Ψ Psychology Press
Taylor & Francis Group

Estimating and Visualizing Nonlinear Relations Among Latent Variables: A Semiparametric Approach

Jolynn Pek, Sonya K. Sterba, Bethany E. Kok, and
Daniel J. Bauer
Department of Psychology
University of North Carolina, Chapel Hill

The graphical presentation of any scientific finding enhances its description, interpretation, and evaluation. Research involving latent variables is no exception, especially when potential nonlinear effects are suspect. This article has multiple aims. First, it provides a nontechnical overview of a semiparametric approach to modeling nonlinear relationships among latent variables using mixtures of linear structural equations. Second, it provides several examples showing how the method works and how it is implemented and interpreted in practical applications. In particular, this article examines the potentially nonlinear relationships between positive and negative affect and cognitive processing. Third, a recommended display format for illustrating latent bivariate relationships is demonstrated. Finally, the article describes an R package and an online utility that generate these displays automatically.

Nonlinear modeling in the context of applied behavioral and social science research has been a useful tool for understanding, describing, and predicting myriad phenomena. One classic example is the Yerkes-Dodson, law which states that with increasing levels of cognitive arousal, performance increases, plateaus, and then decreases (Yerkes & Dodson, 1908). A more contemporary example is the finding by Lubinski and Humphreys (1990) that scores below the mean on a general mathematics aptitude test show little relationship to exceptional

mathematical talent, perhaps reflecting a general lack of prerequisite knowledge for acquiring sophisticated skills, whereas variation above the mean becomes increasingly informative. The result is that the relationship between the two variables is described by a concave up, positively accelerated curve. In this article, we are concerned with identifying and modeling nonlinear relationships such as these when the variables involved are latent rather than observed.

For observed variables, a widely used strategy for diagnosing and modeling nonlinear effects is to examine bivariate scatterplots for evidence of nonlinearity. Such displays can be greatly enhanced by using the LOWESS method (Cleveland, 1981) to produce an estimate of the potentially nonlinear regression function for the two variables (Cohen, Cohen, West, & Aiken, 2003, pp. 110–114). A desirable feature of LOWESS is that the analyst need not specify a particular form for the regression function (e.g., quadratic, cubic, exponential, etc.). The global regression function is obtained by smoothing over locally linear regression estimates computed at each observed value of the predictor. Each locally linear regression is estimated using only neighboring data points weighted by their proximity to the predictor value under consideration. Unfortunately, this approach cannot be applied directly when the relevant variables are latent, given the absence of specific observed values for the predictor or outcome. Recently, however, Bauer (2005) suggested an analogy to LOWESS regression for modeling nonlinear relationships among latent variables.

Bauer's (2005) semiparametric modeling approach involves fitting a Structural Equation Mixture Model (SEMM) to the data (Arminger & Stein, 1997; Arminger, Stein, & Wittenberg, 1999; Dolan & van der Maas, 1998; Jedidi, Jagpal, & DeSarbo, 1997a, 1997b; B. O. Muthén, 2001). The general notion behind the SEMM is that the observed data arises from a mixture of component (or class) distributions where each component is characterized by a distinct structural equation model. Bauer's (2005) approach to modeling nonlinear relationships using SEMMs draws on the fact that the regression of one latent variable on another is linear within each component of the mixture, analogous to the locally linear estimates used in LOWESS. Also like LOWESS, a potentially nonlinear global regression function can be obtained by smoothing over the locally linear regression estimates for the components.[1] In this case, the smoothing is accomplished using weights based on the locations of the component distributions. For instance, Figure 1 shows how this approach can recover a U-shaped relationship between two latent variables using three latent classes. The mean of the latent

---

[1]Note that this use of the SEMM involves no assumption that the component distributions, often referred to as latent classes, reflect true groups within the population. The mixture is estimated only as a statistical expedience to obtain an estimate of the global regression function. This type of mixture application has been referred to as "indirect" by Titterington, Smith, and Makov (1985), in contrast to "direct" applications aimed at modeling population heterogeneity.
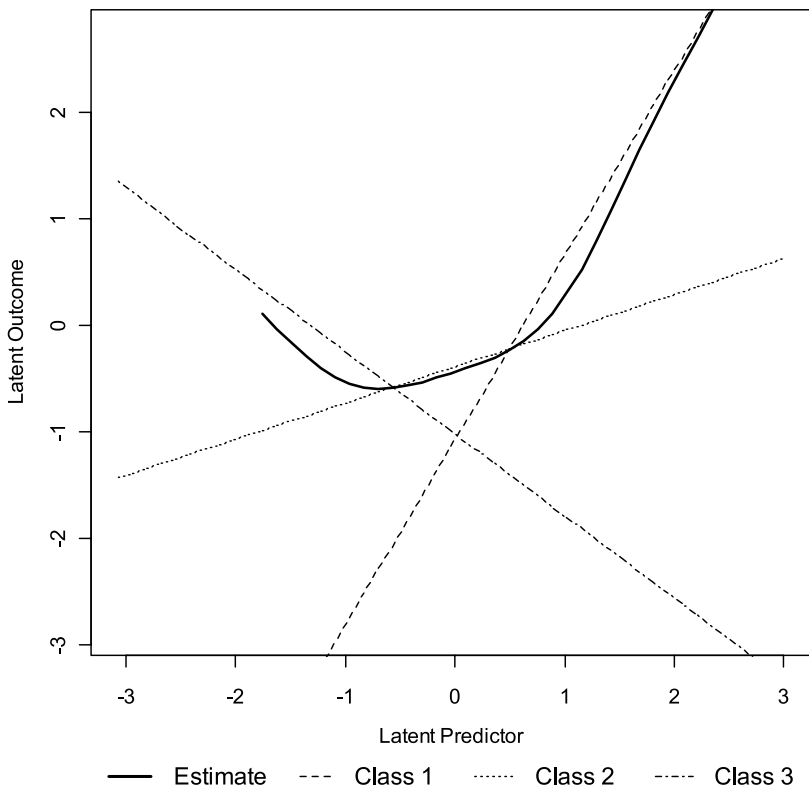
FIGURE 1   Plots of the three locally linear regression estimates and the nonlinear regression of η_2 on η_1.

predictor is lowest in Class 1, followed by Classes 2 and 3. The aggregate function thus hugs the regression line for Class 1 at low levels of the latent predictor, shifts toward the line for Class 2 at the midrange, and then follows the line for Class 3 when the value of the latent predictor is high.

