2011

MEASUREMENT ERROR IN PLS, REGRESSION AND CB-SEM

Dale L. Goodhue
University of Georgia, dgoodhue@uga.edu

William Lewis
, Independent Consultant, Hammond, Louisiana, USA, William.w.lewis@gmail.com

Ron Thompson
Wake Forest University, Winston-Salem, North Carolina, USA, ron.thompson@mba.wfu.edu

MEASUREMENT ERROR IN PLS, REGRESSION AND CB-SEM

Goodhue, Dale, University of Georgia, Athens, Georgia 30606, Brooks Hall 312, USA, dgoodhue@terry.uga.edu.

Lewis, William, Independent Consultant, Hammond, Louisiana, USA, William.w.lewis@gmail.com

Thompson, Ron, Wake Forest University, Winston-Salem, North Carolina, USA, ron.thompson@mba.wfu.edu

Abstract

Partial Least Squares (PLS) has become a very popular statistical technique for analyzing causal path models involving multiple indicator data in MIS research. Proponents of the use of PLS view it as a relatively easy to use "second-generation" statistical analysis technique that, like covariance-based structural equation modeling (CB-SEM), has several advantages over regression. While it isn't frequently explicitly stated, most users seem to believe that PLS somehow takes measurement indicator errors into account, perhaps similarly to the way CB-SEM techniques do. We conducted an extensive Monte Carlo simulation study to compare the accuracy of path estimates for regression, PLS and CB-SEM (employing LISREL) under a variety of conditions. We found that PLS provides path estimates that are much closer to regression than to LISREL. Further, we observed that under most circumstances the amount of bias in the PLS (and regression) estimates could be corrected by applying Nunnally and Bernstein's (xxxx) formula for correcting attenuation. We conclude that our results offer strong evidence that 1) PLS does not take indicator measurement error into account, and 2) under most circumstances the accuracy of PLS (or of regression) path estimates can be improved by applying Nunnally and Bernstein's formula for correcting attenuation. There is also the suggestion that under certain circumstances, PLS may over-estimate path values.

Keywords: Measurement Error, PLS, Regression, Covariance Based Structural Equation Modeling.

1 INTRODUCTION

The Partial Least Squares (PLS) statistical technique for analyzing causal path models using multi-indicator data has seen growing acceptance and even preference in the MIS research community over the last two decade or so. This has in some part been due to its ease of use, and its reputed ability to operate with smaller samples sizes, non-normally distributed data, and to handle formative constructs effectively. It has thus presented researchers an attractive alternative that is in between regression (which is easy to use but in its most accessible form is limited to analyzing single dependent variables) and various Covariance Based Structural Equation Modeling (DB-SEM) techniques such as LISREL, Mplus, etc. (which can handle more complex models but are more demanding in terms of sample size and distribution requirements, and the sophistication required of users).

A large part of the appeal of PLS is that it is presented as being "like" CB-SEM statistical techniques, without the complications. A quote from Chin's (1998) Issues and Opinions piece in MISQ is emblematic of the way PLS and CB-SEM are implicitly equated in terms of their advantages:

> In the past few years, the IS field has seen a substantial increase in the number of submissions and publications using structural equation modeling (SEM) techniques. Part of the reason may be the increase in software packages to perform such covariance-based (e.g., LISREL, EQS, AMOS, SEPATH, RAMONA, MX, and CALIS) and component-based (e.g., PLS-PC, PLS-Graph) analysis. . . . .
>
> When applied correctly, SEM-based procedures have substantial advantages over first-generation techniques such as principal components analysis, factor analysis, discriminant analysis, or multiple regression because of the greater flexibility that a researcher has for the interplay between theory and data. Specifically, SEM provides the researcher with the flexibility to: (a) model relationships among multiple predictor and criterion variables, (b) construct unobservable LVs, (c) model errors in measurements for observed variables, and (d) statistically test *a priori* substantive/theoretical and measurement assumptions against empirical data (i.e., confirmatory analysis). [Chin, 1998, p vii].

In particular, PLS is often presented as being more effective than regression in terms of generating higher reliability in its measurement.

