EDITOR'S COMMENTS

PLS: A Silver Bullet?

By: **George A. Marcoulides**
California State University, Fullerton
gmarcoulides@fullerton.edu

Carol Saunders
University of Central Florida
misq@bus.ucf.edu

We are writing this editorial because it appears to us that some researchers in the Information Systems community view partial least squares modeling (PLS; also referred to as path analysis with composites or soft modeling) as some type of magical silver bullet. These researchers are less critical about the use of PLS than they should be. In spite of cautiously proposed rules of thumb available in the PLS literature, we are frustrated by sweeping claims made by some researchers that PLS modeling can or should be used (often instead of the covariance-based approach) because it makes no sample size assumptions or because somehow "Sample size is less important in the overall model" (Falk and Miller 1992, p. 93). We are seeing an increasing number of such claims in papers submitted for review. It would be nice to think that such statements would be weeded out in the review process. However, more and more studies across a number of disciplines are creeping into the literature in which the samples are dwindling to ridiculously small sizes, despite the inferential intentions of the studies and the magnitude of parent populations. The use of small samples in these studies is frequently legitimized by references to the original developers of the PLS approach. Even *MIS Quarterly* has published at least one such study, which incorrectly states that "the PLS approach does not impose sample size restrictions for the underlying data." While many *MIS Quarterly* readers use PLS correctly, we are writing this editorial to combat the mistaken belief held by some in the IS community that PLS may be used in all cases when the sample size is small. Rather, we wish to stress the importance of adequate sample size, as well as the related issue of standard errors, in PLS research.

The claim about the desirability of larger sample sizes when using PLS is not new. For example, Hui and Wold (1982) determined that PLS estimates improved and their average absolute error rates diminished as sample sizes increased. Similarly, Chin and Newsted (1999) determined that small sample sizes (e.g., N = 20) do not permit a researcher to detect low valued structural path coefficients (e.g., 0.20) until much larger sample sizes (i.e., between N = 150 and N = 200) are reached. Small sample sizes could only be used with higher valued structural path coefficients (e.g., 0.80), and even then will result in "reasonably large standard errors" (Chin and Newsted 1999, p. 333). These results corroborate earlier writings and theorems, which indicated that PLS estimates "are asymptotically correct in the joint sense of consistency (large number of cases) and consistency at large (large number of indicators for each latent variable)" (Jöreskog and Wold 1982, p. 266), implying in the statistical sense that estimation error decreases as N increases (i.e., as $N \to \infty$, the estimation error tends to 0), or simply that any estimated PLS coefficients will converge on the parameters of the model as both sample size and number of indicators in the model become infinite (see also Falk and Miller 1992; McDonald 1996). This same statistical interpretation and recommendation is provided by Hui and Wold (1982), who indicate that PLS "estimates will in the limit tend to the true values as the sample size N increases indefinitely, while at the same time the block sizes increase indefinitely but remain small relative to N" (p. 123).

There is no doubt that sample size plays an important role in almost every statistical technique applied in practice. Although there is universal agreement among researchers that the larger the sample the more stable the parameter estimates, there is no agreement as to what constitutes large. This topic has received a considerable amount of attention in the broad statistical literature, but no easily applicable and clear-cut rules of thumb have been proposed. To give only an idea of the issue involved, a cautious attempt at a rule of thumb in SEM (the term is used generically to refer to path analysis with latent variables or covariance-based models) applications suggests that the sample size should always be more than 10 times the number of free model parameters (see Bentler 1995; Hu et al. 1992). For example, using this rule, the minimum sample size needed for a simple confirmatory factor analysis model with two correlated factors (ϕ_{21}), each of which has three continuous factor indicators and the following a priori proposed factor loading $\Lambda = [\lambda_{11}, \lambda_{21}, \lambda_{31}, \lambda_{42}, \lambda_{52}, \lambda_{62}]$ and error variance $\Theta = [\theta_{11}, \theta_{22}, \theta_{33}, \theta_{44}, \theta_{55}, \theta_{66}]$ matrix structures to be estimated, would be equal to $p(p + 1)/2 = 6(6 + 1)/2 = 21 - 13$[1] $= 8$ free model parameters × 10 = 80 observations. Nevertheless, numerous researchers (e.g., Boomsma 1982; Cudeck and Henly 1991; Jackson 2003; MacCallum et al. 1996; Muthén and Muthén 2002) have indicated that neither the above-illustrated rule nor any other rule of thumb can be applied indiscriminately to all situations. This is because the appropriate size of a sample depends on many factors, including the psychometric properties of the variables, the strength of the relationships among the variables considered, the complexity and size of the model, the amount of missing data, and the distributional characteristics of the variables considered. When these issues are considered, samples of varying magnitude may be needed to obtain reasonable parameter estimates.

