



# Which method is more powerful in testing the relationship of theoretical constructs? A meta comparison of structural equation modeling and path analysis with weighted composites

Lifang Deng<sup>1</sup> · Ke-Hai Yuan<sup>2,3</sup>

Accepted: 16 March 2022 / Published online: 2 June 2022  
© The Psychonomic Society, Inc. 2022

## Abstract

Structural equation modeling (SEM) has been deemed as a proper method when variables contain measurement errors. In contrast, path analysis with composite scores is preferred for prediction and diagnosis of individuals. While path analysis with composite scores has been criticized for yielding biased parameter estimates, recent literature pointed out that the population values of parameters in a latent-variable model depend on artificially assigned scales. Consequently, bias in parameter estimates is not a well-grounded concept for models involving latent constructs. This article compares path analysis with composite scores against SEM with respect to effect size and statistical power in testing the significance of the path coefficients, via the  $z$ - or  $t$ -statistics. The data come from many sources with various models that are substantively determined. Results show that SEM is not as powerful as path analysis even with equally weighted composites. However, path analysis with Bartlett-factor scores and the partial least-squares approach to SEM perform the best with respect to effect size and power.

**Keywords** Effect size · Factor scores · Partial least-squares SEM · Robust method · Measurement reliability

## Introduction

Data in social and behavioral sciences are commonly acquired through questionnaires, and constructs of interests are tapped by indicators that typically contain measurement errors. Because of its capability of modeling measurement errors, structural equation modeling (SEM) is regarded as a proper method for the analysis of such data (Bentler, 1980). While SEM can yield important information about the structural relationship of the manifest and latent variables, regression analysis using composite scores directly serves the purpose of prediction and classification of individuals. For a correctly specified model with two latent variables, a recent study showed that regression analysis via weighted composites yields larger effect size

than SEM for the regression coefficient between the two constructs (Yuan & Fang, 2021). However, substantive studies often involve more constructs and it is also hard to have a correctly specified model with real data. Because path analysis<sup>1</sup> with composite scores is a widely used method for data analysis, it is important to find out whether the method will deliver estimates with greater effect sizes than SEM in the general case. This article aims to address this issue via meta comparison. In particular, we will compare the two classes of methods with respect to effect size<sup>2</sup> and/or statistical power corresponding to the estimates of the path coefficients of the structural model, and the square root of the sample size times the effect size plays the role of non-centrality parameter in the  $z$ - or  $t$ -statistics. The datasets come from

✉ Ke-Hai Yuan  
kyuan@nd.edu

<sup>1</sup> Beihang University, Beijing, China

<sup>2</sup> University of Notre Dame, Notre Dame, IN, USA

<sup>3</sup> Nanjing University of Posts and Telecommunications, Nanjing, China

<sup>1</sup>Because a composite can be a dependent variable in one equation and an independent variable in another equation, path analysis might be a more proper terminology than regression analysis, although least-squares method is used in estimating the regression coefficients with each endogenous construct.

<sup>2</sup>Parallel to the effect size  $\delta = \mu/\sigma = E(\hat{\mu})/[N\text{Var}(\hat{\mu})]^{1/2}$  in testing  $\mu = 0$  corresponding to the estimator  $\hat{\mu} = \bar{x}$ , the effect size corresponding to a regression coefficient  $\beta$  is defined as  $\tau_{\beta} = E(\hat{\beta})/[N\text{Var}(\hat{\beta})]^{1/2}$ , where  $N$  is the sample size.

textbooks, software manuals and published studies where SEM has been used in the original analyses.

Composite scores are not free of measurement errors when variables/indicators contain measurement errors. A criticism against path analysis using composite scores is that the estimates of the path coefficients are biased (Croon, 2002; Devlieger & Rosseel, 2017; Hoshino & Bentler, 2013). However, recent literature pointed out that the population values of factor loadings and path coefficients of the structural model under SEM are artificially determined (Yuan & Deng, 2021). This is because latent variables do not naturally carry units and we have to assign a scale to each latent variable under the SEM methodology. For an exogenous latent variable, this can be done by fixing the factor loading for one of its indicators at 1.0 or by fixing the variance of the latent variable at 1.0, but the choice between the two is arbitrary. The scale of an endogenous latent variable is typically aligned to one of its indicators by fixing the factor loading at 1.0. However, the choice of the indicator as well as the value 1.0 are arbitrary. Note that the population values of the factor loadings as well as those of the path coefficients change with the change of the scaling method, without affecting the population covariance matrix of the indicators. Thus, for models involving latent variables, the particular population values of parameters are neither an integrated part of the observed data nor that of the population distribution.

In addition, the scales of composite scores are also artificially assigned. For example, the total scores contain the same information as the average scores. It can be shown that for a set of given values of the path coefficients among latent variables, one can obtain identical values for these coefficients under path analysis with composite scores by adjusting the scales of the composites (e.g., Croon, 2002; Devlieger, Mayer & Rosseel, 2016), and vice versa. Consequently, bias<sup>3</sup> in the estimates by path analysis with composite scores is artificial rather than substantively grounded. Such a fact leaves more space for researchers to choose methods according to the purpose of a study. In particular, a key step of path analysis with either composites or latent variables is to test the null hypothesis on the path coefficients or to confirm the role of the theoretical predictors. We should choose an approach that delivers the greatest effect size or signal-to-noise ratio in order to conduct this step most efficiently.

<sup>3</sup>Bias is defined as the difference between the expected value of an estimate and the true population value, the value of bias becomes artificial when the “true population” value is artificial. We can always make the composites estimates unbiased by rescaling the composites. For example, with  $\hat{\eta}$  and  $\hat{\xi}$  being the composites, regression coefficient  $\gamma$  can be rescaled to  $\gamma_* = c\gamma$  via the simple algebra  $\hat{\eta} = \gamma\hat{\xi} + \zeta = (c\gamma)(\hat{\xi}/c) + \zeta = \gamma_*\hat{\xi}_* + \zeta$ , and we can always choose a value of  $c$  so that  $\gamma_*$  equals that of its SEM counterpart.

Both path analysis with composite scores and SEM can be used for the purpose of prediction, but they are from different perspectives and for different purposes. The predictive relationship under SEM is among latent variables that represent the population distribution, and all individuals are equivalent under such a relationship. In practice, when observed scores are used for prediction or diagnosis, individuals are no longer equivalent. A person with greater test scores is expected to perform better on the criterion variable, and such a relationship is directly characterized by the path coefficients under path analysis with the composite scores. In addition, even when predictors contain measurement errors, regression model with the LS estimates still yields the best linear unbiased predictor for a future value (see Fuller, 1987, p. 75). Consequently, the issue of whether path analysis with composite scores corresponds to larger signal-to-noise ratios than SEM deserves further study. If the answer is yes, then path analysis with composites not only enjoys more prediction accuracy but also facilitates more efficient selection of predictors. Note that it is the LS estimates of the regression coefficients that yield a predicted value of the outcome variable with the smallest mean-square error (MSE), not parameter estimates under SEM or bias-corrected parameter estimates (Fuller, 1987). We will further discuss the issue in the concluding section.

We may believe that SEM will yield greater signal-to-noise ratios for the path coefficients of the structural model since the effect of measurement errors is removed via the measurement model. However, this is not the case for models with two-latent variables according to the results in Yuan and Fang (2021). With more latent variables, the issues become rather complicated. This is because both measurement errors and correlations among the composites affect both the values of the regression coefficients and those of the standard deviations of their estimates (Buonaccorsi, 2010, p. 109). Both measurement and structural models might be misspecified, which also affect the signal-to-noise ratios of the parameter estimates under both SEM and path analysis with composite scores. It is impossible to quantify the combined effect of these sources on parameter estimates analytically. The Monte Carlo method also becomes impractical because the location and size of model misspecification are typically unknown, and each empirical study tends to have its own features in both model and data. Meta comparison is consequently a viable and informative approach for us to find out the strength of different methods in practice. While the distribution of the data/sample also affects the performances of different methods, we will use a robust technique so that the effects of nonnormally distributed data or data contamination on both parameter estimates and their standard errors are properly regulated.

In the following section, we will review the properties of several weighted composites, including those under the partial least-squares approach to SEM. A robust-transformation technique will be described in a subsequent section, and parameter estimates by both path analysis with composite scores and SEM become more efficient when applying normal-distribution-based methods to the transformed data. Datasets, models, and results will be presented in a follow-up section. Recommendation and discussion will be provided in the concluding section. An [Appendix](#) contains the information on statistical packages and programs that are used to carry out the procedures described in this article.

## Weighted composites

There are an infinite number of ways to formulate composite scores. We can only study a few of them, including the sum score and those possessing certain optimal properties. They are briefly described in this section.

The most widely used composite is the sum score or the simple average, and we will refer to it as the equally weighted composite (EWC). EWCs are easy to formulate but they do not use any statistical/psychometric information of the indicators/variables. It is known that both the path coefficients and  $R^2$  values under path analysis are affected by measurement reliabilities of the composites. More reliable composites correspond to more efficient estimates and greater  $R^2$  values, but an EWC can be less reliable than a single indicator in the composition of the EWC. That is, the sum score ( $x_1 + x_2 + \dots + x_m$ ) can be less reliable than an individual score  $x_j$ .

The Bartlett- and regression-factor scores are among the most widely used composites for subscales (see e.g., Bartholomew, 2009; Lawley & Maxwell, 1971; Schuster & Lubbe, 2020). While the literature has demonstrated that Bartlett-factor scores possess different statistical properties from those of the regression-factor scores (Croon, 2002; Devlieger et al., 2016), Yuan and Deng (2021) showed that their differences in path analysis are only for the values of the estimated path coefficients. They also showed that, for regression analysis, the two types of factor scores yield identical standardized path coefficients, identical effect size for testing the significance of the path coefficients, and identical  $R^2$  value. In addition, the two types of factor scores possess the property of attaining the maximum reliability among all weighted composites (e.g., Yuan & Bentler, 2002). Thus, Bartlett-factor scores are equivalent to regression-factor scores for the purpose of path analysis, and we will only need to include one of them in our comparison of weighted-composite path analysis against SEM.

A method that systematically generates weighted composites is the partial least-squares approach to structural equation modeling (PLS-SEM), which has established its identity in many disciplines. Systematic introductions to PLS-SEM can be found in Wold (1980), Hair, Hult, Ringle, and Sarstedt, (2017) and Henseler (2021). We will briefly review the key elements of this method to make our presentation self-contained. To distinguish the commonly used SEM methodology from PLS-SEM, the former is referred to as covariance-based SEM (CB-SEM). PLS-SEM is essentially path analysis with weighted composites via the method of least-squares (LS) regression, and the method consists of two stages. Weighted composites are computed in the first stage, and path analysis with the obtained composites is conducted in the second stage.

Under PLS-SEM, indicators are unidimensional, that is, each indicator only loads on a single construct. The indicators for a particular construct form a block. The first stage of PLS-SEM offers two ways to compute weighted composites. One is termed as mode A and the other is termed as mode B (Wold, 1980, 1982). Both are computed by LS regression iteratively via the so-called environmental variables, which are temporary variables that are derived according to the model structure. Suppose the model has  $m$  constructs, and a construct  $\eta$  is directly connected<sup>4</sup> to other  $q$  constructs out of the remaining  $m - 1$ . Let the block-wise average scores be the starting values for the composites, and the composite of  $\eta$  is denoted as  $\tilde{\eta}$ . Then the environment variable of  $\eta$  (denoted as  $\eta_{ev}$ ) is defined as a linear combination of the  $q$  composites whose latent-variable counterparts are directly connected to  $\eta$ . The coefficients of the  $q$  composites in  $\eta_{ev}$  can be either the sign of the correlation between  $\tilde{\eta}$  and the directly connected construct (termed as the centroid scheme) or the value of the correlation itself (termed as the factorial scheme, Henseler, 2021, pp. 90–91). Let  $y_j$  ( $j = 1, 2, \dots, p_\eta$ ) be the block of indicators of  $\eta$ . Under mode A, the weight  $w_j$  in the composite  $\tilde{\eta} = w_1y_1 + w_2y_2 + \dots + w_{p_\eta}y_{p_\eta}$  is updated iteratively by the slope parameter of LS regression of  $y_j$  on the environmental variable  $\eta_{ev}$ . Under mode B, the weights ( $w_1, w_2, \dots, w_{p_\eta}$ ) in the composite  $\tilde{\eta}$  are updated iteratively by the coefficients of multiple regression of  $\eta_{ev}$  on  $y_j$ , ( $j = 1, 2, \dots, p_\eta$ ). The iteration alternates across the  $m$  blocks of indicators until all the weights are stabilized, and the  $m$  composites are consequently obtained. In our analysis of the empirical data, we will use the centroid scheme in the formulation of  $\eta_{ev}$ , while analytical studies by

<sup>4</sup>A direct connection to  $\eta$  is defined as a latent variable that either directly predicts  $\eta$  or is directly predicted by  $\eta$ . A covariance between two exogenous latent constructs is not regarded as a direct connection in PLS-SEM.