Drawing further on the analogy to LOWESS, the number of component distributions specified for the SEMM corresponds roughly to setting the smoothing parameter in the LOWESS procedure; adding too many components is like setting the smoothing parameter too small—the regression line may be jagged due to overfitting of chance variation in the data—whereas choosing too few components is akin to setting the smoothing parameter too high, which may obscure the nonlinear trends of interest. To determine the optimal number of components for the SEMM, one can compare models with different numbers of components using information criteria. Additionally, the within-component

variances for the latent predictors relate to the neighborhood of points used in LOWESS; larger variances imply that a larger neighborhood of points contributes to the estimation of the locally linear relationship. An important difference between the LOWESS and SEMM methods, however, is that LOWESS is fully nonparametric whereas the SEMM method is semiparametric, based on a very flexible statistical model. The SEMM method (unlike LOWESS) can thus be used to formally test for nonlinear effects in return for minimal distributional assumptions.[2]

Advantages of using SEMMs for modeling nonlinear relationships between latent variables are that it can recover global relationships of unknown form and does so without assuming that the latent variables are normally distributed. In contrast, other methods for modeling nonlinear relationships among latent variables require specification of the global regression function (i.e., as a low-order polynomial such as a quadratic), and many also assume normality for the distribution of the latent predictor (e.g., Jaccard & Wan, 1995; Jöreskog & Yang, 1996; Kenny & Judd, 1984; Klein & Muthén, 2007; Mooijaart & Bentler, 1986; Ping, 1996). In addition to being of value on its own, the SEMM approach may thus also be a useful precursor to implementing these other methods, as it would enable evaluation of these assumptions.

Despite these advantages of the semiparametric approach, we are aware of no applications of this approach since the publication of Bauer (2005). We suspect that there are four primary reasons for this lack of implementation in applied work, all of which we seek to address in this article. First, the description of this method by Bauer (2005) may be excessively technical for many researchers. In this article, we thus provide a more expository account of the method. Second, Bauer (2005) provided relatively few empirical examples of the technique upon which others might model their own applications. Here we provide a number of examples aimed at illustrating both how the technique works and how it might be implemented and interpreted in practical analyses. In particular, we examine the potentially nonlinear relationship between affect and cognitive processing using data from a study of emotion. Third, Bauer (2005) did not provide a clear recommendation for how the information obtained from the semiparametric approach should be displayed. Here we demonstrate what we regard to be an optimal display format for depicting bivariate relationships. Fourth, and perhaps most important, there has been no software available to

---

[2]A second analogy may be drawn with nonparametric regression splines for observed data where a global curve is obtained by joining knot points together with a piecewise polynomial function (Smith, 1979; Wold, 1974). The number of knot points functions similar to the number of components in the SEMM approach. Spline functions, however, partition the data between user-defined knots into nonoverlapping sets to fit pieces of the function, whereas in SEMM the components overlap and their locations (means) are estimated rather than fixed. The overlap permits smoothing from one within-component linear function to another, unlike, for instance, a piecewise linear model.

generate such displays. Although software is readily available for fitting SEMMs (e.g., Mx; Neale, Boker, Xie, & Maes, 2003; M*plus;* L. K. Muthén & Muthén, 2007; MECOSA in Arminger, Wittenberg, & Schepers, 1996), the evaluation of nonlinear effects requires considerable post-processing of the model results. Indeed, the manipulation of results required to produce plots of the nonlinear effects may be tedious and daunting even for experienced Structural Equation Model (SEM) users. To overcome this difficulty, we provide two plotting utilities to generate these plots automatically.

The outline of our article is as follows: First, we provide a basic overview of the linear SEM and describe Bauer's (2005) semiparametric modeling approach. Next, we introduce an online calculator and an R package, plotSEMM, which generate and plot the potentially nonlinear regression function as well as aux-iliary information that can facilitate interpretation. Several examples are then provided to illustrate the approach, to demonstrate the recommended display format for the results, and to show how the plotting utilities may be used to produce these displays. In addition to simulated data examples, we also analyze data from a study of emotions and well-being to evaluate the relationship between affect and decision making. Of specific interest are the potentially non-linear relationships between positive and negative affect and cognitive heuristic processing.

THE LINEAR STRUCTURAL EQUATION MODEL

For simplicity, in this section we present the linear structural equation model for one latent predictor, η_1, and one latent outcome, η_2. The following develop-ments, however, would also be applicable for evaluating bivariate relationships embedded in more complex models.

The measurement models for the latent factors may be defined as

$$
\begin{aligned}
\mathbf{y}_{1i} &= \mathbf{v}_1 + \boldsymbol{\lambda}_1 \eta_{1i} + \boldsymbol{\varepsilon}_{1i} \\
\mathbf{y}_{2i} &= \mathbf{v}_2 + \boldsymbol{\lambda}_2 \eta_{2i} + \boldsymbol{\varepsilon}_{2i},
\end{aligned}
\tag{1}
$$

where $\mathbf{y}_1$ and $\mathbf{y}_2$ are vectors of observed variables measuring the latent predictor and latent outcome, respectively. The intercepts and slopes for the regression of the observed variables on the latent variables are contained in the vectors $\mathbf{v}$ and $\boldsymbol{\lambda}$, respectively, with subscripts indicating the referent observed variables. The residuals are represented by $\boldsymbol{\varepsilon}_{1i}$ and $\boldsymbol{\varepsilon}_{2i}$ and have expected values of zero and a joint covariance matrix $\boldsymbol{\Theta}$. Typically, but not necessarily, $\boldsymbol{\Theta}$ is constrained to be diagonal, reflecting the assumption that the observed variables are locally independent, conditioning on the latent variables.

