Statistical power with respect to true sample and true population paths: A PLS-based SEM illustration

Ned Kock
Murad Moqbel

Full reference:

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.

Abstract

Monte Carlo experiments aimed at assessing the statistical power of structural equation modeling (SEM) techniques typically focus on true population path coefficients, ignoring true sample path coefficients. We demonstrate the limitations stemming from such practice in statistical power assessments. This is done in the context of SEM techniques employing the partial least squares (PLS) method, where power claims have led to much recent debate. We show that the sample sizes at which power is greater than .8 differ significantly when we consider true population and true sample paths, and that the difference increases with decreases in the magnitudes of the paths being considered. Finally, we illustrate empirically how these differences affect the conclusions we can draw from the analysis of a relatively small sample of size 193.

Keywords: Information Systems; Statistical Power; Structural Equation Modeling; Latent Variables; Partial Least Squares; Monte Carlo Simulation

Biographical notes

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 holds degrees in electronics engineering (B.E.E.), computer science (M.S.), and management information systems (Ph.D.). Ned has authored and edited several books, including the Sage Publications book titled *Systems Analysis and Design Fundamentals: A Business Process Redesign Approach*. He has published his research in a number of high-impact journals including *Communications of the ACM*, *Decision Support Systems*, *European Journal of Information Systems*, *European Journal of Operational Research*, *IEEE Transactions (various)*, *Information & Management*, *Information Systems Journal*, *Journal of the Association for Information Systems*, *MIS Quarterly*, and *Organization Science*. He is the developer of

WarpPLS, a widely used nonlinear structural equation modeling software, and Founding Editor-in-Chief of the *International Journal of e-Collaboration*. His main research interests are biological and cultural influences on human-technology interaction, nonlinear structural equation modeling, electronic communication and collaboration, action research, ethical and legal issues in technology research and management, and business process improvement.

Murad Moqbel is an Assistant Professor of Health Information Management and Health Informatics at the University of Kansas Medical Center. He holds a Ph.D. degree in International Business Administration and Management Information Systems from Texas A&M International University. He received both a B.S. degree with honors in Business Administration and Computer Information Systems and a MBA with Information Systems concentration from Emporia State University. He is on the editorial board of the *International Journal of e-Collaboration*. He won the best student paper award at the Southwest Decision Science Conference in 2012. He has authored and co-authored several papers that appeared in: *Information Technology and People*, *Public Organization Review*, *AIS Transaction on Replication Research,* the proceedings of the *International Conference in Information Systems (ICIS),* and the proceedings of the *Americas Conference on Information Systems (AMCIS).* His research interests focus on the interaction between human behavior and information technologies including social networking, emerging technologies adoption, information security, trust and privacy, and international business.

1. Introduction

Structural equation modeling (SEM) techniques allow for the construction and analysis of path models (Wright, 1934; 1960) with “latent” variables, which are variables that are measured indirectly through other “observed” or “manifest” variables. The latter are frequently referred to as the “indicators” of the latent variables (Kline, 2010; Schumacker & Lomax, 2004).

Indicators often take the form of numeric variables that store answers to question-statements in questionnaires. Each question-statement is designed to refer to a specific latent variable, measuring it with a certain degree of imprecision. Using multiple indicators to measure a latent variable reduces measurement error.

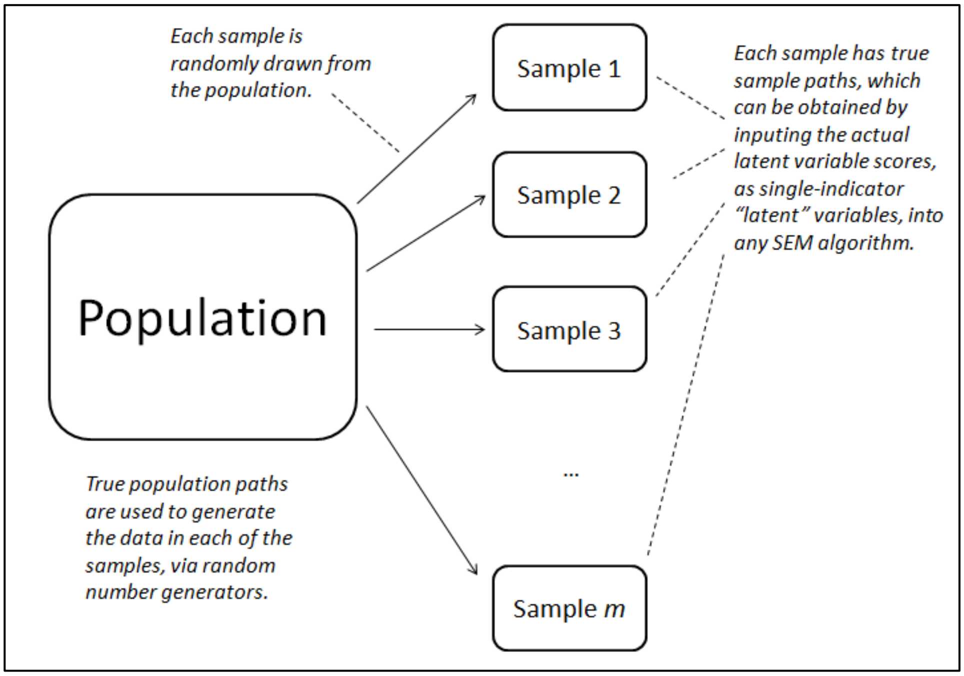
Many SEM techniques have been developed over the years, with two main classes of techniques having gained broader acceptance: covariance-based SEM and PLS-based SEM (Hair et al., 2011; Kline, 2010; Kock & Lynn, 2012). Covariance-based SEM, generally viewed as the classic form of SEM, builds on strong parametric assumptions and relies on the minimization of differences between model-implied and empirical data covariance matrices.

