STRUCTURAL EQUATION MODELING, *12*(4), 513–535

A Semiparametric Approach to Modeling Nonlinear Relations Among Latent Variables

Daniel J. Bauer
The University of North Carolina at Chapel Hill

To date, finite mixtures of structural equation models (SEMMs) have been developed and applied almost exclusively for the purpose of providing model-based cluster analyses. This type of analysis constitutes a direct application of the model wherein the estimated component distributions of the latent classes are thought to represent the characteristics of distinct unobserved subgroups of the population. This article instead considers an indirect application of the SEMM in which the latent classes are estimated only in the service of more flexibly modeling the characteristics of the aggregate population as a whole. More specifically, the SEMM is used to semiparametrically model nonlinear latent variable regression functions. This approach is first developed analytically and then demonstrated empirically through analyses of simulated and real data.

The modeling of nonlinear relations between latent variables has been a topic of long-standing interest. Within the factor-analytic tradition, early contributions to nonlinear latent variable modeling were made by Gibson (1959), McDonald (1967), and Etezadi-Amoli and McDonald (1983). Whereas these approaches focused mainly on nonlinear factor-to-item relations, subsequent contributions have focused specifically on modeling nonlinear effects between latent factors in structural equation models. These include the seminal paper by Kenny and Judd (1984) using products of manifest variables to model latent interactions and quadratic effects, as well as subsequent papers refining and extending this product indicant approach (see Schumacker & Marcoulides, 1998, and references therein).

Problems with the product indicant approach included the tedium of properly specifying the necessary nonlinear constraints of the model and the fact that the

Requests for reprints should be sent to Daniel J. Bauer, Department of Psychology, University of North Carolina, CB#3270 Davie Hall, Chapel Hill, NC 27599–3270. E-mail: dbauer@email.unc.edu

normality assumption of the maximum likelihood (ML) fitting function used to estimate the model was necessarily violated: Even if normality of the exogenous latent variables and their manifest indicators is assumed, the product indicators are then nonnormal, as are the endogenous latent and manifest indicators downstream of the nonlinear effect. Robust and distribution-free methods of estimation for the product indicator approach were proposed (e.g., the weighted least squares and robust ML methods of Jöreskog & Yang, 1996, and the two-stage least squares (2SLS) method of Bollen, 1995, and Bollen & Paxton, 1998), but these have not always performed satisfactorily in finite samples (Schermelleh-Engel, Klein, & Moosbrugger, 1998). In response, several new approaches have been developed that explicitly accommodate the nonnormality implied by the nonlinear effects in the model and that have generally shown superior finite-sample performance to earlier methods, including the ML methods of Klein and Moosbrugger (2000), Klein and Muthén (2003), and Lee and Zhu (2002); the method of moments approach of Wall and Amemiya (2000); and the Bayesian approaches of Arminger and Muthén (1998) and Zhu and Lee (1999).

This article proposes a complementary method for modeling nonlinear latent variable relations. Specifically, finite mixtures of structural equation models are used to flexibly approximate the latent regression function. This approach was presaged by Gibson's (1959) work on the use of latent profile analysis to avoid assumptions of linearity in factor analysis, and was suggested for structural equation models more recently by Bauer and Curran (2004). It is distinguished from the other approaches just described by the fact that it is semiparametric. This feature of the method provides two distinct advantages. First, whereas parametric methods require the functional form of the nonlinear relation to be specified in advance (e.g., as quadratic), the proposed semiparametric approach does not require any prior knowledge of the nature of the modeled relation. The second advantage concerns the distributional assumptions of the model. Unlike many parametric approaches, normality is not required for either the exogenous factors or the disturbances of the endogenous factors, which are instead modeled as mixtures of normal distributions.

To begin, the article describes the structural equation mixture model (SEMM) that forms the basis of the approach. This is followed by a demonstration of how this model can be used to estimate and test nonlinear functions among latent variables. Next, two empirical examples are presented. The first example is a small simulation that both clarifies the mechanics of the method and demonstrates its accuracy and precision within samples of moderate size. The second example is an application to data from the Youth Development Study (Mortimer, 2003) that examines the potentially nonlinear relations among the latent variables of mastery, optimism, and depression. The article concludes with a discussion of promising directions for future research.

THE STRUCTURAL EQUATION MIXTURE MODEL

The SEMM was developed simultaneously and independently by Jedidi, Jagpal, and DeSarbo (1997a, 1997b), Arminger and colleagues (Arminger & Stein, 1997; Arminger, Stein, & Wittenberg, 1999), and Dolan and van der Maas (1998), with later extensions by B. O. Muthén (2001). Prior developments by Blåfield (1980), Yung (1997), and others are reviewed in Arminger et al. (1999) and Bauer and Curran (2004). To facilitate expression of the SEMM, and to establish notation, the standard structural equation model is described first.

The Structural Equation Model

The *measurement model* of the structural equation model relates the continuous observed measures to the underlying continuous latent factors through the equation

$$\mathbf{y}_i = \boldsymbol{\nu} + \boldsymbol{\Lambda}\boldsymbol{\eta}_i + \boldsymbol{\varepsilon}_i; \quad \boldsymbol{\varepsilon}_i \sim N_p(\mathbf{0}, \boldsymbol{\Theta}), \tag{1}$$

where $\mathbf{y}_i$ is a $p \times 1$ vector of observed variables and $\boldsymbol{\eta}_i$ is an $m \times 1$ vector of latent variables for cases $i = 1, 2, \ldots, N$, $\boldsymbol{\nu}$ is a $p \times 1$ vector of measurement intercepts, $\boldsymbol{\Lambda}$ is a $p \times m$ matrix of factor loadings, and $\boldsymbol{\varepsilon}$ is a $p \times 1$ vector of residuals. The *latent variable model* of the structural equation model is then given as

$$\boldsymbol{\eta}_i = \boldsymbol{\alpha} + \mathbf{B}\boldsymbol{\eta}_i + \boldsymbol{\zeta}_i; \quad \boldsymbol{\zeta}_i \sim N_q(\mathbf{0}, \boldsymbol{\Psi}), \tag{2}$$

where $\boldsymbol{\alpha}$ is an $m \times 1$ vector of intercepts for the latent variables, $\mathbf{B}$ is an $m \times m$ matrix of regression coefficients among the latent variables, and $\boldsymbol{\zeta}$ is an $m \times 1$ vector of residuals. It is assumed that $\boldsymbol{\zeta}$ is uncorrelated with $\boldsymbol{\varepsilon}$.[1]

Given Equations 1 and 2, the marginal distribution of $\mathbf{y}$ is multivariate normal with means and covariances parameterized as

$$\boldsymbol{\mu}(\boldsymbol{\theta}) = \boldsymbol{\nu} + \boldsymbol{\Lambda}(\mathbf{I} - \mathbf{B})^{-1}\boldsymbol{\alpha} \tag{3}$$

and

$$\boldsymbol{\Sigma}(\boldsymbol{\theta}) = \boldsymbol{\Lambda}(\mathbf{I} - \mathbf{B})^{-1}\boldsymbol{\Psi}(\mathbf{I} - \mathbf{B})^{-1'}\boldsymbol{\Lambda}' + \boldsymbol{\Theta}, \tag{4}$$

where $\boldsymbol{\theta}$ is the vector of model parameters. Further detail on the standard structural equation model can be obtained from any of a number of excellent sources, includ-

[1]The normality assumptions of Equations 1 and 2 are not always required in the standard structural equation model, but they are essential for extending to the SEMM.

ing Bollen (1989), Kaplan (2000), and Kline (1998). The SEMM expands this model by allowing the estimation of a finite mixture of structural equation models for several latent groups.

