Factor-based structural equation modeling: Going beyond PLS and composites

Ned Kock

Full reference:
Kock (2019). Factor-based structural equation modeling: Going beyond PLS and composites. *International Journal of Data Analysis Techniques and Strategies*, 11(1), 1–28.

Abstract

Partial least squares (PLS) methods offer many advantages for path modeling, such as fast convergence to solutions and relaxed requirements in terms of sample size and multivariate normality. However, they do not deal with factors, but with composites. As a result, they typically underestimate path coefficients and overestimate loadings. Given these, it is difficult to fully justify their use for confirmatory factor analyses or factor-based structural equation modeling (SEM). We addressed this problem through the development of a new method that generates estimates of the true composites and factors, potentially placing researchers in a position where they can obtain consistent estimates of a wide range of model parameters in SEM analyses. A Monte Carlo experiment suggests that this new method represents a solid step in the direction of achieving this ambitious goal.

KEYWORDS: Partial Least Squares; Structural Equation Modeling; Measurement Error; Path Bias; Variation Sharing; Monte Carlo Simulation

Biographical note

Ned Kock is Killam Distinguished Professor and Chair of the Division of International Business and Technology Studies, in the A.R. Sanchez, Jr. School of Business, at Texas A&M International University. He is the developer of WarpPLS, a widely used nonlinear structural equation modeling software.

1. Introduction

Herman Wold, whose seminal contributions include the Cramér–Wold theorem and Wold's decomposition, developed what is known today as the partial least squares (PLS) approach to path modeling (Henseler & Chin, 2010; Kock & Mayfield, 2015; Kock & Moqbel, 2016; Wold, 1980). He intended it as an exploratory and computationally efficient approach that would yield moderately biased model parameters with small samples and with both normal and non-normal data (Kock & Mayfield, 2015; Kock & Moqbel, 2016; Wold, 1980).

This same general approach has recently been gaining an increasing number of adherents as a tool for confirmatory factor analyses and structural equation modeling (SEM), as suggested by the growth of normative discussions clearly aimed at a much broader use of PLS than originally intended (Hair et al., 2011; Richter et al., 2015). Strong criticism by methodological researchers has ensued, and for reasons that are difficult to brush away (McIntosh et al., 2014; Rönkkö & Evermann, 2013).

Historically SEM has been closely aligned with the common factor model, which posits the existence of factors that cause indicators (Kline, 2010; Kock, 2014). PLS methods, on the other hand, are based on composites and not on factors (Kock & Lynn, 2012; Kock & Mayfield, 2015; McDonald, 1996). In part because of this, PLS algorithms yield biased model parameters (Kock, 2015). Notably, path coefficients are generally underestimated, and loadings are more often than not overestimated. Given these problems it is difficult to fully justify the use of PLS, in its classic form, for confirmatory research.

Still, PLS methods have advantages. They virtually always converge to solutions, and fast (Henseler, 2010; Kock & Mayfield, 2015). They make no data or parameter distribution assumptions (Kock, 2016). Since they generate composites, these can be used as “pseudo-factors” or factor approximations, and provide a partial solution to the factor indeterminacy problem of covariance-based SEM (Fornell & Bookstein, 1982; Kline, 2010; Kock & Moqbel, 2016).

We have attempted to address the biases inherent in PLS methods, while preserving most of their advantages, through the development of a new method that deals with composites *and* factors. We refer to this method as “factor-based SEM” (FSEM), and believe it is the first of a future family of related methods. We believe that the FSEM method can be used for confirmatory factor analyses and factor-based SEM analyses, and hope that our discussion in the following pages will make it clear why.

FSEM’s main goal is ambitious: to obtain consistent estimates of *any* model parameters. It aims at doing so by first generating estimates of the true composites, which are then used to produce estimates of the true factors. These estimations also produce loadings, weights, and measurement errors.

FSEM is not a parameter correction method, such as the Cronbach alpha disattenuation (Goodhue et al., 2012) or the consistent PLS (Dijkstra & Schermelleh-Engel, 2014) methods. In general terms, correction methods first estimate parameters (such as path coefficients and loadings) via classic PLS methods. Then they correct those parameters with the goal of obtaining asymptotically unbiased versions of them. Since FSEM yields estimates of composite and factor scores, and those are used as a basis for estimation of parameters, no corrections are needed.

FSEM is more computationally complex than PLS methods in general, as will be seen. However, the FSEM method has the advantage of generating estimates of composite *and* factor scores, from which arguably any model parameter can be estimated. Given this, we see FSEM and these various methods as complementary to one another, as well as to Wold's original PLS path modeling methods and covariance-based SEM methods.

FSEM's underlying mathematics is demonstrated to, for the most part, follow directly from the common factor model. A Monte Carlo experiment suggests that FSEM represents a solid step in the direction of achieving its ambitious goal. This conclusion must be accompanied by the caveat that much more research is needed to further validate the method.

2. The common factor model and measurement error theory

In this section we discuss the mathematics underlying the common factor model (Kline, 2010) and measurement error theory (Nunnally & Bernstein, 1994). This discussion is necessarily narrow, since the mathematical properties of the common factor model and measurement error theory are well known and should not be unnecessarily repeated. Rather, we focus on a set of equations that we will use in later sections, with the required notation to meet our purposes, to derive all of the mathematical equations needed to implement the FSEM method.

We generally use the mathematical notation of classic path analysis, with elements of the notations used in PLS path modeling and covariance-based SEM. Our notation refers to variables as they are typically seen by SEM software users in data tables – e.g., each indicator as a column on a data table. For the sake of simplicity, and without any impact on the generality of the discussion presented here, we assume that variables are standardized – i.e., scaled to have a mean of zero and a standard deviation of 1.

Let F_i be a column vector denoting one of the N_F factors in a common factor model, and x_{ij} be a column vector denoting one of the n_i indicators associated with F_i. Each indicator is associated with its factor according to

$$x_{ij} = F_i\lambda_{ij} + \theta_{ij}, i = 1 \dots N_F, j = 1 \dots n_i, \tag{1}$$

where λ_{ij} denotes the indicator loading and θ_{ij} the indicator error term that accounts for the variance in x_{ij} that is not explained by F_i. This can be simplified through

$$x_i = F_i\lambda_i{}' + \theta_i, i = 1 \dots N_F, \tag{2}$$

where x_i is a matrix with each column referring to one of the indicators associated with F_i; $\lambda_i{}'$ is the transpose of λ_i, a column vector storing the loadings associated with F_i; and θ_i is a matrix with each column storing the indicator error terms.

From measurement error theory we know that each indicator is also associated with its factor according to (3), where ω_{ij} denotes the indicator's weight. The standardized measurement error is denoted by ε_i, and its associated weight as $\omega_{i\varepsilon}$.

$$F_i = \sum_{j=1}^{n_i} x_{ij}\omega_{ij} + \varepsilon_i\omega_{i\varepsilon}. \tag{3}$$

This can be simplified through (4)-(5), where ω_i is a column vector containing the weights, C_i is the composite associated with F_i, and ω_{iC} is the composite's weight. The composite to which we refer here is the *true* composite associated with its corresponding factor. As it will be seen later, its estimation requires a full consideration of the role of the indicators in defining the relationship between weights and loadings.

$$F_i = x_i \omega_i + \varepsilon_i \omega_{i\varepsilon}. \tag{4}$$

$$F_i = C_i \omega_{iC} + \varepsilon_i \omega_{i\varepsilon}. \tag{5}$$

The measurement error ε_i is uncorrelated with its adjacent indicators (i.e. the indicators x_i, which belong to the same factor), and consequently also uncorrelated with its adjacent composite C_i. As demonstrated by Nunnally & Bernstein (1994), the composite and measurement error weights are associated with their factor's true reliability α_i according to

$$\omega_{iC} = \sqrt{\alpha_i}, \tag{6}$$

$$\omega_{i\varepsilon} = \sqrt{1 - \alpha_i}. \tag{7}$$

The above are well known properties, relevant for our discussion, of the measurement component of a common factor model – we refer to it as the "measurement model". The measurement model describes the relationships among factors and their respective indicators.

The structural component of a common factor model, or the "structural model", complements the measurement model, by describing relationships among factors. In it, it is relevant for our discussion to note that

$$F_i = \sum_{j=1}^{N_i} \beta_{ij} F_j + \zeta_i, \tag{8}$$

$$F_i = \Sigma_{F_i F_j} F_j + \delta_{ij}, \tag{9}$$

where: β_{ij} is the standardized partial regression (a.k.a. path) coefficient associated with the criterion-predictor relationship between F_i and F_j; N_i is the number of predictors pointing at F_i in the model; ζ_i is the structural error term accounting for the variance in F_i that is not explained by the factors that point at it in the model; $\Sigma_{F_i F_j}$ is the correlation between F_i and F_j; and δ_{ij} is the correlation error term accounting for the variance in F_i that is not explained by F_j.

3. PLS Mode A

Several PLS algorithms have been developed based on the original design proposed by Herman Wold (see, e.g., Kock & Mayfield, 2015; Kock & Moqbel, 2016; Wold, 1980). Lohmöller (1989) provides what is probably the most extensive discussion to date of PLS algorithms and their use in the analysis of path models.

PLS algorithms do not generate estimates of factors, nor do they explicitly take measurement error into account when estimating model parameters. They estimate composites, which are exact linear combinations of indicators (Kock & Moqbel, 2016; McDonald, 1996). These composites are then used as “pseudo-factors” for the estimation of model parameters (Kock, 2015). These characteristics lead to biased estimates of path coefficients and loadings. Estimated path coefficients are generally lower than the true values, and loadings are generally higher.

The most widely used PLS path modeling algorithm is PLS Mode A (PLSA), which is also seen as compatible with the common factor model because in this algorithm indicators are related to composites in a reflective way (i.e., with arrows pointing from the factors to the indicators).