Similar heuristics have also been proposed in the PLS literature. For example, Chin (1998) suggests that a researcher use a rule of thumb of 10 cases per predictor, whereby the overall sample size is 10 times the largest of two possibilities: (1) the block with the largest number of indicators (i.e., the largest so-called measurement equation) or (2) the dependent variable with the largest number of independent variables impacting it (i.e., the largest so-called structural equation). Although Chin generally favors this rule of thumb when indicating that "under this circumstance it may be possible to obtain stable estimates" (p. 311) he still warns researchers to be cognizant of the fact that, "the stability of the estimates can be affected contingent on the sample size" (p. 305). However, with the exception of Goodhue et al. (2006), we are not aware of any research that has extended Chin's work either by testing the frequently invoked "10 times" rule of thumb for minimum sample sizes or the assertion that PLS has more power than other techniques at small sample size.

There have been, however, studies suggesting the need to consider multiple factors when determining the appropriate sample size. Lu (2004) and Lu et al. (2005) warned researchers about the bias that arises from a failure to use a large number of indicators for each latent variable (i.e., the so-called consistency at large) and labeled it "finite item bias." Dijkstra (1983) and Schneeweiss (1993) discussed the magnitude of standard errors for PLS estimators resulting from not using enough observations (consistency) and indicators for each latent variable (consistency at large). Schneeweiss also provided closed form equations that can be used to determine the magnitude of finite item bias relative to the number of indicators used in a model. Using these equations, Schneeweiss indicated that item bias is generally small when many indicators, "each with a sizeable loading and an error which is small and uncorrelated (or only slightly correlated) with other error variables" (p. 310), are used to measure each factor in the model. These warnings clearly echo those offered in the general SEM literature that a determination of the appropriate sample size depends on many factors, including the psychometric properties of the variables, the strength of the relationships among the variables, the model, and the characteristics of the data.

To use any latent variable methodology to its fullest potential, it is essential that a proposed model meet underlying structural and distributional assumptions (Satorra 1990) for the methodology, and be developed in a manner consistent with all available theoretical and research-accumulated knowledge in a given substantive domain. Structural assumptions demand that no intended (observed or theoretical) variables are omitted from the model under consideration and that no misspecifications are made in the equations underlying the proposed model. An example of such a structural misspecification would be omitting a relevant predictor variable (observed or theoretical) that is correlated with other exploratory variables. The omission of such a predictor variable would bias the estimates. Theoretical knowledge should guide the selection of variables. Distributional assumptions include linearity of relationships, completeness of data, multivariate normality, and adequate sample size. With real data obtained under typical data gathering situations, violations of these distributional assumptions are often inevitable (Marcoulides 2005).

When conceptualizing the model, a typical requirement is that it includes parameters pertaining to the studied research question(s). Frequently aspects of such questions are reflected directly in one or more model parameters, such as structural regression

[1]Thirteen is the number of model parameters that must be estimated (i.e., 6 factor loadings + 6 error variances + 1 factor intercorrelation).

coefficients, factor intercorrelations, or factor loadings. Use of the model then provides the researcher with estimates of these parameters along with indexes of their stability across repeated sampling from the studied population. These indexes—the parameter standard errors—can also play, as is well known, an instrumental role in constructing confidence intervals for particular population parameters of interest (e.g., Denham 1997; Hays 1994; Raykov and Marcoulides 2000; Serneels et al. 2004). Obviously, models estimated using questionable sample sizes with extremely unstable estimates and wielding huge standard errors and confidence intervals should be sufficient evidence for an investigator to question the generalizability of results and validity of conclusions drawn. Questionable sample sizes can also cause standard errors to be either overestimated or underestimated. Overestimated standard errors may result in significant effects being missed, while underestimated standard errors may result in overstating the importance of effects (Muthén and Muthén 2002).