Finite Mixture Structural Equation Model

The SEMM assumes that a structural equation model of the form given in Equations 1 and 2 holds within each of K latent classes. The probability density function (PDF) of the SEMM can then be written as

$$f(\mathbf{y}) = \sum_{k=1}^{K} P(k)\phi_k \left[\mathbf{y}; \boldsymbol{\mu}_k\left(\boldsymbol{\theta}_k\right), \boldsymbol{\Sigma}_k\left(\boldsymbol{\theta}_k\right)\right] \tag{5}$$

where $P(k)$ is the mixing probability (or proportion) for class k = 1, 2, … K, $\sum_{K=1}^{K} P(k) = 1$, and the variables $\mathbf{y}$ within each class follow a multivariate normal PDF (denoted ϕ_k) with mean vector $\boldsymbol{\mu}_k(\boldsymbol{\theta}_k)$ and covariance matrix $\boldsymbol{\Sigma}_k(\boldsymbol{\theta}_k)$ parameterized as

$$\boldsymbol{\mu}_k\left(\boldsymbol{\theta}_k\right) = \boldsymbol{\nu}_k + \boldsymbol{\Lambda}_k\left(\mathbf{I} - \mathbf{B}_k\right)^{-1}\boldsymbol{\alpha}_k \tag{6}$$

and

$$\boldsymbol{\Sigma}_k\left(\boldsymbol{\theta}_k\right) = \boldsymbol{\Lambda}_k\left(\mathbf{I} - \mathbf{B}_k\right)^{-1}\boldsymbol{\Psi}_k\left(\mathbf{I} - \mathbf{B}_k\right)^{-1'}\boldsymbol{\Lambda}_k' + \boldsymbol{\Theta}_k. \tag{7}$$

Each parameter matrix has been subscripted by k to denote that the values of the parameters may vary over classes, although in any given application it is likely that many of these matrices would be constrained to be invariant over classes.[2] In addition, the total number of classes K is often unknown but is seldom estimated directly. In practice, one estimates models with successively greater values for K and compares model fit to determine the optimal value for K, for instance, by selecting the model with minimum Bayes's Information Criterion (BIC).

Finite mixture models like the SEMM are often employed to cluster observations into groups. Unlike heuristic agglomerative or partitioning clustering algorithms, finite mixture models rely on an explicit underlying statistical model, that

[2]Several important extensions to the basic model in Equation 5 are described by Arminger et al. (1999) and B. O. Muthén (2001), including the estimation of conditional SEMMs with nonnormal exogenous regressors (i.e., categorical predictors), the inclusion of categorical latent class indicators, and modeling of the class probabilities as a function of exogenous predictors. These extensions are not detailed here as the simpler model in Equation 5 is sufficient for the purposes of this article. However, all of the techniques developed here could be applied with little or no modification in the context of these more complicated models.

the component distributions of the latent groups are multivariate normal. For the SEMM, the conditional probability of group membership given the data is shown by Bayes' Theorem to be

$$P(k \mid \mathbf{y}) = \frac{P(k)\phi_k\left[\mathbf{y};\boldsymbol{\mu}_k\left(\boldsymbol{\theta}_k\right),\boldsymbol{\Sigma}_k\left(\boldsymbol{\theta}_k\right)\right]}{\sum_{k=1}^{K}P(k)\phi_k\left[\mathbf{y};\boldsymbol{\mu}_k\left(\boldsymbol{\theta}_k\right),\boldsymbol{\Sigma}_k\left(\boldsymbol{\theta}_k\right)\right]}. \tag{8}$$

Estimated posterior probabilities of class membership can be computed from Equation 8 by replacing the population parameter values by their sample estimates. Even when no clustering is desired, the posterior probabilities of class membership figure importantly in the estimation of mixture models via the expectation-maximization (EM) algorithm. In the expectation step, posterior probabilities are estimated for each case given the present values of the parameter estimates. Then, in the maximization step, these posterior probabilities are used as case weights in the maximization of K separate within-group likelihood functions. The two steps are repeated iteratively until convergence criteria are satisfied (L. K. Muthén & Muthén, 2001). Because local solutions are endemic to mixture models, the use of multiple starting values is necessary to ensure that a global solution has been obtained.

It is worth noting that the expressions for the component mean vectors and covariance matrices in Equations 6 and 7 are identical in form to the multisample structural equation model developed by Jöreskog (1971) and Sörbom(1974) for the simultaneous analysis of multiple observed groups. The relation between the two models is particularly salient in the expression of the EM algorithm. If the posterior probabilities for all cases are fixed at 1 or 0, assigning each case to a single class, then each case will contribute to only one within-class likelihood and the model will reduce to the multisample structural equation model (Jedidi et al., 1997b; L. K. Muthén & Muthén, 2001, Appendix 8; Zhu & Lee, 2001). Thus, one way to conceptualize the SEMM is as a multisample structural equation model that allows for measurement error in the classification of groups. Failing to account for such errors may lead to biased estimates (Nagin, 1999). The relation between the two models also implies that the same identification conditions that apply to the multisample SEM (discussed in Sörbom, 1974) are also required for the SEMM (Jedidi et al., 1997a).

Applications of SEMMs

It is perhaps also due to its relation with the multisample SEM for the analysis of observed groups that the SEMM has been applied almost exclusively with the goal of identifying latent groups (Dolan & van der Maas, 1998). Such model-based clustering applications of mixture models have been referred to by Titterington, Smith, and Makov (1985) as direct applications, as the component distributions are interpreted directly as representative of distinct population subgroups. In contrast, in indirect applications, the latent classes are used solely to provide a tractable form of analysis for data that may not obey traditional parametric models and are not given substantive interpretations (Titterington et al., 1985). In this sense, the SEMM may be viewed as intermediate to the traditional parametric structural equation model (assuming that $f(\mathbf{y})$ is a multivariate normal PDF) and a fully nonparametric approach (e.g., kernel methods). Like other mixture models, it can thus be characterized as semiparametric (Escobar & West, 1995; Everitt & Hand, 1981; Ferguson, 1983; McLachlan & Peel, 2000; Nagin, 1999; Roeder & Wasserman, 1997; Silverman, 1986; Verbeke & Lesaffre, 1996).

The purpose to which the SEMM is put in this article is an indirect application. To be explicit, it is not necessary for the latent classes to represent true population subgroups. Neither the number of latent classes nor the estimates obtained within classes will be interpreted substantively. Rather, the classes are estimated only in the service of constructing a semiparametric estimate of the regression function for the latent variables in the population as a whole. Thus, unlike direct applications, for this indirect application of the SEMM, one's theoretical model for the phenomenon need not include latent groups or population heterogeneity—it need only allow for the possibility that the relations among the latent variables are nonlinear in the aggregate population.

SEMIPARAMETRIC LATENT VARIABLE REGRESSION

The basis of the nonlinear approximation afforded by the SEMM is described next. In what follows, metric invariance is assumed to hold across classes so that the continuous latent variables can be regarded as "the same" for all classes (Meredith, 1993). Isomorphism of the latent factors over classes is a logical precondition for the computation of estimates that aggregate over classes. This assumption is not unreasonable given that these matrices are implicitly held invariant for all individuals in the population in traditional (single-group) structural equation models as well. Possible uses of the semiparametric estimator under less stringent conditions are discussed in the concluding section of the article.