In PLSA indicator weight estimates $\hat{w}_{ij}$ are initially set to 1 (or another arbitrary positive real number, assuming that indicators and composites are positively associated), and composite estimates $\widehat{C}_i$ are initialized with a standardized vector of the summed indicators. Then the composites are re-estimated as

$$\hat{C}_i \coloneqq Stdz\left(\sum_{j=1}^{A_i} \hat{v}_{ij}\, \hat{C}_j\right), \tag{10}$$

which is known as the “inside approximation”. Here $Stdz(\cdot)$ is a function that returns a standardized column vector, and A_i is the number of composites $\widehat{C}_j$ $(j = 1 \dots A_i)$ that are “neighbors” of the composite $\widehat{C}_i$. Neighbor composites are those that are linked to a composite by arrows, either by pointing at or being pointed at by the composite. The estimates $\hat{v}_{ij}$ are referred to as the “inner weights” (Lohmöller, 1989, p. 29).

PLSA has three main variations, called “schemes”, which define how the inner weights $\hat{v}_{ij}$ are estimated: centroid, factorial, and path weighting. In the centroid scheme the inner weights are set according to (11), as units with the signs (-1 or $+1$) of the estimated correlations among neighbor composites. In the factorial scheme they are set according to (12), as the correlations among neighbor composites. In the path weighting scheme they are set according to (13), as the path coefficients or correlations among neighbor composites, depending on whether the arrows go in or out respectively.

$$\hat{v}_{ij} \coloneqq Sign\left(\Sigma_{\hat{C}_i \hat{C}_j}\right). \tag{11}$$

$$\hat{v}_{ij} \coloneqq \Sigma_{\hat{C}_i \hat{C}_j}. \tag{12}$$

$$\begin{cases} \hat{v}_{ij} \coloneqq \hat{\beta}_{ij}, & \text{if } \hat{C}_j \text{ points at } \hat{C}_i, \\ \hat{v}_{ij} \coloneqq \Sigma_{\hat{C}_i \hat{C}_j}, & \text{if } \hat{C}_i \text{ points at } \hat{C}_j. \end{cases} \tag{13}$$

There does not seem to exist a solid mathematical basis for any of these inner weight estimation schemes, leading to one of the sources of bias in PLSA. When we say that a mathematical basis is lacking, we mean that none of the equations characterizing these inner weight estimation schemes can be derived from the equations describing the common factor model or measurement error theory, of which several were discussed in the previous section.

This is not meant to be a critical comment as much as it is meant to highlight a clear source of bias, like other similar comments below. Again, it is our belief that Wold never intended PLS to be completely unbiased. In our view he successfully solved an engineering problem, at a time when computing resources were both scarce and expensive, through the development of a moderately biased method.

Broadly speaking, these schemes are based on an intuitive assumption – the so-called "good neighbor assumption" (Kock & Mayfield, 2015). This assumption is described by Adelman & Lohmoller's (1994, p. 357) as the assumption that a latent variable should be estimated "… so that it is a good neighbour in its neighbourhood. That is, estimate the [latent variable] so that it is well predicted by its predecessors in the path diagram and is a good predictor for its followers in the diagram."

The PLSA algorithm proceeds by estimating what are referred to as the "outer weights" by solving (14) for $\widehat{w}_{ij}$, where N_C is the number of composites (the same as the number of factors), and n_i is the number of indicators associated with composite C_i. This yields estimates of loadings for the outer weights.

$$x_{ij} = \hat{C}_i \widehat{w}_{ij} + \hat{\epsilon}_{ij}, i = 1 \dots N_C, j = 1 \dots n_i. \quad (14)$$

Next composites are set as in (15), a step known as "outside approximation", as exact linear combinations of their indicators; i.e., assuming measurement error to be either nil or irrelevant, mathematically speaking. These and the preceding steps are conducted iteratively until the inner weights $\widehat{w}_{ij}$ change by less than a small fraction.

$$\hat{C}_i \coloneqq Stdz\left(\sum_{j=1}^{n_i} \widehat{w}_{ij}\, x_{ij}\right). \quad (15)$$

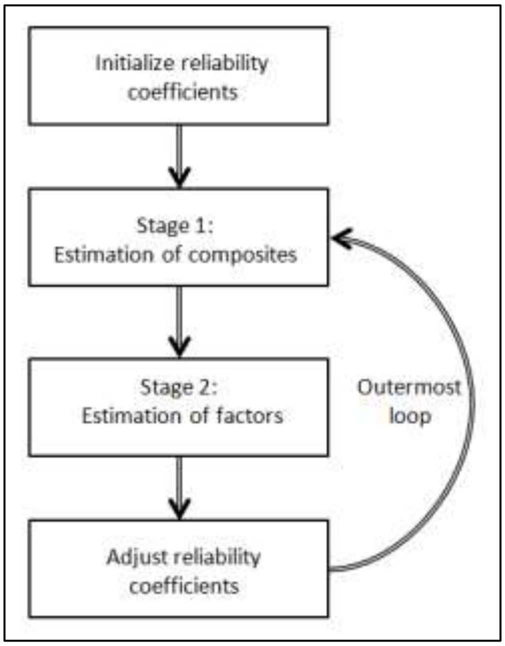
In successive iterations, this ensures that weight estimates are directly proportional to loading estimates. There does not seem to be any mathematical basis for this assumption of proportionality among weights and loadings either. This is another source of bias in PLSA, which is compounded by the assumption of no measurement error. Consequently PLSA is not only restricted to estimates of composites (as opposed to factors), but the composites that are estimated are unlikely to be the *true* composites.

4. Factor-based SEM

The FSEM method described here is made up of two main stages, which can be seen as analogous to the outside and inside approximations of PLSA. Key differences are that separate sets of iterations are carried out for each of these stages, and that the two stages are contained within an outermost loop. This outermost loop is responsible for convergence toward an asymptotically unbiased reliability measure (see Figure 1).

In the first stage we estimate the model's true composites, as well as other elements such as weights and loadings, which are then used as inputs for the second stage. This first stage is analogous to PLSA's outside approximation. In the second stage, analogous to PLSA's inside approximation, we arrive at estimates of the true measurement errors and factors, as well as other elements. In the end, after the second stage is completed, we are provided with estimates of measurement errors, factors, composites, loadings, and weights that are expected to be consistent. These elements in turn allow us to proceed to estimate any other derivative model parameter.

Figure 1. The two stages and the enveloping outermost loop



Notes:
- Both stages 1 and 2 involve iterative steps.
- The outermost loop seeks reliability measures' convergence.

Prior to the first stage, we set $\hat{\varepsilon}_i$ according to (16), where $Rnd(N)$ is a function that returns an independent and identically distributed (i.i.d.) variable as a column vector with N rows, where N is the sample size. In this stochastic approach, the measurement error "stand-in" does not need to be a i.i.d. variable at this point. For example, a normally distributed random error variable could also be used. In software implementations (e.g., WarpPLS), the random seed can be set to a fixed value; this will avoid different results being generated each time an analysis is conducted with the same model and empirical data.

$$\hat{\varepsilon}_i \coloneqq Stdz[Rnd(N)]. \qquad (16)$$

Employing a stochastic approach to estimating the measurement error allows us to solve a longstanding problem that cannot be properly addressed via other approaches that only "assume" the existence of measurement error. The problem is that, without measurement error being actually present, the estimation of the true weights is not possible. The reason for this is that measurement error accounts for the variation in a factor that is not accounted for by the indicators.

Next we set $\hat{\alpha}_i$, $\widehat{\omega}_{iC}$ and $\widehat{\omega}_{i\varepsilon}$ according to (17)-(19), where n_i is the number of indicators of factor F_i, and $\bar{\Sigma}_{x_i x_i}$ is the mean of the non-redundant correlation coefficients among the column vectors that make up x_i (e.g., the mean of the lower triangular version of $\Sigma_{x_i x_i}$).

$$\hat{\alpha}_i \coloneqq \frac{n_i \bar{\Sigma}_{x_i x_i}}{(1+(n_i-1)\bar{\Sigma}_{x_i x_i})}. \tag{17}$$

$$\widehat{\omega}_{iC} \coloneqq \sqrt{\hat{\alpha}_i}. \tag{18}$$

$$\widehat{\omega}_{i\varepsilon} \coloneqq \sqrt{1-\hat{\alpha}_i}. \tag{19}$$

The reliability estimate $\hat{\alpha}_i$ in (17) is the Cronbach's alpha coefficient (Cronbach, 1951; Kline, 2010) associated with the factor F_i. The Cronbach's alpha coefficient is a biased measure of reliability, representing a lower bound estimate (Sijtsma, 2009). Because of this, the FSEM method must iterate through its first and second stages, each time with different measures of reliability, until convergence to the correct measure is achieved.

Convergence to the correct reliability measure occurs when the quantity $\widehat{\omega}_i{}'\hat{\lambda}_i$, which is a direct measure of reliability based on estimated weights and loadings, equals the composite reliability coefficient (Dillon & Goldstein, 1984; Peterson & Yeolib, 2013). The composite reliability coefficient associated with a latent variable indexed by i, denoted as $\hat{\rho}_i$, is calculated based on the latent variable's loadings as:

$$\hat{\rho}_i = \frac{\left(\sum_{j=1}^{n_i} \hat{\lambda}_{ij}\right)^2}{\left(\sum_{j=1}^{n_i} \hat{\lambda}_{ij}\right)^2 + \sum_{j=1}^{n_i}\left(1-\hat{\lambda}_{ij}{}^2\right)}, j = 1 \dots n_i.$$

Should loadings be unbiased, it is generally believed that the composite reliability coefficient will yield the true population reliability for each latent variable. Thus in order to reach the point at which loadings are unbiased, we adjust the reliability coefficient across iterations by making it fall between the quantity $\widehat{\omega}_i{}'\hat{\lambda}_i$ and the estimated composite reliability coefficient $\hat{\rho}_i$:

$$\hat{\alpha}_i \coloneqq \frac{1}{2}\left(\widehat{\omega}_i{}'\hat{\lambda}_i + \hat{\rho}_i\right).$$

Adjusting the reliability coefficient in this way across iterations is analogous to slowly filling a glass G of water with the contents of two other glasses of water G_1 and G_2, where the water in each of these two glasses is stored at different temperatures. The temperature of the water in G_1 is "too low", and in G_2 it is "too high". In this analogy the goal would be to obtain water at the "right" temperature in G; analogously, the "right" reliability in our method. Each time we add some water into G from G_1 and G_2 we measure the temperature in G. We stop this process when we reach the desired temperature in G.