PLS-based SEM, also known as PLS path modeling (Henseler & Chin, 2010), is essentially nonparametric in design, and relies on the estimation of composites that stand in for factors (i.e., the latent variables). Because of its nonparametric nature, it is often claimed that PLS-based SEM can achieve acceptable statistical power with small samples and non-normal data. One of the few ways in which these claims can be tested, and the most widely used, is to conduct simulated experiments employing the Monte Carlo method (Paxton et al., 2001).

Monte Carlo experiments aimed at assessing statistical power of SEM techniques typically focus on true population path coefficients (Cassel et al., 1999; Chin et al., 2003; Chiquoine & Hjalmarsson, 2008; Ferrarii, 2011; Paxton et al., 2001; Qureshi & Compeau, 2009). These coefficients are set by the researcher, and reflect the associations among latent variables and indicators at the population level. Monte Carlo experiments aimed at assessing statistical power of SEM techniques do not usually consider true *sample* path coefficients.

In Monte Carlo experiments, nonzero true population paths above a certain threshold (e.g., .142) are used as a basis for the creation of m multiple samples (a.k.a. replications). This is illustrated in Figure 1.

Figure 1. True sample versus true population paths in a Monte Carlo experiment



Each sample is then analyzed via one or more SEM techniques, leading to rejection or acceptance of the hypotheses that refer to each path in the structural model. These analyses assume that a SEM technique of suitable power will lead to the acceptance of hypotheses associated with nonzero true population paths above a certain threshold for a certain percentage of samples; usually 80 percent. That is, a SEM technique of suitable power will lead to the avoidance of false negatives, or type II errors, 80 percent of the time.

If both indicator and latent variable scores are generated in Monte Carlo experiments, true sample paths can be obtained for each sample by using the sample latent variable scores as inputs to any SEM algorithm or software. These coefficients are the true paths with respect to each sample, which differ from true population paths due to sampling error. Nevertheless, these true sample paths are usually ignored in statistical power assessments via Monte Carlo experiments.

The focus on true population paths in Monte Carlo experiments aimed at testing SEM techniques is widespread and restricts the scope of conclusions to the general case in which: (a) a population is large and samples are comparatively small; (b) random sampling is employed; and (c) samples are small primarily because of sampling limitations; e.g., no access to all the individuals, groups etc. that make up the population.

This can be a problem in many areas of research, and can be particularly severe in some areas, such as information systems research. Information systems research is often concerned with the effects of information and communication technologies on individuals and groups. Part of our discussion builds on an empirical information systems research illustration.

For example, if a researcher decided to study the attitudes toward mobile technologies of billionaire entrepreneurs under the age of 40, the population of the study would probably be very small. In this type of situation, a traditional Monte Carlo experiment focusing on true population paths could lead to a gross overestimation of the minimum sample size required to achieve a certain level of statistical power. The resulting minimum sample size estimate may be too large to be achieved in practice, possibly leading to the abandonment of an interesting research topic.

We argue that, in this type of situation, it is important to try to understand the performance of SEM techniques that are expected to perform well with small samples, such as PLS-based SEM techniques, vis-à-vis true sample paths. When populations are rather small, as in the example above, performance vis-à-vis true sample paths would arguably be a better yardstick for the estimation of minimum required sample sizes.

Even in cases where populations are not small, an assessment of a SEM technique's performance vis-à-vis true sample paths, and the determination of a minimum sample size in this context, would allow researchers to draw conclusions that are sample-specific as opposed to population-general. That is, while a given sample size may not be sufficient for conclusions that generalize the findings of an empirical study to the entire population from which the sample is taken, the sample size in question may well be enough for conclusions that apply to very similar samples. Monte Carlo experiments considering true sample paths allow researchers to determine such a sample size.

We demonstrate here the limitations stemming from a focus on true population paths in Monte Carlo experiments. We conduct a Monte Carlo experiment aimed at testing PLS-based SEM techniques based on an illustrative model containing paths of various magnitudes, whereby we assess statistical power with respect to true population and true sample paths. We show that the sample sizes at which power is greater than .8 differ significantly when we consider true population and true sample paths, and that this difference increases with decreases in the magnitudes of the paths being considered. Finally, we illustrate empirically how these differences affect the conclusions we can draw from the analysis of a small sample.