Dijkstra (1983) and Schneeweiss (1993) imply that different schemes yield essentially the same results.

In the SEM literature, a latent variable can either directly predict or being predicted by a set of manifest variables. The manifest variables being directly predicted by a latent variables are called reflective/effect indicators, and those that directly predict a latent variables are called formative/causal indicators. Starting from Wold (1980, 1982), mode A has been recommended to compute the weights for composites of reflective indicators and mode B to compute the weights for composites of formative indicators. While such recommendations are intuitive, Dijkstra and Henseler (2015) noted that mode A converges faster and yields numerically more stable weights than mode B, but the composite obtained by mode A can be less reliable than the corresponding EWC or a single indicator of the block (Yuan, Wen & Tang, 2020). When applying mode B to models with reflective indicators, Yuan and Deng (2021) showed that, at the level of population, the obtained composites are equivalent to the Bartlett-factor scores, which are known to have the maximum reliability among all weighted composites. However, results in Yuan and Deng (2021) showed that sampling and/or specification errors may cause mode B to yield negative weights even when the items are positively scored, while all the weights by mode A as well as all the factor loadings under CB-SEM are positive for the same dataset. They consequently proposed a procedure to transform mode A to mode B according to the structure of a one-factor model for each block of indicators. They also showed that the weighted composites following the transformed mode (denoted as  $B_A$ ) enjoy the same statistical properties as those directly obtained under mode B. We will use the mode  $B_A$  instead of mode B in the following data analyses because the weights under mode  $B_A$  enjoy the numerical stability of those under mode A and the resulting composites enjoy the statistical properties of those under mode B.

Note that the equivalence of PLS-SEM mode B, mode  $B_A$  and regression analysis with the Bartlett-factor scores is for correctly specified models, while the equivalence between the Bartlett- and regression-factor scores in conducting regression analysis does not need a correctly specified model as long as the two types of factor scores are computed based on the same model (Yuan & Deng, 2021). Since it is unlikely to have a correctly specified model with real data, we will include four types of composites in our comparison of path analysis with weighted composites against CB-SEM. They are equally weighted composites, Bartlett-factor scores, composites by PLS-SEM mode A, and those by PLS-SEM mode  $B_A$ . So there are a total of five different ways in estimating the coefficients of the structural model, one is CB-SEM and the other four are path analyses with different composites.

## Robust transformation with real data

This section describes a robust transformation technique for more efficient parameter estimates with real data. Most classical methods of multivariate analysis are developed for normally distributed data. They are not for real data. The so-called full information maximum likelihood (FIML) method for CB-SEM is normal-distribution-based maximum likelihood (NML). It becomes a pseudo maximum likelihood method unless the true population distribution of the sample is multivariate normal. With real data, a robust method can easily perform better than the pseudo NML method with respect to type I error control and statistical power, as will be further discussed in this section.

Data in social sciences seldom follow normal distributions. Such a fact has been repeatedly confirmed by meta analyses (e.g., Blanca, Arnau, López-Montiel, Bono, & Bendayan, 2013; Cain, Zhang, & Yuan, 2017; Micceri, 1989). However, CB-SEM and path analysis are routinely conducted under the normality assumption. With real data typically having heavier tails than that of the normal distribution, results by NML are most likely misleading or inefficient at least. In particular, the observed significance of the likelihood ratio statistic  $T_{ml}$  for the overall model structure under CB-SEM is partly due to the extra kurtoses of the data in addition to an imperfect model (Hu, Bentler & Kano, 1992), and the true standard errors of the parameter estimates by NML increase with the size of the population kurtoses (Bentler, 1983; Browne, 1982). For such data, robust methods<sup>5</sup> by accounting for the distribution shape of the data not only yield more efficient parameter estimates but also more reliable test statistics and fit indices (Yuan & Bentler, 1998; Yuan & Zhong, 2013). Robust M-estimation is a class of such methods under which the contribution of each observation to the final estimate is rationalized via scalar-valued weights. Cases sitting further away from the center of the data will get smaller weights (Huber, 1981). Yuan, Chan, and Bentler (2000) showed that robust M-estimation can be carried out via a transformation procedure. The sample means and sample covariance matrix of the transformed data are the robust estimates of the population means and covariances, respectively. To use this transformation, the researcher only needs to choose a cutoff value to determine the percent of the outlying cases whose

<sup>5</sup>The robust method here is to compute the M-estimates of the means and covariances first, and these are subsequently fitted by the structural models (Yuan & Bentler, 1998). The method is different from applying a mean-rescaling to the likelihood ratio statistic  $T_{ml}$  and using sandwich-type standard errors following the method of the normal-distribution-based maximum likelihood (Satorra & Bentler, 1994).

effects on the estimated parameters are regulated. In particular, with Huber-type robust M-estimator (Huber, 1981), the cutoff value can be chosen via the quantile of a chi-square distribution. For example, with  $p = 10$  variables and we choose the 95% quantile  $c_{.95}$  of  $\chi_{10}^2$  as the cutoff value, cases sitting further from the center of the data than  $c_{.95}$  will be assigned smaller weights in estimating<sup>6</sup> the means and variances-covariances. The deviation of each case from the center of the data is computed by the Mahalanobis distance, with means and precision matrix being computed iteratively using the robust estimates. The further a case is from the center, the smaller the weights will become so that the contribution/influence of the case on the resulting parameter estimates is regulated. Note that, unless certain values of the observed variables for a case are huge, the weights are not zero so that the case is still counted in the estimation process. Cases with extreme values correspond to tiny weights, which are approximately equivalent to being removed from the analysis.

Following the logic that the sample covariance matrix is most efficient only when data are normally distributed, Yuan et al. (2000) proposed using Mardia's (1970) measure of multivariate kurtosis to facilitate the choice of the cutoff value in conducting robust transformations. They recommended to start with the quantile  $c_{.95}$  as the cutoff value. If the standardized multivariate kurtosis  $M_s$  of the corresponding sample is not significantly different from that of a multivariate normal distribution, then the resulting transformed data will be obtained according to this weighting rule. If  $M_s$  is still significantly different from that of a multivariate normal distribution, then choosing  $c_{.90}$ ,  $c_{.85}$ , ... as the cutoff values for weights until the resulting  $M_s$  is not statistically significant. R and SAS code for conducting the robust transformation can be downloaded at [https://www3.nd.edu/~kyuan/overview\\_robustMethod/](https://www3.nd.edu/~kyuan/overview_robustMethod/) and their use is documented in Yuan and Gomer (2021). Existing studies indicate that the results following a robust transformation are much more reliable in the analyses of real data (e.g., Yuan et al., 2000). We denote the robust transformation using Huber-type weights as the  $H(\alpha)$ -transformation, where  $\alpha$  indicates that the quantile  $c_{1-\alpha}$  of  $\chi_p^2$  is used as the cutoff value for regulating cases.

The scenario for robust M-estimation to work best is for samples whose populations follow elliptically distributions (Maronna, Martin, Yohai, & Salibián-Barrera, 2019). However, outliers or data contamination can cause the observed data skewed. It is difficult to distinguish whether

the skewness in the sample is due to a skewed underlying population distribution or data contamination. Results with real data in Yuan et al. (2000) and Yuan and Bentler (1998) indicate that robust methods yield more reliable model evaluation than NML although the datasets are not symmetrically distributed. Monte Carlo results in Yuan, Chan, and Tian (2016) imply that robust methods yield more accurate parameter estimates than NML for asymmetric distributions with heavier tails than that of the normal distribution. The discussions by Huber (1981, p. 172) and Hampel, Ronchetti, Rousseeuw, and Stahel (1986, pp. 401–402) imply that robust methods are still preferred even when the underlying population distribution is skewed.

Note that raw data are needed for the application of robust transformation. With raw data, we can actually check the normality assumption via the measure of multivariate kurtosis (Mardia, 1970; Cain et al., 2017). While sample skewnesses also reflect departure from normality, it has been pointed out that, without mean structures, results of path analysis and CB-SEM are asymptotically unaffected by skewness of the underlying population distribution (Yuan & Bentler, 2000). Consequently, we do not pay attention to skewness of the observed data in the following section. When only sample means and covariances are available, we will have to assume that the raw data follow a multivariate normal distribution in order to proceed with the analysis, as will be done in the following section.

Like CB-SEM, path analysis with weighted composites can also be conducted using the NML method. Because PLS-SEM uses LS regression to estimate the path coefficients of the structural model, we will also use LS regression in estimating the path coefficients for the method of path analyses with the Bartlett-factor scores and equally weighted composites.

## Data and modeling results

In this section, we compare path analyses with weighted composites against CB-SEM in estimating the coefficients of the structural model. Nine real datasets and 11 models are used in the comparison. These datasets and models were not chosen to obtain particular results, nor were they selected so that a model has to fit a dataset well enough. Rather, they are from SEM textbooks, software demonstrations and authors who were willing to share their raw data with us. These datasets represent different areas where theoretical constructs are modeled for studies of substantive interests. To save space, we will only include the results of the path coefficients of the structural model. Estimates of other model parameters under CB-SEM are reported in supplementary material ([https://www3.nd.edu/~kyuan/SEM\\_PAcomparison/CB\\_SEMtables.pdf](https://www3.nd.edu/~kyuan/SEM_PAcomparison/CB_SEMtables.pdf)). Note that the estimated factor loadings ( $\lambda$ )

<sup>6</sup>All the parameters in CB-SEM or other statistical models can be directly estimated by a robust method (see Yuan & Gomer, 2021). Robust transformation is aligned with the saturated model under which parameters are means and variances-covariances.

and error variances ( $\psi$ ) in the supplementary material are used to compute the Bartlett-factor scores that are used in path analysis.

Each of the 11 models in this section is represented by a path diagram. The path coefficients in each path diagram are labeled using LISREL (Jöreskog & Sörbom, 1993) notation under which  $\gamma$  is used for paths from exogenous variables to endogenous variables, and  $\beta$  is used for paths between endogenous variables. In the presentation we will report the estimates of the path coefficients, their SEs and the corresponding  $z$ -statistics. Because an empirical effect size is just the value of a  $z$ -statistic divided by  $\sqrt{N}$ , we will not separately report the values of the empirical effect size.

### Dataset 1

The first dataset (sample correlations and standard deviations) was originally presented in Wheaton, Muthén, Alwin, and Summers (1977), who studied stability of alienation as predicted by socioeconomic status. This dataset has been used to illustrate the CB-SEM methodology and its applications in textbooks and software manuals (e.g., Bentler, 2006; Jöreskog & Sörbom, 1993). The dataset has  $p = 6$  variables and  $N = 932$  cases. The six variables are anomie 1967, powerlessness 1967, anomie 1971, powerlessness 1971, education 1966, and socioeconomic index (SEI) 1966. Several models have been used to fit this dataset in the literature. We will consider the mediation model in Fig. 1 (the same as the model in Figure 3.3 of Joreskog and Sorbom, 1993, p. 88), which contains three path coefficients:  $\gamma_{11}$ ,  $\gamma_{21}$ ,  $\beta_{21}$ . An alternative model (Figure 2.11 of Bentler, 2006, p. 39) is to have correlated errors between anomie 1967 and anomie 1971, and between powerlessness 1967 and powerlessness 1971. We do not consider the model with

error-covariances because they empirically share the same effect with the values of the path coefficients. In particular, path analysis with weighted composites accounts for all associations through path coefficients. Including error-covariances under CB-SEM does not yield a fair comparison between the two classes of methods.

Because only the sample covariance matrix is available, we will assume that the population distribution of the six variables is multivariate normal. Fitting the model in Fig. 1 to Dataset 1 by NML results in  $T_{ml} = 71.470$  under CB-SEM. The  $p$ -value is essentially zero when compared  $T_{ml}$  against  $\chi^2_6$ . The corresponding value of the root mean square error of approximation (RMSEA, Steiger & Lind, 1980) is .108, which also indicates that the model fit is not adequate. However, the comparative fit index (CFI = .969) implies that the model fits the data reasonably well (Bentler, 1990; Hu & Bentler, 1999). Such an inconsistency between different fit indices is not a surprise since they evaluate the model from different perspectives (Lai & Green, 2016).

