Diagnosing Nonlinearity With Confidence Envelopes for a Semiparametric Approach to Modeling Bivariate Nonlinear Relations Among Latent Variables

Jolynn Pek and R. Philip Chalmers
York University

During the early phases of research, semiparametric models (SPMs) have the advantage of recovering latent nonlinearity over parametric counterparts. Structural equation mixture models (Bauer, 2005) can be applied as SPMs to flexibly recover and describe the form of the unknown latent relationship with minimal distributional assumptions. This short report extends the work on this SPM (Bauer, 2005; Pek, Losardo & Bauer, 2011) by developing approximate simultaneous confidence bands or confidence envelopes (CEs) to evaluate potential nonlinearity of the unknown latent function. A line-finding algorithm to be used in conjunction with these CEs is also developed as an implementation of an informal test to diagnose nonlinearity. Coverage of the CEs and performance of the algorithm in terms of rates of detecting latent nonlinearity are evaluated by Monte Carlo. Recommendations for the use of these CEs and the algorithm for detecting nonlinearity are suggested.

Keywords: confidence envelopes, delta method, diagnostics, nonlinear, parametric bootstrap, structural equation models

Structural equation modeling (SEM) is a highly flexible and powerful framework for estimating relationships among latent variables. The development of structural equation models began from linear structural equations and has been extended to include nonlinear structural equations (see Coenders, Batista-Foguet, & Saris, 2008; Dimitruk, Schermelleh-Engel, Kelava, & Moosbrugger, 2007; Schumacker & Marcoulides, 1998). Nonlinear structural equations were first modeled using parametric approaches (Jaccard & Wan, 1995; Jöreskog & Yang, 1996; Kenny & Judd, 1984; Mooijaart & Bentler, 2010; Ping, 1996), which require the explicit specification of the form between latent predictor and outcome.

During the early phases of research, the form linking latent predictor and latent outcome is typically unknown, raising the need for approaches that can flexibly recover the latent relationship without its explicit specification. One semiparametric approach to recovering the unknown form between latent outcome and latent predictor (Bauer, 2005) involves the use of structural equation mixture models (SEMMs; Arming & Stein, 1997; Arming, Stein, & Wittenberg, 1999; Dolen & van der Maas, 1998; Jedidi, Jagpal, & DeSarbo, 1997a, 1997b; Muthén & Shedd, 1999). In contrast to applying SEMMs for directly discerning population heterogeneity, Bauer’s (2005) semiparametric model (SPM) is an indirect application (Titterington, Smith, & Makov, 1985) where the mixing components serve as a statistical expedience to recover regularity in the aggregate population. This semiparametric modeling approach recovers the global latent relationship by using component probabilities as weights in aggregating locally linear within-component latent relationships.

The semiparametric modeling approach has several notable differences from its parametric counterparts. First, the SPM can flexibly recover global latent relationships of unknown functional form without assuming multivariate normally distributed latent variables. Unlike parametric
approaches, the latent variable distributions in the SPM are approximated by a mixture of normal distributions that can take on a variety of nonnormal forms. Second, the global latent relationship is a nonlinear function of a subset of SPM parameters, and the form of the latent regression cannot be evaluated by tests of specific model parameters unlike parametric approaches (e.g., Klein & Moosbrugger, 2000; Klein & Schermelleh-Engel, 2010). Third, the flexibility of the SPM to recover unspecified nonlinearity, in contrast to parametric approaches, stems from the estimation of more parameters, which results in higher sampling variability. The semiparametric modeling approach to recovering nonlinearity is characteristically exploratory, and tools to evaluate nonlinearity between latent variables are further developed here.

A pair of likelihood ratio tests (LRTs) exist as sufficient conditions for establishing global nonlinearity of the latent function in SPM (see Bauer, 2005, for technical details, and Pek, Sterba, Kok, and Bauer, 2009, for empirical examples). These LRTs test conditions where the global latent function reduces to a linear one; if either LRT is nonsignificant, the null hypothesis of linearity cannot be rejected. However, a lack of evidence for linearity does not imply evidence for linearity (cf. Altman & Bland, 1995); research in equivalence testing (e.g., Rogers, Howard, & Vessey, 1993; Seaman & Serlin, 1998) suggests that concluding linearity from the nonrejection of either LRT reverses the null and alternative hypotheses, resulting in low power and high Type I error rates. As an alternative, we develop confidence envelopes (CEs; Working & Hotelling, 1929) for the SPM and present an algorithm that is designed to detect nonlinearity of the latent relationship.