> PLS treats each indicator separately, allowing each item to differ in the amount of influence on the construct estimate. Therefore, indicators with weaker relationships to related indicators and the latent construct are given lower weightings (Lohmoller 1989; Wold 1982, 1985, 1989), resulting in higher reliability for the construct estimate and thus stronger theoretical development. [Chin, Marcolin and Newsted, 2003, p. 194]
>
> PLS allows [sic] to both specify the relationships among the principal construct and their underlying items, resulting in a simultaneous analysis of both whether the hypothesized relationships at the theoretical level are empirically true and also how well the measures relate to each construct (Chin, 1998). The ability to include multiple measures for each construct provides more accurate estimates of the paths among constructs, which are typically downward biased by measurement error when applying multiple regression analysis. [Pavlou 2002, p. 231]

However, more recently there have been challenges to PLS's claimed advantages. Goodhue, Lewis and Thompson (2006) presented Monte Carlo simulation results suggesting that PLS did not have the reputed advantages at small sample size. Hwang, Malhotra, Kim, Tomiuk and Hong (forthcoming), again using Monte Carlo simulations, presented evidence that PLS estimates of path values were on average less accurate (farther from the true values) than CB-SEM estimates. Evermann and Tate (2010) also used Monte Carlo simulation to show that model misspecification in PLS can be a major concern, leading to incorrect interpretations. They call for a more critical evaluation of the role of PLS in theory testing situations.

Rönkkö and Ylitalo (2010) focused on the accuracy of the estimated construct values, which come out of the first step of a PLS analysis. Again using Monte Carlo simulation, they looked at the estimated construct values from PLS and those obtained from a regression approach of weighting each indicator equally. They examined the correlation between each estimate and the true underlying value from the Monte Carlo model. Rönkkö and Ylitalo conclude that under most circumstance, equal weighting (such as used for regression analysis) produced construct estimates that were more highly correlated to the true value than did PLS's optimized weightings.

Although there are some aspects of Rönkkö and Ylitalo's work that deserve closer attention, their paper adds a new dimension to the critique of PLS by addressing the way in which PLS generates its indicator weights. They present persuasive arguments that rather than increasing the weights for the most reliable indicators, PLS increases the weights for those indicators that explain the most variance. This way of determining weights would be laudable if all measured variance were "useful" variance, but they point out that often there are measurement biases that affect indicators from different constructs equally, and that such biases are (unfairly) exploited by PLS by increasing the weights for the most biased indicators. They develop the equations to show this in algebraic form.

Our work in this paper fits within the general realm of the four above mentioned papers, in an effort to look at the impact of measurement error on the accuracy of path estimates generated by PLS, regression and CB-SEM, under different circumstances. Specifically, we look at differing sample sizes, different effect sizes, different levels of reliability, and different model complexity. To start with, we assert that if PLS's path estimates have somehow taken into account measurement error, then they must have done so by the way in which they have assigned values to the indicator weights that are used to determine the construct scores. This must be the case, since beyond this point in the PLS process, all path estimates are determined by ordinary least squares regression, which does not compensate for measurement error. We would further suggest that under these circumstances, unless at least one indicator measures the construct without error, there is no conceivable weighting scheme that will overcome measurement error.

In our work, in addition to looking to see which techniques are "better" in different situations, we look closer at what conditions affect the quality of the path estimates. As a result of our work we can add something quite new to the discussion of the relative merits of the three techniques. More specifically we argue and present evidence that rather than being merely "less effective" at incorporating measurement error into its estimation process, PLS actually ignores measurement error in its estimation process. In that respect, PLS would seem to be on a par with regression, and far less accurate than CB-SEM. Because of this, the handicap that PLS and regression face relative to CB-SEM is directly calculable based on the measurement reliability of the indicators used. This is consistent with both McDonald (1996, pages 266-267) and Dijkstra (1983, p. 81) who each suggest that with multiple indicators for each construct, the more measurement error there is, the more PLS and regression will suffer in terms of accuracy in comparison with CB-SEM.