Sample Size, Parameter Estimates, and Standard Errors in PLS

How small can a sample be in order to ensure sufficiently stable estimates and provide adequate levels of statistical power in a PLS analysis? A related question is: Does the generic PLS sample size rule of thumb ensure sufficiently stable estimates?

In order to determine the precision of estimation in a particular PLS effort and find standard errors, two approaches can be considered: (1) using analytic approaches such as the delta or Taylor series expansion method (Denham 1997; Raykov and Marcoulides 2004; Serneels et al. 2004) or (2) using computer-intensive resampling methods (Chin 1998; Wold 1982). Unfortunately, finding formulas for standard errors of PLS estimates using the delta or Taylor series expansion method is not a trivial task (Dijkstra 1983), although the recent work by Denham (1997), Raykov and Marcoulides (2004), and Serneels et al. (2004) should prove promising. As a consequence, Monte Carlo simulation resampling methods dominate the field. The principle behind a Monte Carlo simulation is that the behavior of a parameter estimate in random samples can be assessed by the empirical process of drawing many random samples and observing this behavior.

Two kinds of Monte Carlo resampling strategies can be used to examine parameter estimates and related sample size issues. The first strategy can be considered a *reactive* Monte Carlo analysis such as the popular jackknife or bootstrap approaches (Chin 1998; Denham 1997; Wold 1982) in which the performance of an estimator of interest is judged by studying its parameter and standard error bias relative to repeated random samples drawn with replacement from the original observed sample data. This type of Monte Carlo analysis is currently quite popular (particularly the bootstrap approach), despite the fact that it may often give "an unduly optimistic impression of accuracy or stability" of the estimates (Dijkstra 1983, p. 86). There are no generally applicable results as yet of how good the underlying approximation of sampling by pertinent resampling distributions is within the framework of latent variable modeling (Raykov and Marcoulides 2004).

The second, and less commonly known strategy, can be considered a *proactive* Monte Carlo simulation analysis (Marcoulides 1990; Paxton et al. 2001; Steiger 2001). In a proactive Monte Carlo analysis, data are generated from a population with hypothesized parameter values and repeated random samples are drawn to provide parameter estimates and standard errors. A proactive Monte Carlo analysis can be used not only to examine parameter estimate precision, but also to determine the sample size needed to ensure the precision of parameter estimates (Muthén and Muthén 2002). Several commonly available computer programs can be used to design a proactive Monte Carlo simulation to examine sample size needs in PLS (e.g., SAS Macro options; Fan et al. 2002). In addition, general programs that can handle a Cholesky decomposition (Raykov, Marcoulides, and Boyd 2003) or a relatively new program developed for the statistical analysis of latent variables called *Mplus* (Muthén and Muthén 1998) can be used for the proactive Monte Carlo analysis. *Mplus* has a fairly easy-to-use interface and offers researchers a flexible tool to analyze data using all kinds of model choices. Detailed illustrations for using the *Mplus* Monte Carlo simulation options can also be found in Muthén and Muthén (2002), in the *Mplus User's Guide* (Muthén and Muthén 1998), and at the product Web site www.statmodel.com.