2. Composite-based SEM: Three algorithms

PLS-based SEM is a form of composite-based SEM, whereby latent variables are estimated as exact linear combinations of their indicators (Helland, 1988). Our discussion builds on this class of SEM techniques because of their widespread use with small samples, which is often justified based on the nonparametric nature of this class of SEM techniques (Hair et al., 2011; Kock & Lynn, 2012).

Several algorithms can be used in composite-based SEM. We discuss three such algorithms in this section. Two of these are “true” PLS algorithms, whereby weights are assigned to composite indicators according to iterative procedures: PLS Regression and PLS Mode A. One is a simplified algorithm where all weights assume the same value: Path Analysis. This simplified algorithm is included in our Monte Carlo experiment later, with the goal of serving as a baseline algorithm.

All variables are assumed to be scaled to have a mean of zero and a standard deviation of 1; i.e., they are standardized. In all of the algorithms discussed here indicator weight estimates $\widehat{\omega_{ij}}$ are initially set to arbitrary values (usually 1), based on which composite estimates $\widehat{C_i}$ are initialized with standardized vectors of weighted sums of the indicators.

Then the composites are re-estimated based on (1) for PLS Mode A. In this equation the coefficients v_{ij} are referred to as the “inner weights” (Lohmöller, 1989, p. 29), and A_i is the number of composites $\widehat{C_j}$ $(j = 1 \ldots A_i)$ that are structurally adjacent to the composite $\widehat{C_i}$. The function $Stdz(\cdot)$ returns a standardized vector. Composites are structurally adjacent when they are linked by arrows in a model.

$$\widehat{C_i} = Stdz\left(\sum_{j=1}^{A_i} v_{ij}\, \widehat{C_j}\right). \tag{1}$$

In PLS Mode A the inner weights are set according to three main schemes, known as centroid, factorial, and path weighting. Inner weights are not set in PLS Regression or Path Analysis, since in these algorithms the structural model does not influence the estimation of indicator weights.

In the centroid scheme the inner weights are set as the signs of the correlations among structurally adjacent composites; in the factorial scheme, as the correlations among structurally adjacent composites; and in the path weighting scheme, as the path coefficients or correlations among structurally adjacent composites, depending on whether the arrows go in or out, respectively.

The path coefficient estimates $\widehat{\beta_{ij}}$ used in the path weighting scheme of PLS Mode A are obtained through the solution of (2) for endogenous composites; i.e., composites to which other composites point in the structural model. In this equation $\widehat{\zeta_i}$ refers to the estimated residual for each endogenous composite. This residual accounts for the variance that is not explained by the composites $\widehat{C_j}$ ($j = 1 \ldots N_i$) pointing at the endogenous composite $\widehat{C_i}$.

$$\widehat{C_i} = \sum_{j=1}^{N_i} \widehat{\beta_{ij}}\, \widehat{C_j} + \widehat{\zeta_i}. \tag{2}$$

Next (3) is solved for $\widehat{w_{ij}}$ (known as "outer weights") for PLS Regression and PLS Mode A. In this equation n_i is the numbers of indicators x_{ij} ($j = 1 \ldots n_i$) of the composite $\widehat{C_i}$, and $\widehat{\theta_{ij}}$ are the error terms for the indicators.

$$x_{ij} = \widehat{w_{ij}}\widehat{C_i} + \widehat{\theta_{ij}},\ j = 1 \ldots n_i. \tag{3}$$

Following this, the composites are re-estimated based on (4), a step known as "outside approximation". This and the preceding steps are carried out iteratively until the weight estimates $\widehat{w_{ij}}$ change by less than a small fraction.

$$\widehat{C_i} \leftarrow Stdz\left(\sum_{j=1}^{n_i} \widehat{w_{ij}}\, x_{ij}\right). \tag{4}$$