2 DESIGNING A MONTE CARLO SIMULATION

Our goal was to design a Monte Carlo simulation experiment that would allow us to compare four values (the true path, the PLS estimate, the regression estimate and the CB_SEM estimate), and to determine how those comparisons change under a number of relevant conditions. In Monte Carlo simulation the first step is to specify the true model from which to generate the data. Figure 1 shows our model for the first phase of our work.

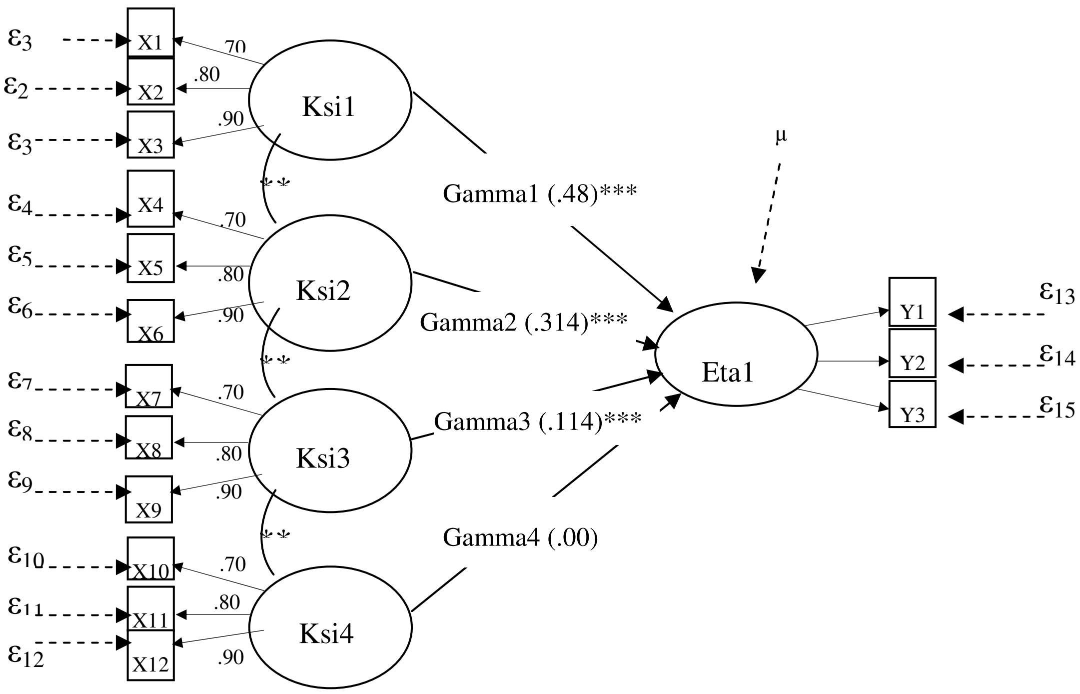


Figure 1. Model #1, Baseline Condition[1]

Each construct in Figure 1 is measured by 3 reflective indicators with weights of .7, .8, and .9, producing a Cronbach's alpha of .84. Random error is added to each indicator scores and to the value for Eta1, so as to give each a variance of 1.0. The relationship between the path coefficients and the effect sizes is given in the footnote below[2], following Cohen (1988).

To compare the four values previously mentioned, we used a technique suggested by Chin et al. (2003), the Mean Relative Bias[3], calculated as follows:

$$MRB = \Sigma_{i=1,500}[(X_i - X_{pop})/ X_{pop}]$$

[1] Model #1 was first used to develop data for a comparison of PLS, LISREL and regression by Goodhue, Lewis and Thompson (2006).