Composite and measurement error weights are re-estimated for use in the first and second stages each time the reliability coefficient is adjusted. The outermost loop enclosing the first and second stages stops when the difference between the direct measure of reliability based on estimated weights and loadings ($\hat{\omega}_i{}'\hat{\lambda}_i$) and the composite reliability coefficient ($\hat{\rho}_i$) falls below a small fraction.

Underlying FSEM's two sequential stages is an important assumption, which follows directly from the mathematical properties of the common factor model and measurement error theory discussed earlier: the true composites are completely determined by their corresponding indicators, and their estimation does not require information about any other true composite in the model.

Therefore, in the first stage we do not need information about how the composites vary with respect to other composites; e.g., we do not need information about correlations among composites. Because of this, we can create and use estimates of measurement error terms that satisfy only a few key properties, notably that they are uncorrelated with their adjacent indicators.

5. Stage 1: Estimation of composites

One of the problems with traditional PLS algorithms, including PLSA, is that they do not incorporate enough information about the relationship between weights and loadings, simply assuming that weights are proportional to loadings. This is addressed through the discussion below. Combining (2) and (4) we arrive at (20), where $\lambda_i{}'$ is the transpose of λ_i.

$$x_i = F_i\lambda_i{}' + \theta_i,\, F_i = x_i\omega_i + \varepsilon_i\omega_{i\varepsilon} \rightarrow$$

$$x_i = (x_i\omega_i + \varepsilon_i\omega_{i\varepsilon})\lambda_i{}' + \theta_i \rightarrow$$

$$x_i = x_i\omega_i\lambda_i{}' + \varepsilon_i\omega_{i\varepsilon}\lambda_i{}' + \theta_i. \tag{20}$$

Applying covariance properties to (20), we obtain (21), where $diag\left(\Sigma_{x_i\theta_i}\right)$ is the matrix of the diagonal elements of $\Sigma_{x_i\theta_i}$, and the superscript $+$ denotes the application of the Moore–Penrose pseudoinverse transformation; e.g., $\lambda_i{}'^{+}$ is the Moore–Penrose pseudoinverse of $\lambda_i{}'$.

$$\Sigma_{x_ix_i} = \Sigma_{x_ix_i}\omega_i\lambda_i{}' + \Sigma_{x_i\varepsilon_i}\omega_{i\varepsilon}\lambda_i{}' + \Sigma_{x_i\theta_i} \rightarrow$$

$$\Sigma_{x_ix_i} = \Sigma_{x_ix_i}\omega_i\lambda_i{}' + diag\left(\Sigma_{x_i\theta_i}\right) \rightarrow$$

$$\Sigma_{x_ix_i}\omega_i\lambda_i{}' = \Sigma_{x_ix_i} - diag\left(\Sigma_{x_i\theta_i}\right) \rightarrow$$

$$\omega_i\lambda_i{}' = \Sigma_{x_ix_i}{}^{-1}\left(\Sigma_{x_ix_i} - diag\left(\Sigma_{x_i\theta_i}\right)\right) \rightarrow$$

$$\omega_i = \Sigma_{x_ix_i}{}^{-1}\left(\Sigma_{x_ix_i} - diag\left(\Sigma_{x_i\theta_i}\right)\right)\lambda_i{}'^{+}. \tag{21}$$

Equation (21) expresses the column vector of weights in terms of loadings and correlations among indicators and indicator error terms. Clearly it incorporates significantly more information about how weights and loadings are interrelated than the PLSA method does. For example, this equation suggests that the relationship between weights and loadings is likely to be nonlinear. With weights we can then obtain the associated composites according to

$$C_i\omega_{iC} = x_i\omega_i \rightarrow$$

$$C_i = \frac{1}{\omega_{iC}} x_i\omega_i. \tag{22}$$

Applying covariance properties to (5) we arrive at (23), which expresses the column vector of loadings in terms of the associated composites and indicators.

$$F_i = C_i\omega_{iC} + \varepsilon_i\omega_{i\varepsilon} \rightarrow$$

$$\Sigma_{F_i x_i} = \Sigma_{C_i x_i}\omega_{iC} + \Sigma_{x_i\varepsilon_i}\omega_{i\varepsilon} \rightarrow$$

$$\Sigma_{F_i x_i} = \Sigma_{C_i x_i}\omega_{iC} \rightarrow$$

$$\lambda_i = \left({C_i}^{+} x_i\right)' \omega_{iC}. \tag{23}$$

The equations above provide all of the elements based on which we can iteratively estimate the composites, one of the main goals of FSEM's first stage. We start by setting weights and loadings as 1, and initializing the composite estimates with a standardized vector of the summed indicators. We set weights and loadings as 1 based on the assumption that all of the indicators measure the factors in a direct, as opposed to reversed, fashion. We can ensure that this assumption is met by appropriately adjusting indicators that measure factors in a reversed fashion.

It is important to emphasize that, in this first stage of our FSEM method, we do not need the measurement errors to incorporate variation from opposite composites and measurement errors, which the true measurement errors do incorporate. However, we do need the measurement errors to be uncorrelated with their adjacent composites, and thus with their adjacent indicators.

We then proceed to iteratively estimate factors, indicator error terms, weights, composites and loadings according to (24)-(28). The iterations continue until the loading estimates stored in the column vector $\hat{\lambda}_i$ change by less than a small fraction.

$$\hat{F}_i \coloneqq Stdz\left(\hat{C}_i\widehat{\omega}_{iC} + \hat{\varepsilon}_i\widehat{\omega}_{i\varepsilon}\right). \tag{24}$$

$$\hat{\theta}_i \coloneqq x_i - \hat{F}_i{\hat{\lambda}_i}'. \tag{25}$$

$$\widehat{\omega}_i \coloneqq {\Sigma_{x_i x_i}}^{-1}\left(\Sigma_{x_i x_i} - diag\left(\Sigma_{x_i\hat{\theta}_i}\right)\right){\hat{\lambda}_i}'^{+}. \tag{26}$$

$$\hat{C}_i \coloneqq \frac{1}{\widehat{\omega}_{iC}}(x_i \widehat{\omega}_i). \tag{27}$$

$$\hat{\lambda}_i \coloneqq \left(\hat{C}_i^{\ +} x_i\right)' \widehat{\omega}_{iC}. \tag{28}$$

The above steps yield estimates that are used in the next stage, where estimates of measurement errors and factors are produced.

6. Stage 2: Estimation of factors

While setting the foundations of measurement error theory, Nunnally & Bernstein (1994) demonstrated that the correlation between any pair of factors $\Sigma_{F_iF_j}$ is related to the correlation between their corresponding true composites $\Sigma_{C_iC_j}$ according to (29), where α_i and α_j are the true reliabilities associated with the factors.

$$\Sigma_{F_iF_j} = \frac{\Sigma_{C_iC_j}}{\sqrt{\alpha_i \alpha_j}}. \tag{29}$$

Since we have estimates of the true composites, this allows us to estimate the elements of the matrix of correlations among factors $\hat{\Sigma}_{FF}$. This matrix will play a key role in this stage, as we will use it as the end point for a fitting algorithm that can be seen as a nonparametric version of the expectation-maximization algorithm (Dempster et al., 1977).

At this point we need to understand how each measurement error and factor varies with respect to other relevant variables in the model. With (5) expressed in terms of any pair of factors we obtain (30). We can also express the measurement error in terms of its factor and a correlation error term $\xi_{i\varepsilon}$, as indicated in (31).

$$F_i = C_i\omega_{iC} + \varepsilon_i\omega_{i\varepsilon},\ F_j = C_j\omega_{jC} + \varepsilon_j\omega_{j\varepsilon},\ F_i = \Sigma_{F_iF_j}F_j + \delta_{ij} \rightarrow$$

$$C_i\frac{\omega_{iC}}{\omega_{i\varepsilon}} + \varepsilon_i = \frac{1}{\omega_{i\varepsilon}}\Sigma_{F_iF_j}\left(C_j\omega_{jC} + \varepsilon_j\omega_{j\varepsilon}\right) + \frac{\delta_{ij}}{\omega_{i\varepsilon}}. \tag{30}$$

$$\varepsilon_i = F_i\omega_{i\varepsilon} + \xi_{i\varepsilon}. \tag{31}$$

We can see that, for each pair of factors F_i and F_j, the measurement error ε_i receives variation from its opposite composite C_j and measurement error ε_j, proportionally to (32). It also receives variation from its adjacent factor F_i, in proportion to (33). Additionally, from (5) we can see that the factor F_i receives variation from its adjacent composite C_i, in proportion to (34).

$$\frac{1}{\omega_{i\varepsilon}}\Sigma_{F_iF_j}\left(C_j\omega_{jC} + \varepsilon_j\omega_{j\varepsilon}\right). \tag{32}$$

$$F_i\omega_{i\varepsilon}. \tag{33}$$

$$C_i\omega_{iC}. \tag{34}$$

The measurement error ε_i should not share any variation with its adjacent composite C_i. Any shared variation between ε_i and C_i, which may exist due to both variables sharing variation with their adjacent factor F_i, should be removed, also proportionally to (34).