APPORXIMATE CONFIDENCE ENVELOPES

Confidence envelopes are a set of interval estimates constructed across the range of the latent predictor \( \eta_1 \) with which the true population function between latent predictor and outcome \( E[\eta_2|\eta_1] \) is expected to fall within \((1 - \alpha)100\% \) of the time over repeated sampling; CEs allow for hypotheses about the form between \( \eta_2 \) and \( \eta_1 \) to be tested. We present two CEs for the SPM based on the delta method and parametric bootstrap approach.

Delta Method

Approximate \((1 - \alpha)100\% \) delta method Wald-type confidence intervals (CIs) and CEs for the SPM are constructed using the formula: estimate \( \pm \) critical value \( \times \) standard error of estimate, where the critical value differs between CIs and CEs (Scheffé, 1953). Given normality of the sampling distribution of \( E[\eta_2|\eta_1] \), CEs are constructed by using the critical value \( \sqrt{\chi^2_{p,1-\alpha}} \), where \( \chi^2_{p,1-\alpha} \) is the \((1 - \alpha)\) quantile of the \( \chi^2 \) distribution with \( p \) degrees of freedom that is the number of parameters used to estimate \( E[\eta_2|\eta_1] \). Details on the delta method standard error of estimate for the SPM are presented in Pek, Losardo, and Bauer (2011).

Parametric Bootstrap

Parametric bootstrap CEs are constructed via a resampling algorithm where some number of bootstrap draws are sampled from \( \hat{\theta}_p \), where \( g(\hat{\theta}_p) = E[\eta_2|\eta_1] \), to obtain bootstrap estimates of the latent function (see Efron & Tibshirani, 1993, for an introduction, and Pek et al., 2011, for parametric bootstrap CIs in the SPM). In the context of CEs for the SPM, the resampled bootstrap functions empirically estimate the distribution \( E[\eta_2|\eta_1] \) such that the CEs are determined from the boundary of the overlap of the bootstrap curves. The number of bootstrap draws is based on the expected value of the range of \( \sqrt{\chi^2_{p,1-\alpha}} [\text{VAR}(\hat{\theta}_p)]^{1/2} \), where \( \text{VAR}(\hat{\theta}_p) \) is the root of the estimated variance–covariance matrix of \( \hat{\theta}_p \) (Thissen & Wainer, 1990). The online appendix provides a table of \( B \) draws determined by \( p \) for \( \alpha = .05 \) and \( \alpha = .10 \), and more details of this method.

DIAGNOSING NONLINEARITY

The developed CEs for the SPM can be used to inform of the nonlinearity of \( E[\eta_2|\eta_1] \) by evaluating the null hypothesis \( H_0: \eta_2 = \beta_0 + \beta_1 \eta_1 \), where \( \beta_0 \) and \( \beta_1 \) denote the intercept and slope, respectively. Nonlinearity is implied when the \((1 - \alpha)100\% \) CE does not contain all possible linear regressions of \( \eta_2 \) onto \( \eta_1 \). Conversely, if a single linear function relating \( \eta_2 \) to \( \eta_1 \) is contained within the CE, there is insufficient evidence for nonlinearity. Later, we describe a line-finding algorithm, based on how CEs and regression lines are graphed, that is developed to detect nonlinearity of \( E[\eta_2|\eta_1] \) by informally testing the stated \( H_0 \).

A CE for \( E[\eta_2|\eta_1] \) is graphed from two sets of \( Q \) points that are interpolated to form the upper and lower confidence bands. Figure 1 depicts a CE represented by black curves that have been interpolated from two sets of \( Q = 50 \) points that are represented by open circles. The line-finding algorithm begins by constructing candidate lines that could fall within the CE by using every point of the set of 2\( Q \) points that form the envelope. There are \( Q^2 \) unique lines that intersect points of the upper to lower bounds of the CE, or a grid of \( Q^2 \) points that form these lines. In Figure 1, four candidate lines are constructed from the 37th point of the upper bound to four different points on the lower bound (5th, 10th, 20th, and 30th points). After constructing candidate lines, the algorithm examines whether each point of the candidate lines lies within or outside the CE. In Figure 1, the solid line is contained by the CE, implying that the latent function could be linear.
FIGURE 1 Construction of a confidence envelope (CE) with \( Q = 50 \) points across \( \eta_1 \), and an illustration of the line-finding algorithm for diagnosing nonlinearity. As the dark solid line is contained within the CE, nonlinearity cannot be concluded.