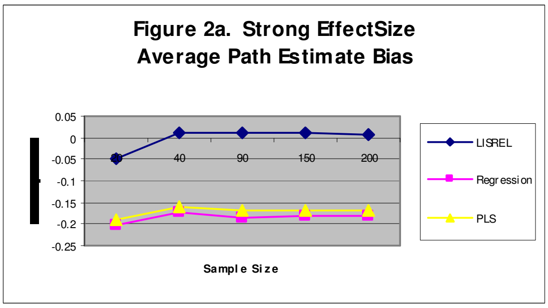
[2] By construction Ksi1 through Ksi4 are independent, and all five constructs are N(0,1) distributed. Therefore the partial R^2 for each Ksi is the square of the partial correlation (i.e. the square of the path coefficient). With path coefficients of .48, .314 and .114, the partial R^2s are .230, .099, and .013. This gives an overall R^2 of .342. Effect size is the partial R^2 divided by the unexplained variance. Since unexplained variance is (1-.342) = .658, this gives: Gamma1 = .230 / (.658) = .350, and similarly Gamma2 = .099 / (.658) = .150, Gamma3 = .013 / (.658) = .020.

[3] We reversed the order of the subtraction from that shown in Chin et al. (2003) to get a MRB that was negative when the estimate was lower than the true value.

3 SIMULATION RESULTS

Figure 2 shows the results of the MRB for the three different statistical techniques, using sample sizes of 20, 40, 90, 150, 200, and generating 500 Monte Carlo samples from the Figure 1 baseline model for each sample size. Figure 2a shows the results for the strong effect size path (Gamma 1 in Figure 1); Figure 2b shows the same for the medium effect size path (Gamma2 in figure 1). These results are consistent with findings of Chin and Newsted (1999) in that PLS seems to have an advantage over regression, but they are also consistent with the findings of Malhotra, Kim, Tomiuk and Hong in that CB-SEM has an advantage over PLS.

However, the advantage of PLS over regression is relatively small – both have about a 15% negative bias, while CB-SEM has close to a zero bias. This pattern changes when sample sizes are reduced to n=20. At this sample size, all techniques seem to suffer. We note that n=20 and n=40 are below any recommended sample size for CB-SEM, and in fact especially at n=20 the CB-SEM technique often failed to converge on an admissible solution.



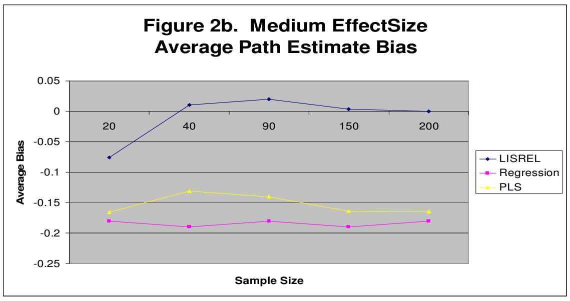


Figure 2a. *Strong EffectSize - Average Path Estimate Bias*

Figure 2b. *Medium EffectSize - Average Path Estimate Bias*

These results offer rather stark hints that PLS is quite a bit more like regression than CB-SEM in terms of accuracy. But they do not address the question of the link between reliability and bias suggested in the introduction of this paper. To address that, we needed to rerun the model in Figure 1 under conditions creating different reliabilities. Figure 1 shows the indicator loadings to be .7, .8. and .9, resulting in a Cronbach's □ of .84. This reliability is what produced the results in Figure 2.

To vary the reliability levels, we created a reliability of .71 by changing the indicator loadings to .6, .7, and .8 and then reran the simulation with another 500 samples in each sample size condition. Then we created a reliability of .91 by going back to the earlier loadings but doubling the number of indicators, giving us the following loadings: .7, .7, .8, .8, .9, .9. Again we ran the analysis with 500 samples for each sample size.

To simplify our results, we averaged across the three larger sample sizes (n=90, n=150, and n=200). Figure 3 shows the results, with the bias at three different levels of reliability (α = .71, .84. .91) with the large effect size shown first on the left side of the figure; then three levels of reliability for the medium effect size results shown in the middle, and the three levels of reliability for the small effect size results shown on the right side of the figure.

Displaying the data in this way suggest a somewhat remarkable pattern – the size of the bias for PLS and regression seems not to depend much on the effect size, but does seem to quite decidedly depend upon the reliability of the measurement of the constructs. CB-SEM continues to display essentially zero bias, regardless of the effect size or the reliability.