The latent variable model is then

$$\eta_{1i} = \alpha_1 + \zeta_{1i}$$
$$\eta_{2i} = \alpha_2 + \beta_{21}\eta_{1i} + \zeta_{2i},$$

(2)

where the mean and variance of the latent predictor are α_1 and $VAR(\zeta_{1i}) = \psi_{11}$, and the intercept, slope, and residual variance for the linear regression of the latent outcome on the latent predictor are α_2, β_{21}, and $VAR(\zeta_{2i}) = \psi_{22}$, respectively. The residuals ζ_1 and ζ_2 are assumed to be uncorrelated with each other and with ε_1 and ε_2.

Equations 1 and 2 imply a specific mean structure and covariance structure for the complete vector of observed variables (Bollen, 1989). By convention, the model-implied mean vector and covariance matrix are designated as $\mu(\theta)$ and $\Sigma(\theta)$, respectively, where θ is the vector of model parameters. If we add to Equations 1 and 2 the assumption that all residuals (i.e., ζ_1, ζ_2, ε_1, and ε_2) are normally distributed, then the linear form of the model implies that the observed variables will also have a joint normal Probability Density Function (PDF), which we designate $\phi[\mathbf{y}; \mu(\theta), \Sigma(\theta)]$. This PDF provides the basis for maximum likelihood estimation of the parameters of the model. A more extensive treatment of the standard structural equation model can be obtained from Bollen (1989), Kaplan (2000), or Kline (2005), among other texts.

For the current purposes, it is worth emphasizing that a key assumption of the standard SEM, embodied in Equation 2, is that the latent predictor and the latent outcome are linearly related. We now show one way to relax this assumption by moving to the SEMM.

SEMIPARAMETRICALLY MODELING NONLINEAR EFFECTS VIA SEMMS

The SEMM assumes that the observed data were obtained from a mixture of K multivariate normal distributions, each parameterized via a linear SEM. It is common to refer to each component normal distribution as a latent class, and we use the two terms interchangeably here without implication that the latent classes represent literal groups. The goal in fitting a SEMM is to estimate both the SEM model parameters for each latent class (some of which may take on unique values and others of which may be constrained to equality over classes) and the mixing probabilities for the classes. In the following, we designate the probability of membership in class $k = 1, 2, \ldots, K$ as π_k, where $\sum_{k=1}^{K} \pi_k = 1$.

For our purposes, we assume that the SEM specified for each class is of the same form (i.e., has the same zero and nonzero paths). Additionally, we constrain the measurement model presented in Equation 1 to be strictly invariant over latent

classes so that the latent variables are equivalently defined for all individuals in the population (Meredith, 1993).[3] Only the parameters involved in the latent variable model must differ over latent classes to model nonlinear effects between latent variables. Specifically, Equation 2 is modified to be

$$\eta_{1i} = \alpha_{1k} + \zeta_{1i}$$
$$\eta_{2i} = \alpha_{2k} + \beta_{12k}\eta_{1i} + \zeta_{2i},$$
(3)

where the class mean and variance of the latent predictor are α_{1k} and $VAR(\zeta_{1i}) = \psi_{11k}$, respectively, and the intercept, slope, and residual variance for the within-class linear regression of the latent outcome on the latent predictor are α_{2k}, β_{21k}, and $VAR(\zeta_{2i}) = \psi_{22k}$, respectively. The only difference from the standard SEM is that these parameters, defining the joint distribution of the latent variables, differ over latent classes. Note that ζ_{1i} and ζ_{2i} are still assumed to be uncorrelated. In addition, across-class equality constraints on the variance parameters ψ_{11k} and ψ_{22k} can be helpful for avoiding problems of estimation.

From Equation 3 it follows that the expected value of the latent outcome η_2 within each latent class is

$$E_k[\eta_2|\eta_1] = \alpha_{2k} + \beta_{21k}\eta_1.$$
(4)

Thus the relationship between the latent variables is locally linear within the latent space characterized by each specific class. The global relationship between the latent variables obtained by aggregating across the mixing components is, however, nonlinear, given the differential weighting of these locally linear relationships over the range of the latent predictor, that is, as the value of η_1 moves across the latent space. More specifically, aggregating across the mixing components, we obtain the global regression function

$$E[\eta_2|\eta_1] = \sum_{k=1}^{K}(\pi_k|\eta_1)E_k[\eta_2|\eta_1],$$
(5)

where

$$\pi_k|\eta_1 = \frac{\pi_k\phi_k(\eta_1;\alpha_{1k},\psi_{11k})}{\sum_{k=1}^{K}\pi_k\phi_k(\eta_1;\alpha_{1k},\psi_{11k})},$$
(6)

---

[3]The assumption of strict measurement invariance over classes is consistent with the standard single-group SEM, which assumes strict measurement invariance for all individuals. Because the classes are not believed to represent distinct groups of individuals in this type of SEMM application, this assumption is retained for parsimony.

$\phi(\cdot)$ again indicates the normal PDF, and $\pi_k|\eta_1$ represents the conditional probability of class membership given a set value for the latent predictor. Note that the numerator of Equation 6 is the weighted marginal distribution of η_1 for Class k, whereas the denominator is the marginal mixture distribution of η_1 across classes. The global regression function in Equation 5 is obtained by smoothing over the local linear regressions within each class as expressed in Equation 4 using the localized weights in Equation 6. These weights are mixing probabilities that reflect the dominance of Class k at varying levels of η_1. When a specific value for η_1 is examined that is located near the mean of a particular component distribution, then that component will usually have the highest conditional mixing probability. The expected value for η_2 computed from Equation 5 will then be determined largely by the locally linear relationship estimated within that class. Likewise, if a value of η_1 is examined that lies between two components then the expected value computed from Equation 5 will generally be a more equitably weighted combination of the within-component expected values. Hence, changes in $\pi_k|\eta_1$ over the range of η_1 allow the global function is Equation 5 to move smoothly from one locally linear relationship to another.