MONTE CARLO EVALUATIONS

Simulation methodology was used to evaluate the coverage of the two types of CEs, and the performance of the line-finding algorithm. The form (linear, quadratic, or exponential) and curvature (low or high) of the true latent relationship, the variance of the latent outcome accounted for by the latent predictor (50% or 25%), and sample size \((N = 250, 500, \text{ or } 1,000)\) were manipulated. These conditions are an extension of those in Pek et al. (2011), and the online appendix provides details of the eight population-generating functions. The study had \(8 \times 3 = 24\) conditions (see Figures 2 and 3), from which 1,000 data sets were generated. Four levels of \( Q = 25, 50, 100, \text{ or } 250\) were evaluated.

All models were estimated with Mplus 6.2 (Muthén & Muthén, 2011), and start values for each replication followed the methodological approach of Bauer, Baldasaro, and Gottfredson (2012). For each replication, the number of mixing components was determined by the Akaike information criterion (AIC; Akaike, 1974) because simulation work by Bauer, et al. (2012) confirms that selecting the number of classes by AIC results in less bias than the Bayesian information criterion (BIC; Schwarz, 1978) in the recovery of latent functions. Following model convergence, we constructed 95% delta method and parametric bootstrap CEs.

Coverage

Coverage assesses the probability that the CEs contain the population function as a whole, across the range of the latent predictor \( \eta_1 \). Coverage rates for the delta method and parametric bootstrap CEs for the eight functions are presented in Figure 2. As the SPM is relatively statistically inefficient, requiring the estimation of more parameters than necessary, coverage rates tend to be above the expected .95 rate for the linear and quadratic population functions. Coverage for the exponential functions with small curvature were close to the nominal rate of .95. The different coverage rates of the exponential functions with different curvature reflects bias in the SPM’s approximation of the global function; large curvature is associated with more bias (see Figure B in online appendix).

Coverage rates did not vary by \( Q \), increased with increasing sample size, and were closer to the nominal rate of .95 for the parametric bootstrap CEs compared to the delta method CEs in all conditions except for the first exponential function (large curvature and variance; Exponential 1 in Figure 2). Coverage rates for the delta method CEs for this exponential function increased and then slightly decreased with sample size, mirroring the sample size effect on bias (Figure B in online appendix). Similar effects were observed for delta method CIs in Pek et al. (2011), which could be due to the plausible violation of the parameter drift assumption (Stroud, 1972; Wald, 1943).

Nonlinearity

The line-finding algorithm was developed as a diagnostic tool, in conjunction with the CEs, to inform whether the latent bivariate relationship of unknown functional form recovered by the SPM is nonlinear. Rates of diagnosing nonlinearity are presented in Figure 3. As expected, the line-finding algorithm correctly and reliably does not detect nonlinearity across all study conditions for the linear population functions for all replications.

For the symmetric or quadratic functions examined, rates of diagnosing nonlinearity increased with increasing sample size (second row of plots in Figure 3). The positive effect of sample size reflects increasing efficiency, mirroring tighter CEs. The \( Q \) number of points had a very small positive effect on these rates; \( Q \geq 50 \) is associated with more precision and was found to be associated with very slightly larger rejection rates compared to \( Q = 25 \). Rates of diagnosing linearity were also observed to be higher for the delta method CEs compared to the bootstrap CEs, which might due to (a) the symmetry versus asymmetry of the delta method versus the bootstrap CEs, and (b) the bootstrap CEs tending to be very slightly wider than delta method CEs (see Figure C in online appendix). Both cases result in the bootstrap CEs containing candidate lines more often than the delta method CEs. Additionally, there were notable effects of curvature on these rates between the two CEs. Curvature had a small negative effect on these rates for delta method CEs, which is likely due to these CEs being larger at the tail ends of \( \eta_1 \) for the larger versus smaller curvature conditions (see Figure
C in online appendix). Conversely, rates for the bootstrap CEs were slightly higher for the higher curvature condition, which could be due to the fact that the larger the curvature of the function, the less it would be considered linear. Finally, rates of diagnosing nonlinearity showed an interaction effect of variance accounted for in $\eta_2$ by $\eta_1$, type of CE, and sample size. With decreasing optimality of the data, rates of diagnosing nonlinearity decreased with lower variance and smaller sample sizes; larger effects of these two factors were also observed for bootstrap CEs compared to delta method CEs.