Capturing Nonlinear Trends

To begin, let us partition the vector of latent factors into a set of exogenous latent factors $\boldsymbol{\eta}_1$ and a set of endogenous latent factors $\boldsymbol{\eta}_2$ such that $\boldsymbol{\eta}' = [\boldsymbol{\eta}_1' \; \boldsymbol{\eta}_2']$. The following partitions then also obtain:

$$\boldsymbol{\alpha}_k' = \begin{bmatrix} \boldsymbol{\alpha}_{1k}' & \boldsymbol{\alpha}_{2k}' \end{bmatrix} \tag{9}$$

$$\mathbf{B}_k = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{B}_{21k} & \mathbf{0} \end{bmatrix} \tag{10}$$

$$\boldsymbol{\Psi}_k = \begin{bmatrix} \boldsymbol{\Psi}_{11k} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Psi}_{22k} \end{bmatrix} \tag{11}$$

For simplicity, these expressions assume that the only causal effects are from the exogenous to the endogenous factors and that the exogenous factors are uncorrelated with the endogenous disturbances, though these assumptions could easily be relaxed.[3] Within each class, the conditional distribution of $\boldsymbol{\eta}_2$ is then normal with residual covariance matrix $\boldsymbol{\Psi}_{22k}$ and expected value

$$E_k(\boldsymbol{\eta}_2 \mid \boldsymbol{\eta}_1) = \boldsymbol{\alpha}_{2k} + \mathbf{B}_{21k}\boldsymbol{\eta}_1, \tag{12}$$

Although the regression of $\boldsymbol{\eta}_2$ on $\boldsymbol{\eta}_1$ is linear within each latent class *k,* nonlinearity in the regression of $\boldsymbol{\eta}_2$ on $\boldsymbol{\eta}_1$ in the total population can still be recovered by allowing the parameters in Equations 9, 10, and 11 to vary over classes.

Application of the law of total expectation to aggregate across the mixing components shows that the expected value of $\boldsymbol{\eta}_2$ given $\boldsymbol{\eta}_1$ is a weighted sum of the within-class expected values, where the weights are the conditional probabilities of class membership given $\boldsymbol{\eta}_1$:

$$E(\boldsymbol{\eta}_2 \mid \boldsymbol{\eta}_1) = \sum_{k=1}^{K} P(k \mid \boldsymbol{\eta}_1) E_k(\boldsymbol{\eta}_2 \mid \boldsymbol{\eta}_1). \tag{13}$$

Because the marginal distribution of $\boldsymbol{\eta}_1$ is approximated by a mixture of normals, $P(k|\boldsymbol{\eta}_1)$ can be expressed as

$$P(k \mid \boldsymbol{\eta}_1) = \frac{P(k)\phi_k(\boldsymbol{\eta}_1; \boldsymbol{\alpha}_{1k}, \boldsymbol{\Psi}_{11k})}{\sum_{k=1}^{K} P(k)\phi_k(\boldsymbol{\eta}_1; \boldsymbol{\alpha}_{1k}, \boldsymbol{\Psi}_{11k})}, \tag{14}$$

Note that changes in $P(k|\boldsymbol{\eta}_1)$ over the range of $\boldsymbol{\eta}_1$ permit the function in Equation 13 to obtain a smooth nonlinear form. For instance, consider a simple case where

[3]This approach is equally applicable to the case where either or both of the variables under study are observed rather than latent. However, in the case where the exogenous predictor is a continuous observed variable (i.e., a so-called fixed *x* model), the predictor must be included in the joint distribution modeled by the mixture. That is, for the purpose of evaluating nonlinearity in the manner described here, one should not model the mixture conditional on the exogenous predictor. Importantly, this implies that the distribution of the observed predictor can be well approximated by a normal mixture, a condition not required by the conditional SEMM. The model can still be conditioned on other exogenous predictors (i.e., dichotomous variables) where this assumption would be unrealistic (see Arminger et al., 1999, for further discussion of the conditional SEMM).

there is a single exogenous factor predicting a single endogenous factor, $K = 2$, the class variances for the exogenous factor are invariant, and the mean of the exogenous factor is lower in Class 1 than Class 2. In this case, as $\boldsymbol{\eta}_1$ moves from small to large values, $P(k = 1|\boldsymbol{\eta}_1)$ will shift in a nonlinear way from a value near 1 to a value near 0 and the complementary change will take place in $P(k = 2|\boldsymbol{\eta}_1)$. Thus $E(\boldsymbol{\eta}_2|\boldsymbol{\eta}_1)$ will also shift away from $E_1(\boldsymbol{\eta}_2|\boldsymbol{\eta}_1)$ and toward $E_2(\boldsymbol{\eta}_2|\boldsymbol{\eta}_1)$ in a nonlinear way.

Testing for Nonlinear Trends

These observations may also be used to construct a test of whether the functional relation between $\boldsymbol{\eta}_1$ and $\boldsymbol{\eta}_2$ is linear or not. Clearly, if the best fitting model includes only one class then this is implicit support for the linear model. Of greater interest is the case where $K > 1$. Rewriting Equation 13 as

$$E(\boldsymbol{\eta}_2 \mid \boldsymbol{\eta}_1) = \left[\sum_{k=1}^{K} P(k \mid \boldsymbol{\eta}_1)\boldsymbol{\alpha}_{2k}\right] + \left[\sum_{k=1}^{K} P(k \mid \boldsymbol{\eta}_1)\mathbf{B}_{21k}\right]\boldsymbol{\eta}_1, \tag{15}$$

it can be seen that linearity is implied by two conditions. One necessary condition for $E(\boldsymbol{\eta}_2|\boldsymbol{\eta}_1)$ to exhibit a nonlinear trend is that $P(k|\boldsymbol{\eta}_1)$ must vary with $\boldsymbol{\eta}_1$. If instead $P(k|\boldsymbol{\eta}_1) = P(k)$ then the two bracketed terms will correspond to the intercept and slope of a linear function, as the weights used to combine the class intercepts and slopes will be constant over $\boldsymbol{\eta}_1$. By definition, for $P(k|\boldsymbol{\eta}_1)$ to equal $P(k)$, class membership must be independent of $\boldsymbol{\eta}_1$, implying that the distribution of $\boldsymbol{\eta}_1$ is invariant over classes. Thus for the latent variable regression to be nonlinear, the means or covariances among the exogenous factors must differ across classes. A test of this condition can be performed by evaluating the tenability of invariance constraints on $\hat{\alpha}_{1k}$ and $\hat{\boldsymbol{\Psi}}_{11k}$ with a likelihood ratio test. If the invariance constraints are not rejected, then the null hypothesis of linearity also cannot be rejected, and it may be concluded the latent classes in the model serve to capture some other aspect of the data (i.e., nonnormality of the conditional distribution of $\boldsymbol{\eta}_2$).

Conversely, rejection of these invariance constraints is not sufficient to imply that $E(\boldsymbol{\eta}_2|\boldsymbol{\eta}_1)$ is nonlinear. This is because the weights $P(k|\boldsymbol{\eta}_1)$ are immaterial if the within-class regressions are identical (i.e., $\boldsymbol{\alpha}_{2k} = \boldsymbol{\alpha}_2$ and $\mathbf{B}_{21k} = \mathbf{B}_{21}$). Invariance of the within-class regressions permits Equation 15 to be rewritten as

$$E(\boldsymbol{\eta}_2 \mid \boldsymbol{\eta}_1) = \boldsymbol{\alpha}_2\left[\sum_{k=1}^{K} P(k \mid \boldsymbol{\eta}_1)\right] + \left[\mathbf{B}_{21}\sum_{k=1}^{K} P(k \mid \boldsymbol{\eta}_1)\right]\boldsymbol{\eta}_1, \tag{16}$$

Because $\sum_{k=1}^{K} P(k \mid \boldsymbol{\eta}_1) = 1$, this equation can be reduced to the linear function

$$E(\boldsymbol{\eta}_2 \mid \boldsymbol{\eta}_1) = \boldsymbol{\alpha}_2 + \mathbf{B}_{21}\boldsymbol{\eta}_1, \tag{17}$$

Thus, to reject the null hypothesis that the regression of $\boldsymbol{\eta}_2$ on $\boldsymbol{\eta}_1$ is linear in the aggregate data, invariance constraints on $\hat{\boldsymbol{\alpha}}_{2k}$ and $\hat{\mathbf{B}}_{21k}$ must also be rejected. Failure to reject these invariance constraints again suggests that the latent classes serve to capture some other feature of the data than nonlinearity (i.e., nonnormality of the distribution of $\boldsymbol{\eta}_1$). If both sets of invariance constraints are rejected, this is support for the hypothesis that the regression of $\boldsymbol{\eta}_2$ on $\boldsymbol{\eta}_1$ is nonlinear in the aggregate population.