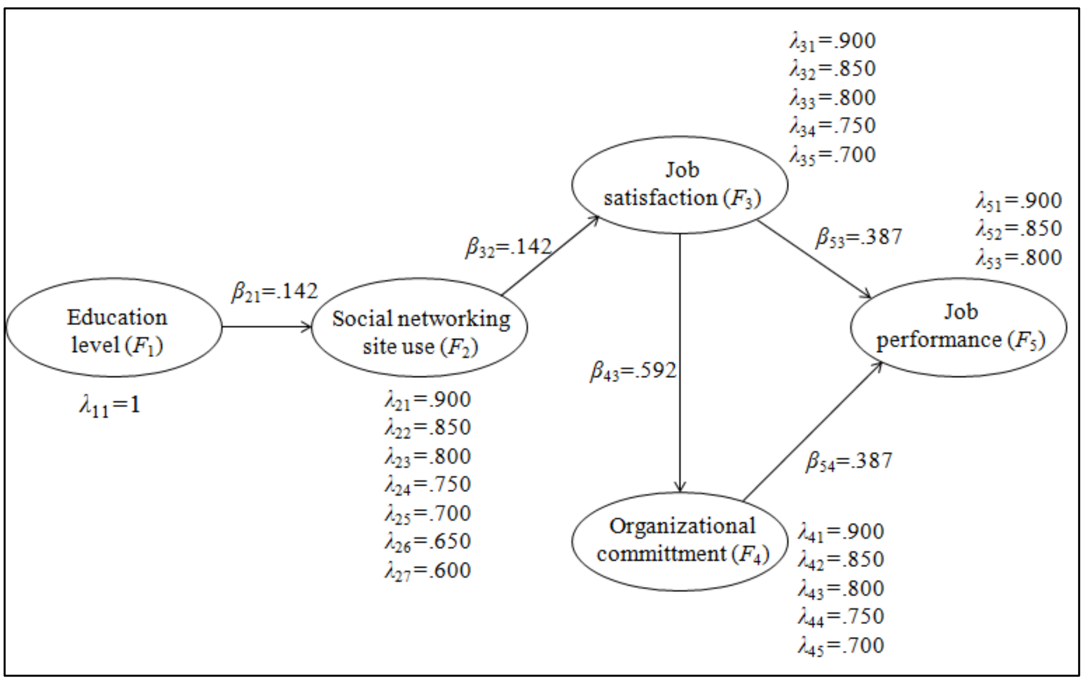
In the Path Analysis algorithm all of the weights $\widehat{w_{ij}}$ are set to 1 and composites are estimated based on (4) only once. That is, Path Analysis is a simplified non-iterative algorithm. Like PLS Regression and PLS Mode A, it also leads to the estimation of composites as exact linear combinations of their respective indicators.

After stable weights $\widehat{w_{ij}}$ are obtained, the final values of the path coefficient estimates $\widehat{\beta_{ij}}$ are obtained via (2). This applies to all of the algorithms discussed in this section. Since in PLS Regression and PLS Mode A indicator weights are proportional to indicator loadings, these algorithms are generally believed to perform better than Path Analysis when indicator loadings are unevenly distributed.

3. Creating data via the Monte Carlo method for power analyses

The power of a particular technique, such as Path Analysis, is normally assessed through a Monte Carlo experiment where data is created based on a pre-specified true population model. Such model contains true population path coefficients and loadings. To illustrate this, let us assume we have a population whose behavior in the context of social networking site use is described by the path model in Figure 2.

Figure 2. Illustrative model



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

This illustrative model is based on an actual study of the overall effect of social networking site use (F_2), such as use of Facebook, on job performance (F_5). This overall effect is hypothesized to be mediated by job satisfaction (F_3) and organizational commitment (F_4). Education level (F_1) is hypothesized to influence social networking site use (F_2). This study is revisited later, with a discussion of related empirical results.

True population path coefficients are set to assume three values ($.142$, $.387$ and $.592$), which when squared yield the thresholds for small, medium and large effect sizes ($.02$, $.15$ and $.35$) proposed by Cohen (1988; 1992). The number of indicators associated with each latent variable is based on the actual empirical study used as a basis for the simulated experiment. True population loadings associated with the indicators start with the value of $.9$ and go down in increments of $.05$. True population path coefficients and loadings are similar in magnitude to those in the empirical study used as a basis for the simulation.

For 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 ζ_i be the error variable that accounts for the variance in an endogenous latent variable F_i that is not explained by the exogenous latent variables that point at F_i. Let F_j be one of the N_i exogenous latent variables that point at an endogenous latent variable F_i. And let θ_{ij} be the standardized error variable that accounts for the variance in the indicator x_{ij} that is not explained by its latent variable F_i.

In a Monte Carlo experiment where multiple replications of a model are created, error variables and exogenous variables can be created according to equations (5) to (7). In these equations $Rndn(\mathrm{N})$ is a function that returns a different normal random variable each time it is invoked, as a vector with N elements, and $Stdz(\cdot)$ is a function that returns a standardized variable.

$$\zeta_i \leftarrow Stdz\big(Rndn(\mathrm{N})\big). \tag{5}$$

$$F_j \leftarrow Stdz\big(Rndn(\mathrm{N})\big). \tag{6}$$

$$\theta_{ij} \leftarrow Stdz\big(Rndn(\mathrm{N})\big). \tag{7}$$

The above assumes that simulated datasets that follow normal distributions are desired. To obtain non-normal data, transformations based on the normally-distributed variables can be employed. For example, equations (8) to (10) transform the normal variables into corresponding non-normal variables that follow a χ^2 distribution with 1 degree of freedom. This is well known non-normal distribution, with theoretical skewness and kurtosis (a.k.a. excess kurtosis) values of $\sqrt{8} \cong 2.828$ and 12 respectively.

$$\zeta_i \leftarrow Stdz({\zeta_i}^2). \quad (8)$$

$$F_j \leftarrow Stdz({F_j}^2). \quad (9)$$

$$\theta_{ij} \leftarrow Stdz({\theta_{ij}}^2). \quad (10)$$

Once the error variables and exogenous latent variables are created, endogenous latent variables are created based on the true population path coefficients. This is indicated in (11), where R_{ij} are the correlations among the linked latent variables. Finally, indicators are created based on the true population loadings, as indicated in (12).