For the exponential functions examined, rates of detecting nonlinearity were observed to be stable when $Q \geq 50$. Curvature, variance accounted for in $\eta_2$ due to $\eta_1$, sample size, and type of CE interacted to influence rates of detecting nonlinearity for these asymmetric functions. Rates tended to be higher for delta method CEs compared to bootstrap CEs, larger curvature, larger sample sizes, and larger variance accounted for in $\eta_2$ by $\eta_1$ (third row of plots in Figure 3). Similar to the quadratic conditions, these results reflect the tendency for bootstrap CEs to contain candidate lines more often than delta method CEs due to the former’s asymmetry and lower efficiency. Additionally the positive effect of curvature on these rates reflect effect size differences in the form of greater departure from linearity. Overall, the CEs used in conjunction with the line-finding algorithm do not detect asymmetric nonlinearity as well as symmetric nonlinearity.

**CONCLUSIONS AND RECOMMENDATIONS**

Latent variable regressions of unknown functional form can be adequately recovered with the SPM, and delta method and parametric bootstrap CEs can reasonably depict estimate precision of the latent function as a whole. In addition, the developed line-finding algorithm could be a useful diagnostic tool for detecting nonlinearity of the unknown function in an exploratory analysis, especially when symmetric nonlinearity is present, sample size is large, curvature is large, variance accounted for in $\eta_2$ by $\eta_1$ is large, and $Q \geq 50$. 
The choice of constructing delta method versus parametric bootstrap CEs depends on the goals of the analysis, as both types of CEs differ somewhat on certain properties. In terms of coverage, the delta method CEs tended to be more conservative compared to the bootstrap CEs. Additionally, the bootstrap CEs had coverage rates that were closer to nominal levels compared to the delta method CEs. Bootstrap CEs therefore have a small advantage over delta method CEs in terms of coverage. In terms of diagnosing nonlinearity, delta method CEs had higher rates of detection compared to bootstrap CEs. Both CEs did not detect nonlinear asymmetric functions well. In general, delta method CEs can more often ascertain nonlinearity in the targeted conditions examined.

In practice, parametric bootstrap CEs require much computational power and some “tuning” to construct. For instance, although computational power is readily available with modern computers, it took an average of 20 min to construct one bootstrap CE about a function with \( p = 16 \) SPM parameters on a computer with 3.40 GHz of clock speed. With parallel processing, computing efficiency would improve. When constructing bootstrapped CEs, inadmissible bootstrapped values for latent predictor variances would either have to be truncated at some arbitrary value or censored, technically increasing the number of \( B \) draws required to construct the CE. Despite the noted differences between delta method and bootstrap CEs in terms of coverage and rates of diagnosing nonlinearity, these two types of CEs are very similar for any given replication. With the availability of both types of CEs, a sensitivity analysis could be conducted to assess whether results are robust to the type of CE constructed. The availability of these exploratory tools (Pek, Chalmers, Kok, & Losardo, 2015) should encourage the diagnostic assessment of nonlinearity among latent variables.
ACKNOWLEDGMENTS

We thank Daniel J. Bauer for helpful comments on an earlier draft and Jifeng Li for computing support. Part of this work was made possible by the facilities of the Shared Hierarchical Academic Research Computing Network (SHARCNET: www.sharcnet.ca) and Compute/Calcul Canada.

FUNDING

This work was supported by the Faculty of Health Minor Research Grant, York University and the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant, both awarded to Jolynn Pek.

SUPPLEMENTAL MATERIAL

Supplemental data for this article can be accessed at the publisher’s website.

REFERENCES

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19, 716–723.

Altman, D. G., & Bland, J. M. (1995). Statistics notes: Absence of evidence is not evidence of absence. *British Medical Journal*, 311, 485.

Armingier, G., & Stein, P. (1997). Finite mixtures of covariance structure models with regressors: Loglikelihood function, minimum distance estimation, fit indices, and a complex example. *Sociological Methods & Research*, 26, 148–182.

Armingier, G., Stein, P., & Wittenberg, J. (1999). Mixtures of conditional mean- and covariance-structure models. *Psychometrika*, 64, 475–494.

Bauer, D. J. (2005). A semiparametric approach to modeling nonlinear relations among latent variables. *Structural Equation Modeling*, 12, 513–535.

Bauer, D. J., Baldasaro, R. E., & Gottfredson, N. C. (2012). Diagnostic procedures for detecting nonlinear relationships between latent variables. *Structural Equation Modeling*, 19, 157–177.

Coenders, G., Batista-Foguet, J. M., & Saris, W. E. (2008). Simple, efficient and distribution-free approach to interaction effects in complex structural equation models. *Quality & Quantity*, 42, 369–396.