It is important to recognize that this approach does not presume a nonlinear relationship must exist. Bauer (2005) noted two alternative sufficient conditions for Equation 5 to reduce to a linear function (aside from the trivial case when $K = 1$):

Condition A: The distribution of the latent predictor is the same in each latent class, that is, α_{1k} and ψ_{11k} are equal across classes.
Condition B: The within-class regression line obtained for each class is the same, that is, α_{2k} and β_{21k} are equal across classes.

Each condition can be evaluated by a likelihood ratio test. If either likelihood ratio test is nonsignificant, then the null hypothesis of a linear relationship cannot be rejected. If both are significant, this suggests that the relationship may be nonlinear and it will be of interest to graphically evaluate its form.[4]

In practice, computing and plotting the semiparametric regression function defined in Equation 5 involves nontrivial programming to post-process the results obtained from fitting SEMMs to the data. Moreover, the interpretation of this relationship can be enhanced by plotting the model-implied marginal and bivariate distributions of the latent variables. Such plots can reveal areas

---

[4]Although there are no other obvious ways for Equation 5 to reduce to a strictly linear function, we have observed in practice that the aggregate function implied by Equation 5 can sometimes appear roughly linear even when conditions A and B are rejected. For instance, one component may be so small that the aggregate function is minimally influenced by the presence of this component.

of sparseness in the latent space where one would exercise caution in making interpretations of the regression function or they can show discontinuities that would preclude using a smoothing function like Equation 5. For these reasons, we provide a set of utilities that produce these plots automatically given input of the estimated mixing probabilities and latent variable model parameters for the latent classes.

PLOTTING UTILITY

We have developed two tools for users wishing to use the technique described earlier for assessing potentially nonlinear relationships among latent variables. Each utility produces the same graphs, and both are based on code written in R, but the first is geared toward R users and the second toward non R users. The first option is an R package called plotSEMM, which is a set of functions used to generate the plots described earlier given input of some of the parameter estimates from an SEMM analysis. This package has to be run within the R environment. R is a free software environment for statistical computing and graphics and can be downloaded from http://cran.r-project.org/ For R users, the add-on plotSEMM package has the advantage of saving plots in file types supported by R, which include portable document format (PDF), postscript, and Metafile, among others. The plotSEMM package is also more flexible than the online utility, providing more graphics options, and the code is open source, permitting user modifications. For users unfamiliar with R, however, the online utility, located at http://www.unc.edu/psychology/dbauer/plotSEMM.htm, may be more appealing. The online interface dynamically generates R code based on user inputs and submits it to Rweb, a Web-based interface to R. This utility is readily accessible and circumvents the need to install R on one's computer or learn how to write R commands.

With both utilities, two plots will be generated. The first plot superimposes the smoothed regression function obtained by Equation 5 on the model-implied bivariate contour plot for the two latent variables. The model-implied marginal distributions of the latent predictor and outcome are shown above and to the right of the bivariate plot. Additionally, the locally linear (within-class) regression functions and within-class marginal distributions for the latent variables are also plotted, although these features can be suppressed if desired. The second plot shows the mixing probabilities conditional on the value of the latent predictor, as generated by Equation 6. The marginal distribution of the latent predictor is shown on this plot as well. This second plot is not essential but can aid in understanding how the global regression function shown in the first plot is generated. The only information required by plotSEMM and the online utility is a subset of the parameter estimates generated by the program originally used

to fit the SEMM to the data. We now describe each plotting option in greater detail.

plotSEMM

User input.   For plotSEMM, the user needs to specify six vectors of model estimates, each containing K elements (where K is the number of classes estimated in the user's final SEMM model). These vectors are specified to include (a) the marginal class probabilities, π_k; (b) the class means for the latent predictor, α_{1k}; (c) and (d) the intercepts and slopes from the within-class regression of the latent outcome on the latent predictor, α_{2k} and β_{21k}, respectively; (e) the within-class variances of the latent predictor, ψ_{11k}; and (f) the within-class residual variances of the latent outcome, ψ_{22k}.

R functions and controlling output.   To invoke the functions in plotSEMM, the user first calls the setup function, which requires six arguments that correspond to the six vectors. These vectors must be entered in the order specified previously. The user may then call two plotting functions as well as the legend function. plotSEMM_contour plots the global regression function, the locally linear within-class regression lines, and the model-implied bivariate and marginal distributions of the latent variables. plotSEMM_probability plots the mixing probabilities as a function of the latent predictor (i.e., the weights generated through Equation 6 for aggregating over the within-class regression lines) and again presents the marginal distribution of the latent predictor for reference. plotSEMM_legend generates the legend associated with these plots.

The plotting functions allow the user to specify labels for the latent variables. By default, the latent predictor and outcome are labeled "Eta1" and "Eta2," respectively, and the class-specific regression lines and marginal distributions will be shown. However, users can choose to suppress class-specific information in plotSEMM_contour. An example script file for inputting data in vectors, calling the setup function and using the plotting functions, is provided in Appendix A. The plotSEMM source code, documentation, and example files are readily downloadable at http://cran.r-project.org/

Online Utility

User input.   The online utility provides almost the same functionality as the plotSEMM package within a graphical user interface that obviates the need for experience using R. Figure 2 presents a screen shot of the utility. Unlike plotSEMM, which calls for the specification of parameter estimates in vector form, the online utility takes a step-by-step approach to obtaining the same information. First, the number of classes for the model should be specified and

Nonlinear Relations among Latent Variables

Jolynn Pek, Sonya K. Sterba, Bethany E. Kok and Daniel J. Bauer
University of North Carolina at Chapel Hill

This web page generates semiparametric estimates of the regression function for two latent variables. A mixture of linear structural equation models must first be fit to the data. The mixing probabilities (π) and latent variable model parameters (labeled in the diagram below) are then input into the boxes provided.