As an illustration of using a proactive Monte Carlo simulation analysis, consider the following simple confirmatory factor analysis (CFA) model with two correlated factors (φ_{21}), each of which has three continuous factor indicators and the following factor loading $\Lambda = [\lambda_{11}, \lambda_{21}, \lambda_{31}, \lambda_{42}, \lambda_{52}, \lambda_{62}]$ and error variance $\Theta = [\theta_{11}, \theta_{22}, \theta_{33}, \theta_{44}, \theta_{55,} \theta_{66}]$ matrix structures. Based on this model, data were generated with varying values for the model parameters, factor intercorrelations (φ_{21} between .1 and .9), each factor loading (λ between .4 and .9), for the error variances (θ between .84 and .19) and, consequently, for the indicator reliabilities (between .16 and .81). The various CFA model configurations were examined with normally distributed data, and no missing data patterns were considered. To ensure stability of results, the number of sample replications was set at 5,000. To simplify matters, the proactive Monte Carlo simulation concerning sample size needs focused only on the factor correlation parameter (φ_{21}), although any other model parameter could be similarly examined. For example, if a researcher were interested in examining the value of the factor loading λ_{11}, a similar simulation analysis could be easily conducted. An example input file using the *Mplus* command language for a normally distributed data set with N = 100 observations and all the factor loadings λ specified as equal to .9, which implies each error variance is .19 and the reliability of each factor indicator is equal to .81, and the factor intercorrelation φ_{21} specified as .1 is provided in Appendix A (for a detailed description of each command see the M*plus User's Guide*, Muthén and Muthén 1998).

Based upon this simple model, what sample sizes would be needed to achieve a sufficient level of power, say equal to .80 (considered by most researchers as acceptable power) to reject the hypothesis that the factor correlation in the population is zero (i.e., the probability of rejecting the null hypothesis when it is false)? The results of such a proactive Monte Carlo analysis under conditions of normality are provided in Table 1. The boldfaced column-heading values correspond to the various factor loadings considered (i.e., the values of λ), while the boldfaced row values correspond to the considered factor intercorrelations (i.e., the values of φ_{21}). The entries provided in Table 1 correspond to the sample size needed to achieve power equal to .80. As can be seen, relatively small sample sizes can be used when psychometrically sound factor indicators (i.e., indicators with sizeable factor loadings) are available to examine high valued factor intercorrelations. However, when trying to examine low valued factor intercorrelations using poor quality indicators, much larger sample sizes are needed. For example, using indicators with 0.7 valued factor loadings, a sample of size 1261 would be required in order to examine a factor intercorrelation equal to 0.1.[2]

It is important to note that the results presented in Table 1 corroborate those presented by Hui and Wold (1982), Chin and Newsted (1999), and Schneeweiss (1993) that small sample sizes do not permit a researcher to detect low valued model coefficients until much larger sample sizes are reached. However, the problem can be much more disconcerting than these researchers reported when moderately non-normal data are considered. When moderately non-normal data are considered, a markedly larger sample size is needed despite the inclusion of highly reliable indicators in the model.

Determining the appropriate sample size even for a simple CFA depends on a number of factors, not just the magnitude of the relationship and the desired level of power. It is also important to note that for only a limited number of normally distributed data conditions would the PLS rule of thumb of 10 cases per indicator really suffice in the example CFA models considered. As such, it is evident that a simple application of the generic PLS rule of thumb does not always ensure sufficiently stable estimates and cannot be applied indiscriminately to all situations. Indeed, a researcher must consider the distributional characteristics of the data, potential missing data, the psychometric properties of the variables examined, and the magnitude of the relationships considered before deciding on an appropriate sample size to use or to ensure that a sufficient sample size is actually available to study the phenomena of interest.

Guidelines to Consider Prior to PLS Modeling

So what does a researcher need to consider when using PLS? The following necessary (but certainly not sufficient) list can be used as an aid for researchers to follow before proceeding with PLS modeling (as well as when conducting analysis using any latent variable modeling technique):

- Propose a model that is consistent with all currently available theoretical knowledge and collect data to test that theory. Research begins with the premise of a theory and then facts are accumulated to test the theory. Popper (1962) called this

[2]In the statistical literature, testing such low factor intercorrelations is a well-known problem related to testing a boundary solution (e.g., Andrews 1999, 2001).