Dimitruk, P., Schermelleh-Engel, K., Kelava, A., & Moosbrugger, H. (2007). Challenges in nonlinear structural equation modeling. *Methodology*, 3, 100–114.

Dolan, C. V., & van der Maas, H. L. (1998). Fitting multivariate normal finite mixtures subject to structural equation modeling. *Psychometrika*, 63, 227–253.

Efron, B., & Tibshirani, R. (1993). An introduction to the bootstrap. New York, NY: Chapman & Hall.

Jaccard, J., & Wan, C. K. (1995). Measurement error in the analysis of interaction effects between continuous predictors using multiple regression: Multiple indicator and structural equation approaches. *Psychological Bulletin*, 117, 348–357.

Jedidi, K., Jagpal, H. S., & De Sarbo, W. S. (1997a). Finite-mixture structural equation models for response-based segmentation and unobserved heterogeneity. *Marketing Science*, 16, 39–59.

Jedidi, K., Jagpal, H. S., & De Sarbo, W. S. (1997b). STEMM: A general finite mixture structural equation model. *Journal of Classification*, 14, 23–50.

Jöreskog, K. G., & Yang, F. (1996). Nonlinear structural equation models: The Kenny–Judd model with interaction effects. In G. A. Marcoulides & R. E. Schumacker (Eds.), *Advanced structural equation modeling: Issues and techniques* (pp. 57–88). Mahwah, NJ: Erlbaum.

Kenny, D. A., & Judd, C. M. (1984). Estimating the nonlinear and interactive effects of latent variables. *Psychological Bulletin*, 96, 201–210.

Klein, A. G., & Moosbrugger, H. (2000). Maximum likelihood estimation of latent interaction effects with the LMS method. *Psychometrika*, 65, 457–474.

Klein, A. G., & Schermelleh-Engel, K. (2010). Introduction of a new measure for detecting poor fit due to omitted nonlinear terms in SEM. *AStA Advances in Statistical Analysis*, 94, 157–166.

Mooijaart, A., & Bentler, P. M. (2010). An alternative approach for nonlinear latent variable models. *Structural Equation Modeling*, 17, 357–373.

Muthén, B. O., & Shedden, K. (1999). Finite mixture modeling with mixture outcomes using the EM algorithm. *Biometrics*, 55, 463–469.

Muthén, L. K., & Muthén, B. O. (2011). *Mplus user’s guide* (6th ed.). Los Angeles, CA: Muthén & Muthén.

Pek, J., Chalmers, R. P., Kok, B. E. & Losardo, D. (2015). Visualizing confidence bands for semiparametrically estimated nonlinear relations among latent variables. *Journal of Educational and Behavioral Statistics*. Manuscript submitted for publication.

Pek, J., Losardo, D., & Bauer, D. J. (2011). Confidence intervals for a semiparametric approach to modeling nonlinear relations among latent variables. *Structural Equation Modeling*, 18, 537–553.

Pek, J., Sterba, S. K., Kok, B. E., & Bauer, D. J. (2009). Estimating and visualizing nonlinear relations among latent variables: A semiparametric approach. *Multivariate Behavioral Research*, 44, 407–436.

Ping, R. A. (1996). Latent variable interaction and quadratic effect estimation: A two-step technique using structural equation analysis. *Psychological Bulletin*, 119, 166–175.

Rogers, J. L., Howard, K. I., & Vessey, J. T. (1993). Using significance tests to evaluate equivalence between two experimental groups. *Psychological Bulletin*, 113, 553.

Scheffé, H. (1953). A method for judging all contrasts in the analysis of variance. *Biometrika*, 40, 87–410.

Schumacker, R. E., & Marcoulides, G. A. (1998). *Interaction and nonlinear effects in structural equation modeling*. Mahwah, NJ: Erlbaum.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6, 461–464.

Seaman, M. A., & Serlin, R. C. (1998). Equivalence confidence intervals for two-group comparisons of means. *Psychological Methods*, 3, 403.

Stroud, T. W. F. (1972). Fixed alternatives and Wald’s formulation of the noncentral asymptotic behavior of the likelihood ratio statistic. *The Annals of Mathematical Statistics*, 43, 447–454.

Thissen, D., & Wainer, H. (1990). Confidence envelopes for item response theory. *Journal of Educational and Behavioral Statistics*, 15, 113–128.

Titterington, D. M., Smith, A. F., & Makov, U. E. (1985). *Statistics and Engineering Research Council of Canada (NSERC)*

Working, H., & Hotelling, H. (1929). Applications of the theory of error to the interpretation of trends. *Journal of the American Statistical Association*, 24, 73–85.