We also know, based on our discussion of the common factor model and measurement error theory presented earlier, that the measurement error is uncorrelated with its adjacent composite (as noted above), that the correlation between a factor and its composite equals the composite's weight, and that the correlation between a factor and its measurement error equals the measurement error's weight:

$$\Sigma_{C_i\varepsilon_i} = 0. \tag{35}$$

$$\Sigma_{F_iC_i} = \omega_{iC}. \tag{36}$$

$$\Sigma_{F_i\varepsilon_i} = \omega_{i\varepsilon}. \tag{37}$$

The equations above provide the mathematical foundation on which we can iteratively estimate measurement error terms and factors that incorporate all of the variation necessary for the consistent estimation of model parameters.

We now proceed to estimate the matrix of correlations among factors $\hat{\Sigma}_{F_iF_j}$ through (38), which follows from (29). This matrix will be our target in this second stage, to which we will iteratively fit the matrix of correlations among estimated factors, whose elements are denoted by $\Sigma_{\hat{F}_i\hat{F}_j}$. At this point we also initialize the factors as indicated in (39).

$$\hat{\Sigma}_{F_iF_j} \coloneqq \frac{\Sigma_{\hat{C}_i\hat{C}_j}}{\sqrt{\hat{\alpha}_i\hat{\alpha}_j}}. \tag{38}$$

$$\hat{F}_i \coloneqq Stdz\left(\hat{C}_i\widehat{\omega}_{iC} + \hat{\varepsilon}_i\widehat{\omega}_{i\varepsilon}\right). \tag{39}$$

Next each measurement error $\hat{\varepsilon}_i$ and factor $\hat{F}_i$ are iteratively adjusted by having variation added to or removed from them as indicated in (40)-(42). The main sources of the variation are $\hat{C}_j$, $\hat{\varepsilon}_j$, $\hat{C}_i$ and $\hat{F}_i$. Clearly $\hat{F}_j$ is also a source of variation, via $\hat{C}_j$ and $\hat{\varepsilon}_j$, with this already being incorporated in (40). The variation is added or removed according to (32)-(34).

$$\hat{\varepsilon}_i \coloneqq Stdz\left[\hat{\varepsilon}_i + \left(\hat{\Sigma}_{F_iF_j} - \Sigma_{\hat{F}_i\hat{F}_j}\right)\frac{1}{\widehat{\omega}_{i\varepsilon}}\hat{\Sigma}_{F_iF_j}\left(\hat{C}_j\widehat{\omega}_{jC} + \hat{\varepsilon}_j\widehat{\omega}_{j\varepsilon}\right)\right]. \tag{40}$$

$$\hat{F}_i \coloneqq Stdz\left[\hat{F}_i + \left(\widehat{\omega}_{iC} - \Sigma_{\hat{F}_i\hat{C}_i}\right)\hat{C}_i\widehat{\omega}_{iC}\right]. \tag{41}$$

$$\hat{\varepsilon}_i \coloneqq Stdz\left[\hat{\varepsilon}_i - \Sigma_{\hat{C}_i\hat{\varepsilon}_i}\hat{C}_i\widehat{\omega}_{iC} + \left(\widehat{\omega}_{i\varepsilon} - \Sigma_{\hat{F}_i\hat{\varepsilon}_i}\right)\hat{F}_i\widehat{\omega}_{i\varepsilon}\right]. \tag{42}$$

The first assignment equation above adds or removes variation in $\hat{\varepsilon}_i$ with respect to the opposite variables, via the estimates $\hat{C}_j$ and $\hat{\varepsilon}_j$. Whether variation is added or removed depends on the sign of $\left(\hat{\Sigma}_{F_iF_j} - \Sigma_{\hat{F}_i\hat{F}_j}\right)$, with the end goal being that $\hat{\Sigma}_{F_iF_j} = \Sigma_{\hat{F}_i\hat{F}_j}$ for all pairs of factors.

The second assignment equation above adds or removes variation in $\hat{F}_i$ with respect to the adjacent composite $\hat{C}_i$. Whether variation is added or removed depends on the sign of $\left(\widehat{\omega}_{iC} - \Sigma_{\hat{F}_i\hat{C}_i}\right)$, with the end goal being that $\Sigma_{\hat{F}_i\hat{C}_i} = \widehat{\omega}_{iC}$ for all factors.

The third assignment equation above adds or removes variation in $\hat{\varepsilon}_i$ with respect to the adjacent variables, via the estimates $\hat{C}_i$ and $\hat{F}_i$. Similarly, whether variation is added or removed depends on the signs of $-\Sigma_{\hat{C}_i\hat{\varepsilon}_i}$ and $\left(\widehat{\omega}_{i\varepsilon} - \Sigma_{\hat{F}_i\hat{\varepsilon}_i}\right)$, with the end goals being that $\Sigma_{\hat{C}_i\hat{\varepsilon}_i} = 0$ and $\Sigma_{\hat{F}_i\hat{\varepsilon}_i} = \widehat{\omega}_{i\varepsilon}$ for all factors. Note that the term $-\Sigma_{\hat{C}_i\hat{\varepsilon}_i}$ appears to be an exception here, but it is not since it stands for $\left(0 - \Sigma_{\hat{C}_i\hat{\varepsilon}_i}\right)$.

In each iteration in this second stage, after the steps above are carried out, estimates of factors and measurement errors are obtained according to (43)-(44).

$$\hat{F}_i \coloneqq Stdz\left(\hat{C}_i\widehat{\omega}_{iC} + \hat{\varepsilon}_i\widehat{\omega}_{i\varepsilon}\right). \tag{43}$$

$$\hat{\varepsilon}_i \coloneqq Stdz\left[\frac{1}{\widehat{\omega}_{i\varepsilon}}\left(\hat{F}_i - \hat{C}_i\widehat{\omega}_{iC}\right)\right]. \tag{44}$$

These iterative steps in Stage 2 are conducted until the sum of the absolute differences $\hat{\Sigma}_{F_iF_j} - \Sigma_{\hat{F}_i\hat{F}_j}$ falls below a small fraction, or until the sum of the absolute differences between successive estimates $\Sigma_{\hat{F}_i\hat{F}_j}$ changes by less than a small fraction. When convergence is achieved, final estimates of the composites, weights and loadings are generated according to (45)-(47).

$$\hat{C}_i \coloneqq Stdz\left[\frac{1}{\widehat{\omega}_{iC}}\left(\hat{F}_i - \hat{\varepsilon}_i\widehat{\omega}_{i\varepsilon}\right)\right]. \tag{45}$$

$$\widehat{\omega}_i \coloneqq {x_i}^{+}\hat{C}_i\widehat{\omega}_{iC}. \tag{46}$$

$$\hat{\lambda}_i \coloneqq {x_i}'{\hat{F}_i}'^{+}. \tag{47}$$

Once convergence is achieved at the outermost loop, which envelops stages 1 and 2, path coefficients can then be generated by solving (48) for $\hat{\beta}_{ij}$, which follows from (8), and where $\hat{\beta}_{ij}$ is the estimated standardized partial regression (a.k.a. path) coefficient associated with the criterion-predictor relationship between $\hat{F}_i$ and $\hat{F}_j$, N_i is the number of predictors pointing at F_i in the model, $\hat{\zeta}_i$ is the structural residual accounting for the variance in $\hat{F}_i$ that is not explained by the estimates of the factors that point at it in the model.

$$\hat{F}_i = \sum_{j=1}^{N_i} \hat{\beta}_{ij} \hat{F}_j + \hat{\zeta}_i. \tag{48}$$

Through the FSEM method described above we obtain estimates of factors, measurement errors, composites, loadings and weights that can arguably serve as the basis for the estimation of any other SEM model parameter. We also obtain path estimates, which like weights and loadings, are expected to be consistent under common factor model assumptions. Moreover, the FSEM method is essentially nonparametric, making no data or parameter distribution assumptions.

7. Monte Carlo experiment

A Monte Carlo experiment (Paxton et al., 2001) was conducted to test FSEM's performance in terms of its ability to reproduce the true values of path coefficients and loadings, and also in terms of the standard errors of the estimates of these true values. The Monte Carlo experiment was conducted with MATLAB.

Assessing performance in terms of standard errors of estimates is important because a method that is more complex may induce more variation in its estimates of model parameters, which would in turn reduce the method's statistical power (i.e., its ability to avoid false negatives). As we have seen in the previous sections, FSEM is a more computationally complex method than PLSA.

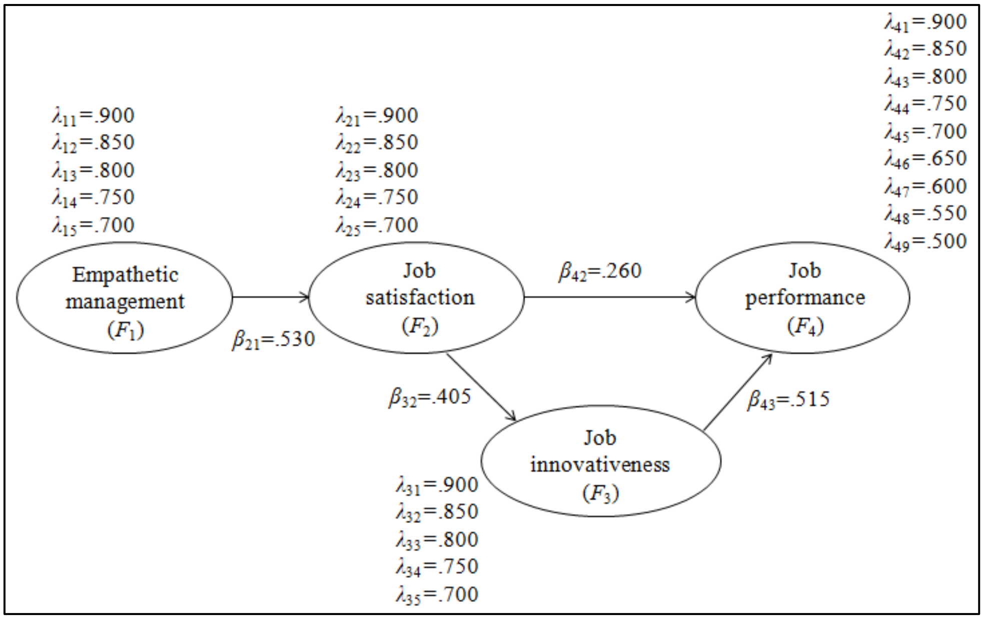
We created 1,000 samples, 500 of normal data and 500 of non-normal data, based on the true population model in Figure 2 for each of the following sample sizes: 50, 100, 200, 500 and 1,000. Kock's (2016) error-based technique for proper non-normality propagation was employed in the creation of the non-normal data samples.