To summarize, the SEMM can be used to both capture and inferentially test nonlinear latent variable regressions without making strong assumptions about either the functional form of the relations or the distributions of the latent variables or their disturbances. Importantly, although the SEMM does not even require that the relation between latent variables be strictly continuous, the aggregation over classes that takes place in Equation 13 should only be performed if continuity can reasonably be assumed. To verify this assumption, one can first examine the estimated bivariate mixture density of the latent variables for evidence of discontinuities. This approach is demonstrated in the first empirical example described in the next section. Univariate mixture density estimates may also aid in interpretation (or be of substantive interest in their own right), as illustrated in the second empirical example.

EMPIRICAL EXAMPLES

This section presents two empirical examples with the dual goals of further explicating the analytical basis of this new indirect application of the SEMM and providing a preliminary evaluation of the performance of the method. Estimation was carried out using the commercial software program M*plus* 3.11, which implements an accelerated EM algorithm and permits the inclusion of cases with partially missing data under the missing at random assumption (L. K. Muthén & Muthén, 2001). As finite normal mixtures are well known to have poorly behaved likelihood surfaces, often with multiple optima and singularities, an exhaustive examination of the sensitivity of the model results to variations in start values was conducted for each model to ensure that the ML solution was obtained in each case. Start values were varied through the author's own program rather than via the internal algorithm of M*plus* (details available on request).

Example 1: Simulated Nonlinear Latent Variable Regression

The ability of the SEMM to semiparametrically model nonlinear bivariate latent variable distributions was tested by drawing on a model originally presented by Bauer and Curran (2004). Namely, the population generating model was a linear two-factor structural equation model of the type given in Equations 1 and 2, including three indicators of the exogenous latent factor, y_1, y_2, and y_3, and three indicators of the endogenous latent factor, y_4, y_5, and y_6, with a total measurement model defined by

$$\boldsymbol{\nu} = \begin{pmatrix} 0^* \\ 0 \\ 0 \\ 0^* \\ 0 \\ 0 \end{pmatrix}, \quad \boldsymbol{\Lambda} = \begin{pmatrix} 1^* & 0^* \\ 1 & 0^* \\ 1 & 0^* \\ 0^* & 1^* \\ 0^* & 1 \\ 0^* & 1 \end{pmatrix}, \quad \boldsymbol{\Theta} = (.333)\mathbf{I}. \tag{18}$$

where the residuals were drawn from independent normal distributions. The population generating model departed from a typical structural equation model in the latent variable model, given by

$$\begin{aligned} \eta_{1i} &= \zeta_{1i} \\ \eta_{2i} &= -.5 + .5\eta_{1i} + .5\eta_{1i}^2 + \zeta_{2i}, \end{aligned} \tag{19}$$

where the conditional mean vector and covariance matrix for $\boldsymbol{\eta}_i$ were parameterized as

$$\boldsymbol{\alpha} = \begin{pmatrix} 0 \\ -.5 \end{pmatrix}, \quad \boldsymbol{\Psi} = \begin{pmatrix} 1.0 & 0 \\ 0 & .25 \end{pmatrix}, \tag{20}$$

and ζ_{1i} and ζ_{2i} were drawn from independent normal distributions.

In this study, 500 replications with $N = 500$ cases each were generated from this population model. Then, one-, two-, three- and four-class SEMMs of the type in Equation 5 were fit to the data. To assess for and avoid local optima, the first 11 replications were run with 500 random starts for all models. Starting values for the parameter estimates were sampled randomly from independent uniform distributions with specified ranges. The estimates obtained from the highest log-likelihood solutions were then averaged across these 11 replications and used as starting values for the full set of replications. Further details on this aspect of the simulation study are available on request.

Each estimated SEMM implemented a measurement model equal to Equation 18 for all classes. The starred parameters in Equation 18 were fixed at their population values to properly define the measurement structure and scale the latent variables, and the remaining parameters in Equation 18 were freely estimated but constrained to be invariant over classes (i.e., $\boldsymbol{\nu}_k = \boldsymbol{\nu}$, $\boldsymbol{\Lambda}_k = \boldsymbol{\Lambda}$, and $\boldsymbol{\Theta}_k = \boldsymbol{\Theta}$ for all k). The within-class latent variable model for the SEMMs was a main-effect-only model such that

$$\boldsymbol{\alpha}_k = \begin{pmatrix} \alpha_{1k} \\ \alpha_{2k} \end{pmatrix},\ \mathbf{B}_k = \begin{pmatrix} 0 & 0 \\ \beta_{21k} & 0 \end{pmatrix},\ \boldsymbol{\Psi}_k = \begin{pmatrix} \psi_{11} & 0 \\ 0 & \psi_{22} \end{pmatrix}. \tag{21}$$

Note that whereas $\boldsymbol{\alpha}_k$ and $\mathbf{B}_k$ were permitted to vary over classes, $\boldsymbol{\Psi}_k$ was held invariant over classes. Without this constraint, a number of improper solutions were obtained (i.e., negative variances or correlations greater than 1), increasing with the number of classes in the model. In contrast, with this constraint, no improper solutions were obtained for the estimated models, and likelihood ratio tests failed to reject invariance of $\boldsymbol{\Psi}_k$ in 85% to 90% of replications across models, suggesting that holding this matrix invariant over classes minimally impacted the recovery of the characteristics of the observed data. Given the invariance constraints placed on the model, there were 19 parameters to be estimated in the one-class model and four new parameters to be estimated with each additional latent class (α_{1k}, α_{2k}, β_{21k}, and $P(k)$). The minimum BIC was obtained with two classes in just two replications, with three classes in 388 replications, and with four classes in 110 replications.

In the population the functional relation between η_1 and η_2 is continuous in nature, so application of Equation 13 is entirely justified. In practice, however, the nature of the relation is often unknown. Earlier it was suggested that a plot of the bivariate mixture density could be used to better judge the continuity of latent variable relations. Accordingly, Figure 1 presents the estimated bivariate density of η_1

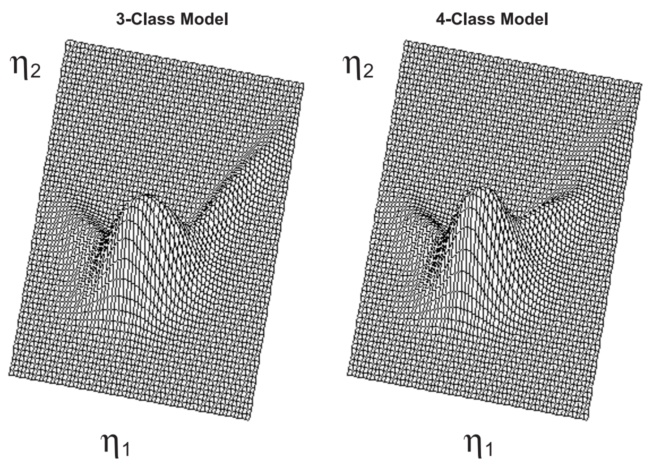

FIGURE 1 Bivariate latent variable density estimates afforded by SEMMs with three or four latent classes for a single randomly selected replication.

and η_2 obtained by fitting three- and four-class SEMMs to a single randomly selected replication. These plots correctly suggest that there is one particularly distinct mode and that the tails of the bivariate distribution stretch into the first and second quadrants, reflecting the general form of the nonlinear relation. There is little substantive difference between the density estimates of the three- and four-class SEMMs, with the exception that the four-class model differs slightly in the first quadrant. In neither plot are there widely separated dense regions, so an assumption of continuity for the relation between η_1 and η_2 appears reasonable.