Table 1. Sample Sizes Needed to Achieve Power = .80 with Normally Distributed Data and No Missing Values

	λ					
φ_{21}	**0.9**	**0.8**	**0.7**	**0.6**	**0.5**	**0.4**
0.1	916	1053	1261	1806	2588	4927
0.2	256	292	371	457	764	1282
0.3	96	99	147	223	317	672
0.4	46	57	71	98	186	343
0.5	25	34	43	66	111	220
0.6	16	20	23	44	78	175
0.7	15	15	17	33	61	134
0.8	15	15	17	25	46	109
0.9	15	15	17	25	42	99

process falsification, the logical consequence of which is that no theory can ever be proved true from data; it can only be corroborated. So it must be recognized that the accumulation of knowledge may provide a basis for refuting a theory at some future time. If a study's finding is important, other researchers may elaborate on it or extend it by defining moderators, boundary conditions, etc. As knowledge accumulates, findings that are flukes of a sample or otherwise not robust will tend to fall by the wayside (Blair and Zinkhan 2006). Consequently, the process of empirical research rests on the premise that researchers will build upon the knowledge that is presently available to attempt to explain why certain phenomena occur in their findings.

- Perform data screening. This step involves both essential data screening (e.g., accuracy of input, outliers, etc.) and an examination of the distributional characteristics (e.g., normality, missing data, etc.) of the sample(s) used in the study. Evaluation of the normality of variable distributions is particularly relevant whenever any statistical inferences are planned. Normality of variables is generally assessed by either statistical methods (i.e., skewness or kurtosis) or graphical methods. Any departures from normality and missing data will influence the sample size requirements of a study and potentially deteriorate power in a study.

- Examine the psychometric properties of all variables in the model. This step involves an assessment of the variable, indicator, scale, factor, or composite reliability. Assessment of reliability utilizes primarily the model of congeneric measures, which can be readily calculated by using the formula, $\lambda^2 /(\lambda^2 + \theta)$ for each variable or indicator, or by using the formula, $(\Sigma\lambda)^2 /[(\Sigma\lambda)^2 + \Sigma\theta]$ for scales, factors, or composites (where λ is the factor loading, and θ is the error variance) (for further details, see Chin 1998; Raykov and Marcoulides 2006). Poorly measured variables, indicators, scales, factors, or composites influence sample size requirements and potentially deteriorate power in a study.

- Examine the magnitude of the relationships and effects between the variables being considered in the proposed model. When using poorly measured variables, the ability to reject the hypothesis that the coefficients in the population are zero deteriorates with low valued coefficients and influences the sample size requirements of a study.

- Examine the magnitude of the standard errors of the estimates considered in the proposed model and construct confidence intervals for the population parameters of interest. Unstable coefficients with large standard errors and wide confidence intervals are an indication of inadequate sample size. Problematic sample sizes can cause standard errors to be either overestimated or underestimated.

- And last but not least, assess and report the power of your study.

Of course, this editorial does not address what a researcher should do if the power is not adequate to test the model or how a researcher should go about getting a large enough sample in the first place. That is fodder for several other editorials. Rather, our message is quite simple: *Please make sure that your sample size is large enough to support your conclusions and do not allow perceived fancy modeling techniques to overwhelm common sense.* The PLS rule of thumb might work well in some instances, but in others it might fail miserably. PLS is not a silver bullet to be used with samples of any size!

Finally, we are calling for research to explore the many factors involved in selecting an appropriate sample size with PLS modeling. These factors include the psychometric properties of the variables considered, the strength of the relationship among the variables, the complexity and size of the model, the amount of missing data, and distributional characteristics of the variables, especially non-normality. Examining all these issues should provide researchers with some insight concerning the stability and power of the parameter estimates that would be obtained across repeated sampling from the studied population.

References

Andrews, D. W. K. "Estimation When a Parameter Is on a Boundary," *Econometrica* (67), 1999, pp. 1341-1383.

Andrews, D. W. K. "Testing When a Parameter Is on the Boundary of the Parameter Space," *Econometrica* (69), 2001, pp. 683-734.