Table 1 contains the estimates of the three path coefficients, their standard errors (SEs) and the corresponding  $z$ -statistics. As we noted earlier that the estimates of the coefficients are not directly comparable, since they depend on the scales of the predictors as well as that of the outcome variable. While the  $z$ -statistics following path analyses with the four types of composites are rather close, CB-SEM yields uniformly the smallest  $z$ -statistics for each of the coefficients. In particular, the  $z$ -statistic for  $\gamma_{21}$  under CB-SEM is close to half of those under path analyses with weighted composites. The results in Table 1 indicate that PLS-SEM mode B<sub>A</sub> yields the largest  $z$ -statistic for  $\gamma_{11}$ , while path analysis with EWCs yields the largest  $z$ -statistic for  $\gamma_{21}$ , and PLS-SEM mode A corresponds to the largest  $z$ -statistic for  $\beta_{21}$ . Because, for correctly specified models and conditional on the population weights, path analysis

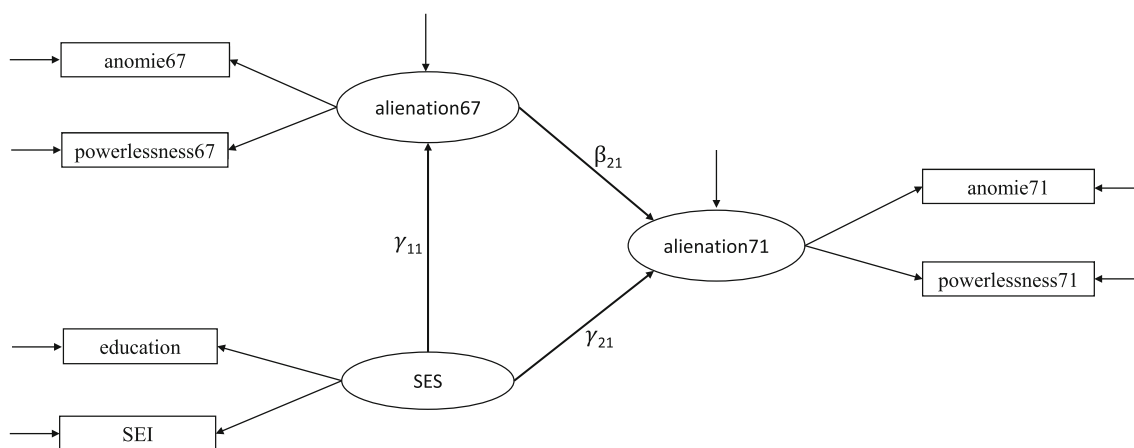


Fig. 1 A model for stability of alienation (Bentler, 2006; Jöreskog & Sörbom, 1993; Wheaton et al., 1977;  $N = 932$ )

**Table 1** Parameter estimates (est), their SEs (se) and z-statistics for the path coefficients in Fig. 1 ( $p = 6$ ,  $N = 932$ ,  $T_{ml} = 71.470$ ,  $df = 6$ ,  $p$ -value = .000; RMSEA = .108, and CFI = .969)

$\theta$	CB-SEM			BFS-PA			EWC-PA			PLS-SEM A			PLS-SEM B <sub>A</sub>		
	est	se	z	est	se	z	est	se	z	est	se	z	est	se	z
$\gamma_{11}$	-1.585	0.120	-13.219	-1.188	0.080	-14.837	-1.166	0.080	-14.628	-0.435	0.030	-14.727	-0.438	0.029	-14.874
$\gamma_{21}$	-0.450	0.136	-3.303	-0.522	0.082	-6.374	-0.523	0.081	-6.443	-0.183	0.029	-6.414	-0.183	0.029	-6.389
$\beta_{21}$	0.705	0.054	13.163	0.549	0.030	18.175	0.548	0.030	18.209	0.520	0.029	18.227	0.518	0.029	18.109

BFS = Bartlett-factor score, EWC = equally weighted composite, PA = path analysis

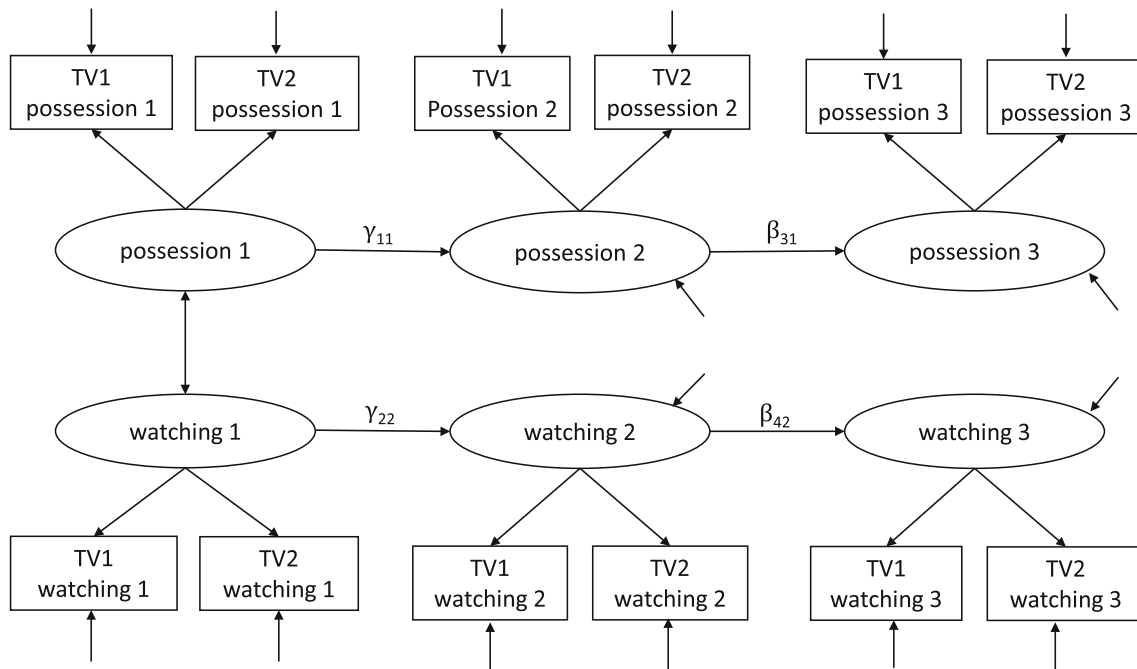
using the Bartlett-factor scores and PLS-SEM mode B<sub>A</sub> are equivalent, the observed differences between the results of the two methods in Table 1 can be due to model misspecification and sampling errors.

## Dataset 2

The second dataset (sample covariance matrix) is from Table 7.4 of the LISREL manual (Jöreskog & Sörbom, 1993, p. 193), which was from Wiley and Hornik (1973) who studied communication processes with data collected from El Salvador. The study consists of two constructs, television watching by children and television possession by their family. Each construct was taped by two congeneric measures at three points in time. With a total of  $p = 12$  manifest variables, the sample covariance matrix is from a dataset with  $N = 189$  participants. The path diagram

presented in Fig. 2 is for a model given by Joreskog and Sorbom (1993, p. 194). It is an auto regression model for each construct, and the two constructs at time 1 are correlated. There are four path coefficients:  $\gamma_{11}$ ,  $\gamma_{22}$ ,  $\beta_{31}$  and  $\beta_{42}$ . Again, we will regard the raw data as being normally distributed in the analysis.

Fitting the model in Fig. 2 to the television dataset by NML yields  $T_{ml} = 185.378$ , and the p-value is essentially 0 according to  $T_{ml} \sim \chi_{49}^2$ . The corresponding RMSEA = .122 also indicates that the model fit is poor, but, with CFI = .938, the model can be regarded as acceptable in real data analysis. The estimates of the four coefficients by the five methods are given in Table 2. Again, the z-statistics under CB-SEM are uniformly the smallest. While PLS-SEM mode B<sub>A</sub> yields uniformly the largest values of the z-statistic, the results under path analyses by the four types of weighted composites are close.

**Fig. 2** A longitudinal model for two versions of television possession and watching at three occasions (Jöreskog & Sörbom, 1993; Wiley & Hornik, 1973;  $N = 189$ )

**Table 2** Parameter estimates (est), their SEs (se) and z-statistics for the path coefficients in Fig. 2 ( $p = 12, N = 189, T_{ml} = 185.378, df = 49, p\text{-value} = 0; RMSEA = .122, \text{ and } CFI = .938$ )

$\theta$	CB-SEM			BFS-PA			EWC-PA			PLS-SEM A			PLS-SEM B <sub>A</sub>		
	est	se	z	est	se	z	est	se	z	est	se	z	est	se	z
$\gamma_{11}$	0.421	0.029	14.282	0.399	0.019	20.946	0.398	0.019	20.868	0.837	0.040	20.877	0.839	0.040	21.048
$\gamma_{22}$	1.123	0.106	10.574	0.876	0.068	12.933	0.880	0.066	13.337	0.696	0.053	13.238	0.698	0.052	13.338
$\beta_{31}$	0.881	0.047	18.906	0.846	0.042	20.385	0.846	0.042	20.272	0.829	0.041	20.283	0.830	0.041	20.361
$\beta_{42}$	0.936	0.095	9.827	0.682	0.056	12.246	0.686	0.056	12.205	0.667	0.054	12.249	0.667	0.054	12.254

BFS = Bartlett-factor score, EWC = equally weighted composite, PA = path analysis

**Dataset 3**

The third dataset is from Table 10.1 of Schumacker and Lomax (2010, p. 202), which is for a study of home resource and educational achievement. With  $p = 9$  variables, the sample covariance matrix is based on a sample of size  $N = 200$ . The model, represented by Fig. 3, is as given in Figure 10.1 of Schumacker and Lomax (2010, p. 196), which implies that the effects of both home resource and ability on achievement are mediated by aspiration. The model contains five path coefficients:  $\gamma_{11}, \gamma_{21}, \gamma_{12}, \gamma_{22}, \beta_{21}$ . Because only the sample covariance matrix is available, we regard the raw data as being normally distributed in the analysis.

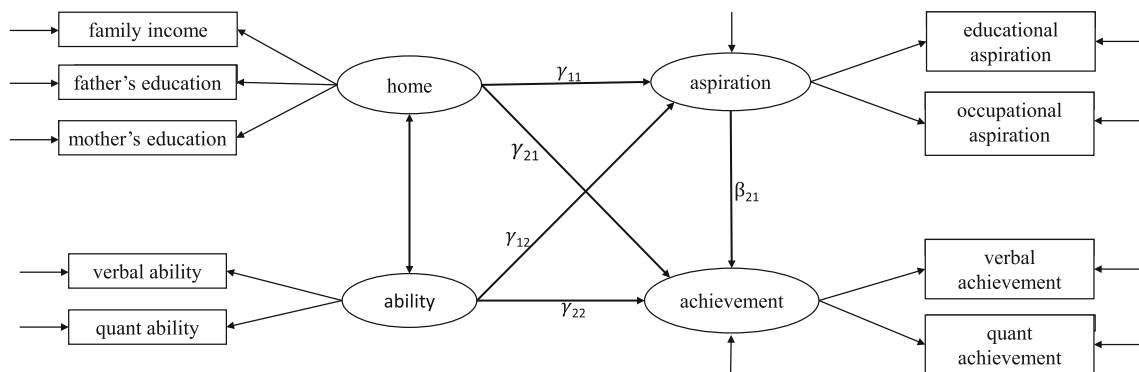
Fitting the model in Fig. 3 to the educational-achievement data by NML results in  $T_{ml} = 57.167$ , which corresponds to  $p\text{-value} = 3.39 \times 10^{-5}$  when referred to  $\chi^2_{21}$ . With  $CFI = .974$  and  $RMSEA = .093$ , the model might be regarded as adequate in practice. Table 3 contains the estimates of the five path coefficients by each of the five methods. While the z-statistics following path analyses are comparable across the four different weighted composites, those following CB-SEM are uniformly the smallest. In particular, the estimate of  $\gamma_{21}$  under CB-SEM is not statistically significant ( $z = 1.900$ ) at the level of .05. The smallest z-statistic for  $\gamma_{21}$  by path analyses is 3.959 and with the

Bartlett-factor scores. Again, the largest z-statistics for the five parameters are either by PLS-SEM mode B<sub>A</sub> ( $\gamma_{11}, \gamma_{21}$ ) or by path analyses using the Bartlett-factor scores ( $\gamma_{12}, \gamma_{22}, \beta_{21}$ ).

Note that an estimate that is not statistically significant at the level of .05 does not imply that its population counterpart is 0. Many factors contribute to effect size and power in statistical modeling, as was discussed earlier in this article.

**Dataset 4**

The fourth dataset is from Table 8.1 of Kline (1998, p. 254), which originated from Worland, Weeks, Janes, and Strock (1984). The authors studied the effect of familial risk on academic performance and classroom adjustment with junior and senior high school students. Figure 4 contains the path diagram for the model as presented on page 255 of Kline (1998), which suggests that the effect of familial risk on classroom adjustment is completely mediated by cognitive ability and scholastic achievement. With one exogenous and three endogenous constructs, the model contains three path coefficients:  $\gamma_{11}, \beta_{21}, \text{ and } \beta_{32}$ . However, only a 12-variable correlation matrix based on a sample of size  $N = 158$  is available according to Kline (1998). We will treat the sample correlation matrix as a



**Fig. 3** A model of home resources and educational achievement (Schumacker & Lomax, 2010,  $N = 200$ )



**Table 3** Parameter estimates (est), their SEs (se) and  $z$ -statistics for the path coefficients in Fig. 3 ( $p = 9$ ,  $N = 200$ ,  $T_{ml} = 57.167$ ,  $df = 21$ ,  $p$ -value =  $3.39 \times 10^{-5}$ ; RMSEA = .093, and CFI = .974)

$\theta$	CB-SEM			BFS-PA			EWC-PA			PLS-SEM A			PLS-SEM B <sub>A</sub>		
	est	se	$z$	est	se	$z$	est	se	$z$	est	se	$z$	est	se	$z$
$\gamma_{11}$	0.299	0.091	3.294	0.321	0.061	5.262	0.348	0.063	5.563	0.379	0.067	5.643	0.388	0.068	5.700
$\gamma_{21}$	0.177	0.093	1.900	0.251	0.064	3.959	0.286	0.067	4.273	0.232	0.053	4.370	0.238	0.054	4.422
$\gamma_{12}$	0.480	0.094	5.135	0.395	0.060	6.619	0.372	0.061	6.081	0.397	0.067	5.915	0.391	0.068	5.748
$\gamma_{22}$	0.611	0.113	5.414	0.491	0.064	7.630	0.492	0.066	7.413	0.393	0.053	7.358	0.369	0.054	6.851
$\beta_{21}$	0.548	0.113	4.846	0.520	0.069	7.485	0.468	0.071	6.605	0.350	0.052	6.709	0.368	0.052	7.061

BFS = Bartlett-factor score, EWC = equally weighted composite, PA = path analysis

sample covariance matrix and the underlying population distribution as normally distributed in our analysis.