Given these findings, Equation 13 was applied to approximate the nonlinear regression of η_2 on η_1. Figure 2 contrasts the nonlinear approximations afforded by SEMMs with varying numbers of classes. The bold line plots the average estimate of $E(\eta_2|\eta_1)$ over replications as given by Equation 13 and the dashed line represents the expected value in the population based on the generating function in Equation 19. Clearly, the approximation becomes more accurate, especially in the extremities of the relation, with more latent classes. Given that η_1 was generated from a standard normal distribution, roughly 95% of the individual data points fall within the interval from –2 to 2. Within this interval, the nonlinear approximations afforded by the three-class model and four-class model show negligible bias, whereas the one- and two-class models are comparatively poor. The improvements seen in the four-class model largely take place outside of this interval. It is likely for this reason that the minimum BIC was most often obtained with three classes.

The analytical basis of the nonlinear approximation can be more readily discerned by evaluating the results more thoroughly for the single replication selected previously. For this one replication, the first panel of Figure 3 plots the nonlinear function given by Equation 13 along with the within-class linear regressions given by Equation 12. It can be seen that the three linear regressions serve as basis functions for the nonlinear approximation. The weights applied to the basis functions are the mixing probabilities given by Equation 14, which are plotted in the second panel of Figure 3. The continuous nonlinear changes in these mixing probabilities permit the nonlinear function in the first panel to smoothly shift from one linear basis function to the other over the range of η_1. Because $\hat{\psi}_{11}$ was held invariant, the conditional probabilities plotted in the second panel exclusively reflect differences in the class means of η_1, which were estimated as 1.35, –.07, and –1.22.

Comparing the two panels of Figure 3 also highlights that the function obtained from Equation 13 will be linear only if either the mixing probabilities are constant, requiring invariance of α_{1k} and ψ_{11k}, or the within-class linear regressions are identical, requiring equality of α_{2k} and β_{21k}. As previously noted, an inferential test of nonlinearity can be performed by testing these two sets of constraints. For the three-class model, invariance constraints on $\hat{\alpha}_{1k}$ were rejected in 100% of replications (note that $\hat{\psi}_{11}$ had already been constrained to be invariant), as were invariance constraints on $\hat{\alpha}_{2k}$ and $\hat{\beta}_{21k}$. The universal rejection of both sets of invariance constraints correctly implies that the linear model does not hold for

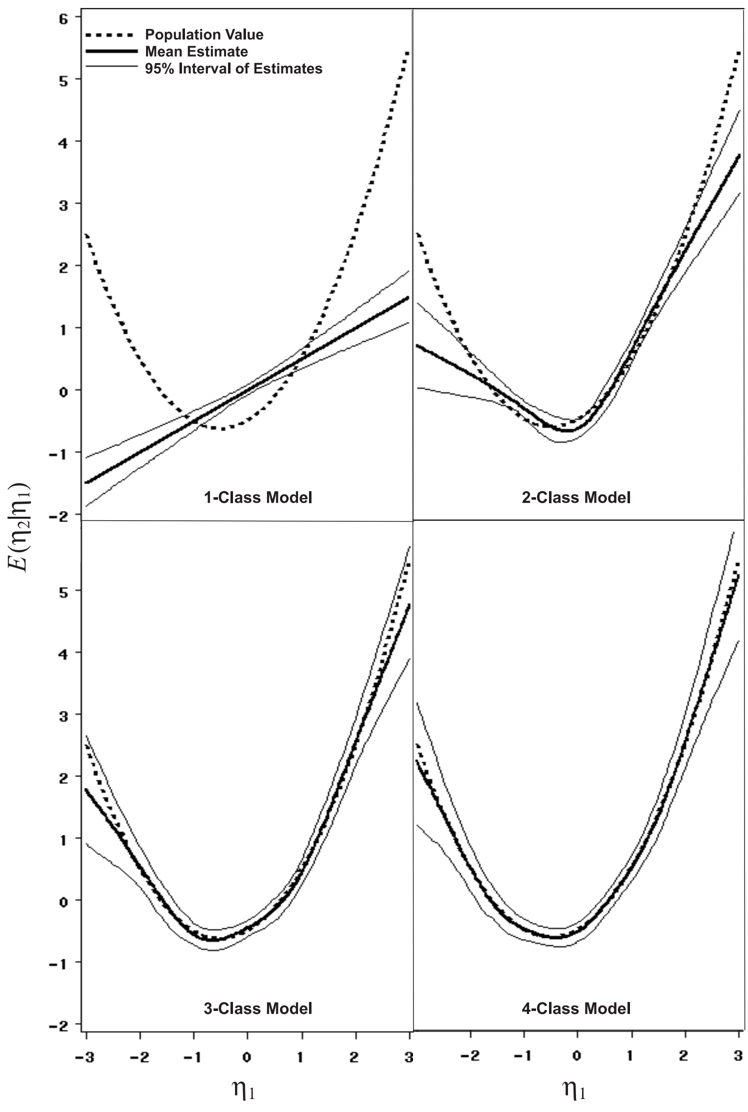

FIGURE 2 The nonlinear approximation afforded by SEMMs with one to four latent classes.

these data. The high power of these tests in this case likely reflects the pronounced nature of the nonlinear trend.

One additional finding of interest for this example concerns the empirical 95% intervals of the sample estimates in Figure 2. Interestingly, the nonlinear approximations given by the two-, three- and four-class models have intervals of roughly the same width as the intervals for the linear one-class model. This suggests that,

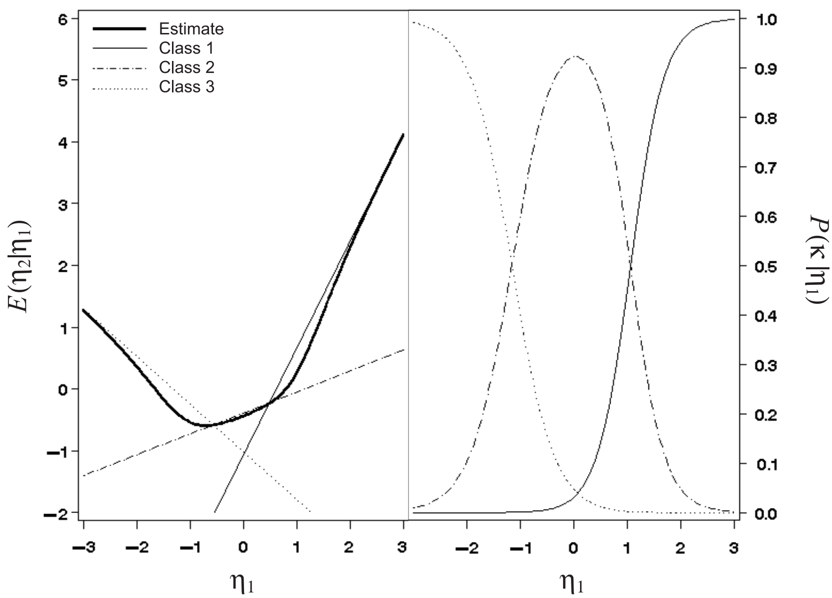

FIGURE 3 The linear within-class basis functions of a three-class SEMM superimposed by the nonlinear regression afforded by weighting the within-class functions by the conditional mixing probabilities for a single randomly selected replication.

despite being semiparametric, the SEMM approximation is a reasonably efficient method for estimating nonlinear latent variable regressions. Last, it is worth noting that although the data were simulated using a quadratic function and assuming normality of η_1, a distinct advantage of the SEMM approach is that neither of these conditions is at all necessary.

Example 2: The Effect of Mastery on Depression and Optimism in Adolescence

The second application concerns data from 1,010 high school students from the St. Paul Public School District who were selected from the Youth Development Study (Mortimer, 2003).[4] Students were surveyed annually from 9th grade (1988) to 12th grade (1992) with excellent panel retention (92.7%). All cases, even those with partially missing data, contributed to these analyses. The estimated models center on three constructs—mastery, depression, and optimism—each measured with multi-

[4]Data were provided courtesy of Jeylan Mortimer, principal investigator of the Youth Development Study, and were collected with funding from the National Institute of Child Health and Human Development (Grant HD44138) and the National Institute of Mental Health (Grant MH42843).