Empirical example on positive emotions and heuristic processing Mplus output Mx output

Empirical example on negative emotions and heuristic processing Mplus output Mx output

ψ_{11k} α_{1k} η_1 y_1 y_2 ... y_l ε_1 ε_2 ε_l

β_{21k}

α_{2k} η_2 ζ ψ_{22k} y_{l+1} y_{l+2} ... y_n ε_{l+1} ε_{l+2} ε_n

Example 1   Example 2   Example 3   Example 4

Label for η_1: Latent Predictor   Label for η_2: Latent Outcome

Number of classes: 1   Change Class Number

Class 1: π = α_1 = α_2 = β_{21} = ψ_{11} = ψ_{22} =

☑ Show class specific distributions, regression lines and mixing probabilities

Submit to Rweb   Clear all values

FIGURE 2   Screen shot of the online plotting utility.

updated by clicking on the "Change Class Number" button. Then, for each class in the model, users are required to input estimates for the probability of class membership (labeled π), mean of the latent predictor (labeled α_1), intercept and slope for the regression of the latent outcome on the latent predictor (labeled α_2 and β_{21}, respectively), the variance of the latent predictor (labeled ψ_{11}), and the residual variance of the latent outcome (labeled ψ_{22}).

Controlling output.   By default, class information will be shown. However, users can choose to suppress class information by unchecking the option "Show class specific distributions, regression lines, and mixing probabilities." Once all the required information is input, clicking the "Submit to Rweb" button will open another window displaying the two sets of plots. Several examples are also provided so that users can become familiar with how to input the model estimates and interpret the generated plots. Clicking on the "Example 1" and "Example 2" buttons will display the inputs used for the simulated examples; the former was presented in Bauer (2005). Clicking on the "Example 3" and "Example 4" buttons will display the inputs for the real data examples to be described in the next section.

EMPIRICAL EXAMPLES

This section presents several simulated and real data examples of how the utilities described previously can be used to facilitate and enhance modeling of nonlinear relationships among latent variables. For all models, estimation was carried out with M*plus* 5 (L. K. Muthén & Muthén, 2007). For reference, both M*plus* and Mx scripts for fitting some of the models are provided in Appendices B and C. Each SEMM was estimated using 100 random starts to avoid interpreting local solutions. To keep the presentation compact, we suppress the class information for the artificial data examples, showing this only for the real data examples.

Artificial Data Examples

Two data sets of 500 cases each were simulated to demonstrate what the results of the semiparametric modeling approach look like when the latent variable regression is nonlinear versus linear in form. The measurement model used to generate the data in each case followed Equation 1 and included three measured variables for the latent predictor and three measured variables for the latent outcome. The parameter values were

$$\boldsymbol{v}_1 = \boldsymbol{v}_2 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \boldsymbol{\lambda}_1 = \boldsymbol{\lambda}_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \boldsymbol{\varepsilon}_{1i}, \boldsymbol{\varepsilon}_{2i} \sim N\left(0, \frac{1}{3}\right). \tag{7}$$

The latent variables thus explained 75% of the variance of their respective measured variables.

For the nonlinear data example, the latent variable regression was quadratic in form, specifically,

$$\eta_{2i} = -.5 + .5\eta_{1i} + .5\eta_{1i}^2 + \zeta_{2i}, \tag{8}$$

where $\eta_{1i} \sim N(0, 1)$ and $\zeta_{2i} \sim N(0, .25)$, matching an example in Bauer & Curran (2004). For the linear data example, the latent variable model was given by

$$\eta_{2i} = .866\eta_{1i} + \zeta_{2i}, \tag{9}$$

where η_{1i} was generated from a $\chi^2(3)$ distribution standardized to have a mean of zero and unit variance and $\zeta_{2i} \sim N(0, .25)$. In both models, 75% of the variance in η_{2i} is explained by η_{1i}.

For each data set, it should be possible to estimate an SEMM with multiple latent classes. For the nonlinear data, the mixture should function to capture the nonlinear relationship between the latent variables. Likelihood ratio tests should then indicate rejection of both Conditions A and B described earlier (i.e., invariance of the predictor distribution across class and invariance of the within-class regression lines, respectively). In contrast, for the linear data, the mixture should function to accommodate the nonnormal distribution of the latent predictor (see Bauer & Curran, 2004) but should not imply a nonlinear relationship between the latent variables. For this data set, we expect to reject Condition A but not Condition B.

For the nonlinear data, a three-class model with across-class equality constraints on the variance parameters showed the best fit to the data, as determined by comparison of model fit indices (e.g., Bayes's Information Criterion or BIC). Because the variance parameters were already held constant over classes, Condition A was evaluated by further constraining the means of η_1 (i.e., α_1) to be equal, resulting in $\Delta\chi^2(2) = 128.22$, $p < .001$. Condition B was evaluated by removing the previous constraint and instead applying invariance constraints on the intercepts and slopes for the within-class regression lines (i.e., α_2 and β_{21}), resulting in $\Delta\chi^2(4) = 277.982$, $p < .001$. The rejection of both sets of constraints correctly suggests that the relationship between the latent variables is not linear. Figure 3 shows the globally nonlinear regression function generated by the software utilities for this example, superimposed on a contour plot of the model-implied bivariate distribution for the latent variables. The model-implied marginal distributions of the latent predictor and outcome are depicted to the top and right of the main panel, respectively.