$$F_i = \sum_{j=1}^{N_i} \beta_{ij} F_j + \left(\sqrt{\left(1 - \sum_{j=1}^{N_i} \beta_{ij} R_{ij}\right)}\right) \zeta_i. \quad (11)$$

$$x_{ij} = \lambda_{ij} F_i + \left(\sqrt{1 - {\lambda_{ij}}^2}\right) \theta_{ij}, j = 1 \dots n_i. \quad (12)$$

Generally a set of samples is generated for each sample size, with sample sizes varying incrementally to produce multiple sets of samples. The larger the set of samples created in connection with each sample size, the more precise is the measure of statistical power obtained via a Monte Carlo experiment for that particular sample size. Typically 250 or more replications per sample size are used.

4. Power with respect to true population paths

Let K be the set $\{\widehat{k_1}, \widehat{k_2} \dots \widehat{k_m}\}$, where $\widehat{k_j}$ is the vector of parameter estimates generated for replication j for a given model, based on an analysis technique, and m is the total number of replications. Let Φ denote the analysis technique used (e.g., Path Analysis), and N the sample size used, in replication j. Let β_P be true population path coefficient considered (e.g., $\beta_P =$ .142), and $\hat{\beta}$ the corresponding path coefficient estimated with Φ. Let K_P be the subset of K that satisfies $|\hat{\beta}| > 0$ for a particular combination of Φ, N, and β_P. The power of the technique Φ in this scenario, with respect to the true population paths that assume the value β_P, is denoted as $W_P(\Phi, \mathrm{N}, \beta_P)$ and is given by (13), where $Card(K_P)$, $Card(K)$, and $Card(m)$ are the cardinalities of K_P, K, and m, respectively.

$$W_P(\Phi, \mathrm{N}, \beta_P) = \frac{Card(K_P)}{Card(K)} \text{ as } Card(m) \to \infty. \tag{13}$$

This power estimation method relies on true population paths only, and is the most commonly used in power assessments. In it the condition expressed by the inequality $|\hat{\beta}| > 0$ is tested through an assessment of statistical significance, which can be conducted via the calculation of a chance probability (a.k.a. P value) or confidence interval.

5. Power with respect to true sample paths

Let K, Φ, N, β_P, and $\hat{\beta}$ be defined as in the section above. Let β_S be the true sample path corresponding to the true population path β_P, calculated based on the latent variable scores (F_i) for replication j. Let $\dot{K}$ be the subset of K that satisfies the condition $\beta_P \leq \beta_S < \beta'_P$, where β'_P is the next threshold greater than β_P from the set of thresholds used – in our case: $\{.142, .387, .592, +\infty\}$. Let K_S be the subset of $\dot{K}$ that satisfies $|\hat{\beta}| > 0$, for a particular combination of Φ, N, and β_P. The power of the technique Φ in this scenario, with respect to the true sample paths corresponding to β_P, is denoted as $W_S(\Phi, \mathrm{N}, \beta_P)$ and is given by (14), where $Card(K_S)$ and $Card(\dot{K})$ are the cardinalities of K_S and $\dot{K}$, respectively.

$$W_S(\Phi, \mathrm{N}, \beta_P) = \frac{Card(K_S)}{Card(\dot{K})} \text{ as } Card(m) \to \infty. \tag{14}$$

Note that generally $W_S(\Phi, \mathrm{N}, \beta_P) > W_P(\Phi, \mathrm{N}, \beta_P)$ because the calculation of $W_S(\Phi, \mathrm{N}, \beta_P)$ excludes those instances in which the condition $\beta_P \leq \beta_S < \beta'_P$ is not satisfied due to sampling error. Those instances are not excluded in the calculation of $W_P(\Phi, \mathrm{N}, \beta_P)$.

This is a novel way of conducting a statistical power test that can be employed in the estimation of alternative minimum sample sizes that are more compatible with sample-specific than population-general conclusions. The condition $\beta_P \leq \beta_S < \beta_P'$ refers to a subset of $\mathbb{R}$ that arises from the fact that $\beta_S \neq \beta_P$ due to sampling error.

An instance in which the condition $\beta_P \leq \beta_S < \beta_P'$ is not satisfied is one in which sampling error leads to a true sample path whose value is markedly different from the corresponding true population path. Notably, this can lead to instances where $\beta_P > \beta_S$, which would be more likely to lead to false negatives for true population paths of low magnitude. Therefore, it is reasonable to expect that the gap between $W_S(\Phi, \mathrm{N}, \beta_P)$ and $W_P(\Phi, \mathrm{N}, \beta_P)$, where $W_S(\Phi, \mathrm{N}, \beta_P) > W_P(\Phi, \mathrm{N}, \beta_P)$, will increase as $|\beta_P|$ decreases.