Estimates of skewness and excess kurtosis were obtained for all factors and indicators in all samples analyzed. In the non-normal data, skewness ranged from -.286 to 5.109 and excess kurtosis from -1.260 to 46.923. Two tests of normality were conducted for each factor and indicator in each non-normal sample to ensure that non-normality propagation from factors and error terms to indicators occurred as expected. The tests used were the classic Jarque-Bera test (Jarque & Bera, 1980; Bera & Jarque, 1981) and Gel & Gastwirth's (2008) robust modification of this test. Both tests indicated statistically significantly non-normality in all non-normal sample factors and indicators.

Our true population model is based on an actual empirical study of the overall effect of empathetic management (F_1) on job performance (F_4), via intermediate effects on job satisfaction (F_2) and job innovativeness (F_3). While the results of the base empirical study are not our focus here, readers may find useful to know that those results suggest that employee performance is significantly associated with the degree of use by supervisors of a management style that demonstrates care about the employees' well being.

Factor and error scores were generated directly based on the true population model, and indicator scores were subsequently generated based on those factor and error scores. With the non-normal data, both factors and errors were created based on independent non-normal distributions. Each sample was analyzed with PLSA and FSEM.

Figure 2. True population model



Notes:

- β_{ij} = true population coefficient for path pointing from factor F_j to factor F_i.
- λ_{ij} = true population loading for the *j*th indicator of factor F_i.

Samples were screened for instances of multicollinearity and Simpson's paradox, as well as for Heywood cases. Based on this screening, 21 samples of normal data and 17 samples of non-normal data were discarded. All of these discarded samples occurred at N=50. As a result, a total of 479 normal and 483 non-normal data samples of size 50 were analyzed, of the original 1,000 created. No samples had to be discarded at sample sizes greater than 50.

Tables 1 and 2 show a summarized set of results, for sample sizes 50 and 500 only. True values, mean parameter estimates, and standard errors are shown next to one another. Results for all four structural paths in the model are shown.

Loadings only for empathetic management (F_1) are shown, to avoid crowding, because the pattern of results in terms of loading biases and standard errors repeats itself for all factors. Complete results are available in appendices A and B.

These summarized results illustrate the performance of the FSEM method vis-à-vis PLSA. The path coefficient and loading estimates produced by FSEM were virtually unbiased at N=100 and above with normal data, and at N>200 with non-normal data. Broadly speaking, FSEM yielded path coefficient and loading estimates that were closer to the true values for small and large sample sizes, and with normal and non-normal data. As expected, some degradation of performance occurred with non-normal data, with FSEM generally performing better than PLSA.

Figure 3 provides a further summarization of the results, in terms of root-mean-square errors (RMSEs) for path coefficients and loadings. These RMSEs have been calculated based on the true and mean path coefficients and loadings summarized through the tables above. They visually highlight the better performance of FSEM compared with PLSA, arguably in remarkable fashion.

Table 1: Summarized Monte Carlo experiment results for path coefficients

	Normal data				Non-normal data			
	PLSA	FSEM	PLSA	FSEM	PLSA	FSEM	PLSA	FSEM
N	50	50	500	500	50	50	500	500
β_{21}	.530	.530	.530	.530	.530	.530	.530	.530
$\hat{\beta}_{21}$	.498	.522	.482	.528	.492	.519	.482	.529
$SE(\hat{\beta}_{21})$	.092	.102	.056	.032	.102	.109	.056	.032
β_{32}	.405	.405	.405	.405	.405	.405	.405	.405
$\hat{\beta}_{32}$	.372	.392	.369	.406	.376	.395	.367	.403
$SE(\hat{\beta}_{32})$	.120	.127	.051	.039	.115	.121	.052	.038
β_{42}	.260	.260	.260	.260	.260	.260	.260	.260
$\hat{\beta}_{42}$	.260	.266	.255	.262	.258	.266	.252	.259
$SE(\hat{\beta}_{42})$	.109	.118	.036	.040	.116	.125	.038	.041
β_{43}	.515	.515	.515	.515	.515	.515	.515	.515
$\hat{\beta}_{43}$	.483	.507	.473	.518	.480	.503	.474	.518
$SE(\hat{\beta}_{43})$	.104	.114	.051	.034	.101	.106	.053	.037

Notes: β_{ij} = true path coefficient; $\hat{\beta}_{ij}$ = mean path coefficient estimate; $SE(\cdot)$ = standard error of estimate (a.k.a. root-mean-square error and standard deviation of estimate).

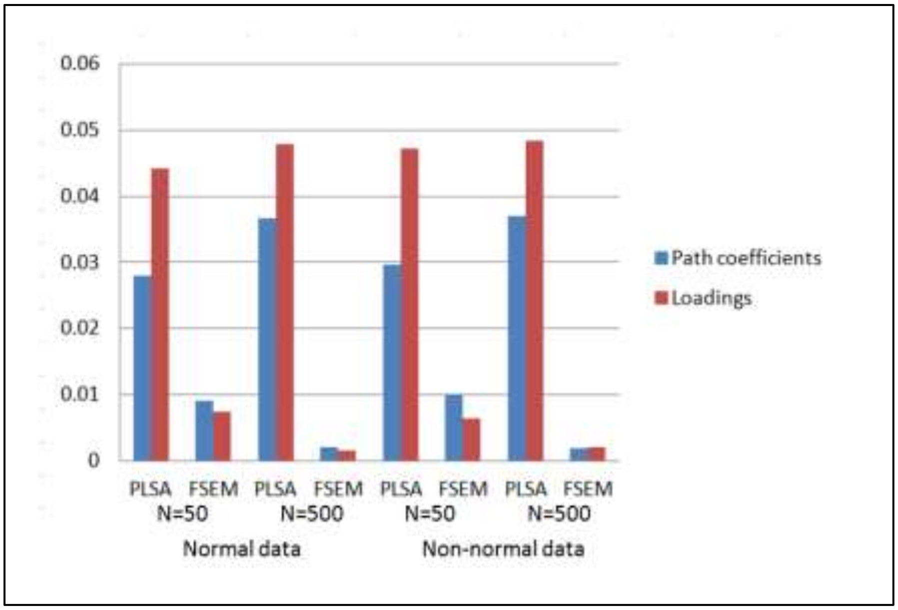
Table 2: Summarized Monte Carlo experiment results for empathetic management (F_1) loadings

	Normal data				Non-normal data			
	PLSA	FSEM	PLSA	FSEM	PLSA	FSEM	PLSA	FSEM
N	50	50	500	500	50	50	500	500
λ_{11}	.900	.900	.900	.900	.900	.900	.900	.900
$\hat{\lambda}_{11}$	.905	.891	.907	.897	.906	.890	.907	.897
$SE(\hat{\lambda}_{11})$	.020	.034	.009	.011	.021	.036	.009	.011
λ_{12}	.850	.850	.850	.850	.850	.850	.850	.850
$\hat{\lambda}_{12}$	.876	.843	.879	.850	.875	.841	.878	.849
$SE(\hat{\lambda}_{12})$	.038	.046	.030	.014	.035	.046	.029	.015
λ_{13}	.800	.800	.800	.800	.800	.800	.800	.800
$\hat{\lambda}_{13}$	.842	.796	.847	.801	.844	.796	.846	.800
$SE(\hat{\lambda}_{13})$	.055	.056	.048	.017	.055	.052	.047	.017
λ_{14}	.750	.750	.750	.750	.750	.750	.750	.750
$\hat{\lambda}_{14}$	.809	.747	.810	.751	.809	.753	.810	.752
$SE(\hat{\lambda}_{14})$	.075	.065	.062	.020	.075	.067	.061	.020
λ_{15}	.700	.700	.700	.700	.700	.700	.700	.700
$\hat{\lambda}_{15}$	.762	.689	.769	.701	.771	.701	.772	.703
$SE(\hat{\lambda}_{15})$	.086	.078	.071	.023	.089	.071	.074	.023

Notes λ_{ij} = true loading; $\hat{\lambda}_{ij}$ = mean loading estimate; $SE(\cdot)$ = standard error of estimate (a.k.a. root-mean-square error and standard deviation of estimate).

The RMSEs for path coefficients provide an aggregated view of the results for the four path coefficients shown in the first table. Analogously, the RMSEs for loadings provide an aggregated view of the results for the five loadings show in the second table. Like the tables, the RMSEs are grouped based on sample size and data distribution (normal versus non-normal).

Figure 3. Root-mean-square errors for path coefficients and loadings



Somewhat surprisingly, given that FSEM is significantly more computationally complex than PLSA, relatively small standard errors were generated via FSEM. Standard errors were typically higher for FSEM at *N*=50, and lower at higher sample sizes.

The slightly greater standard errors at *N*=50 were generally offset by the less biased path coefficients generated by FSEM, which also tended to be of greater magnitude than those generated by PLSA. This generally bodes well in terms of statistical power. Greater standard errors would tend to decrease statistical power in the absence of any bias correction. In FSEM's case, the bias correction appears to make up for the standard error increase at *N*=50.