Bentler, P. M. *EQS Structural Equation Program Manual*, Multivariate Software Inc., Encino, CA, 1995.

Blair, E., and Zinkhan, G. M. "Nonresponse and Generalizability in Academic Research," *Journal of the Academy of Marketing Science* (34:1), 2006, pp. 4-7.

Boomsma, A. "The Robustness of LISREL Against Small Sample Sizes in Factor Analysis Models," in *Systems Under Indirect Observation, Part I*, K. G. Jöreskog and H. Wold (eds.), North Holland, Amsterdam, 1982, pp. 149-174.

Chin, W. W. "The Partial Least Squares Approach for Structural Equation Modeling," in *Modern Methods for Business Research*, G. A. Marcoulides (ed.), Lawrence Erlbaum Associates, Mahwah, NJ, 1998, pp. 295-336.

Chin, W. W., and Newsted, P. R. "Structural Equation Modeling Analysis with Small Samples Using Partial Least Squares," in *Statistical Strategies for Small Sample Research*, R. H. Hoyle (ed.), Sage Publications, Thousand Oaks, CA, 1999, pp. 307-341

Cudeck, R., and Hensly, S. J. "Model Selection in Covariance Structure Analysis and the 'Problem' of Sample Size: A Clarification," *Psychological Bulletin* (109), 1991, pp. 512-519.

Denham, M. C. "Prediction Intervals in Partial Least Squares," *Journal of Chemometrics* (11), 1997, pp. 39-52.

Dijkstra, T. "Some Comments on Maximum Likelihood and Partial Least Squares Methods," *Journal of Econometrics* (22), 1983, pp. 67-90.

Falk, R. F., and Miller, N. B. *A Primer of Soft Modeling*, The University of Akron Press, Akron, OH, 1992.

Fan, X., Felsöváįlyi, A., Sivo, S. A., and Keenan, S. C. *SAS for Monte Carlo Studies: A Guide for Quantitative Researchers*, SAS Institute, Inc., Cary, NC, 2002.

Goodhue, D., Lewis, W., and Thompson, R. "PLS, Small Sample Size and Statistical Power in MIS Research," in *Proceedings of the 39th Hawaii International Conference on System Sciences*, R. Sprague Jr. (ed.), Kauii, Hawaii, January 4-7, 2006 (CD-Rom Proceedings).

Hays, W. L. *Statistics*, Harcourt Brace Jovanovic Fort Worth, TX, 1994.

Hu, L. T., Bentler, P. M., and Kano, Y. "Can Test Statistics in Covariance Structure Analysis Be Trusted?," *Psychological Bulletin* (112), 1992, pp. 351-362.

Hui, B. S., and Wold, H. "Consistency and Consistency at Large in Partial Least Squares Estimates," in *Systems Under Indirect Observation, Part II,* K. G. Jöreskog and H. Wold (eds.), North Holland, Amsterdam, 1982, pp. 119-130.

Jackson, D. L. "Revisiting Sample Size and the Number of Parameter Estimates: Some Support for the N:q Hypothesis," *Structural Equation Modeling* (10), 2003, pp. 128-141.

Jöreskog, K. G., and Wold, H. *Systems Under Indirect Observation, Part I and II*, North Holland, Amsterdam, 1982.

Lu, I. R. R. *Latent Variable Modeling in Business Research: A Comparison of Regression Based on IRT and CTT Scores with Structural Equation Models*, unpublished doctoral dissertation, Carleton University, Canada, 2004.

Lu, I. R. R., Thomas, D. R., and Zumbo, B. D. "Embedding IRT in Structural Equation Models: A Comparison with Regression Based on IRT Scores," *Structural Equation Modeling* (12), 2005, pp. 263-277.

MacCallum, R. C., Browne, M. W., and Sugawara, H. M. "Power Analysis and Determination of Sample Size for Covariance Structure Modeling," *Psychological Methods* (1), 1996, pp. 130-149.