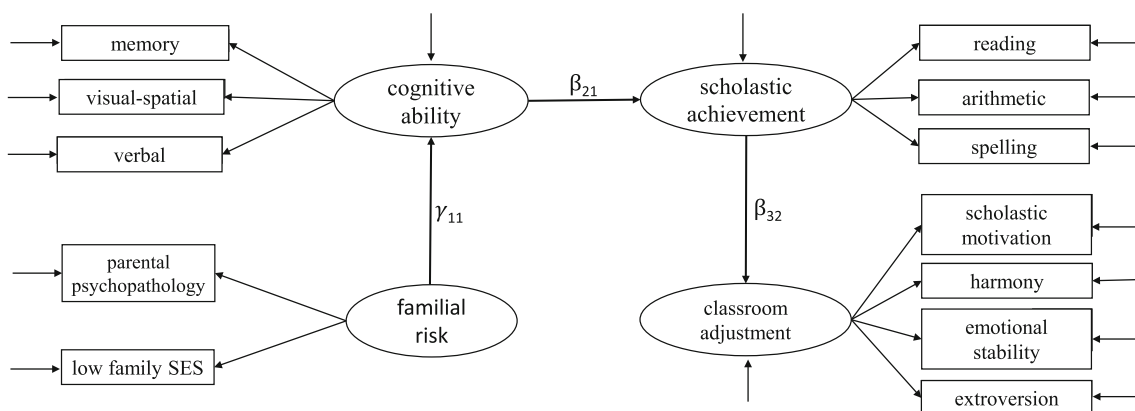
Fitting the structural model in Fig. 4 to the  $12 \times 12$  correlation matrix by NML results in  $T_{ml} = 181.424$  ( $df = 51$ , RMSEA = .128, CFI = .892). The model cannot be regarded as achieving an adequate fit according to established norms (Hu & Bentler, 1999), but it serves the purpose for comparing the different methods when the model fit is poor. Table 4 contains the estimated coefficients and the corresponding  $z$ -statistics. CB-SEM yields the smallest  $z$ -statistics for  $\gamma_{11}$  and  $\beta_{21}$  but path analysis with EWCs yields the smallest  $z$ -statistic for  $\beta_{32}$ . In contrast, PLS-SEM mode B<sub>A</sub> yields the largest  $z$ -statistics for all the three parameters across the five methods.

## Dataset 5

The fifth dataset is from Table 14.1 of Kline (2016, p. 342), which contains the sample correlations and standard deviations of  $p = 12$  variables. The data were from Houghton and Jinkerson (2007) who studied constructive thought strategies and job satisfaction, based on a sample of size  $N = 263$  full-time university employees. With three indicators for each latent variables, the 12 variables

respectively measure four constructs: 1) constructive thinking, 2) dysfunctional thinking, 3) subjective well-being, and 4) job satisfaction. Two models were estimated for this dataset in Kline (2016). The first model, as represented by only the solid arrows in Fig. 5, hypothesizes that constructive thinking reduces dysfunctional thinking, which leads to an enhanced sense of well-being, which in turn results in greater job satisfaction. This model has four path coefficients:  $\gamma_{11}$ ,  $\beta_{21}$ ,  $\beta_{31}$ , and  $\beta_{32}$ . The second model includes both the solid and dashed arrows, which hypothesizes that constructive thinking also has direct effects on sense of well-being and job satisfaction. This second model has six path coefficients:  $\gamma_{11}$ ,  $\gamma_{21}$ ,  $\gamma_{31}$ ,  $\beta_{21}$ ,  $\beta_{31}$ ,  $\beta_{32}$ .

The NML likelihood ratio statistic for the first model is  $T_{ml} = 66.061$ , corresponding to a  $p$ -value of .064 when referred to  $\chi^2_{50}$ . With RMSEA = .035 and CFI = .984, the model would be regarded as having achieved close fit in practice. Table 5 contains the parameter estimates for this model by the five different methods. Different from the previous datasets, CB-SEM yields the largest  $z$ -statistic for parameter  $\gamma_{11}$  ( $z = -1.813$ ), although it is still not significant at the level of .05. For the other three coefficients ( $\beta_{21}$ ,  $\beta_{31}$ , and  $\beta_{32}$ ), CB-SEM again yields uniformly the

**Fig. 4** A model of familial risk for psychopathology and child adjustment (Kline, 1998; Worland et al., 1984,  $N = 158$ )

**Table 4** Parameter estimates (est), their SEs (se) and z-statistics for the path coefficients in Fig. 4 ( $p = 12, N = 158, T_{ml} = 181.424, df = 51, p\text{-value} = 0; RMSEA = .128, \text{ and } CFI = .892$ )

$\theta$	CB-SEM			BFS-PA			EWC-PA			PLS-SEM A			PLS-SEM B <sub>A</sub>		
	est	se	z	est	se	z	est	se	z	est	se	z	est	se	z
$\gamma_{11}$	-0.650	0.088	-7.364	-0.406	0.048	-8.516	-0.406	0.048	-8.384	-0.560	0.066	-8.442	-0.566	0.066	-8.581
$\beta_{21}$	0.991	0.073	13.623	0.862	0.048	18.017	0.834	0.050	16.641	0.809	0.047	17.189	0.826	0.045	18.322
$\beta_{32}$	0.642	0.071	9.004	0.595	0.065	9.083	0.612	0.069	8.907	0.594	0.064	9.228	0.618	0.063	9.830

BFS = Bartlett-factor score, EWC = equally weighted composite, PA = path analysis

smallest z-statistics. In particular,  $z = -1.980$  for  $\beta_{31}$  under CB-SEM is only marginally significant, the z-statistics for this parameter by the other four methods are all larger than 2.997 in absolute value.

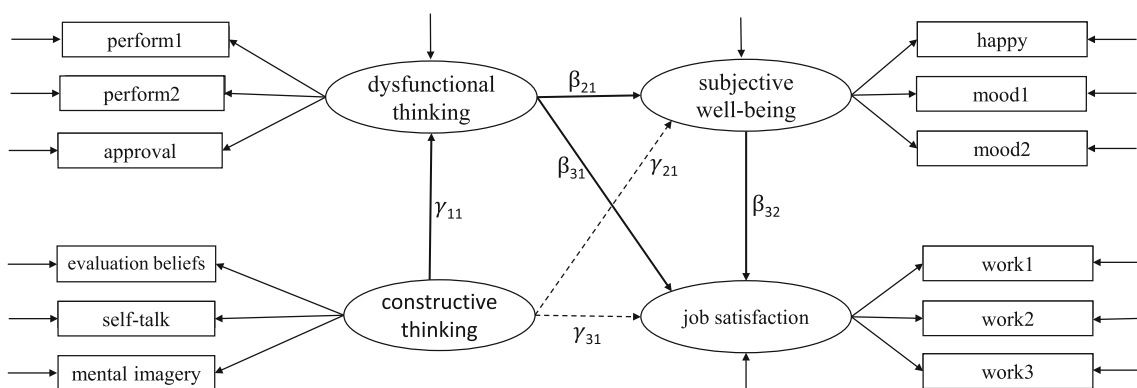
Fitting the second model to the job satisfaction data by NML yields  $T_{ml} = 62.231$ , corresponding to a p-value = .081 when referred to  $\chi^2_{48}$ . The corresponding fit indices (RMSEA = .034, CFI = .986) also imply that the model has achieved close fit in practice. Table 6 contains the estimates of the six coefficients of the structural model and their corresponding z-statistics. None of the z-statistics for the first three coefficients ( $\gamma_{11}, \gamma_{21}, \gamma_{31}$ ) is significant at the level of .05. The z-statistic for  $\beta_{31}$  under CB-SEM (-1.936) is not significant either while those by the methods of path analysis with the four types of composite scores are. Actually, for the three coefficients among the endogenous constructs ( $\beta_{21}, \beta_{31}, \beta_{32}$ ), the z-statistics under CB-SEM are uniformly the smallest.

**Dataset 6**

Weston and Gore (2006, Table 2) presented a sample covariance matrix for a dataset with  $N = 403$  cases and  $p = 12$  variables. The data were from a survey of college students who participated in a vocational psychology research project. The indicator variables were

based on a subset of items that measure social dimensions (Holland, 1997). With three indicators for each construct, the 12 variables respectively tap 1) self-efficacy beliefs, 2) outcome expectations, 3) career-related interests, and 4) occupational considerations. Weston and Gore Jr considered two structural models, as represented in Fig. 6. The model represented by the solid arrows (model 1) suggests that the effect of self-efficacy beliefs on career-related interests is partially mediated by outcome expectations, while the effect of self-efficacy beliefs on occupational considerations is completely through the two mediator variables (outcome expectations and career-related interests). This first model has four path coefficients in the structural model:  $\gamma_{11}, \gamma_{21}, \beta_{21}, \beta_{32}$ . The second model, represented by both the solid and dashed arrows in Fig. 6, suggests that the effects of self-efficacy beliefs on career-related interests and on occupational considerations are partially mediated by outcome expectations, and the effect of outcome expectations on occupational considerations is also partially mediated by career-related interests. This second model has six path coefficients in the structural model:  $\gamma_{11}, \gamma_{21}, \gamma_{31}, \beta_{21}, \beta_{31}, \beta_{32}$ .

Fitting the first model in Fig. 6 to the vocational-psychology dataset by NML results in  $T_{ml} = 416.061$ , which corresponds to a p-value that is essentially 0 when referred to  $\chi^2_{50}$ . With CFI = .913, and RMSEA = .135, the



**Fig. 5** Two models of thought strategies and job satisfaction (Houghton & Jinkerson, 2007; Kline, 2016,  $N = 263$ )

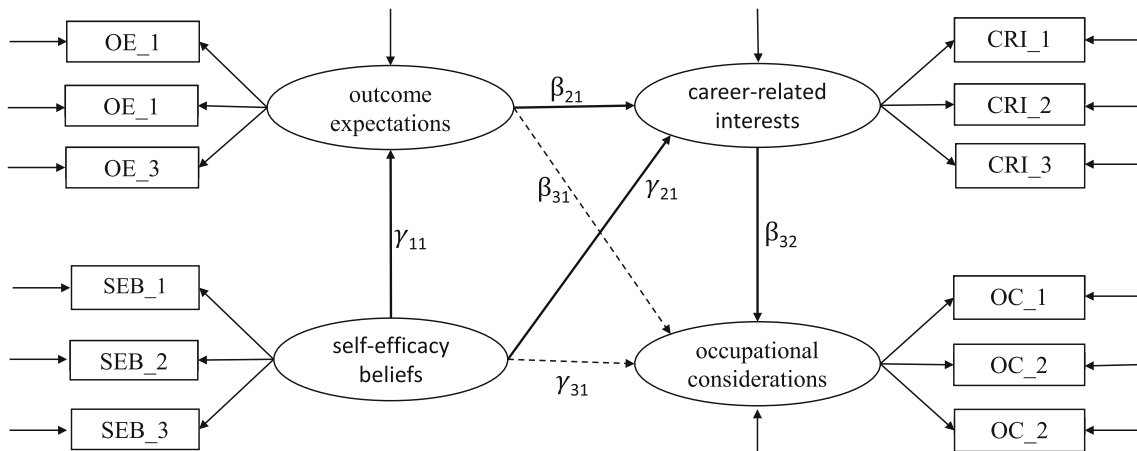
**Table 5** Parameter estimates (est), their SEs (se) and z-statistics for the path coefficients in Fig. 5, excluding the two dashed arrows ( $p = 12$ ,  $N = 263$ ,  $T_{ml} = 66.061$ ,  $df = 50$ ,  $p$ -value = .064; RMSEA = .035, and CFI = .984)