Both methods converged to viable solutions in all samples, including those that were not deemed usable (i.e., the samples that were discarded). On average FSEM converged after 10 iterations, counted as iterations within each of the stages – the first stage, which conducts the estimation of composites; and the second stage, which conducts the estimation of factors. PLSA converged over twice as fast for this model, after 4 iterations on average.

8. Discussion and conclusion

The positive results from the Monte Carlo experiment, and the fact that the equations used in estimations are consistent with the common factor model and measurement error theory, give us confidence that the FSEM method is capable of generating asymptotically unbiased model parameters. Of course, this should apply to models that meet the assumptions of the common factor model and measurement error theory. Much more research needs to be conducted to see how deviations from the those assumptions influence model parameter bias in FSEM.

The two-stage estimation process underlying the FSEM method could be seen as combining elements of the classic PLSA and covariance-based SEM (CSEM) iterative estimation processes. Like PLSA, the FSEM method generates factor scores, with the difference that FSEM provides estimates of the true factor scores while PLSA uses composites as factor approximations. This difference accounts for the FSEM estimates being asymptotically unbiased, with those yielded by PLSA being asymptotically biased. In PLSA path coefficients tend to be underestimated, and loadings overestimated, as sample sizes grow to infinity. While CSEM yields asymptotically unbiased estimates of path coefficients and loadings, in it factors are not estimated as part of the iterative parameter estimation process.

Like CSEM, FSEM accounts for measurement error. However, FSEM does so explicitly, estimating measurement error terms as part of the estimation process. CSEM accounts for measurement error implicitly, as it does not iteratively estimate measurement error terms. (Measurement error terms should not to be confused with endogenous factor or indicator residuals, both of which are estimated by CSEM.) Like CSEM, FSEM minimizes differences between model-implied and empirical covariance matrices, with the difference that CSEM focuses on matrices calculated based on indicators, and FSEM on matrices calculated based on factors. These key differences are summarized in Table 3.

Table 3: FSEM versus classic SEM approaches

	FSEM	**PLSA**	**CSEM**
Factor estimation	Estimates the true factors during the iterative convergence process.	Estimates composites during the iterative convergence process, as approximations of factors.	Does not produce factor estimates during the iterative process.
Measurement error	Explicitly accounts for measurement error.	Does not account for measurement error.	Implicitly accounts for measurement error.
Minimization criterion	Explicitly minimizes the difference between factor covariance matrices.	Implicitly minimizes the variance explained by measurement error.	Explicitly minimizes the difference between indicator covariance matrices.
Path coefficient and loading bias	Produces asymptotically unbiased estimates of path coefficients and loadings.	Tends to underestimate path coefficients and overestimate loadings.	Produces asymptotically unbiased estimates of path coefficients and loadings.

We see our study as a first step in the development of asymptotically unbiased composite-to-factor methods and related algorithms that can give researchers access to estimates of any model parameter. Further tests of the FSEM method must be conducted for us to understand its limitations and how to properly address them. Tests with more complex models are needed and planned for the future, and are also suggested to other methodological investigators as future research.

A standardized variation sharing approach, used in the second stage or the estimation of factors, serves as a pillar for the FSEM method. To the best of our knowledge, this rather simple approach is novel, although it could be seen as a "soft" form of the expectation-maximization algorithm used in CSEM. Through it measurement errors and factors receive the variation that was not, by definition, incorporated into their corresponding true composites.

The FSEM method described in this study allows researchers to obtain estimates of factors, composites, loadings and weights that can serve as the basis for the estimation of any derivative SEM model parameter. This includes a variety of fit indices. For instance, given that the elements of the indicator errors matrix $\hat{\theta}$ are essentially uncorrelated disturbances, and that the elements of the matrices of factor scores $\hat{F}$ and loadings $\hat{\lambda}$ are estimated, the model-implied indicator correlation matrix can consequently be easily estimated.

Corresponding covariance matrices can be obtained via unstandardization of indicators, based on the indicators' means and standard deviations. Therefore, versions of the classic covariance-based SEM fit indices can be obtained by employing this model-implied indicator correlation matrix $\hat{\Sigma}_{XX}$ in conjunction with the actual indicator correlation matrix Σ_{XX}. Classic covariance-based SEM fit indices are essentially a measure of the extent to which $\hat{\Sigma}_{XX}$ fits Σ_{XX}.

If in the future fit indices based on covariance matrices are developed based on FSEM, which we consider a worthy pursuit and recommend, it seems reasonable to focus on the fit between $\Sigma_{\hat{F}\hat{F}}$ and $\hat{\Sigma}_{FF}$, since the method explicitly attempts to fit these two matrices. These are, respectively, the model-implied matrix of correlations among factors (estimated based on model parameters), and the more direct sample estimate of the population matrix generated at the beginning of FSEM's second stage.

In addition to allowing researchers to estimate any model parameter, the FSEM method is designed to be nonparametric, making no data or parameter distribution assumptions. We believe that these characteristics make the FSEM method a worthy addition to the portfolio of SEM techniques available to researchers.

References

Adelman, I., & Lohmoller, J.-B. (1994). Institutions and development in the nineteenth century: A latent variable regression model. *Structural Change and Economic Dynamics*, 5(2), 329-359.

Bera, A.K., & Jarque, C.M. (1981). Efficient tests for normality, homoscedasticity and serial independence of regression residuals: Monte Carlo evidence. *Economics Letters*, 7(4), 313-318.

Cronbach, L.J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297–334.

Dempster, A.P., Laird, N.M., & Rubin, D.B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, 39 (1), 1-38.

Dijkstra, T.K., & Schermelleh-Engel, K. (2014). Consistent partial least squares for nonlinear structural equation models. *Psychometrika*, 79(4), 585-604.

Dillon, W.R., & Goldstein, M.. (1984). *Multivariate analysis: Methods and applications*. New York, NY: Wiley.

Fornell, C., & Bookstein, F.L. (1982). Two structural equation models: LISREL and PLS applied to consumer exit-voice theory. *Journal of Marketing Research*, 19(4), 440-452.

Gel, Y.R., & Gastwirth, J.L. (2008). A robust modification of the Jarque–Bera test of normality. *Economics Letters*, 99(1), 30-32.

Goodhue, D.L., Lewis, W., and Thompson, R. (2012). Does PLS have advantages for small sample size or non-normal data? *MIS Quarterly*, 36(3), 981-1001.

Hair, J.F., Ringle, C.M., & Sarstedt, M. (2011). PLS-SEM: Indeed a silver bullet. *The Journal of Marketing Theory and Practice*, 19(2), 139-152.

Henseler, J. (2010). On the convergence of the partial least squares path modeling algorithm. *Computational Statistics*, 25(1), 107-120.

Henseler, J., & Chin, W.W. (2010). A comparison of approaches for the analysis of interaction effects between latent variables using partial least squares path modeling. *Structural Equation Modeling*, 17(1), 82-109.

Jarque, C.M., & Bera, A.K. (1980). Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics Letters*, 6(3), 255-259.

Kline, R.B. (2010). *Principles and practice of structural equation modeling*. New York, NY: The Guilford Press.

Kock, N. (2014). Advanced mediating effects tests, multi-group analyses, and measurement model assessments in PLS-based SEM. *International Journal of e-Collaboration*, 10(3), 1-13.

Kock, N. (2015). One-tailed or two-tailed P values in PLS-SEM? *International Journal of e-Collaboration*, 11(2), 1-7.

Kock, N. (2016). Non-normality propagation among latent variables and indicators in PLS-SEM simulations. *Journal of Modern Applied Statistical Methods*, 15(1), 299-315.

Kock, N., & Lynn, G.S. (2012). Lateral collinearity and misleading results in variance-based SEM: An illustration and recommendations. *Journal of the Association for Information Systems*, 13(7), 546-580.

Kock, N., & Mayfield, M. (2015). PLS-based SEM algorithms: The good neighbor assumption, collinearity, and nonlinearity. *Information Management and Business Review*, 7(2), 113-130.

Kock, N., & Moqbel, M. (2016). Statistical power with respect to true sample and true population paths: A PLS-based SEM illustration. *International Journal of Data Analysis Techniques and Strategies*, 8(4), 316-331.

Lohmöller, J.-B. (1989). *Latent variable path modeling with partial least squares*. Heidelberg, Germany: Physica-Verlag.

McIntosh, C. N., Edwards, J. R., & Antonakis, J. (2014). Reflections on partial least squares path modeling. *Organizational Research Methods*, 7(2), 210-251.

McDonald, R.P. (1996). Path analysis with composite variables. *Multivariate Behavioral Research*, 31(2), 239-27.

Nunnally, J.C., & Bernstein, I.H. (1994). *Psychometric theory*. New York, NY: McGraw-Hill.

Paxton, P., Curran, P.J., Bollen, K.A., Kirby, J., & Chen, F. (2001). Monte Carlo experiments: Design and implementation. *Structural Equation Modeling*, 8(2), 287-312.

Peterson, R.A., & Yeolib, K. (2013). On the relationship between coefficient alpha and composite reliability. *Journal of Applied Psychology*, 98(1), 194-198.

Richter, N. F., Cepeda, G., Roldán, J. L., & Ringle, C. M. (2015). European management research using partial least squares structural equation modeling (PLS-SEM). *European Management Journal*, 33(1), 1-3.

Rönkkö, M., & Evermann, J. (2013). A critical examination of common beliefs about partial least squares path modeling. *Organizational Research Methods*, 16(3), 425-448.

Sijtsma, K. (2009). On the use, the misuse, and the very limited usefulness of Cronbach’s alpha. *Psychometrika*, 74(1), 107-120.

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

Appendix A: Full Monte Carlo results with normal data

Notes: XX>YY = link from variable XX to YY; EM = empathetic management; JI = job innovativeness; JS = job satisfaction; JP = job performance; XX1 … XXn = indicators associated with factor XX; TruePath = true path coefficient; AvgPath = mean path coefficient estimate; SEPath = standard error of path estimate; TrueLoad = true loading; AvgLoad = mean loading estimate; SELoad = standard error of loading estimate.