Marcoulides, G. A. "Evaluation of Confirmatory Factor Analytic and Structural Equation Models Using Goodness-of-Fit Indices," *Psychological Reports* (67), 1990, pp. 669-671.

Marcoulides, G. A. "Nonstandard Cases in Structural Equation Modeling," in *Encyclopedia of Statistics, Volume 4*, B. Everitt and D. Howell (eds.), Wiley, Chichester, UK, 2005, pp. 1850-1855.

McDonald, R. P. "Path Analysis with Composite Variables," *Multivariate Behavioral Research* (31), 1996, pp. 239-270.

Muthén, L. K., and Muthén, B. O. "How to Use a Monte Carlo Study to Decide on Sample Size and Determine Power," *Structural Equation Modeling* (9), 2002, pp. 599-620.
Muthén, L. K., and Muthén, B. O. *Mplus User's Guide*, Muthén and Muthén, Los Angeles, CA, 1998.
Paxton, P., Curran, P. J., Bollen, K. A., Kirby, J., and Chen, F. "Monte Carlo Experiments: Design and Implementation," *Structural Equation Modeling* (8), 2001, pp. 287-312.
Popper, K. *On the Structure of Scientific Revolution*, Chicago University Press, Chicago, IL, 1962.
Raykov, T., and Marcoulides, G. A. "Estimation of Generalizability Coefficients via a Structural Equation Modeling Approach to Scale Reliability Evaluation," *International Journal of Testing* (6), 2006, pp. 81-95.
Raykov, T., and Marcoulides, G. A. *A First Course in Structural Equation Modeling*, Lawrence Erlbaum Associates, Mahwah, NJ, 2000.
Raykov, T., and Marcoulides, G. A. "Using the Delta Method for Approximate Interval Estimation of Parameter Functions in SEM," *Structural Equation Modeling* (11), 2004, pp. 659-675.
Raykov, T., Marcoulides, G. A., and Boyd, J. "Using SEM Programs to Perform Matrix Manipulations and Data Simulation," *Structural Equation Modeling* (10), 2003, pp. 312-322.
Satorra, A. "Robustness Issues in Structural Equation Modeling: A Review of Recent Developments," *Quality and Quantity* (24), 1990, pp. 367-386.
Schneeweiss, H. "Consistency at Large in Models with Latent Variables," in *Statistical Modelling and Latent Variables*, K. Haagen, D. J. Bartholomew, and M. Deistler (eds.), Elsevier, Amsterdam, 1993.
Serneels, S., Lemberge, P., and Van Espen, P. J. "Calculation of PLS Prediction Intervals Using Efficient Recursive Relations for the Jacobian Matrix," *Journal of Chemometrics* (18), 2004, pp. 76-80.
Steiger, J. H. "Driving Fast in Reverse: The Relationship Between Software Development, Theory, and Education in Structural Equation Modeling," *Journal of the American Statistical Association* (96), 2001, pp. 331-338.
Wold, H. "Soft Modeling: Intermediate Between Traditional Model Building and Data Analysis," *Mathematical Statistics* (6), 1982, pp. 333-346.

Appendix A

Example *Mplus* Setup File

```
TITLE:                  EXAMPLE MPLUS SETUP FILE
MONTECARLO:             NAMES ARE y1-y6;
                        NOBSERVATIONS = 100;
                        NREPS = 5000;
                        SEED = 2999;
                        nCLASSES = 1;
                        GCLASSES = 1;
                        SAVE = output.sav;
ANALYSIS:               TYPE = MIXTURE;
                        ESTIMATOR = ML;
MODEL MONTECARLO:       %OVERALL%
                        F1 BY y1-y3*.9;
                        F2 BY y4-y6*.9;
                        F1@1 F2@1;
                        y1-y6*.19;
                        F1 WITH F2*.1;
MODEL:                  %OVERALL%
                        F1 BY y1-y3*.9;
                        F2 BY y4-y6*.9;
                        F1@1 F2@1;
                        y1-y6*.19;
                        F1 WITH F2*.1;
OUTPUT:                 TECH9;
```