For the linear data example, a two-class model with class-varying variance parameters fit the data best. As expected, likelihood ratio tests indicated rejection
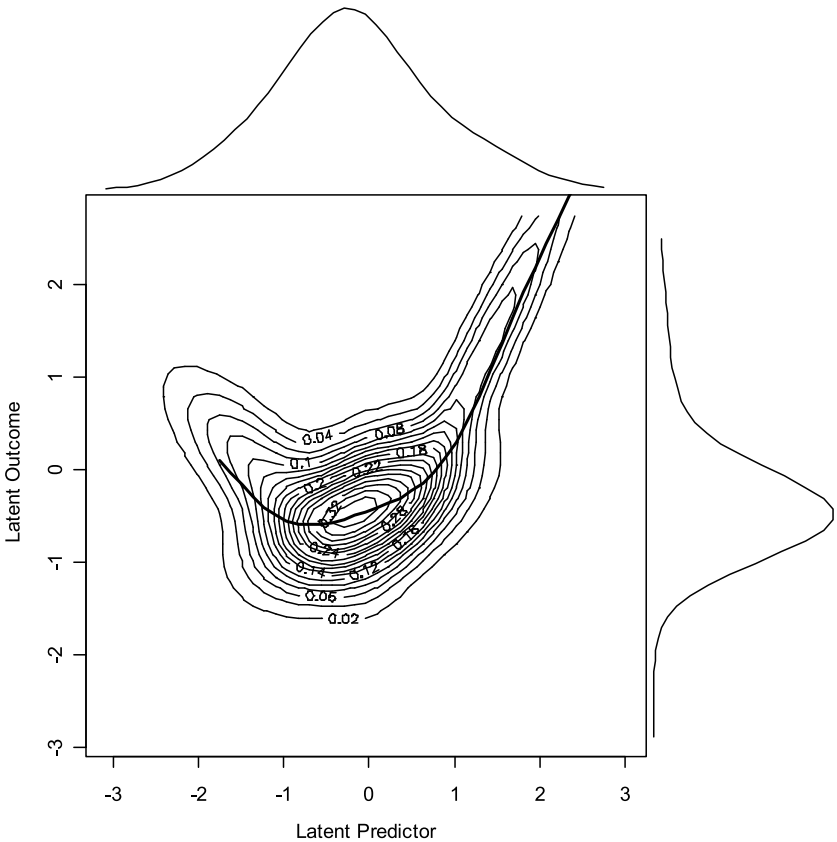
FIGURE 3   Plots of the marginal mixture densities for η_1 and η_2, the bivariate contour plot, and the nonlinear regression of η_2 on η_1.

of Condition A, $\Delta\chi^2(2) = 67.072$, $p < .001$, but not Condition B, $\Delta\chi^2(2) = 1.280$, *ns*. That is, the distribution of the latent predictor differed in location and scale across components of the mixture, implying nonnormality, but the intercepts and slopes of the within-class regression lines did not significantly differ from one another, so that the null hypothesis of linearity could not be rejected. This pattern of results is clearly seen in Figure 4.

Empirical Application: Affect and Heuristic Processing

The following two examples examine the relationship between affect and cognitive heuristic processing. Prior research has indicated that individuals in a positive mood are more likely to adopt a heuristic processing strategy characterized
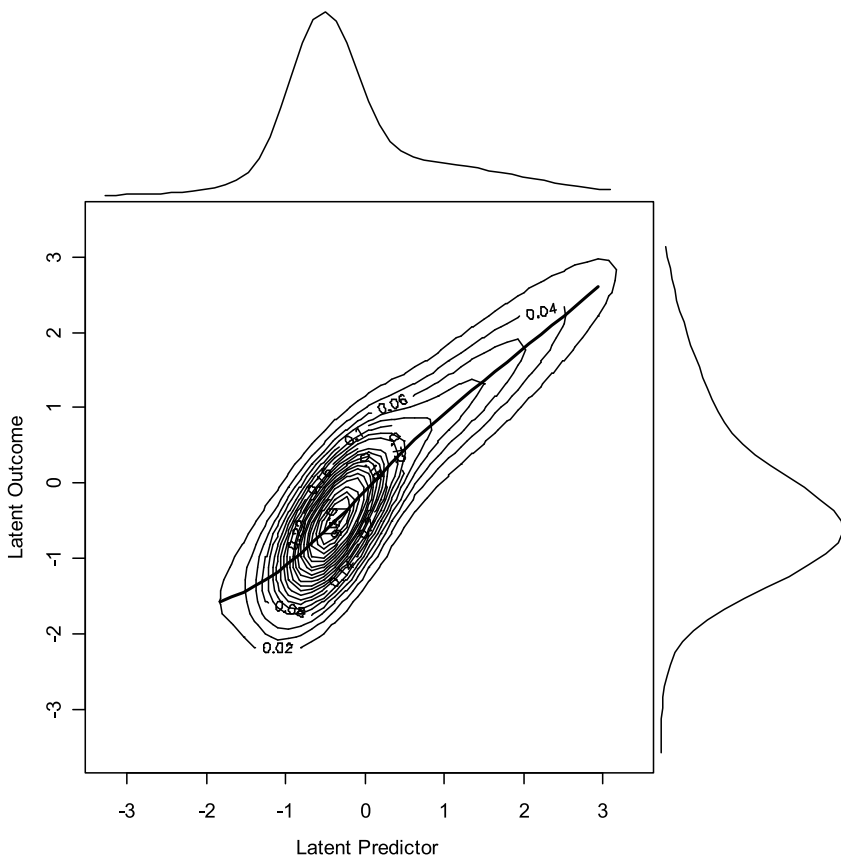
FIGURE 4   Plots of the marginal mixture densities for η_1 and η_2, the bivariate contour plot, and the linear regression of η_2 on η_1.

by relatively little attention to detail and increased reliance on general knowledge structures (see Schwarz, 2000, for a review). In contrast, individuals who are in a negative mood are highly likely to adopt a less heuristic processing strategy characterized by systematic cognition and considerable attention to detail (Schwarz & Clore, 1996). We thus anticipated that positive and negative emotions would both be monotonically related to heuristic processing (in opposite directions) but that these relationships might not be strictly linear.