Each true sample path β_S is calculated directly based on the latent variable scores generated based on the Monte Carlo method, via a standard path analysis. As such, the true sample paths for a given replication j are the same regardless of the analysis technique used. Each estimated sample path $\hat{\beta}$, on the other hand, is calculated based on the composite scores estimated with the indicator scores. Thus estimated sample paths differ based on the analysis technique employed.

6. Monte Carlo experiment

This section summarizes the results of our power assessment for the illustrative model presented earlier with respect to true sample and true population paths. We analyzed 500 simulated samples generated for each sample size point, with the following 11 sample sizes: 20, 40, 60, 80, 100, 150, 200, 250, 300, 350, and 400. Each simulated sample was analyzed with the Path Analysis, PLS Regression, and PLS Mode A techniques.

We started with a fairly small sample size and went up to N=400 to ensure that we captured the behavior of the techniques we studied up to the point where all techniques achieved a statistical power of .8 or greater. Thus a total of 5,500 samples of simulated data were analyzed. The data was generated and analyzed with MATLAB, a widely used numeric computing tool with which we implemented the equations discussed in the previous sections.

Table 1 shows power values with respect to true sample and true population paths for each of the three path levels. Table 2 shows the sample sizes at which each of the techniques achieved power values equal to or above .8.

Table 1: Power with respect to true sample and true population paths

				True sample				True population		
Technique		**N**		β_P **= .142**	β_P **= .387**	β_P **= .592**		β_P **= .142**	β_P **= .387**	β_P **= .592**
Path Analysis		20		.080	.564	.945		.153	.439	.802
Path Analysis		40		.263	.882	.995		.213	.736	.986
Path Analysis		60		.440	.968	1		.272	.895	1
Path Analysis		80		.584	.991	1		.325	.959	1
Path Analysis		100		.674	.999	1		.370	.986	1
Path Analysis		150		.805	1	1		.465	.999	1
Path Analysis		200		.930	1	1		.592	1	1
Path Analysis		250		.967	1	1		.697	1	1
Path Analysis		300		.984	1	1		.727	1	1
Path Analysis		350		.994	1	1		.794	1	1
Path Analysis		400		1	1	1		.850	1	1
PLS Regression		20		.125	.596	.910		.180	.451	.790
PLS Regression		40		.325	.877	.992		.229	.749	.978
PLS Regression		60		.483	.967	1		.287	.905	1
PLS Regression		80		.629	.994	1		.341	.966	1
PLS Regression		100		.702	.999	1		.384	.987	1
PLS Regression		150		.814	1	1		.475	.999	1
PLS Regression		200		.940	1	1		.596	1	1
PLS Regression		250		.974	1	1		.697	1	1
PLS Regression		300		.990	1	1		.732	1	1
PLS Regression		350		.996	1	1		.797	1	1
PLS Regression		400		.998	1	1		.856	1	1
PLS Mode A		20		.103	.528	.904		.153	.413	.782
PLS Mode A		40		.293	.877	.992		.211	.746	.978
PLS Mode A		60		.445	.966	1		.263	.906	1
PLS Mode A		80		.594	.993	1		.323	.961	1
PLS Mode A		100		.673	.998	1		.361	.985	1
PLS Mode A		150		.816	1	1		.470	.999	1
PLS Mode A		200		.918	1	1		.572	1	1
PLS Mode A		250		.977	1	1		.691	1	1
PLS Mode A		300		.986	1	1		.727	1	1
PLS Mode A		350		.996	1	1		.782	1	1
PLS Mode A		400		.998	1	1		.842	1	1

Table 2: Sample sizes at which power is equal to or greater than .8

		True sample				True population		
Technique		β_P **= .142**	β_P **= .387**	β_P **= .592**		β_P **= .142**	β_P **= .387**	β_P **= .592**
Path Analysis		149	35	< 20		356	49	< 20
PLS Regression		144	35	< 20		353	47	22
PLS Mode A		145	36	< 20		365	47	22

The values in the second table (sample sizes at which power is equal to or greater than .8) were obtained by simple interpolation and rounding to the next greater integer based on the values in the first (power values). For example, the sample size of 149, corresponding to the

Path Analysis technique and paths whose corresponding true population value was $\beta_P = .142$ was obtained by rounding:

$$100 + [(150 - 100)/(.805 - .674)](.8 - .674) = 148.092.$$

The statistical significance for any estimated path coefficient $\hat{\beta}$ was assessed through a two-tailed P value calculated based on (15), where $\widehat{\mathrm{SE}}$ is the estimated standard error, N is the sample size used, and $I(\cdot)$ is the two-tailed incomplete beta function. For PLS Regression and PLS Mode A, $\widehat{\mathrm{SE}}$ was estimated via bootstrapping with 500 resamples. For Path Analysis, $\widehat{\mathrm{SE}}$ was estimated based on the simplified assumption that the path estimates would be distributed according to the central limit theorem, as $1/\sqrt{\mathrm{N}}$; this likely led to a slight overestimation of $\widehat{\mathrm{SE}}$. The threshold for P used was $.05$.

$$\mathrm{P} = I\left(\hat{\beta}, \widehat{\mathrm{SE}}, \mathrm{N}\right). \tag{15}$$

As we can see from the results, the sample sizes at which power reaches the .8 threshold differed significantly with respect to true sample and true population paths, even though they did not differ nearly as much across analysis techniques (i.e., Path Analysis, PLS Regression, and PLS Mode A). As expected, the gap between minimum required sample sizes (for the .8 power threshold) with respect to true sample and true population paths was wider for paths of small magnitude, and progressively narrowed as path magnitude increased.