	PLSA	FSEM	PLSA	FSEM	PLSA	FSEM	PLSA	FSEM	PLSA	FSEM
Sample size	**50**	**50**	**100**	**100**	**200**	**200**	**500**	**500**	**1000**	**1000**
EM>JS(TruePath)	.530	.530	.530	.530	.530	.530	.530	.530	.530	.530
EM>JS(AvgPath)	.498	.522	.491	.528	.482	.524	.482	.528	.480	.527
EM>JS(SEPath)	.092	.102	.073	.069	.067	.052	.056	.032	.054	.023
JS>JI(TruePath)	.405	.405	.405	.405	.405	.405	.405	.405	.405	.405
JS>JI(AvgPath)	.372	.392	.369	.396	.370	.404	.369	.406	.367	.404
JS>JI(SEPath)	.120	.127	.084	.082	.064	.059	.051	.039	.045	.026
JS>JP(TruePath)	.260	.260	.260	.260	.260	.260	.260	.260	.260	.260
JS>JP(AvgPath)	.260	.266	.258	.265	.255	.261	.255	.262	.254	.262
JS>JP(SEPath)	.109	.118	.083	.091	.062	.070	.036	.040	.026	.029
JI>JP(TruePath)	.515	.515	.515	.515	.515	.515	.515	.515	.515	.515
JI>JP(AvgPath)	.483	.507	.477	.514	.472	.514	.473	.518	.471	.517
JI>JP(SEPath)	.104	.114	.083	.083	.067	.059	.051	.034	.049	.027
EM1>EM(TrueLoad)	.900	.900	.900	.900	.900	.900	.900	.900	.900	.900
EM1>EM(AvgLoad)	.905	.891	.908	.898	.906	.896	.907	.897	.907	.898
EM1>EM(SELoad)	.020	.034	.015	.023	.011	.018	.009	.011	.008	.008
EM2>EM(TrueLoad)	.850	.850	.850	.850	.850	.850	.850	.850	.850	.850
EM2>EM(AvgLoad)	.876	.843	.878	.846	.878	.850	.879	.850	.878	.849
EM2>EM(SELoad)	.038	.046	.034	.032	.030	.022	.030	.014	.029	.010
EM3>EM(TrueLoad)	.800	.800	.800	.800	.800	.800	.800	.800	.800	.800
EM3>EM(AvgLoad)	.842	.796	.844	.796	.845	.798	.847	.801	.846	.802
EM3>EM(SELoad)	.055	.056	.050	.038	.048	.028	.048	.017	.047	.012
EM4>EM(TrueLoad)	.750	.750	.750	.750	.750	.750	.750	.750	.750	.750
EM4>EM(AvgLoad)	.809	.747	.809	.747	.810	.751	.810	.751	.810	.751
EM4>EM(SELoad)	.075	.065	.067	.047	.064	.032	.062	.020	.061	.014
EM5>EM(TrueLoad)	.700	.700	.700	.700	.700	.700	.700	.700	.700	.700
EM5>EM(AvgLoad)	.762	.689	.769	.699	.770	.701	.769	.701	.771	.701
EM5>EM(SELoad)	.086	.078	.079	.052	.075	.035	.071	.023	.072	.016
JS1>JS(TrueLoad)	.900	.900	.900	.900	.900	.900	.900	.900	.900	.900
JS1>JS(AvgLoad)	.906	.891	.906	.893	.907	.897	.907	.898	.907	.897
JS1>JS(SELoad)	.020	.034	.014	.024	.011	.017	.009	.011	.008	.008
JS2>JS(TrueLoad)	.850	.850	.850	.850	.850	.850	.850	.850	.850	.850
JS2>JS(AvgLoad)	.877	.845	.878	.849	.879	.849	.878	.849	.879	.851
JS2>JS(SELoad)	.036	.046	.033	.030	.031	.021	.030	.014	.029	.010
JS3>JS(TrueLoad)	.800	.800	.800	.800	.800	.800	.800	.800	.800	.800
JS3>JS(AvgLoad)	.845	.796	.847	.802	.846	.800	.846	.802	.846	.801
JS3>JS(SELoad)	.055	.052	.052	.040	.049	.026	.047	.017	.047	.012
JS4>JS(TrueLoad)	.750	.750	.750	.750	.750	.750	.750	.750	.750	.750
JS4>JS(AvgLoad)	.809	.750	.810	.751	.809	.750	.809	.750	.811	.753
JS4>JS(SELoad)	.072	.061	.067	.044	.063	.031	.061	.019	.062	.014
JS5>JS(TrueLoad)	.700	.700	.700	.700	.700	.700	.700	.700	.700	.700
JS5>JS(AvgLoad)	.770	.700	.768	.695	.771	.700	.771	.702	.771	.703
JS5>JS(SELoad)	.086	.070	.077	.052	.075	.036	.073	.023	.072	.016
JI1>JI(TrueLoad)	.900	.900	.900	.900	.900	.900	.900	.900	.900	.900
JI1>JI(AvgLoad)	.905	.889	.907	.895	.907	.897	.907	.897	.907	.898
JI1>JI(SELoad)	.019	.034	.014	.023	.011	.016	.009	.010	.008	.007
JI2>JI(TrueLoad)	.850	.850	.850	.850	.850	.850	.850	.850	.850	.850

	PLSA	FSEM	PLSA	FSEM	PLSA	FSEM	PLSA	FSEM	PLSA	FSEM
Sample size	**50**	**50**	**100**	**100**	**200**	**200**	**500**	**500**	**1000**	**1000**
JI2>JI(AvgLoad)	.876	.843	.876	.845	.878	.849	.878	.850	.879	.850
JI2>JI(SELoad)	.036	.046	.032	.033	.031	.022	.030	.015	.029	.009
JI3>JI(TrueLoad)	.800	.800	.800	.800	.800	.800	.800	.800	.800	.800
JI3>JI(AvgLoad)	.843	.796	.845	.798	.845	.798	.846	.800	.847	.802
JI3>JI(SELoad)	.054	.055	.050	.038	.048	.026	.047	.018	.047	.012
JI4>JI(TrueLoad)	.750	.750	.750	.750	.750	.750	.750	.750	.750	.750
JI4>JI(AvgLoad)	.810	.748	.808	.750	.809	.748	.810	.751	.810	.752
JI4>JI(SELoad)	.073	.062	.066	.048	.062	.032	.062	.019	.061	.014
JI5>JI(TrueLoad)	.700	.700	.700	.700	.700	.700	.700	.700	.700	.700
JI5>JI(AvgLoad)	.769	.698	.770	.698	.767	.696	.771	.704	.772	.703
JI5>JI(SELoad)	.088	.074	.079	.050	.072	.037	.073	.022	.073	.016
JP1>JP(TrueLoad)	.900	.900	.900	.900	.900	.900	.900	.900	.900	.900
JP1>JP(AvgLoad)	.896	.885	.895	.887	.897	.891	.896	.890	.896	.890
JP1>JP(SELoad)	.019	.033	.014	.025	.009	.017	.007	.014	.006	.012
JP2>JP(TrueLoad)	.850	.850	.850	.850	.850	.850	.850	.850	.850	.850
JP2>JP(AvgLoad)	.860	.834	.860	.840	.862	.842	.862	.843	.863	.844
JP2>JP(SELoad)	.028	.045	.021	.031	.017	.021	.015	.014	.014	.011
JP3>JP(TrueLoad)	.800	.800	.800	.800	.800	.800	.800	.800	.800	.800
JP3>JP(AvgLoad)	.825	.792	.823	.788	.825	.792	.825	.794	.825	.794
JP3>JP(SELoad)	.043	.053	.033	.038	.030	.026	.027	.017	.026	.013
JP4>JP(TrueLoad)	.750	.750	.750	.750	.750	.750	.750	.750	.750	.750
JP4>JP(AvgLoad)	.778	.732	.782	.739	.785	.744	.786	.745	.785	.745
JP4>JP(SELoad)	.056	.066	.045	.044	.042	.031	.038	.019	.037	.014
JP5>JP(TrueLoad)	.700	.700	.700	.700	.700	.700	.700	.700	.700	.700
JP5>JP(AvgLoad)	.739	.687	.741	.692	.741	.693	.745	.696	.743	.695
JP5>JP(SELoad)	.065	.067	.056	.048	.049	.034	.048	.022	.045	.017
JP6>JP(TrueLoad)	.650	.650	.650	.650	.650	.650	.650	.650	.650	.650
JP6>JP(AvgLoad)	.689	.634	.698	.644	.700	.646	.699	.646	.700	.646
JP6>JP(SELoad)	.080	.083	.065	.055	.059	.038	.053	.024	.052	.019
JP7>JP(TrueLoad)	.600	.600	.600	.600	.600	.600	.600	.600	.600	.600
JP7>JP(AvgLoad)	.647	.589	.649	.593	.654	.596	.653	.596	.654	.596
JP7>JP(SELoad)	.091	.090	.070	.059	.066	.044	.058	.028	.057	.021
JP8>JP(TrueLoad)	.550	.550	.550	.550	.550	.550	.550	.550	.550	.550
JP8>JP(AvgLoad)	.600	.545	.601	.542	.602	.543	.605	.546	.606	.548
JP8>JP(SELoad)	.106	.100	.080	.069	.069	.049	.062	.032	.060	.021
JP9>JP(TrueLoad)	.500	.500	.500	.500	.500	.500	.500	.500	.500	.500
JP9>JP(AvgLoad)	.553	.500	.557	.496	.552	.492	.557	.498	.557	.499
JP9>JP(SELoad)	.115	.104	.092	.079	.072	.052	.065	.033	.061	.024

Appendix B: Full Monte Carlo results with non-normal data

Notes: XX>YY = link from variable XX to YY; EM = empathetic management; JI = job innovativeness; JS = job satisfaction; JP = job performance; XX1 … XXn = indicators associated with factor XX; TruePath = true path coefficient; AvgPath = mean path coefficient estimate; SEPath = standard error of path estimate; TrueLoad = true loading; AvgLoad = mean loading estimate; SELoad = standard error of loading estimate.