The data were obtained from a convenience sample of 507 adults (41.89% males) between 19 and 60 years of age. The variables of interest here are positive and negative emotions, as measured by the modified Differential Emotions Scale (Fredrickson, Tugade, Waugh, & Larkin, 2003), and cognitive heuristic process-

ing, as measured by the Marlowe-Crowne Social Desirability Scale (Crowne & Marlowe, 1960). The latter measure has been shown to reflect cognitive heuristic processing by Shrauger (1972) as well as Evans and Forbach (1982). Indicators within each construct were combined by taking a weighted mean score to form three-item parcels for each latent variable. All parcels were scored in the direction where higher scores indicate more endorsement of the construct.

Of interest here are the bivariate relationships between positive and negative emotions and heuristic processing. Two sequences of models were thus estimated, the first predicting heuristic processing as a function of positive emotion and the second predicting heuristic processing as a function of negative emotion. In each case, SEMMs with one to five classes were estimated. For contrast, prior to discussing the results of the SEMMs, we first present the standard SEM for each model.

The effect of positive emotions on heuristic processing.   The linear effect of positive emotions on heuristic processing obtained from the standard SEM is in the expected direction, as seen in Figure 5. Yet there is little empirical basis for making the assumption that this relationship is linear, motivating our use of the semiparametric modeling approach. Fitting SEMMs to the data, the BIC, and other considerations suggested that a two-class model with across-class equality constraints on the variance parameters fit the data best. Likelihood ratio tests rejected both Condition A (i.e., an identical distribution for positive emotions in each class), $\Delta\chi^2(1) = 7.568$, $p < .01$, and Condition B (i.e., equal intercepts and slopes for the regression of heuristic processing on positive emotions in each class), $\Delta\chi^2(2) = 8.054$, $p < .05$, suggesting that heuristic processing is nonlinearly related to positive affect. Inputting the parameter estimates for the two-class model, shown in Table 1, into either the plotSEMM package or the online utility produces the plots shown in Figures 6 and 7. Note

TABLE 1
Parameter Estimates for the Effects of Affect on Cognitive Heuristic Processing

Parameters	π	α_1	α_2	β_{21}	ψ_{11}	ψ_{22}	BIC
Positive emotions							
One class model	1.000	3.047	0.234	0.093	0.621	0.024	1885.679
Class 1	0.602	3.529	0.020	0.152	0.265	0.023	1895.013
Class 2	0.398	2.317	0.336	0.053	0.265	0.023	
Negative emotions							
One class model	1.000	1.913	0.680	−0.084	0.459	0.026	1690.591
Class 1	0.308	2.670	0.392	0.019	0.492	0.023	1473.685
Class 2	0.692	1.573	1.124	−0.364	0.066	0.021	

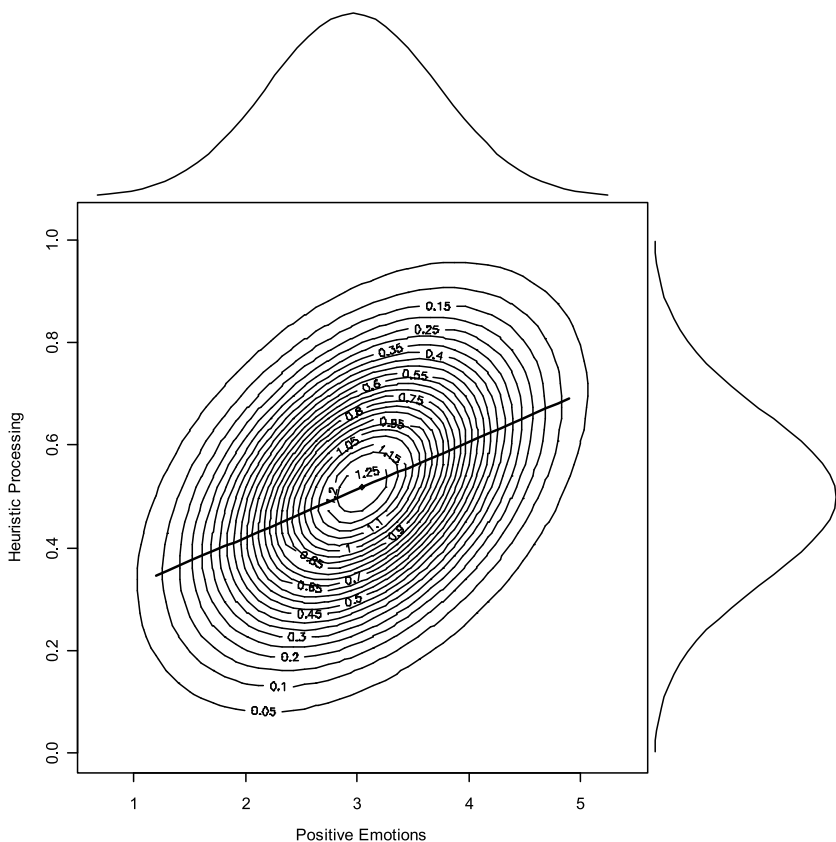*Note.*   BIC = Bayes Information Criterion.

FIGURE 5   Plots of the marginal densities for heuristic processing and positive emotions, the bivariate contour plot, and the linear regression of heuristic processing on positive emotions.

that, unlike Figures 3 and 4, we elected to include class information in Figures 6 and 7. Let us first consider the information presented in the two figures, followed by a summary of the substantive implications of the results.