For example, for PLS Mode A and the paths whose corresponding true population value was $\beta_P = .142$ power levels equal to or greater than .8 were achieved at N=145 with respect to true sample paths, whereas N=365 was required to achieve the same power levels with respect to true population paths. For the same analysis technique (i.e., PLS Mode A), but for the higher true population value of $\beta_P = .387$, the corresponding sample sizes were N=36 and N=47 respectively.

7. Empirical illustration

Table 3 shows the results of our analyses of one sample of empirical data of size 193, obtained from individuals who were employed by organizations in the USA and who used social networking sites to different degrees. We used the illustrative model presented earlier, in terms of the hypothesized relationships, to conduct the analysis. The question-statements used for data collection are listed in Appendix A. These results are not presented here as part of a stand-alone empirical contribution; they are used to illustrate how the conclusions from the Monte Carlo experiment can be interpreted in the context of an empirical study.

Table 3. Empirical illustration results

Path	Path Analysis	PLS Regression	PLS Mode A
Education level (C_1) → social networking site use (C_2)	$\widehat{\beta_{21}} = .202$ **	$\widehat{\beta_{21}} = .207$ **	$\widehat{\beta_{21}} = .130$ NS
Social networking site use (C_2) → job satisfaction (C_3)	$\widehat{\beta_{32}} = .179$ **	$\widehat{\beta_{32}} = .147$ *	$\widehat{\beta_{32}} = .445$ ***
Job satisfaction (C_3) → organizational commitment (C_4)	$\widehat{\beta_{43}} = .644$ ***	$\widehat{\beta_{43}} = .651$ ***	$\widehat{\beta_{43}} = .635$ ***
Job satisfaction (C_3) → job performance (C_5)	$\widehat{\beta_{53}} = .288$ ***	$\widehat{\beta_{53}} = .290$ **	$\widehat{\beta_{53}} = .284$ **
Organizational commitment (C_4) → job performance (C_5)	$\widehat{\beta_{54}} = .253$ ***	$\widehat{\beta_{54}} = .252$ **	$\widehat{\beta_{54}} = .260$ **

Notes: * P < .05; ** P < .01; *** P < .001; NS not significant; N=193.

From our Monte Carlo experiment, we know that all three composite-based SEM techniques achieved a power greater than .8 for the small true population path of .142 at sample sizes 150 or greater, with respect to true sample paths. However, sample sizes greater than 350 were required for that power level to be reached with respect only to the true population path of .142 and disregarding true sample paths.

Given this, we are pressed to derive different conclusions regarding the sample and the population from which the sample is taken. We can generally conclude based on the empirical results summarized above that our sample size of 193 is sufficient, in terms of statistical power, to provide general support for our model with respect to this sample and very similar samples. The only exception to this is the non-significant coefficient obtained via PLS Mode A for the path education level (C_1) → social networking site use (C_2).

That is, given the combination of sample and path magnitudes in this analysis, the results are generally expected to extend to contexts that are very similar to this study's (i.e., very similar samples), but not necessarily to the entire population from which the sample was taken. In other words, given the relatively small N=193 we can derive only sample-specific conclusions.

Nevertheless, we cannot conclude based on the empirical results summarized above that our sample size of 193 is sufficient to support our model with respect to the entire population from which the sample was taken. For that we would need a sample of size 350 or greater. In other words, we cannot derive population-general conclusions.

8. Discussion and conclusion

It is strongly recommended that researchers conduct statistical power analyses based on Monte Carlo experiments prior to collecting and analyzing empirical data. The main goal of these power analyses is to find out what the minimum required sample sizes are for related empirical studies. Such power analyses build on true population models, which in turn are constructed by researchers based on expectations from past empirical research and theory.

Statistical power values estimated based on Monte Carlo experiments differ significantly when we consider true population and true sample paths. This has an effect on the estimation of minimum sample sizes required to achieve a certain power level (e.g., .8). If we consider only true population paths, minimum sample sizes required may be grossly overestimated for paths of low expected magnitude.

The reason for this is that not calculating or using the true sample paths in statistical power assessments leads to an unwanted effect of sampling error in the estimation of minimum required sample sizes. Sampling error may lead to instances in which true sample paths are statistically indistinguishable from zero when the corresponding population paths are not.

If these instances are allowed to affect the results of a power analysis and a researcher does not plan on generalizing analysis results from an empirical study to an entire population, using the same model as that used in the Monte Carlo experiment, the minimum required sample size will be overestimated. The same applies to the analysis of a small population, where the sample may encompass the entire population, because in this case sampling error is by definition nonexistent. In these cases our proposed power assessment method, relying on true sample paths, will lead to more accurate estimates of minimum required sample sizes.