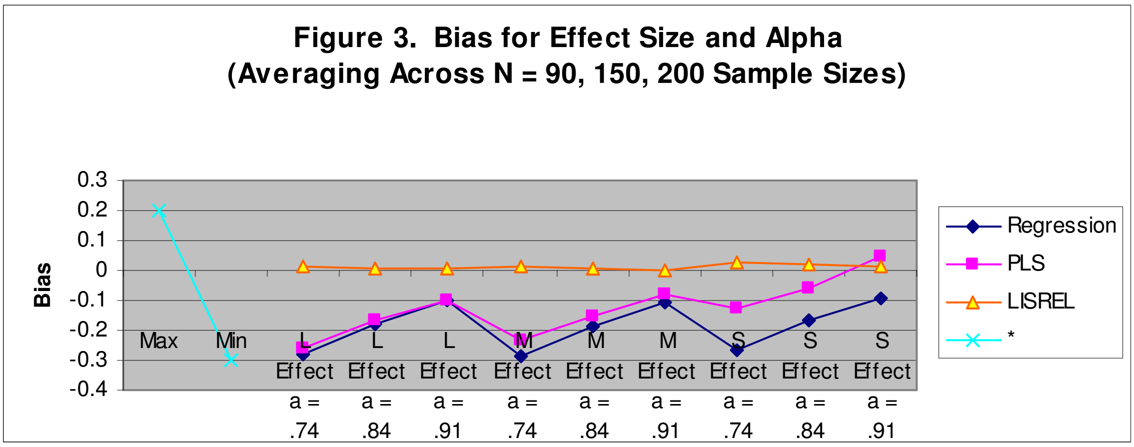


Figure 3. *Bias for Effect Size and Alpha (Averaging Across N = 90, 150, 200 Sample Sizes)*

It is generally accepted that regression does not account for measurement reliability in its path estimates, but that LISREL does. Regression path estimates are "attenuated" by measurement error, according to the following equation from Nunnally and Bernstein (1994, pp. 241, 257):

Apparent Correlation_{XY} =

Actual Correlation_{XY} * Square Root(Reliability_X * Reliability_Y)

Given this equation, and knowing the reliabilities of all the constructs in our various experiments, we should be able to "adjust" the attenuated regression path estimates to the correct value using the reliability of the constructs in our model. Figure 4 shows both PLS and regression path estimates corrected for measurement error and LISREL estimates unchanged.

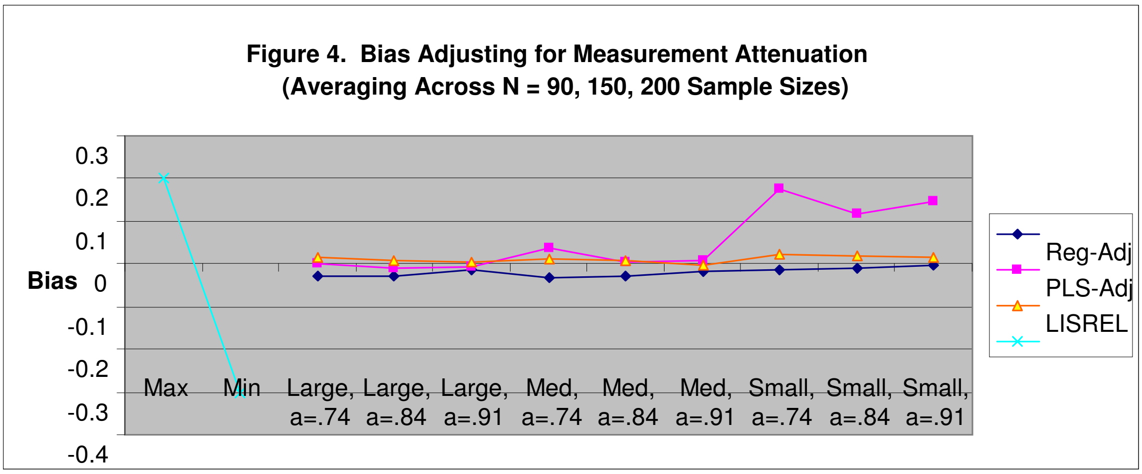


Figure 4. *Bias Adjusting for Measurement Attenuation (Averaging Across N = 90, 150, 200 Sample Sizes)*