$\theta$	CB-SEM			BFS-PA			EWC-PA			PLS-SEM A			PLS-SEM B <sub>A</sub>		
	est	se	z	est	se	z	est	se	z	est	se	z	est	se	z
$\gamma_{11}$	-0.065	0.036	-1.813	-0.049	0.029	-1.706	-0.038	0.027	-1.385	-0.096	0.062	-1.557	-0.099	0.062	-1.610
$\beta_{21}$	-0.332	0.062	-5.371	-0.291	0.044	-6.586	-0.290	0.043	-6.711	-0.380	0.057	-6.641	-0.396	0.057	-6.974
$\beta_{31}$	-0.259	0.131	-1.980	-0.304	0.101	-3.010	-0.298	0.099	-3.021	-0.190	0.061	-3.104	-0.185	0.062	-2.997
$\beta_{32}$	0.907	0.220	4.116	0.640	0.131	4.877	0.638	0.131	4.881	0.299	0.061	4.881	0.301	0.062	4.892

**Table 6** Parameter estimates (est), their SEs (se) and z-statistics for the path coefficients in Fig. 5, including the two dashed arrows ( $p = 12$ ,  $N = 263$ ,  $T_{ml} = 62.231$ ,  $df = 48$ ,  $p$ -value = .081; RMSEA = .034, and CFI = .986)

$\theta$	CB-SEM			BFS-PA			EWC-PA			PLS-SEM A			PLS-SEM B <sub>A</sub>		
	est	se	z	est	se	z	est	se	z	est	se	z	est	se	z
$\gamma_{11}$	-0.061	0.036	-1.690	-0.049	0.029	-1.700	-0.038	0.027	-1.385	-0.100	0.062	-1.618	-0.103	0.062	-1.673
$\gamma_{21}$	0.036	0.026	1.387	0.031	0.020	1.496	0.011	0.019	0.573	0.100	0.057	1.743	0.095	0.057	1.655
$\gamma_{31}$	0.066	0.056	1.194	0.062	0.044	1.424	0.050	0.040	1.236	0.086	0.057	1.505	0.096	0.057	1.684
$\beta_{21}$	-0.317	0.061	-5.202	-0.281	0.044	-6.386	-0.285	0.043	-6.629	-0.367	0.057	-6.388	-0.378	0.057	-6.605
$\beta_{31}$	-0.251	0.130	-1.936	-0.296	0.101	-2.930	-0.290	0.099	-2.937	-0.186	0.061	-3.043	-0.174	0.061	-2.846
$\beta_{32}$	0.885	0.221	4.011	0.628	0.133	4.734	0.637	0.132	4.839	0.289	0.061	4.720	0.297	0.061	4.828

BFS = Bartlett-factor score, EWC = equally weighted composite, PA = path analysis



**Fig. 6** Two mediated models for self-efficacy belief on occupational considerations (Weston & Gore, 2006,  $N = 403$ )

**Table 7** Parameter estimates (est), their SEs (se) and z-statistics for the path coefficients in Fig. 6, excluding the two dashed arrows ( $p = 12$ ,  $N = 403$ ,  $T_{ml} = 416.061$ ,  $df = 50$ ,  $p$ -value = .000; RMSEA = .135, and CFI = .913)

$\theta$	CB-SEM			BFS-PA			EWC-PA			PLS-SEM A			PLS-SEM B <sub>A</sub>		
	est	se	z	est	se	z	est	se	z	est	se	z	est	se	z
$\gamma_{11}$	1.186	0.096	12.364	1.090	0.077	14.151	1.097	0.077	14.232	0.580	0.041	14.251	0.581	0.041	14.287
$\gamma_{21}$	0.046	0.009	5.161	0.038	0.009	4.047	0.037	0.010	3.797	0.187	0.048	3.906	0.206	0.047	4.364
$\beta_{21}$	0.057	0.006	10.003	0.053	0.005	10.619	0.053	0.005	10.279	0.498	0.048	10.394	0.499	0.047	10.571
$\beta_{32}$	10.368	0.822	12.615	6.480	0.392	16.523	6.044	0.396	15.246	0.614	0.039	15.593	0.649	0.038	17.091

BFS = Bartlett-factor score, EWC = equally weighted composite, PA = path analysis

**Table 8** Parameter estimates (est), their SEs (se) and z-statistics for the path coefficients in Fig. 6, including the two dashed arrows ( $p = 12$ ,  $N = 403$ ,  $T_{ml} = 361.848$ ,  $df = 48$ ,  $p$ -value = .000; RMSEA = .128, and CFI = .926)

$\theta$	CB-SEM			BFS-PA			EWC-PA			PLS-SEM A			PLS-SEM B <sub>A</sub>		
	est	se	z	est	se	z	est	se	z	est	se	z	est	se	z
$\gamma_{11}$	1.187	0.096	12.381	1.091	0.077	14.175	1.098	0.077	14.232	0.580	0.041	14.263	0.583	0.041	14.381
$\gamma_{21}$	0.039	0.011	3.462	0.038	0.009	4.065	0.037	0.010	3.797	0.187	0.048	3.904	0.206	0.047	4.347
$\gamma_{31}$	0.415	0.096	4.320	0.438	0.079	5.568	0.494	0.079	6.238	0.251	0.040	6.333	0.238	0.039	6.043
$\beta_{21}$	0.055	0.006	8.510	0.053	0.005	10.621	0.053	0.005	10.279	0.498	0.048	10.383	0.496	0.047	10.482
$\beta_{31}$	0.390	0.061	6.347	0.442	0.046	9.573	0.457	0.046	9.909	0.425	0.044	9.700	0.410	0.043	9.451
$\beta_{32}$	3.968	0.715	5.550	2.772	0.411	6.744	2.264	0.400	5.668	0.235	0.041	5.783	0.270	0.041	6.667

BFS = Bartlett-factor score, EWC = equally weighted composite, PA = path analysis

model might not be regarded as fitting the data adequately although it is substantively sound (see Weston & Gore, 2006 and references therein). Such a discrepancy between theory and goodness of model-fit is not unusual in empirical modeling. Table 7 contains the estimates of the coefficients represented by the four solid arrows. CB-SEM yields the greatest z-statistic for parameter  $\gamma_{21}$ , but for the other three parameters the z-statistics following CB-SEM are uniformly the smallest. The methods yield the greatest z-statistics for  $\gamma_{11}$ ,  $\beta_{21}$  and  $\beta_{32}$  are PLS-SEM mode B<sub>A</sub>, path analysis using the Bartlett-factor scores, and PLS-SEM mode B<sub>A</sub>, respectively.

Table 8 contains the results of fitting the second model in Fig. 6 to the vocational-psychology data. With two additional parameters the statistic  $T_{ml}$  drops to 361.848, which corresponds to RMSEA = .128, and CFI = .926. The partial mediation model still cannot be regarded as adequate according to the established norms on fit indices (Hu & Bentler, 1999). However, the value of  $T_{ml}$  as well as those of the derived RMSEA and CFI are affected by other factors in addition to the structural model. In particular, extra kurtosis of the observed sample alone can cause the significance of  $T_{ml}$  even if the model is literally correct (Hu et al., 1992), but we do not have such information for this dataset, because only a sample covariance matrix is available. In Table 8, CB-SEM yields uniformly the smallest z-statistics for all the six parameters. The greatest z-statistics for  $\gamma_{11}$ ,  $\gamma_{11}$ ,  $\beta_{21}$ ,  $\beta_{31}$ ,  $\beta_{32}$  are obtained by PLS-SEM mode B<sub>A</sub>, PLS-SEM mode B<sub>A</sub>, PLS-SEM mode A, path analyses via Bartlett-factor scores, equally weighted composites, and Bartlett-factor scores, respectively.

**Dataset 7**

The seventh is a raw dataset from Neumann (1994), who studied the relationship of psychopathology and alcoholism. The dataset consists of  $p = 10$  variables and  $N = 335$

cases. The single exogenous construct (family history) is tapped by two indicators: family history of psychopathology and family history of alcoholism. The other eight indicators are respectively the age of 1st problem with alcohol, age of 1st detoxification from alcohol, alcohol severity score, alcohol use inventory, SCL-90 psychological inventory, the sum of the Minnesota Multiphasic Personality Inventory scores, the lowest level of psychosocial functioning during the past year, and the highest level of psychosocial functioning during the past year. With two indicators for each latent construct, these eight indicators respectively measure: age of onset, alcohol symptoms, psychopathology symptoms, and global functioning. Neumann’s (1994) theoretical model for this data set is represented by Fig. 7, which has six path coefficients:  $\gamma_{11}$ ,  $\beta_{21}$ ,  $\beta_{31}$ ,  $\beta_{32}$ ,  $\beta_{41}$ ,  $\beta_{42}$ .

With raw data, we can check the distribution properties of this psychopathology-alcoholism sample, which corresponds to a standardized multivariate kurtosis  $M_s = 14.763$ . Consequently, the sample cannot be regarded as normally distributed. Using the 95% quantile ( $c_{.95}$ ) of  $\chi^2_{10}$  as the cutoff value for regulating observations according to Huber-type weights, the resulting kurtosis of the transformed sample is  $M_s = 2.248$ , still statistically significant when referred to  $N(0, 1)$ . Further tuning with  $c_{.90}$  as the cutoff value for regulating observations results in  $M_s = -.110$ . Thus, we can regard the transformed data by H(.10) as approximately normally distributed. Our analysis and comparison below will be for this transformed sample.

Fitting the model in Fig. 7 to the transformed sample by NML results in  $T_{ml} = 40.985$ , corresponding to a p-value = .069 when referred to  $\chi^2_{29}$ . Fit indices (RMSEA = .035, CFI = .987) also indicate that the model reached close fit. The estimates of the six coefficients are reported in Table 9. Estimates of  $\beta_{31}$  by the 5 methods are consistently not statistically significant at .05 level although the z-statistic under CB-SEM is the largest. For the other 5 parameters,

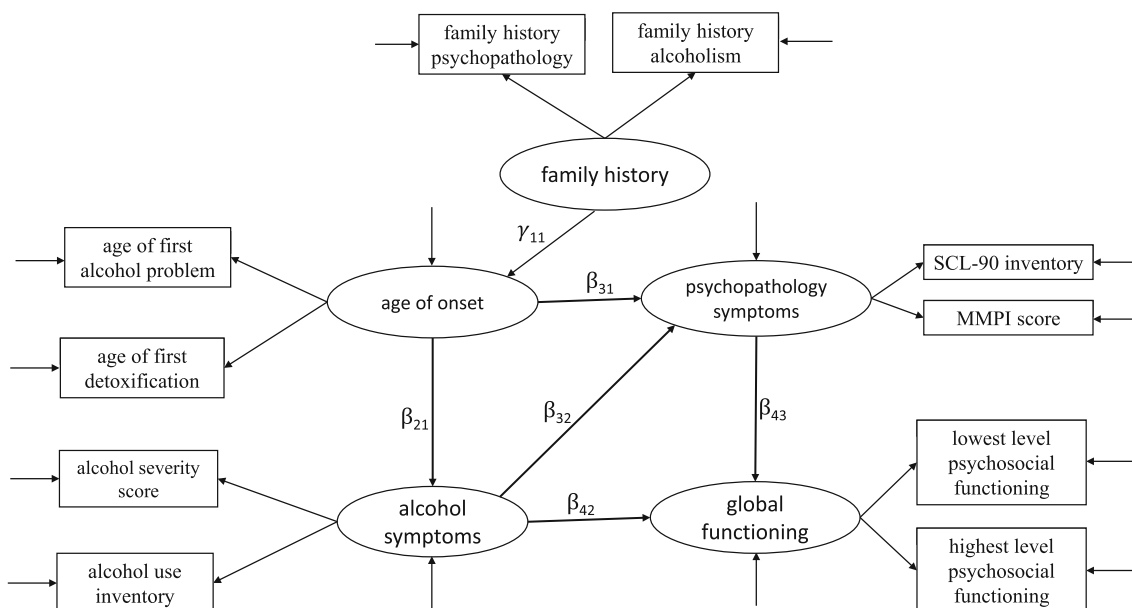


Fig. 7 A model of symptoms of alcoholism and psychopathology (Neumann, 1994,  $N = 335$ )

PLS-SEM mode  $B_A$  yields the largest  $z$ -statistics for  $\gamma_{11}$  and  $\beta_{41}$ , PLS-SEM mode A yields the largest  $z$ -statistics for  $\beta_{21}$  and  $\beta_{42}$ , and path analysis with the Bartlett-factor scores yields the largest  $z$ -statistic for  $\beta_{32}$ . Again, a non-significant  $z$ -statistic at the .05 level does not mean that we can regard the parameter  $\beta_{31}$  as 0. Many factors contribute to the significance of a parameter estimate in statistical modeling in general.