TABLE 1
Indicators of Mastery, Depression, and Optimism

Indicators of mastery[a]	
m_1	I mostly feel helpless in dealing with the problems in life.
m_2	There is really no way I can solve some of the problems that I have.
m_3	Sometimes I feel that I am being pushed around in life.
m_4	I have little control over the things that happen to me.
m_5	There is little I can do to change many important things in my life.
Indicators of depression[b]	
d_1	How often have you felt depressed?
d_2	How often have you been in low or very low spirits?
d_3	How often have you felt downhearted and blue?
d_4	How often have you been moody or brooded about things?
Indicators of optimism[b]	
o_1	How often have you generally enjoyed the things you like?
o_2	How often have you felt that the future looks hopeful?
o_3	How often have you felt cheerful or lighthearted?

[a]Range = 1–4; reverse coded. [b]Range = 1–5.

ple indicators as defined in Table 1. Of interest is the regression of depression and optimism on mastery. Moreover, although it is anticipated that mastery will negatively predict depression and positively predict optimism, and that these relations will be monotonic, it is not known whether they will be strictly linear. The analyses were stratified by wave of assessment both to gauge the stability of the results and to evaluate any developmental trends that might emerge. That is, each model was fitted independently to the data from each wave of assessment.

The measurement model, held invariant over latent classes, was specified as

$$\boldsymbol{\nu} = \begin{pmatrix} 0 \\ \nu_{m2} \\ \nu_{m3} \\ \nu_{m4} \\ \nu_{m5} \\ 0 \\ \nu_{d2} \\ \nu_{d3} \\ \nu_{d4} \\ 0 \\ \nu_{o2} \\ \nu_{o3} \end{pmatrix}, \quad \boldsymbol{\Lambda} = \begin{pmatrix} 1 & 0 & 0 \\ \lambda_{m2} & 0 & 0 \\ \lambda_{m3} & 0 & 0 \\ \lambda_{m4} & 0 & 0 \\ \lambda_{m5} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \lambda_{d2} & 0 \\ 0 & \lambda_{d3} & 0 \\ 0 & \lambda_{d4} & 0 \\ 0 & 0 & 1 \\ 0 & 0 & \lambda_{o2} \\ 0 & 0 & \lambda_{o3} \end{pmatrix}, \quad \boldsymbol{\Theta} = \text{diag}\begin{pmatrix} \theta_{m1} \\ \theta_{m2} \\ \theta_{m3} \\ \theta_{m4} \\ \theta_{m5} \\ \theta_{d1} \\ \theta_{d2} \\ \theta_{d3} \\ \theta_{d4} \\ \theta_{o1} \\ \theta_{o2} \\ \theta_{o3} \end{pmatrix}, \tag{22}$$

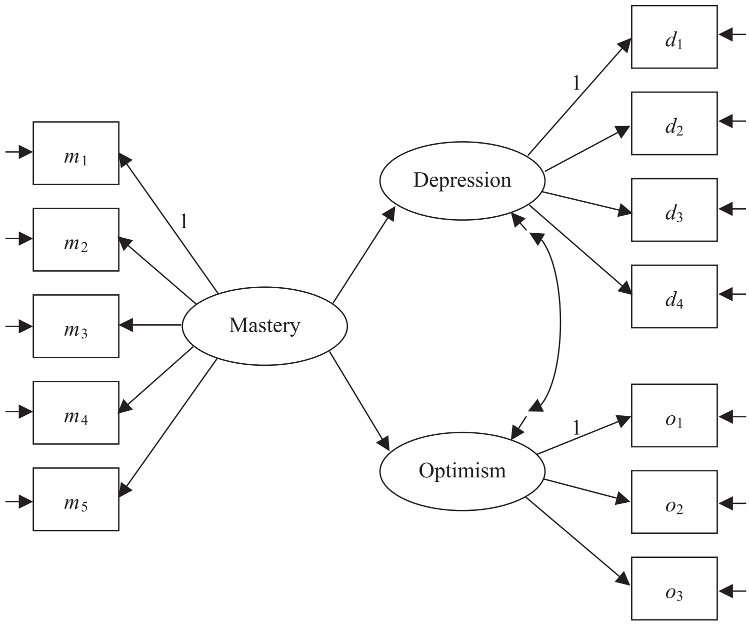

FIGURE 4 Path diagram of the model fit to the Youth Development Study data.

where the subscripts refer to the variable indexes given in Table 1, such that the first five items measure the construct mastery, the next four measure depression, and the last three measure optimism. In contrast to the measurement model, the latent variable model was permitted to vary over classes and was defined as

$$\alpha_k = \begin{pmatrix} \alpha_{M_k} \\ \alpha_{D_k} \\ \alpha_{O_k} \end{pmatrix}, \quad \mathrm{B}_k = \begin{pmatrix} 0 & 0 & 0 \\ \beta_{DM_k} & 0 & 0 \\ \beta_{OM_k} & 0 & 0 \end{pmatrix}, \quad \Psi_k = \begin{pmatrix} \psi_{M_k} & 0 & 0 \\ 0 & \psi_{D_k} & \psi_{DO_k} \\ 0 & \psi_{DO_k} & \psi_{O_k} \end{pmatrix}, \qquad (23)$$

where the capital letters *M, D,* and *O* reference the latent variables mastery, depression, and optimism. A path diagram of the within-class structural model is displayed in Figure 4.

SEMMs with one to four latent classes were estimated both with and without invariance constraints on $\boldsymbol{\Psi}_k$ for each wave of data. Each model was estimated from 100 random starts to guard against the interpretation of local optima. Comparison of the BIC and other considerations suggested that the two-class model with class-varying $\boldsymbol{\Psi}_k$ was the best fit to the data in each wave. An example input file for fitting this model is provided in the Appendix (representing one random start). For the sake of brevity, results are reported only for this model with attention

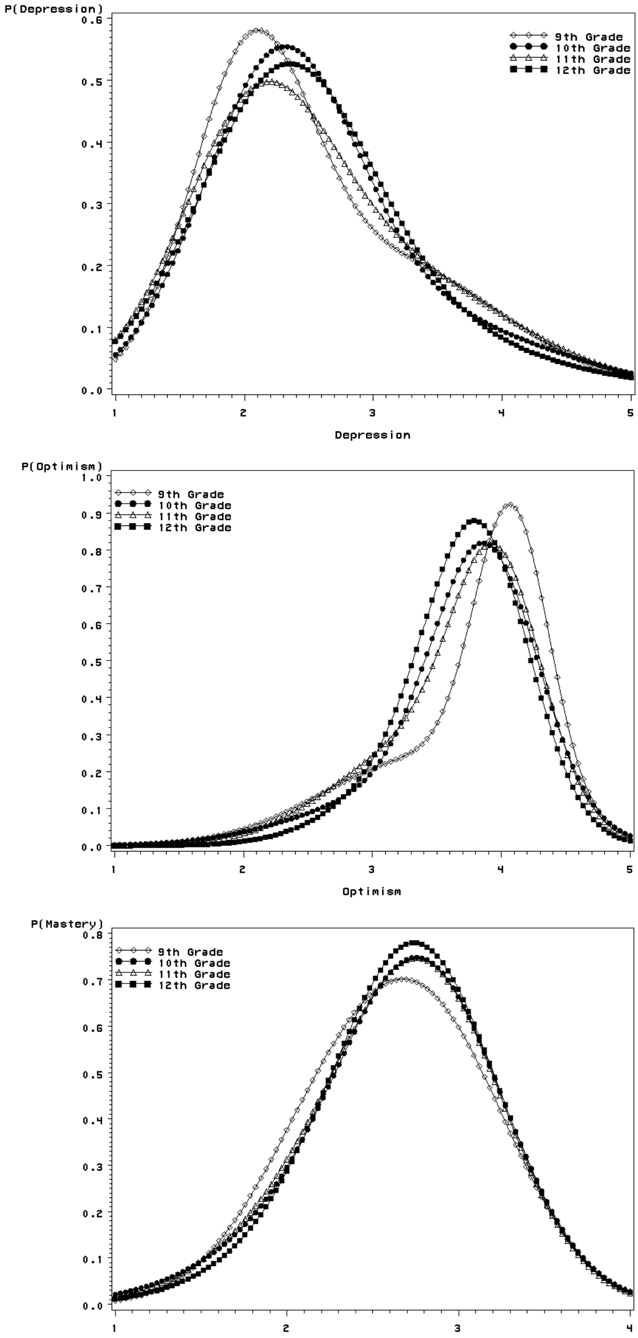

FIGURE 5 Estimated densities for depression, optimism, and mastery.

given only to the latent variable regressions (detailed results on the parameter estimates of this and the other fitted models are available from the author on request).