Looking at Figures 3 and 4 suggests the following. The negative bias of PLS and regression is proportional to the reliability of the constructs, and is essentially corrected when Nunnally and Bernstein's equation is used. The only time this is not true is when a small effect size is involved, in which case the correction sometimes seems to over correct for PLS[4]. PLS is apparently more similar to regression than it is to LISREL in the way in which it compensates for measurement error. In fact, our evidence seems to suggest that PLS path estimates (like regression path estimates) *do not compensate for measurement error at all*, while LISREL path estimates do.

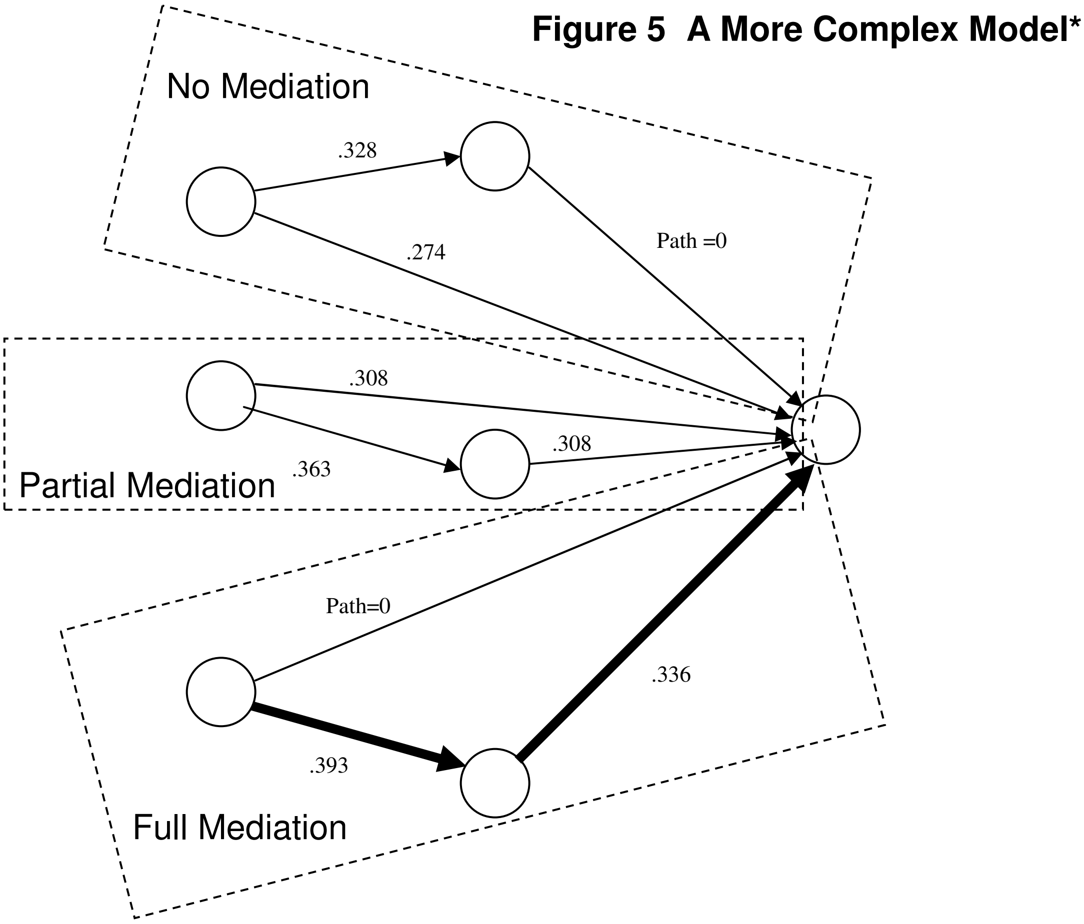


Figure 5 *A More Complex Model*[5]