**Dataset 8**

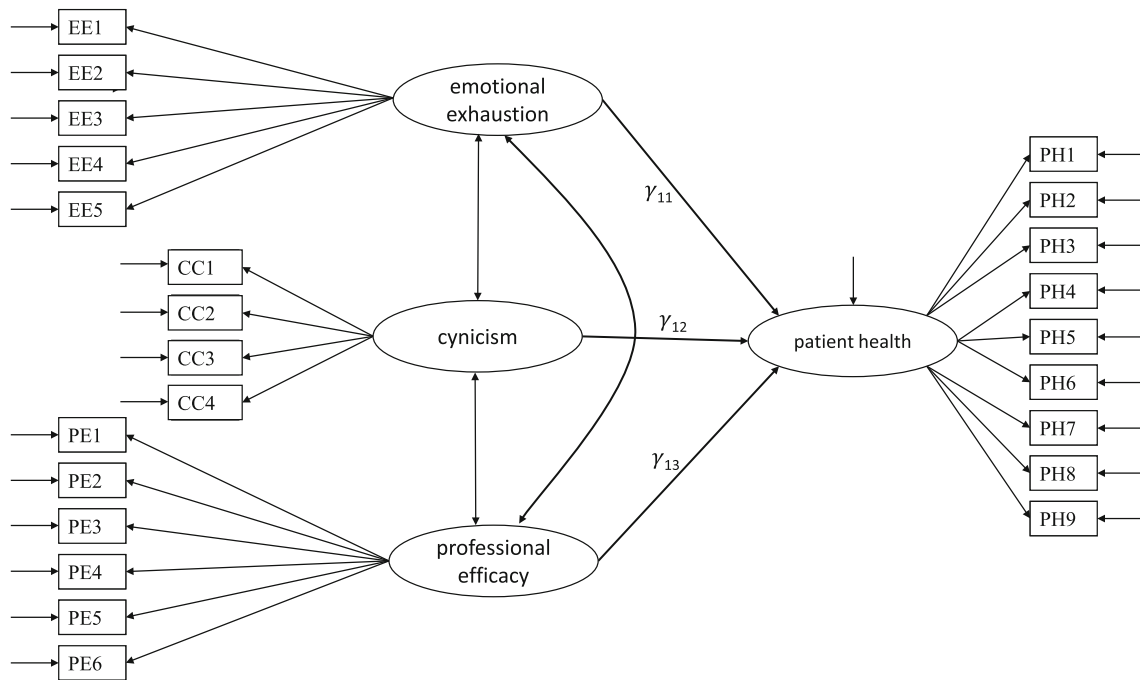
The eighth is a raw dataset from a study of health and stress, and was examined in Yuan and Deng (2021), with  $p = 24$  and  $N = 264$ . The 264 participants were recruited from both high-pressure professionals and from a

psychiatric hospital. The 24 variables are indicators of four constructs. The three exogenous constructs are respectively: 1) emotional exhaustion with 5 indicators, 2) cynicism with 4 indicators, and 3) professional efficacy with 6 indicators. The single endogenous construct (depression) has 9 indicators obtained from patient health questionnaire (PHQ9, Kroenke, Spitzer & Williams, 2001). The path diagram for the structural model is in Fig. 8, which has three path coefficients:  $\gamma_{11}$ ,  $\gamma_{12}$ , and  $\gamma_{13}$ .

As noted in Yuan and Deng (2021), with a standardized multivariate kurtosis  $M_s = 23.292$ , we cannot regard the sample as following a multivariate normal distribution. Using Huber-type weights and cutoff value  $c_{.95}$  (95% quantile of  $\chi^2_{24}$ ) in the robust transformation, the standardized multivariate kurtosis of the transformed sample is  $M_s =$

Table 9 Parameter estimates (est), their SEs (se) and  $z$ -statistics for the path coefficients in Fig. 7 ( $p = 10$ ,  $N = 335$ ,  $T_{ml} = 40.985$ ,  $df = 29$ ,  $p$ -vale = .069; RMSEA = .035, and CFI = .987)

$\theta$	CB-SEM			BFS-PA			EWC-PA			PLS-SEM A			PLS-SEM $B_A$		
	est	se	$z$	est	se	$z$	est	se	$z$	est	se	$z$	est	se	$z$
$\gamma_{11}$	-0.342	0.071	-4.819	-0.226	0.042	-5.403	-0.175	0.038	-4.573	-0.250	0.053	-4.703	-0.285	0.053	-5.427
$\beta_{21}$	-0.483	0.067	-7.217	-0.402	0.045	-8.924	-0.407	0.045	-9.035	-0.444	0.049	-9.047	-0.435	0.049	-8.811
$\beta_{31}$	0.087	0.074	1.182	0.020	0.054	0.362	0.056	0.056	1.002	0.042	0.055	0.767	0.001	0.054	0.015
$\beta_{32}$	0.627	0.086	7.324	0.495	0.059	8.398	0.483	0.061	7.926	0.448	0.055	8.101	0.453	0.054	8.335
$\beta_{41}$	-0.222	0.093	-2.381	-0.243	0.066	-3.668	-0.222	0.068	-3.288	-0.192	0.055	-3.491	-0.209	0.056	-3.732
$\beta_{42}$	-0.418	0.101	-4.135	-0.318	0.061	-5.200	-0.335	0.062	-5.423	-0.301	0.055	-5.462	-0.279	0.056	-4.986



**Fig. 8** A model of burnout and depression (Yuan & Deng, 2021,  $N = 264$ )

8.420, which is still highly significant. We subsequently tuned the cutoff value and end up with  $M_s = -.003$  under the transformation H(.22). The results presented below are based on fitting the model in Fig. 8 to this robustly transformed sample by NML.

The likelihood ratio statistic for the overall model is  $T_{ml} = 469.579$ , highly significant when referred to  $\chi^2_{246}$ . However, fit indices CFI = .948 and RMSEA = .059 suggest that the model in Fig. 8 fits the sample reasonably well. The estimates of the three coefficients by the five methods are given in Table 10, where the  $z$ -statistics for both  $\gamma_{11}$  and  $\gamma_{13}$  under CB-SEM are not statistically significant at the .05 level. The  $z$ -statistics for  $\gamma_{13}$  are not significant under path analyses with the four weighted composites. But the four estimates of  $\gamma_{11}$  under path analyses are significant. Path analysis using the Bartlett-factor scores yields the

largest  $z$ -statistic for  $\gamma_{12}$  while PLS-SEM mode B<sub>A</sub> yields the largest  $z$ -statistic for  $\gamma_{11}$ .

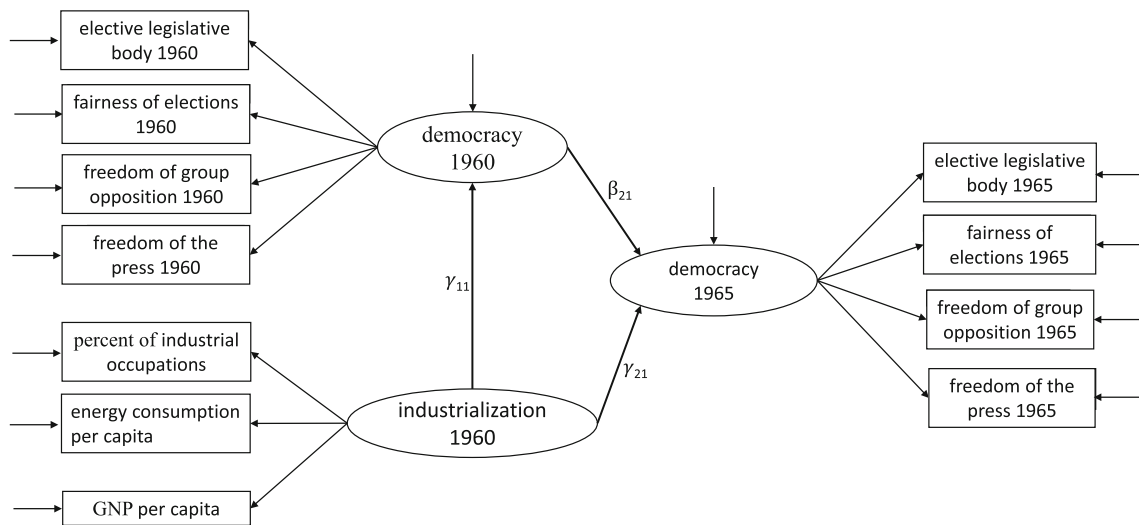
**Dataset 9**

The ninth dataset is from Table 8.2 of Bollen (1989, p. 334), which contains a sample covariance matrix of  $p = 11$  industrialization-democracy variables based on a sample of  $N = 75$  developing countries. These 11 variables are indicators of three constructs. Industrialization in 1960 is an exogenous construct and has three indicators: gross national product (GNP) per capita, energy consumption per capita, and the percent of the labor force in industrial occupations. Democracy 1960 is an endogenous construct and has four indicators: freedom of the press, freedom of group opposition, fairness of elections, and the elective

**Table 10** Parameter estimates (est), their SEs (se) and  $z$ -statistics for the path coefficients in Fig. 8 ( $p = 24, N = 264, T_{ml} = 469.579, df = 246, p\text{-value} = .000; RMSEA = .059, \text{ and CFI} = .948$ )

$\theta$	CB-SEM			BFS-PA			EWC-PA			PLS-SEM A			PLS-SEM B <sub>A</sub>		
	est	se	$z$	est	se	$z$	est	se	$z$	est	se	$z$	est	se	$z$
$\gamma_{11}$	0.066	0.036	1.831	0.072	0.032	2.245	0.075	0.032	2.341	0.148	0.058	2.539	0.161	0.059	2.725
$\gamma_{12}$	0.269	0.040	6.783	0.243	0.032	7.622	0.239	0.032	7.527	0.436	0.058	7.500	0.435	0.059	7.432
$\gamma_{13}$	-0.044	0.032	-1.365	-0.044	0.030	-1.490	-0.044	0.030	-1.487	-0.079	0.054	-1.474	-0.084	0.053	-1.575

BFS = Bartlett-factor score, EWC = equally weighted composite, PA = path analysis



**Fig. 9** A model of industrialization and democracy (Bollen, 1989,  $N = 75$ ), each of the four factor loadings on democracy 1960 equals that on democracy 1965

nature and effectiveness of the legislative body. Democracy 1965 is another endogenous construct and has the same set of indicators as democracy 1960 but measured at a later time. Bollen (1989) considered multiple models for this dataset. We will estimate the model in Fig. 9, which has three path coefficients:  $\gamma_{11}$ ,  $\gamma_{21}$ , and  $\beta_{21}$ .

The difference of this model from the previous ones is that each of the four factor loadings on democracy 1960 is constrained to be equal to that on democracy 1965, due to the nature of panel data. The constraints were also part of the model in Figure 8.1 of Bollen (1989, p.324), which also contains six pairs of correlated errors. We do not include the 6 error-covariances in the model since they empirically share the responsibility of the path coefficients. In particular, under CB-SEM error-covariances and the path coefficients in the structural model jointly account for the association of the indicators, whereas path analysis via weighted composites and PLS-SEM account for the association only through the path coefficients in the structural model. Thus, it is not a fair comparison between the two classes of methods when models under CB-SEM have correlated errors.

Fitting the model (with three extra constraints on factor loadings) in Fig. 9 to the  $p = 11$  industrialization-democracy variables by NML yields  $T_{ml} = 72.709$ , which corresponds to a  $p$ -value = .004 when referred to  $\chi^2_{44}$ . Fit indices CFI = .957 and RMSEA = .094 indicate that the model fits the data reasonably well. Table 11 contains the estimates of the three path coefficients by the five methods. Note that the Bartlett-factor scores are computed using the factor loadings and error variances estimated by NML, and they implicitly used the equality constraints. However, the other three types of composites do not use the equality constraints. PLS-SEM mode  $B_A$  yields the smallest  $z$ -statistic for  $\beta_{21}$  while CB-SEM yields the smallest  $z$ -statistic for  $\gamma_{21}$ . The largest  $z$ -statistics for  $\gamma_{11}$ ,  $\gamma_{21}$  and  $\beta_{21}$  are obtained by PLS-SEM mode  $B_A$ , path analysis with EWCs, and path analysis with the Bartlett-factor scores, respectively.