As a preliminary to plotting the regression functions, the estimated univariate mixture densities afforded by the two-class model were evaluated, shown in Figure 5. These plots are useful for showing where the data are dense versus sparse, aiding interpretation of the semiparametric regression functions. Figure 5 shows that the density estimates for each latent variable are fairly stable from one wave to the next, with the possible exception of 9th grade. In each wave, the distributions for mastery and optimism are unimodal and negatively skewed, whereas the distribution for depression is unimodal and positively skewed.

The estimated latent variable regressions are plotted in Figure 6. Again, relative stability is seen from one wave of assessment to the next. The regression of depression on mastery appears roughly linear, whereas the regression of optimism on mastery shows some evidence of nonlinearity, flattening out at especially low and high levels of mastery (in actuality, likelihood ratio tests rejected the linear model in both cases, likely due to exceptionally high power with $N = 1{,}010$ cases). That is, individual differences in mastery at the high and low ends of the scale do not relate as strongly to differences in optimism as do differences in mastery within the middle of the scale. Given the skew of the mastery distribution, these results are probably most stable for high levels of mastery.

CONCLUSIONS AND DIRECTIONS FOR FUTURE RESEARCH

This article has shown how indirect applications of the SEMM can be used to provide semiparametric regression estimates for latent variable models. Advantages of the method include that the functional form of the latent relation need not be known and the distributional assumptions of the method are weak, making it a semiparametric estimator. This method has been developed primarily for exploratory and descriptive analyses, although some inferential tests have also been described. In addition to the need to further refine and investigate the methodology set forth here both analytically and through additional simulation studies, there are at least three promising directions for extending the current developments.

First, the measurement model was assumed throughout to be invariant over classes, such that the factor loadings, item intercepts, and residual variances could be constrained to be equal for all classes. This assumption was made both to simplify development of the methodology and to ensure that the latent variables estimated in each class had the same intrinsic meaning. One consequence of this assumption is that it implies that the item residuals are normally distributed across classes, hence any nonnormality of the observed variables must be indicative of nonnormality of the latent factors that affect them. This may be an overly restric-

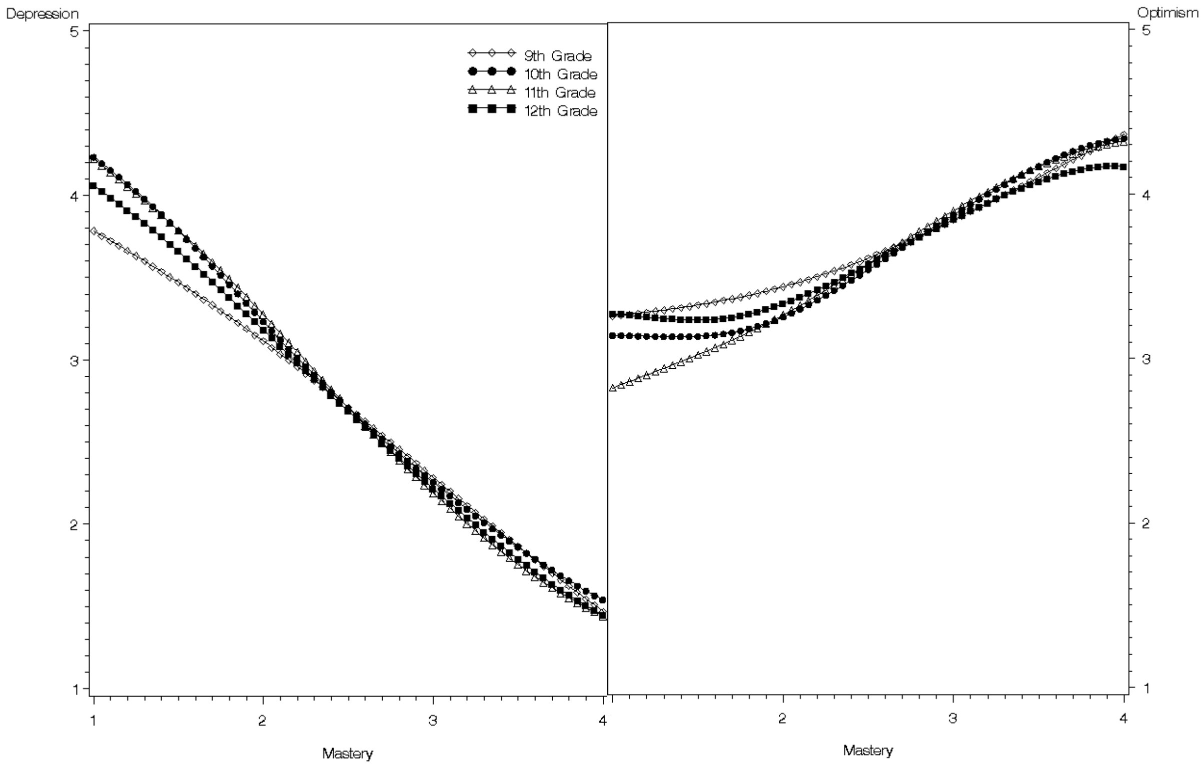

FIGURE 6 Semiparametric regressions of depression and optimism on mastery.

tive requirement. Freeing the item intercepts and residual variances over classes would permit some portion of the observed nonnormality to be unique to the individual observed variables. Similarly, the assumption of invariant intercepts and factor loadings implies that the effects of the latent variables on the observed indicators are linear. Releasing these invariance constraints would permit the factors to nonlinearly affect the indicators and this nonlinearity could be investigated using the same tools presented here for evaluating nonlinear latent variable regressions.

A second area deserving further investigation is the estimation of nonlinear effects involving two or more exogenous latent factors. Although the analytical expressions given here are sufficiently general to encompass these more complex models, the accuracy of the semiparametric approximation afforded by Equation 13 has not yet been assessed beyond the bivariate case. Nevertheless, the present method holds the potential to model interactions that are neither bilinear nor symmetric in form without a priori specification. A preliminary foray into the modeling of nonlinear interactions via mixtures may be found in Bauer and Shanahan (in press), but much work remains to be done.

Third, it would be useful to systematically compare this approach for semiparametrically modeling nonlinear latent variable relations to the flurry of new parametric approaches recently presented in the structural equation modeling literature (e.g., Arminger & Muthén, 1998; Klein & Moosbrugger, 2000; Klein & Muthén, 2003; Lee & Zhu 2002; Wall & Amemiya, 2000; Zhu & Lee, 1999). In general, there is a trade-off to be made between efficiency and bias when selecting between parametric and semiparametric approaches. Parametric approaches typically yield less biased and more efficient estimates than semiparametric approaches when the assumptions of the parametric approaches are met. If these assumptions are incorrect, however, this may lead to bias in the estimates obtained from parametric approaches that would not be incurred using a semiparametric approach. Preliminary Monte Carlo studies comparing this semiparametric method to the quasi-maximum likelihood approach of Klein and Muthén(2003) bear out these expectations. Although such a comparison inevitably suggests that the two methods are in competition, it should be stressed that the semiparametric approach has been designed as a complement to (rather than a replacement for) parametric approaches and is best suited to situations in which the functional form of the relation between latent variables is unknown.