4 A MORE COMPLEX MODEL

Some readers might be concerned that our model in Figure 1 is too simplistic, and that PLS cannot demonstrate its strengths without a more complex model. With this in mind we developed a more complex model, as shown in Figure 5. On this model we will only look at two paths: the Ksi3 to Eta3

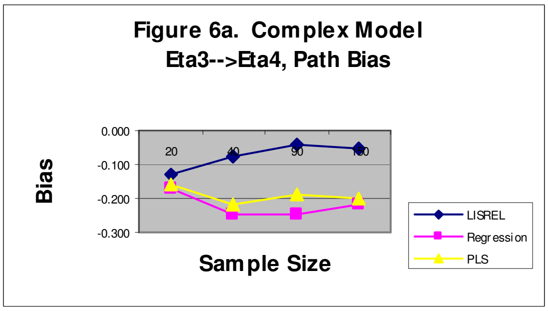
---

[4] We note that with small effect sizes and our simple model, none of the three techniques had even close to acceptable power even at N=200, with the highest power value of 42% for LISREL and six indicators.

[5] The more complex model was first used to develop data for a comparison of PLS, LISREL and regression by Goodhue, Lewis and Thompson (2009).

path and the Eta3 to Eta4 path, as shown in bold in the figure. Here we are looking at two medium effect size paths, with all constructs measured with reliability .84.

Figures 6a and 6b show the mean relative bias as a function of sample size for the three statistical techniques. Figure 6a shows the results for the path between two endogenous constructs. It is quite is reminiscent of Figure 2b, the medium effect size in the simple model, though CB-SEM is not as free from bias as in the simpler model.



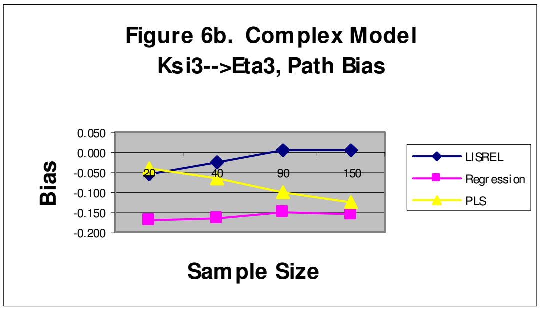
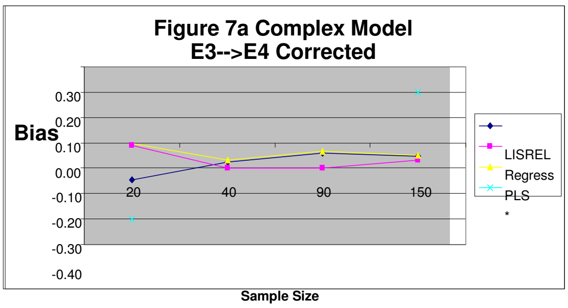


Figure 6a.	*Complex Model*	*(Eta3 →Eta4, Endogenous → Endogenous (Path Bias)*
Figure 6b.	*Complex Model*	*(Ksi3 →Eta3, Exogenous → Endogenous (Path Bias)*

Figure 6b however, with a path from an exogenous to an endogenous construct, gives a different picture. There we see the bias of the PLS path estimate seems to decrease (gets closer to zero) as sample size decreases. However, if we believe what was suggested from our analysis of Figure 2, 3 and 4, (that PLS and regression estimates do not compensate for measurement error and that we therefore should correct them with Nunnally and Bernstein's equation for attenuation), then we would want to correct the results in figure 6a and 6b as well. This produces Figures 7a and 7b. Again, Figure 7a seems quite reasonable, but Figure 7b shows PLS generating a significant positive bias, overestimating the true path value by as much as about 14% with a sample size of 20.