We have compared CB-SEM against path analysis with four types of composites in estimating the path coefficients of the structural models. Eleven models over nine datasets were examined, with a total of 47 parameters. CB-SEM yielded the smallest  $z$ -statistics 38 times, and largest

**Table 11** Parameter estimates (est), their SEs (se) and  $z$ -statistics for the path coefficients in Fig. 9 ( $p = 11$ ,  $N = 75$ ,  $T_{ml} = 72.709$ ,  $df = 44$ ,  $p$ -value = .004; RMSEA = .094, and CFI = .957)

$\theta$	CB-SEM			BFS-PA			EWC-PA			PLS-SEM A			PLS-SEM $B_A$		
	est	se	$z$	est	se	$z$	est	se	$z$	est	se	$z$	est	se	$z$
$\gamma_{11}$	0.981	0.267	3.670	0.947	0.244	3.876	0.870	0.247	3.528	0.384	0.108	3.549	0.455	0.104	4.367
$\gamma_{21}$	0.312	0.152	2.052	0.417	0.135	3.080	0.431	0.136	3.173	0.188	0.062	3.015	0.184	0.065	2.838
$\beta_{21}$	0.887	0.074	11.978	0.766	0.059	12.973	0.757	0.060	12.687	0.783	0.062	12.536	0.773	0.065	11.919

BFS = Bartlett-factor score, EWC = equally weighted composite, PA = path analysis

**Table 12** Counts for each method to yield the largest and smallest  $z$ -statistics as well as the mean rank (1 to 5) and mean value of the  $z$ -statistics across the 47 parameters

	CB-SEM	BFS-PA	EWC-PA	PLS-SEM A	PLS-SEM B <sub>A</sub>
Count of largest $z$	3	11	5	7	21
Count of smallest $z$	38	0	7	0	2
Average rank of $ z $	1.468	3.468	2.872	3.383	3.809
Average value of $ z $	6.341	7.801	7.640	7.718	7.840

$z$ -statistics 3 times.<sup>7</sup> Path analysis with Bartlett-factor scores never had the smallest  $z$ -statistics nor PLS-SEM mode A. The frequencies for each method to yield the largest and smallest  $z$ -statistics are reported in Table 12. For each parameter, we also ranked the five  $z$ -statistics from the smallest (rank 1) to the largest (rank 5) in absolute values. The average value of the rank for each method across the 47 parameters is obtained and so is the average of the absolute value of the 47  $z$ -statistics. They are included in Table 12 as well. All the results in Table 12 indicate that CB-SEM has the smallest power<sup>8</sup> and/or effect size in testing the significance of the path coefficients of the structural models. In contrast, PLS-SEM mode B<sub>A</sub> yielded the largest average  $z$ -statistics and average rank, followed by path analysis with Bartlett-factor scores. In particular, the results in Table 12 indicate that path analysis with equally weighted composites also corresponds to greater effect size and statistical power in testing the path coefficients than CB-SEM.

Because the focus of the article is on the effect size and statistical power in estimating/testing path coefficients, we did not include  $R^2$  values in the comparison. But measurement errors are known to affect the  $R^2$  values. In particular, the  $R^2$  for regression analysis with composite scores will become smaller when either the outcome variable or the predictors contain more measurement errors. The value of the resulting  $R^2$  is proportional to the reliability of the outcome variable and a weighted average of the reliabilities of the predictors (Cochran, 1970). There are 30 endogenous constructs within the 11 models we examined. CB-SEM yields the largest  $R^2$  values for 29 out of the 30 cases. The only exception is for the

regression model of democracy 1960 being predicted by industrialization 1960 (the left part of Fig. 9), where PLS-SEM mode B<sub>A</sub> corresponds to a greater value of  $R^2$  than CB-SEM, due to a substantial change in weights from PLS-SEM mode A to PLS-SEM mode B<sub>A</sub>. However, the  $R^2$  value under CB-SEM is not achievable when the model is used for prediction or classification of individuals. This is because we have to use the individual scores to predict the individual outcome variables, and measurement errors cannot be avoided with the observed scores. The reliabilities of the observed scores affect not only the value of  $R^2$  but also the prediction accuracy. Among path analyses with weighted composites, PLS-SEM mode B<sub>A</sub> yields the largest  $R^2$  in most cases, followed by path analysis using the Bartlett-factor scores, PLS-SEM mode A, and path analysis using equally weighted composites. Thus, PLS-SEM mode B<sub>A</sub> and path analyses using Bartlett-factor scores are also preferred from the perspective of prediction of individuals with the smallest MSE.

## Conclusions and discussion

In social and behavioral sciences, theoretical constructs are commonly measured by indicators that contain measurement errors. CB-SEM is regarded as a proper method for modeling the relationship among the constructs because of its function in separating measurement errors and the latent constructs. However, the estimates of the path coefficients under CB-SEM may contain sizable sampling errors due to simultaneously estimating many parameters. In practice, the substantive interests are often prediction and/or classification of individuals. Then path analysis with composite scores has the advantage of directly estimating the relationship of the observed scores. This article shows that path analysis via weighted composites has an additional advantage of yielding path coefficients with less relative errors, as reflected by greater effect size and statistical power. The most interesting finding is that even path analysis with EWCs by the LS method yields estimates with greater signal-to-noise ratio than CB-SEM on average.

<sup>7</sup>They are respectively for the estimates of  $\gamma_{11}$  in Table 5,  $\gamma_{21}$  in Table 7, and  $\beta_{31}$  in Table 9. However, two of the three estimates ( $\hat{\gamma}_{11}$  and  $\hat{\beta}_{31}$ ) are not statistically significant at the level .05.

<sup>8</sup>All the results of path analysis with weighted composites presented in this article are conditional on the estimated weights, as is typically done in practice (see e.g., DiStefano, Zhu & Mindrila, 2009). The method is even more powerful when sampling errors in weights are considered (Yuan & Fang, 2021)



The results of this article imply that path analyses with Bartlett-factor scores or PLS-SEM mode  $B_A$  tend to yield greater effect size for the path coefficients of the structural model than PLS-SEM mode A or path analysis with EWCs. Such a pattern is closely related to measurement reliability of the composites. In particular, Bartlett-factor scores and the weighted composites under PLS-SEM mode  $B_A$  are most reliable among all weighted combinations of the indicators (Yuan & Bentler, 2002; Yuan & Deng, 2021). Consequently, we recommend mode  $B_A$  instead of mode A for models with reflective indicators. Because formative indicators typically do not contain measurement errors or they do not share a common construct, mode B is still a viable option for computing the weights of formative indicators.

Although latent variables are error-free and CB-SEM directly models their relationship, measurement errors still interfere with the accuracy of the resulting parameter estimates. For a model containing two constructs ( $\eta$ ,  $\xi$ ) with a structural model  $\eta = \gamma\xi + \zeta$ , let  $\hat{\gamma}_{ls}$  be the LS estimate of  $\gamma$  when  $\xi$  and  $\eta$  were directly used in the estimation, which is an ideal scenario. Then, the NML estimate of  $\gamma$  under CB-SEM is mathematically less accurate than  $\hat{\gamma}_{ls}$  even as the numbers of the indicators for both  $\xi$  and  $\eta$  go to infinity. This is because both the number of factor loadings and the number of error variances increase as the number of indicators increases, and CB-SEM has to face the issue of increasing number of model parameters. In contrast, as the number of indicators increases and under proper conditions, the reliabilities of all the composites  $\tilde{\xi}$  and  $\tilde{\eta}$  approach 1.0 (Hayashi, Yuan, & Sato, 2021). Let  $\hat{\gamma}_{ls}^*$  be the LS estimate for the model  $\tilde{\eta} = \gamma^*\tilde{\xi} + \zeta^*$ . The signal-to-noise ratio in  $\hat{\gamma}_{ls}^*$  then converges to that of the LS estimate  $\hat{\gamma}_{ls}$ . That is, as the number of indicators increases, the accuracy of the LS estimate of the regression coefficient under path analysis with composite scores approaches that under the ideal scenario. The results imply that, as the number of indicators increases, the advantage of path analysis with weighted composites over CB-SEM becomes more obvious, and the difference between the different types of composites will be less clear.

While path analysis with composite scores yields path coefficients with greater signal-to-noise ratio, the method itself is unable to offer the information such as reliability of individual items or the goodness of the overall model structure. These are important features of measurement and modeling, and can be directly obtained from CB-SEM. Other aspects of measurement such as unidimensionality as well as convergent and discriminant validity are also better evaluated under CB-SEM. Thus, while path analysis with composites has extra strength in prediction, classification and diagnosis for individuals, CB-SEM still offers valuable information that path analysis via composite scores does not provide. We recommend CB-SEM analysis being conducted

first and followed by path analysis using Bartlett-factor scores or PLS-SEM mode  $B_A$  even if prediction is of primary interest.

Under the assumption of normality on the distribution of the composite scores, the mechanism for path analysis to yield prediction with the smallest MSE is due to the fact that the predicted value is the conditional mean of the outcome variable. When data are not normally distributed, the prediction is most accurate in the class of linear predictors (Fuller, 1987, p.75). In addition to the properties inherited from conditional means, composites under PLS-SEM mode  $B_A$  also possess the property of maximum reliability, which also directly contributes to minimizing the MSE of the prediction. However, path analysis with weighted composites or PLS-SEM is a two-step procedure for modeling the relationship of the constructs. We can also obtain the conditional mean of an endogenous construct directly from the covariance matrices under the CB-SEM model. While this direct approach may seem more expedient, it is also a two-step procedure since parameters under CB-SEM have to be estimated before conditional means can be computed. With a correctly formulated model structure, the direct approach may yield equivalent predictors as those under path analysis using the Bartlett-factor scores or PLS-SEM mode  $B_A$ . The matter becomes complicated with misspecified models. Regardless of whether the model is correctly specified or not, PLS-SEM has two mechanisms to maximize the predictive relationship among constructs: 1) The estimated weights under PLS-SEM maximize the correlation between the indicators in different blocks by regression analysis via the environmental variables (see e.g., Boardman, Hui & Wold, 1981); and 2) the estimated path coefficients maximize the correlation between the predictors and the outcome variable according to LS regression. Similarly, with misspecified models, the parameter estimates under CB-SEM also alter their values in order to account for the extra association of the observed indicators (Yuan, Marshall & Bentler, 2003). Additional studies are needed for comparing the two approaches.

The literatures of regression analysis with weighted composites and measurement-error models contain proposals for correcting the regression coefficients so that they become consistent with those under CB-SEM or regression analysis with true scores (e.g., Croon, 2002; Devlieger et al., 2016; Devlieger & Rosseel, 2017; Dijkstra & Henseler, 2015; Fuller, 1987; Hoshino & Bentler, 2013; Yuan et al., 2020). However, the corrected estimates may not work well when they are teamed with the composites for prediction or diagnosis of individuals. This is because the composites may not be on the same scales as those of the latent variables under CB-SEM or the true scores. Even if they are on the same scales, the corrected estimates will break the property

of the “best in the class of linear predictors” of the intact LS estimates (Fuller, 1987, p. 75). In addition, the estimated path coefficients under two different types of composites are not exchangeable for the purpose of prediction since they may have different scales (Croon, 2002; Devlieger et al., 2016).

**Acknowledgements** This work was supported by Grant 31971029 from the Natural Science Foundation of China, Grant 7202101 from Beijing Municipal Natural Science Foundation, and a grant from the Department of Education (R305D210023). However, the contents of the study do not necessarily represent the policy of the funding agencies, and you should not assume endorsement by the federal government.

## Appendix: Software implementation

The results in this article are computed using SAS IML programs, coded by the authors. The programs can be downloaded at [https://www3.nd.edu/~kyuan/SEM\\_PAcomparison/](https://www3.nd.edu/~kyuan/SEM_PAcomparison/), and are named according to the tables they generate (Table1.sas to Table11.sas). The structures of the SAS IML programs are about the same, as documented via the comments within each program. Because Dataset 7 is not publicly available, only the SAS IML program is provided for the results in Table 9. For comparison purpose, EQS (Bentler, 2006) programs are also provided for computing the results under CB-SEM, which are identical to those obtained by the SAS IML programs. The website also contains an R package that computes the results in Table 7. Researchers who are familiar with the R language (R Core Team, 2021) can modify the code to compute the results in the other tables. Both the SAS IML and R programs are designed according to subroutines and functions, respectively. The purpose of each subroutine or function is explained via comments within the programs.

The results under CB-SEM can also be obtained by other free and commercial SEM software (LISREL, Jöreskog & Sörbom, 1993; Mplus, Muthén & Muthén, 2007; lavaan, Rosseel, 2012). However, there might be differences between the outputs of different programs. As implied by the SAS IML source code, our results for CB-SEM are computed by fitting the structural model to the unbiased sample covariance matrices by minimizing the normal-distribution-based discrepancy function  $F_{ml}$  (e.g., equation 4.67 of Bollen, 1989). EQS also implements the NML method the same way (see e.g., equation 5.13 of Bentler, 2016). Programs that do not operate according to the above scheme may generate different results, which might correspond to even smaller effect sizes than the CB-SEM estimates presented in this article.

There are multiple free and commercial programs for conducting PLS-SEM (e.g., ADANCO, Henseler &

Dijkstra, 2016; semPLS, Monecke & Leisch, 2012; cSEM, Rademaker & Schuberth, 2020; SmartPLS, Ringle, Wende, & Becker, 2015). They seem to require the input of raw data, and will not work with sample covariance matrices. In addition, these programs compute the weights for standardized item scores. However, CB-SEM is typically conducted using the sample covariance matrix instead of the sample correlation matrix. For a fair comparison between CB-SEM and PLS-SEM, the PLS-SEM methodology in our SAS IML and the R programs ([https://www3.nd.edu/~kyuan/SEM\\_PAcomparison/](https://www3.nd.edu/~kyuan/SEM_PAcomparison/)) is implemented using the sample covariance matrix instead of the sample correlation matrix. Previous PLS-SEM programs do not contain the option for computing the weights under mode  $B_A$ . A subroutine for computing this mode is in each of our SAS IML programs and the R package. Note that PLS-SEM mode  $B_A$  involves estimating the error variance for each reflective indicator. The program sets by default the boundary values of negative error variances (Heywood cases) to .05, and users can change it to another small positive number.