It is not difficult to conduct a Monte Carlo experiment to assess the power of one or more SEM techniques considering both true population and true sample paths. All that is needed is an additional analysis for each simulated sample, whereby latent variable scores are used instead of indicator scores. Based on the discussion presented here, it seems reasonable to contend that researchers should conduct these additional analyses whenever possible, which will enable them to establish which sample sizes are needed for conclusions that are population-general and sample-specific.

The thresholds used in our analysis, namely $\{.142, .387, .592, +\infty\}$, were chosen for illustration purposes. Arguably, the more thresholds, the better the precision of the minimum required sample sizes obtained. Nevertheless, the more thresholds, the greater will be the number of simulated samples generated per sample size point needed, so that a sizeable number of simulated samples lead to path coefficients that are within each interval.

Power assessments with respect to true sample paths are likely to be particularly useful in certain fields where constant change makes it difficult to find samples of appropriate size for population generalizations, and where such generalizations are unlikely to be of practical relevance.

One field where change is the norm, often being of a fast pace, is that of information systems; where the effects of information and communication technologies on individuals and groups are studied. Such technologies are constantly changing. Therefore, the results of any study are likely to be sample-specific in terms of their practical relevance, because the technologies that are actually used in practice, by the time the results of a study are published, are likely to be different from those studied.

We could not find any discussion in the statistics literature on the issue addressed here – different statistical power values in Monte Carlo experiments aimed at assessing PLS-based SEM techniques, depending on whether true population and true sample paths are considered. We hope that the discussion presented here contributes to bringing light to this issue, and stimulate future research on this and related issues.

References

Cassel, C., Hackl, P., & Westlund, A. (1999). Robustness of partial least-squares method for estimating latent variable quality structures. *Journal of Applied Statistics*, 26(4), 435-446.

Chin, W.W., Marcolin, B.L., & Newsted, P.R. (2003). A partial least squares latent variable modeling approach for measuring interaction effects: Results from a Monte Carlo simulation study and an electronic-mail emotion/adoption study. *Information Systems Research*, 14(2), 189-218.

Chiquoine, B., & Hjalmarsson, E. (2008). *Jackknifng stock return predictions*. Washington, DC: Federal Reserve Board.

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

Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155-159.

Ferrarini, A. (2011). A fitter use of Monte Carlo simulations in regression models. *Computational Ecology and Software*, 1(4), 240-243.

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.

Helland, I.S. (1988). On the structure of partial least squares regression. *Communications in statistics-Simulation and Computation*, 17(2), 581-607.

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.

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

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.

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

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.

Qureshi, I., & Compeau, D.R.. (2009). Assessing between-group differences in information systems research: A comparison of covariance- and component-based SEM. *MIS Quarterly*, 33(1), 197-214.

Schumacker, R.E., & Lomax, R.G. (2004). *A beginner's guide to structural equation modeling*. Mahwah, NJ: Lawrence Erlbaum.

Wright, S. (1934). The method of path coefficients. *The Annals of Mathematical Statistics,* 5(3), 161-215.

Wright, S. (1960). Path coefficients and path regressions: Alternative or complementary concepts? *Biometrics*, 16(2), 189-202.

Appendix A: Questionnaire used in empirical study

The question-statements below were used for data collection related to the indicators of the latent variables in the illustrative empirical study. They were answered on Likert-type scales going from “1 – Strongly disagree” to “5 – Strongly agree”.

Education level (F_1)

- F_{11}: Education level choices: High school, two-year college degree, four-year college degree, master’s degree, and doctoral degree.

Social networking site use (F_2)

- F_{21}: In the past week, on average, approximately how much time per day have you spent on social network sites’ accounts?
- F_{22}: My SNSs’ account/s are/is a part of my everyday activity.
- F_{23}: I am proud to tell people I’m on SNSs such as Facebook.
- F_{24}: SNSs have become part of my daily routine.
- F_{25}: I feel out of touch when I haven’t logged onto SNSs for a while.
- F_{26}: I feel I am part of the SNSs community.
- F_{27}: I would be sorry if SNSs shut down.

Job satisfaction (F_3)

- F_{31}: I am very satisfied with my current job.
- F_{32}: My present job gives me internal satisfaction.
- F_{33}: My job gives me a sense of fulfillment.
- F_{34}: I am very pleased with my current job.
- F_{35}: I will recommend this job to a friend if it is advertised /announced.

Organizational commitment (F_4)

- F_{41}: I would be very happy to spend the rest of my career with this organization.
- F_{42}: I feel a strong sense of belonging to my organization.
- F_{43}: I feel ‘emotionally attached’ to this organization.
- F_{44}: Even if it were to my advantage, I do not feel it would be right to leave my organization.
- F_{45}: I would feel guilty if I left my organization now.

Job performance (F_5)

- F_{51}: My performance in my current job is excellent.
- F_{52}: I am very satisfied with my performance in my current job.
- F_{53}: I am very happy with my performance in my current job.