ACKNOWLEDGMENTS

I thank Patrick Curran and Michael Shanahan for helpful discussions of this material and Jeylan Mortimer for providing the Youth Development Study data set for analysis.

REFERENCES

Arminger, G., & Muthén, B. (1998). A Bayesian approach to nonlinear latent variable models using the Gibbs sampler and the Metropolis–Hastings algorithm. *Psychometrika, 63,* 271–300.

Arminger, G., & Stein, P. (1997). Finite mixtures of covariance structure models with regressors. *Sociological Methods and Research, 26,* 148–182.

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

Bauer, D. J., & Curran, P. J. (2004). The integration of continuous and discrete latent variable models: Potential problems and promising opportunities. *Psychological Methods, 9,* 3–29.

Bauer, D. J., & Shanahan, M. J. (in press). Modeling complex interactions: Person-centered and variable-centered approaches. In T. D. Little, J. Bovaird, & N. A. Card (Eds.), Modeling ecological and contextual effects in longitudinal studies of human development. Mahwah, N.J.: Lawrence Erlbuam.

Blåfield, E. (1980). *Clustering of observations from finite mixtures with structural information* (Jyvaskyla Studies in Computer Science, Economics & Statistics, No. 2). Jyvaskyla, Finland: Jyvaskyla University Press.

Bollen, K. A. (1989). *Structural equations with latent variables.* New York: Wiley.

Bollen, K. A. (1995). Structural equation models that are nonlinear in latent variables: A least squares estimator. *Sociological Methodology, 25,* 223–251.

Bollen, K. A., & Paxton, P. (1998). Interactions of latent variables in structural equation models. *Structural Equation Modeling, 5,* 266–293.

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

Escobar, M. D., & West, M. (1995). Bayesian density estimation and inference using mixtures. *Journal of the American Statistical Association, 90,* 577–588.

Etezadi-Amoli, J., & McDonald, R. P. (1983). A second generation nonlinear factor analysis. *Psychometrika, 48,* 315–342.

Everitt, B. S., & Hand, D. J. (1981). *Finite mixture distributions.* London: Chapman & Hall.

Ferguson, T. S. (1983). Bayesian density estimation via mixtures of normal distributions. In M. H. Rizvi, J. S. Rustagi, & D. Siegmund (Eds.), *Recent advances in statistics* (pp. 287–302). New York: Academic.

Gibson, W. A. (1959). Three multivariate models: Factor analysis, latent structure analysis, and latent profile analysis. *Psychometrika, 24,* 229–252.

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

Jöreskog, K. G. (1971). Simultaneous factor analysis in several populations. *Psychometrika, 36,* 409–426.

Jöreskog, K. G., & Yang, F. (1996). Non-linear 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: Lawrence Erlbaum Associates, Inc.

Kaplan, D. (2000). *Structural equation modeling: Foundations and extensions.* Thousand Oaks, CA: Sage.

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

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

Klein, A., & Muthén, B. O. (2003). *Quasi maximum likelihood estimation of structural equation models with multiple interaction and quadratic effects.* Unpublished manuscript, University of California, Los Angeles.

Kline, R. B. (1998). *Principles and practice of structural equation modeling.* New York: Guilford.

Lee, S. Y., & Zhu, H. T. (2002). Maximum likelihood estimation of nonlinear structural equation models. *Psychometrika, 67,* 189–210.

McDonald, R. P. (1967). *Nonlinear factor analysis* (Psychometric Monograph No. 15). Richmond, VA: William Byrd Press.

McLachlan, G., & Peel, D. (2000). *Finite mixture models.* New York: Wiley.

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika, 58,* 525–543.

Mortimer, J. T. (2003). *Working and growing up in America.* Cambridge, MA: Harvard University Press.

Muthén, B. O. (2001). Second-generation structural equation modeling with a combination of categorical and continuous latent variables: New opportunities for latent class/latent growth modeling. In A. Sayer & L. Collins (Eds.), *New methods for the analysis of change* (pp. 291–322). Washington, DC: American Psychological Association.

Muthén, L. K., & Muthén, B. O. (2001). *M*plus *user's guide* (2nd ed.). Los Angeles: Muthén & Muthén.

Nagin, D. (1999). Analyzing developmental trajectories: A semi-parametric, group-based approach. *Psychological Methods, 4,* 139–157.

Roeder, K., & Wasserman, L. (1997). Practical Bayesian density estimation using mixtures of normals. *Journal of the American Statistical Association, 92,* 894–902.

Schermelleh-Engel, K., Klein, A., & Moosbrugger, H. (1998). Estimating nonlinear effects using a latent moderated structural equations approach. In R. E. Schumacker & G. A. Marcoulides (Eds.), *Interaction and nonlinear effects in structural equation modeling* (pp. 203–238). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

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

Silverman, B. W. (1986). *Density estimation for statistics and data analysis.* New York: Chapman & Hall.

Sörbom, D. (1974). A general method for studying differences in factor means and factor structure between groups. *British Journal of Mathematical and Statistical Psychology, 27,* 229–239.

Titterington, D. M., Smith, A. F. M., & Makov, U. E. (1985). *Statistical analysis of finite mixture distributions.* Chichester, UK: Wiley.

Verbeke, G., & Lesaffre, E. (1996). A linear mixed-effects model with heterogeneity in the random-effects population. *Journal of the American Statistical Association, 91,* 217–221.

Wall, M. M., & Amemiya, Y. (2000). Esitmation for polynomial structural equations. *Journal of the American Statistical Association, 95,* 929–940.

Yung, Y.-F. (1997). Finite mixtures in confirmatory factor-analysis models. *Psychometrika, 62,* 297–330.

Zhu, H.-T., & Lee, S.-Y. (1999). Statistical analysis of nonlinear factor analysis models. *British Journal of Mathematical and Statistical Psychology, 52,* 225–242.

Zhu, H.-T., & Lee, S.-Y. (2001). A Bayesian analysis of finite mixtures in the LISREL model. *Psychometrika, 66,* 133–152.

APPENDIX
Example M*plus* Input File for Youth Development Study Analysis (Heteoroscedastic 2-Class Model)

```
TITLE: Mastery -> Dep & Opt;
DATA: FILE IS D:\YDS\DATA\Wave1.dat;
VARIABLE: NAMES ARE nocon noso push help dolit
    deprs lospt down mood cheer hope enjoy;
    CLASSES = Class(2);
    MISSING = .;
ANALYSIS: TYPE = MIXTURE MISSING;
    ESTIMATOR=ML;
    ALGORITHM=EMA;
    STARTS = 0;
MODEL:

%OVERALL%
DEP by Deprs@1 lospt*.956 down*.855 mood*.713;
OPT by enjoy@1 hope*.781 cheer*.886;
MAST by help@1 noso*.911 push*.658 nocon*.703 dolit*.680;
[Deprs@0 Enjoy@0 Help@0];
[lospt*-.082] (1);
[down*.362] (2);
[mood*.871] (3);
[cheer*.039] (4);
[hope*.247] (5);
[push*.666] (6);
[nocon*.996] (7);
[noso*.398] (8);
[dolit*1.059] (9);
Deprs*.458 lospt*.353 down*.394 mood*.466;
Enjoy*.362 hope*.751 cheer*.520;
help*.285 noso*.391 push*.497 nocon*.467 dolit*.504;
[Class#1*-1.84055];

%Class#1%
[DEP*5.10586 OPT*6.30694 MAST*0.16141];
DEP*0.28932 OPT*0.24326 MAST*0.20193;
DEP WITH OPT*0.12566;
DEP ON MAST*0.01454;
OPT ON MAST*1.02325;

%Class#2%
[DEP*0.20396 OPT*6.87203 MAST*3.61495];
DEP*0.24183 OPT*0.30999 MAST*0.41116;
DEP WITH OPT*-0.05010;
DEP ON MAST*-0.55951;
OPT ON MAST*-0.24420;

OUTPUT: STANDARDIZED RES SAMPSTAT TECH1;
```