## References

- Bartholomew, D. J. (2009). The origin of factor scores: Spearman, Thomson and Bartlett. *British Journal of Mathematical and Statistical Psychology*, 62, 569–582. <https://doi.org/10.1348/000711008X365676>
- Bentler, P. M. (1980). Multivariate analysis with latent variables: Causal modeling. *Annual Review of Psychology*, 31, 419–456. <https://doi.org/10.1146/annurev.ps.31.020180.002223>
- Bentler, P. M. (1983). Some contributions to efficient statistics in structural models: Specification and estimation of moment structures. *Psychometrika*, 48, 493–517. <https://doi.org/10.1007/BF02293875>
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107, 238–246. <https://doi.org/10.1037/0033-2909.107.2.238>
- Bentler, P. M. (2006). *EQS 6 Structural equations program manual*. Encino: Multivariate Software.
- Blanca, M. J., Arnau, J., López-Montiel, D., Bono, R., & Bendayan, R. (2013). Skewness and kurtosis in real data samples. *Methodology*, 9(2), 78–84. <https://doi.org/10.1027/1614-2241/a000057>
- Boardman, A. E., Hui, B. S., & Wold, H. (1981). The partial least squares-fix point method of estimating interdependent systems with latent variables. *Communications in Statistics-Theory and Methods*, 10(7), 613–639. <https://doi.org/10.1080/03610928108828062>
- Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley.
- Browne, M. W. (1982). Covariance structures. In D. M. Hawkins (Ed.) *Topics in applied multivariate analysis* (pp. 72–141). London: Cambridge University Press.
- Buonaccorsi, J. P. (2010). *Measurement error: Models, methods, and applications*. Boca Raton: Chapman & Hall.
- Cain, M., Zhang, Z., & Yuan, K.-H. (2017). Univariate and multivariate skewness and kurtosis for measuring nonnormality: Prevalence, influence and estimation. *Behavior Research Methods*, 49(5), 1716–1735. <https://doi.org/10.3758/s13428-016-0814-1>

- Cochran, W. G. (1970). Some effects of errors of measurement on multiple correlation. *Journal of the American Statistical Association*, 65(329), 22–34. <https://doi.org/10.1080/01621459.1970.10481059>
- Croon, M. (2002). Using predicted latent scores in general latent structure models. In G. A. Marcoulides, & I. Moustaki (Eds.) *Latent variable and latent structure modeling* (pp. 195–223). Mahwah: Lawrence Erlbaum.
- 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>
- Devlieger, I., & Rosseel, Y. (2017). Factor score path analysis: An alternative for SEM? *Methodology*, 13, 31–38. <https://doi.org/10.1027/1614-2241/a000130>
- Dijkstra, T. (1983). Some comments on maximum likelihood and partial least squares estimates. *Journal of Econometrics*, 22, 67–90. [https://doi.org/10.1016/0304-4076\(83\)90094-5](https://doi.org/10.1016/0304-4076(83)90094-5)
- Dijkstra, T. K., & Henseler, J. (2015). Consistent and asymptotically normal PLS estimators for linear structural equations. *Computational Statistics and Data Analysis*, 81, 10–23. <https://doi.org/10.1016/j.csda.2014.07.008>
- DiStefano, C., Zhu, M., & Mindrila, D. (2009). Understanding and using factor scores: Considerations for the applied researcher. *Practical Assessment, Research, and Evaluation*, 14, Article 20. <https://doi.org/10.7275/da8t-4g52>
- Fuller, W. A. (1987). *Measurement error models*. New York: Wiley.
- Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2017). *A primer on partial least squares structural equation modeling (PLS-SEM)* (2nd edn). Thousand Oaks: Sage.
- Hampel, F. R., Ronchetti, E., Rousseeuw, P. J., & Stahel, W. (1986). *Robust statistics: The approach based on influence functions*. New York: Wiley.
- Hayashi, K., Yuan, K.-H., & Sato, R. (2021). On the coefficient alpha in high-dimensions. In M. Wiberg, D. Molenaar, J. González, U. Böckenholt, & J.-S. Kim (Eds.) *Quantitative psychology: The 85th annual meeting of the psychometric society* [https://doi.org/10.1007/978-3-030-74772-5\\_12](https://doi.org/10.1007/978-3-030-74772-5_12) (pp. 127–139). New York: Springer.
- Henseler, J. (2021). *Composite-based structural equation modeling*. New York: Guilford.
- Henseler, J., & Dijkstra, T. K. (2016). *ADANCO 2.0.1 User manual Kleve*. Germany: Composite Modeling. <http://www.composite-modeling.com>
- Holland, J. L. (1997). *Making vocational choices: A theory of vocational personalities and work environments* (3rd edn). Psychological Assessment Resources: Odessa.
- Hoshino, T., & Bentler, P. M. (2013). Bias in factor score regression and a simple solution. In: A. R. de Leon & K. C. Chough (Eds.) *Analysis of mixed data: Methods & applications* (pp. 43–61). Boca Raton: Chapman & Hall. <https://doi.org/10.1201/B14571-11>
- Houghton, J. D., & Jinkerson, D. L. (2007). Constructive thought strategies and job satisfaction: A preliminary examination. *Journal of Business Psychology*, 22, 45–53. <https://doi.org/10.1007/s10869-007-9046-9>
- Hu, L.-T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1–55. <https://doi.org/10.1080/10705519909540118>
- Hu, L., Bentler, P. M., & Kano, Y. (1992). Can test statistics in covariance structure analysis be trusted? *Psychological Bulletin*, 112, 351–362. <https://doi.org/10.1037/0033-2909.112.2.351>
- Huber, P. J. (1981). *Robust statistics*. New York: Wiley.
- Jöreskog, K. G., & Sörbom, D. (1993). *LISREL 8: Structural equation modeling with the SIMPLIS command language*. Lincolnwood: Scientific Software International.
- Kline, R. B. (1998). *Principles and practice of structural equation modeling*. New York: Guilford Press.
- Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th edn). New York: Guilford Press.
- Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine*, 16, 606–613. <https://doi.org/10.1046/j.1525-1497.2001.016009606.x>
- Lai, K., & Green, S. B. (2016). The problem with having two watches: Assessment of fit when RMSEA and CFI disagree. *Multivariate Behavioral Research*, 51(2–3), 220–239. <https://doi.org/10.1080/00273171.2015.1134306>
- Lawley, D. N., & Maxwell, A. E. (1971). *Factor analysis as a statistical method*. New York: Macmillan Publishing.
- Mardia, K. V. (1970). Measure of multivariate skewness and kurtosis with applications. *Biometrika*, 57, 519–530. <https://doi.org/10.1093/biomet/57.3.519>
- Maronna, R. A., Martin, R. D., Yohai, V. J., & Salibián-Barrera, M. (2019). *Robust statistics: Theory and methods (with R)* (2nd edn.). Hoboken: Wiley.
- Micceri, T. (1989). The unicorn, the normal curve, and other improbable creatures. *Psychological Bulletin*, 105, 156–166. <https://doi.org/10.1037/0033-2909.105.1.156>
- Monecke, A., & Leisch, F. (2012). semPLS: Structural equation modeling using partial least squares. *Journal of Statistical Software*, 48(3), 1–32. <https://doi.org/10.18637/jss.v048.i03>
- Muthén, L. K., & Muthén, B. O. (2007). *Mplus user's guide* (5th edn). Los Angeles: Muthén & Muthén.
- Neumann, C. S. (1994). *Structural equation modeling of symptoms of alcoholism and psychopathology Dissertation*. Lawrence: University of Kansas.
- R Core Team. (2021). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. <https://www.R-project.org/>
- Rademaker, M. E., & Schubert, F. (2020). cSEM: Composite-based structural equation modeling (Computer software manual). Retrieved from <https://m-e-rademaker.github.io/cSEM/>.
- Ringle, C. M., Wende, S., & Becker, J.-M. (2015). SmartPLS 3. Böningstedt: SmartPLS. Retrieved from <https://www.smartpls.com>.
- Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48, 1–36. <https://doi.org/10.18637/jss.v048.i02>
- Satorra, A., & Bentler, P. M. (1994). Corrections to test statistics and standard errors in covariance structure analysis. In A. von Eye, & C. Clogg (Eds.) *Latent variables analysis: Applications for developmental research* (pp. 399–419). Thousand Oaks: Sage.
- Schneeweiss, H. (1993). In K. Haagen, D. J. Bartholomew, M. Deistler, & H. Schneeweiss (Eds.) *Consistency at large in models with latent variables* (pp. 299–320). Elsevier Science Publishers.
- Schumacker, R. E., & Lomax, R. G. (2010). *A beginner's guide to structural equation modeling* (3rd edn.) New York: Routledge.
- Schuster, C., & Lubbe, D. (2020). A note on residual M-distances for identifying aberrant response patterns. *British Journal of Mathematical and Statistical Psychology*, 73, 164–169. <https://doi.org/10.1111/bmsp.12161>
- Steiger, J. H., & Lind, J. M. (1980). Statistically based tests for the number of common factors. Paper presented at the Annual Meeting of the Psychometric Society, Iowa City.
- Weston, R., & Gore, P. A. Jr. (2006). A brief guide to structural equation modeling. *The Counseling Psychologist*, 34(5), 719–751. <https://doi.org/10.1177/0011000006286345>
- Wheaton, B., Muthén, B., Alwin, D., & Summers, G. (1977). Assessing reliability and stability in panel models. *Sociological Methodology*, 8, 84–136. <https://doi.org/10.2307/270754>

- Wiley, D. E., & Hornik, R. (1973). Measurement error and the analysis of panel data. Mehrlicht! Studies of educative processes. Report No.5. University of Chicago.
- Wold, H. (1980). Model construction and evaluation when theoretical knowledge is scarce. In J. Kmenta, & J. B. Ramsey (Eds.) *Evaluation of econometric models* (pp. 47–74). New York: Academic Press.
- Wold, H. (1982). Soft modeling: The basic design and some extensions. In K. G. Jöreskog, & H. Wold (Eds.) *Systems under indirect observation, Part II* (pp. 1–54). Amsterdam: North-Holland.
- Worland, J., Weeks, G. G., Janes, C. L., & Strock, B. D. (1984). Intelligence, classroom behavior, and academic achievement in children at high and low risk for psychopathology: A structural equation analysis. *Journal of Abnormal Child Psychology*, *12*, 437–454. <https://doi.org/10.1007/BF00910658>
- Yuan, K.-H., & Bentler, P. M. (1998). Structural equation modeling with robust covariances. *Sociological Methodology*, *28*, 363–396. <https://doi.org/10.1111/0081-1750.00052>
- Yuan, K.-H., & Bentler, P. M. (2000). Inferences on correlation coefficients in some classes of nonnormal distributions. *Journal of Multivariate Analysis*, *72*, 230–248. <https://doi.org/10.1006/jmva.1999.1858>
- Yuan, K.-H., & Bentler, P. M. (2002). On robustness of the normal-theory based asymptotic distributions of three reliability coefficient estimates. *Psychometrika*, *67*, 251–259. <https://doi.org/10.1007/BF02294845>
- Yuan, K.-H., Chan, W., & Bentler, P. M. (2000). Robust transformation with applications to structural equation modeling. *British Journal of Mathematical and Statistical Psychology*, *53*, 31–50. <https://doi.org/10.1348/000711000159169>
- Yuan, K.-H., Chan, W., & Tian, Y. (2016). Expectation-robust algorithm and estimating equations for means and dispersion matrix with missing data. *Annals of the Institute of Statistical Mathematics*, *68*, 329–351. <https://doi.org/10.1007/s10463-014-0498-1>
- Yuan, K.-H., & Deng, L. (2021). Equivalence of partial-least-squares SEM and the methods of factor-score regression. *Structural Equation Modeling*, *28*(4), 557–571. <https://doi.org/10.1080/10705511.2021.1894940>
- Yuan, K.-H., & Fang, Y. (2021). Which method delivers greater signal-to-noise ratio: Structural equation modeling or regression analysis with composite scores? Under review.
- Yuan, K.-H., & Gomer, B. (2021). An overview of applied robust methods. *British Journal of Mathematical and Statistical Psychology*, *74*, 199–246. <https://doi.org/10.1111/bmsp.12230>
- Yuan, K.-H., Marshall, L. L., & Bentler, P. M. (2003). Assessing the effect of model misspecifications on parameter estimates in structural equation models. *Sociological Methodology*, *33*, 241–265. <https://doi.org/10.1111/j.0081-1750.2003.00132.x>
- 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*(3), 333–350. <https://doi.org/10.1080/10705511.2019.1647107>
- Yuan, K.-H., & Zhong, X. (2013). Robustness of fit indices to outliers and leverage observations in structural equation modeling. *Psychological Methods*, *18*, 121–136. <https://doi.org/10.1037/a0031604>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.