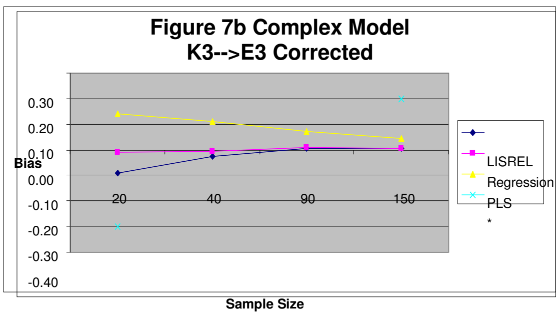


Figure 7a.	*Complex Model*	*(Eta3 →Eta4, Endogenous → Endogenous (Path Bias Corrected)*
Figure 7b.	*Complex Model*	*(Ksi3 →Eta3, Exogenous → Endogenous (Path Bias Corrected)*

We now have some persuasive evidence (in the left hand side of Figure 4 and in the Figure 7a) that PLS (like regression) for the most part does not compensate for measurement error in its path estimates.   On the other hand we have some evidence (in the right hand side of Figure 4 where there is a small effect size, and in Figure 7b with an endogenous construct) that could be interpreted in either of two ways. Either in these specific cases PLS does a better job of compensating for measurement error, or alternatively, PLS doesn't compensate for measurement error any more here than before, but instead tends to overestimate the path values under these specific circumstances, which appears to partly compensate for the lack of sensitivity to measurement error.

5 CONCLUSION

PLS has become a very popular statistical technique for MIS researchers when they are analyzing causal path models involving multiple indicators for constructs. It appears that many users of PLS believe it to be similar to CB-SEM statistical techniques in many respects. Our results demonstrate that, like regression, PLS produces path estimates that are attenuated by measurement error. In that respect, PLS is actually much more similar to regression than it is to CB-SEM techniques.

We also demonstrated that, under many circumstances, applying Nunnally and Bernstein's (1994) formula for attenuation can improve the accuracy of the PLS (and regression) path estimates, essentially removing the bias caused by the technique not accounting for measurement error.

We hope this study will encourage other researchers to take a closer look at the efficacy of PLS, and subsequently assist the MIS research community in becoming more informed users of this, and other, statistical analysis techniques.

References

Chin, W.W. "The Partial Least Squares Approach to Structural Equation Modeling," in G.A. Marcoulides (Ed.) *Modern Methods for Business Research*, London, 1998, pp. 295-336.

Chin, W.W., Marcolin, B. and P. Newsted. 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) 189-217.

Chin, W. W., and Newsted, P. R. "Structural Equation Modeling analysis with Small Samples Using Partial Least Squares"**.** In Rick Hoyle (Ed.), *Statistical Strategies for Small Sample Research*, Sage Publications, 1999, pp. 307-341.

Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*, L. Erlbaum Associates, Hillside, NJ, 1988.

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

Evermann, J., and Tate, M. "Testing Models or Fitting Models? Identifying Model Misspecification in PLS," proceedings, *International Conference on Information Systems*, St. Louis, 2010.

Goodhue, D., Lewis, W. and R. Thompson (2006) "Small Sample Size and Statistical Power in MIS Research," Proceedings of the 39th Hawaii International Conference on Systems Sciences, (CD), R. Sprague (Ed) IEEE Computer Society Press, Los Alamitos, CA, January 4-7, pg. 1-10.

Goodhue, D., Lewis, W., and Thompson, R. "Is PLS A Silver Bullet For Small Sample Size and Non-Normal Data? Testing Statistical Techniques With Monte Carlo Simulation" Working paper, 2009.

Hwang, Malhotra, Kim, Tomiuk and Hong, "A Comparative Study on Parameter Recovery of Three Approaches to Structural Equation Modeling," *Journal of Marketing Research*, forthcoming

Marcoulides, G.A. and Saunders, C. "Editor's Comments: PLS: A Silver Bullet?" *MIS Quarterly*, 30(2), June 2006, pp. iii-ix.

McDonald, R.P. "Path Analysis with Composite Variables," *Multivariate Behavioral Research*, 31 (2), 1996, 239-270.

Nunnally, J.C. and Bernstein, I.H. (1994). *Psychometric theory,(3rd ed.).* , New York, NY: McGraw-Hill.

Rönkkö, M., Ylitalo, J. "Construct Validity in Partial Least Squares Path Modeling," proceedings, *International Conference on Information Systems*, St. Louis, 2010.