2001). Regression among factor scores. *Psychometrika*, *66*(4), 563–575. https://doi.org/10.1007/BF02296196
Taguchi, G., Chowdhury, S., & Wu, Y. (Eds.). (2005). Introduction to the signal-to-noise ratio. In *Taguchi's quality engineering handbook* (pp. 221–238). Wiley.
Takane, Y., & Hwang, H. (2018). Comparisons among several consistent estimators of structural equation models. *Behaviormetrika*, *45*, 157–188. https://doi.org/10.1007/s41237-017-0045-5
Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B*, *58*(1), 267–288. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x
Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics with S* (4th ed.). Springer.
Vilares, M. J., Almeida, M. H., & Coelho, P. S. (2010). Comparison of likelihood and PLS estimators for structural equation modeling: A simulation with customer satisfaction data. In V. E. Vinzi, W. W. Chin, J. Henseler, & H. Wang (Eds.), *Handbook of partial least squares* (pp. 289–305). Springer.
Wencheko, E. (2000). Estimation of the signal-to-noise in the linear regression model. *Statistical Papers*, *41*(3), 327–343. https://doi.org/10.1007/bf02925926
Wold, H. (1975). Path models with latent variables: The NIPALS approach. In H. Blalock, A. Aganbegian, F. Borodkin, R. Boudon, & V. Capecchi (Eds.), *Quantitative sociology* (pp. 307–357). Academic Press.
Yu, X., Schuberth, F., & Henseler, J. (2023). Specifying composites in structural equation modeling: A refinement of the Henseler–Ogasawara specification. *Statistical Analysis and Data Mining*. https://doi.org/10.1002/sam.11608
Yuan, K.-H., & Deng, L. (2021). Equivalence of partial-least-squares SEM and the methods of factor-score regression. *Structural Equation Modeling: A Multidisciplinary Journal*, *28*(4), 557–571. https://doi.org/10.1080/10705511.2021.1894940
Yuan, K.-H., & Fang, Y. (2023). Which method delivers greater signal-to-noise ratio: Structural equation modelling or regression analysis with weighted composites? *British Journal of Mathematical and Statistical Psychology*. https://doi.org/10.1111/bmsp.12293

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Schuberth, F., Schamberger, T., Rönkkö, M., Liu, Y., Henseler, J. (2023). Premature conclusions about the signal-to-noise ratio in structural equation modeling research: A commentary on Yuan and Fang (2023). *British Journal of Mathematical and Statistical Psychology*, *00*, 1–13. https://doi.org/10.1111/bmsp.12304