Algoritm	PLSA	FSEM	PLSA	FSEM	PLSA	FSEM	PLSA	FSEM	PLSA	FSEM
Sample size	**50**	**50**	**100**	**100**	**200**	**200**	**500**	**500**	**1000**	**1000**
EM>JS(TruePath)	.530	.530	.530	.530	.530	.530	.530	.530	.530	.530
EM>JS(AvgPath)	.492	.519	.481	.519	.484	.527	.482	.529	.481	.528
EM>JS(SEPath)	.102	.109	.084	.077	.063	.047	.056	.032	.053	.023
JS>JI(TruePath)	.405	.405	.405	.405	.405	.405	.405	.405	.405	.405
JS>JI(AvgPath)	.376	.395	.370	.399	.368	.402	.367	.403	.367	.405
JS>JI(SEPath)	.115	.121	.084	.083	.066	.059	.052	.038	.045	.027
JS>JP(TruePath)	.260	.260	.260	.260	.260	.260	.260	.260	.260	.260
JS>JP(AvgPath)	.258	.266	.257	.264	.255	.263	.252	.259	.252	.259
JS>JP(SEPath)	.116	.125	.081	.090	.058	.065	.038	.041	.027	.029
JI>JP(TruePath)	.515	.515	.515	.515	.515	.515	.515	.515	.515	.515
JI>JP(AvgPath)	.480	.503	.477	.514	.477	.518	.474	.518	.474	.520
JI>JP(SEPath)	.101	.106	.083	.083	.064	.059	.053	.037	.047	.027
EM1>EM(TrueLoad)	.900	.900	.900	.900	.900	.900	.900	.900	.900	.900
EM1>EM(AvgLoad)	.906	.890	.906	.895	.907	.897	.907	.897	.907	.898
EM1>EM(SELoad)	.021	.036	.015	.024	.012	.016	.009	.011	.008	.007
EM2>EM(TrueLoad)	.850	.850	.850	.850	.850	.850	.850	.850	.850	.850
EM2>EM(AvgLoad)	.875	.841	.877	.847	.878	.849	.878	.849	.879	.851
EM2>EM(SELoad)	.035	.046	.033	.032	.031	.022	.029	.015	.029	.010
EM3>EM(TrueLoad)	.800	.800	.800	.800	.800	.800	.800	.800	.800	.800
EM3>EM(AvgLoad)	.844	.796	.846	.798	.844	.798	.846	.800	.846	.801
EM3>EM(SELoad)	.055	.052	.052	.039	.047	.027	.047	.017	.046	.012
EM4>EM(TrueLoad)	.750	.750	.750	.750	.750	.750	.750	.750	.750	.750
EM4>EM(AvgLoad)	.809	.753	.807	.747	.811	.752	.810	.752	.810	.752
EM4>EM(SELoad)	.075	.067	.064	.043	.064	.031	.061	.020	.061	.014
EM5>EM(TrueLoad)	.700	.700	.700	.700	.700	.700	.700	.700	.700	.700
EM5>EM(AvgLoad)	.771	.701	.770	.701	.771	.701	.772	.703	.772	.703
EM5>EM(SELoad)	.089	.071	.081	.051	.075	.036	.074	.023	.073	.016
JS1>JS(TrueLoad)	.900	.900	.900	.900	.900	.900	.900	.900	.900	.900
JS1>JS(AvgLoad)	.905	.890	.906	.895	.907	.897	.907	.898	.907	.898
JS1>JS(SELoad)	.019	.034	.014	.025	.011	.017	.009	.011	.008	.008
JS2>JS(TrueLoad)	.850	.850	.850	.850	.850	.850	.850	.850	.850	.850
JS2>JS(AvgLoad)	.875	.842	.878	.847	.879	.849	.879	.850	.879	.851
JS2>JS(SELoad)	.034	.045	.032	.032	.031	.022	.029	.013	.030	.009
JS3>JS(TrueLoad)	.800	.800	.800	.800	.800	.800	.800	.800	.800	.800
JS3>JS(AvgLoad)	.845	.794	.845	.797	.845	.798	.847	.801	.846	.802
JS3>JS(SELoad)	.055	.056	.050	.038	.048	.027	.048	.017	.047	.011
JS4>JS(TrueLoad)	.750	.750	.750	.750	.750	.750	.750	.750	.750	.750
JS4>JS(AvgLoad)	.806	.747	.810	.750	.810	.752	.810	.751	.810	.751
JS4>JS(SELoad)	.072	.067	.067	.042	.063	.031	.061	.020	.061	.014
JS5>JS(TrueLoad)	.700	.700	.700	.700	.700	.700	.700	.700	.700	.700
JS5>JS(AvgLoad)	.767	.694	.767	.696	.768	.697	.772	.703	.771	.702
JS5>JS(SELoad)	.086	.075	.075	.048	.073	.036	.074	.023	.072	.016
JI1>JI(TrueLoad)	.900	.900	.900	.900	.900	.900	.900	.900	.900	.900
JI1>JI(AvgLoad)	.906	.891	.905	.893	.907	.897	.907	.897	.907	.897
JI1>JI(SELoad)	.019	.033	.014	.024	.011	.016	.009	.011	.008	.008
JI2>JI(TrueLoad)	.850	.850	.850	.850	.850	.850	.850	.850	.850	.850
JI2>JI(AvgLoad)	.877	.844	.880	.851	.879	.848	.879	.850	.879	.851

Algoritm	PLSA	FSEM	PLSA	FSEM	PLSA	FSEM	PLSA	FSEM	PLSA	FSEM
Sample size	**50**	**50**	**100**	**100**	**200**	**200**	**500**	**500**	**1000**	**1000**
JI2>JI(SELoad)	.036	.043	.034	.031	.031	.022	.030	.014	.029	.010
JI3>JI(TrueLoad)	.800	.800	.800	.800	.800	.800	.800	.800	.800	.800
JI3>JI(AvgLoad)	.841	.792	.846	.798	.845	.799	.846	.801	.847	.802
JI3>JI(SELoad)	.053	.054	.050	.036	.048	.026	.047	.017	.047	.012
JI4>JI(TrueLoad)	.750	.750	.750	.750	.750	.750	.750	.750	.750	.750
JI4>JI(AvgLoad)	.807	.745	.807	.745	.810	.750	.811	.752	.811	.752
JI4>JI(SELoad)	.071	.063	.064	.047	.063	.033	.062	.020	.061	.014
JI5>JI(TrueLoad)	.700	.700	.700	.700	.700	.700	.700	.700	.700	.700
JI5>JI(AvgLoad)	.764	.696	.769	.700	.770	.699	.771	.702	.772	.703
JI5>JI(SELoad)	.084	.074	.077	.048	.074	.037	.073	.022	.073	.016
JP1>JP(TrueLoad)	.900	.900	.900	.900	.900	.900	.900	.900	.900	.900
JP1>JP(AvgLoad)	.894	.883	.897	.888	.896	.890	.896	.891	.896	.891
JP1>JP(SELoad)	.021	.035	.012	.023	.010	.018	.007	.013	.006	.011
JP2>JP(TrueLoad)	.850	.850	.850	.850	.850	.850	.850	.850	.850	.850
JP2>JP(AvgLoad)	.859	.834	.861	.838	.861	.841	.862	.842	.862	.843
JP2>JP(SELoad)	.028	.043	.022	.032	.017	.021	.014	.015	.013	.011
JP3>JP(TrueLoad)	.800	.800	.800	.800	.800	.800	.800	.800	.800	.800
JP3>JP(AvgLoad)	.820	.785	.823	.790	.824	.791	.826	.794	.825	.793
JP3>JP(SELoad)	.040	.052	.033	.035	.029	.025	.028	.017	.026	.013
JP4>JP(TrueLoad)	.750	.750	.750	.750	.750	.750	.750	.750	.750	.750
JP4>JP(AvgLoad)	.784	.742	.785	.744	.787	.745	.785	.744	.786	.745
JP4>JP(SELoad)	.054	.059	.046	.043	.042	.029	.038	.019	.037	.014
JP5>JP(TrueLoad)	.700	.700	.700	.700	.700	.700	.700	.700	.700	.700
JP5>JP(AvgLoad)	.739	.690	.741	.690	.743	.695	.744	.696	.744	.695
JP5>JP(SELoad)	.069	.072	.056	.051	.050	.034	.047	.021	.046	.016
JP6>JP(TrueLoad)	.650	.650	.650	.650	.650	.650	.650	.650	.650	.650
JP6>JP(AvgLoad)	.692	.637	.702	.648	.699	.646	.700	.646	.700	.646
JP6>JP(SELoad)	.080	.080	.069	.055	.057	.038	.054	.025	.052	.019
JP7>JP(TrueLoad)	.600	.600	.600	.600	.600	.600	.600	.600	.600	.600
JP7>JP(AvgLoad)	.654	.596	.650	.593	.651	.594	.653	.595	.653	.595
JP7>JP(SELoad)	.092	.084	.075	.063	.064	.044	.058	.029	.055	.019
JP8>JP(TrueLoad)	.550	.550	.550	.550	.550	.550	.550	.550	.550	.550
JP8>JP(AvgLoad)	.598	.540	.599	.541	.607	.547	.605	.546	.605	.547
JP8>JP(SELoad)	.107	.103	.084	.074	.072	.047	.062	.032	.058	.022
JP9>JP(TrueLoad)	.500	.500	.500	.500	.500	.500	.500	.500	.500	.500
JP9>JP(AvgLoad)	.543	.485	.549	.491	.554	.496	.557	.500	.557	.498
JP9>JP(SELoad)	.113	.109	.088	.078	.074	.053	.066	.034	.061	.023