The primary panel of Figure 6 depicts the globally nonlinear regression function obtained from Equation 5, in bold, superimposed on a contour plot of the model-implied bivariate distribution of the latent variables. To the top and right of the panel are the model-implied marginal distributions of the latent predictor and latent outcome, respectively. The marginal and bivariate distributions may be of interest in their own right, but they also serve two important purposes for interpreting the global regression function. First, clear evidence
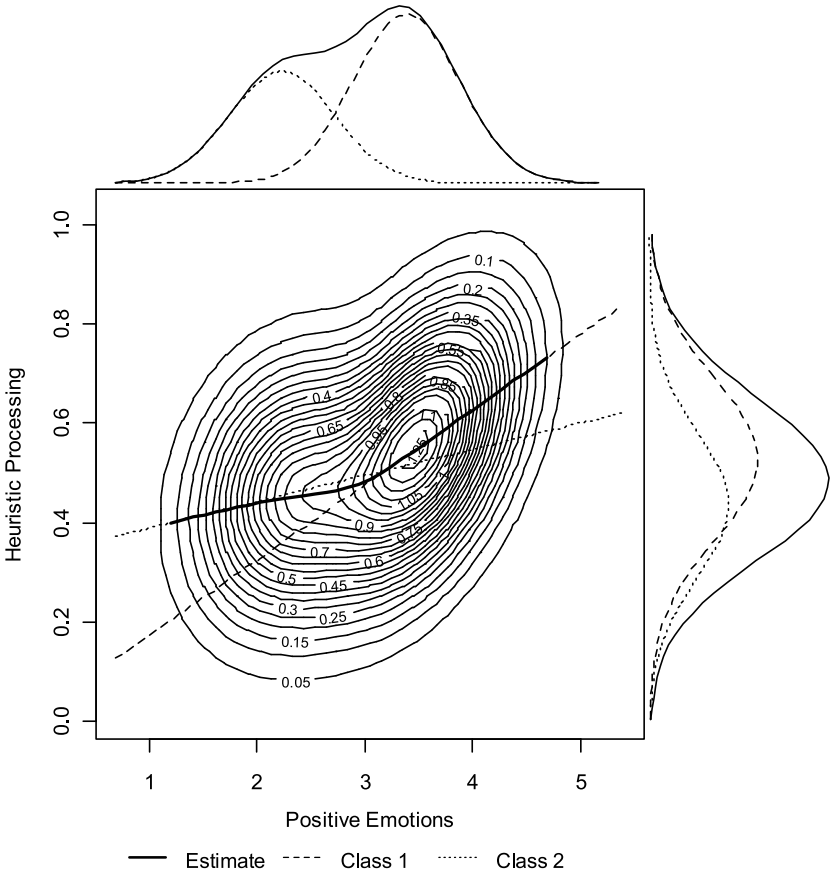
FIGURE 6   Plots of the marginal mixture densities for heuristic processing and positive emotions, the two locally linear regression estimates, the bivariate contour plot, and the nonlinear regression of heuristic processing on positive emotions.

of multimodality or discontinuities in the latent variable distributions may be indicative that smoothing over the within-class estimates is inadvisable. No such evidence is found in Figure 6. Second, the distribution plots indicate densely and sparsely populated regions of the latent space. The shape of the regression function should be interpreted with caution within sparse regions, whereas it may be interpreted more confidently within dense regions. For instance, the bend in the regression function seen in Figure 6 at positive emotion scores of about 3 takes place within a region of relatively high density.

Additionally, the within-class regression lines and the within-class marginal distributions are plotted in Figure 6 via different line types. This within-class
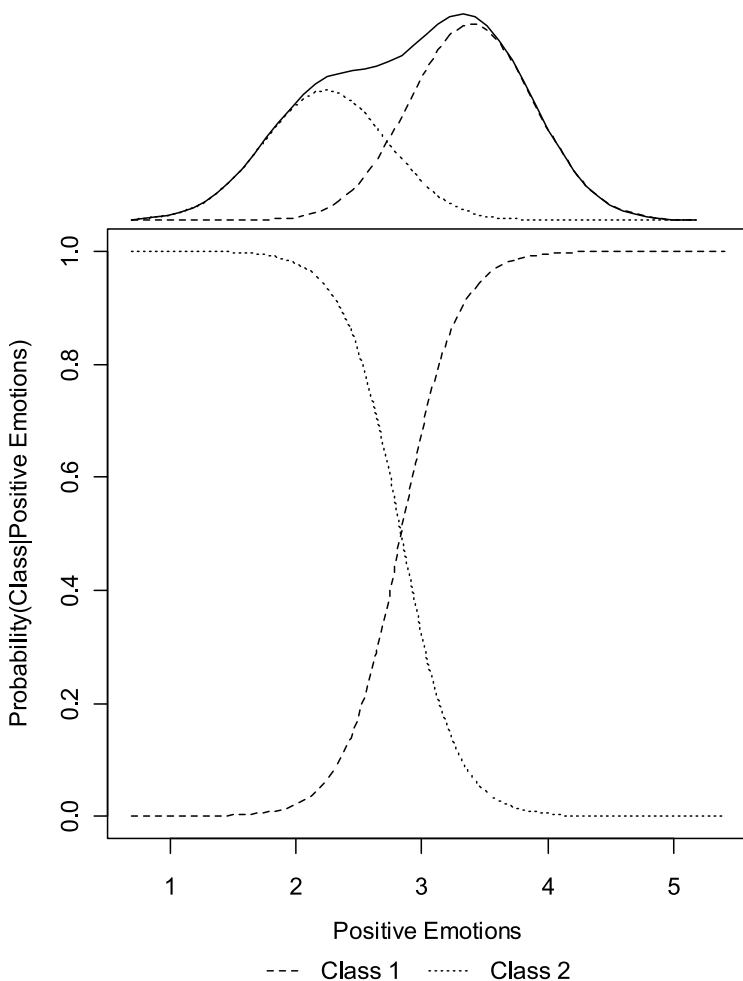
FIGURE 7   Plots of the marginal densities of positive emotions and the class probabilities across the range of positive emotions.

information is useful for interpreting the smoothing that takes place in computing the global regression function, particularly when viewed in conjunction with the plot of conditional mixing probabilities provided in Figure 7 (with similar line types). The within-class estimates shown in Figure 6 correspond to Equation 4 and indicate that the locally linear relationship between positive emotions and heuristic processing is weaker in Class 2 than Class 1. Class 2 is also character-ized by a lower mean level of positive emotions, as can be seen in the marginal

distribution of positive emotions plotted at the tops of Figures 6 and 7 (see also Table 1). The within-class marginal distributions for the latent predictor (dashed lines) contribute to the numerator of Equation 6 used to compute the conditional mixing probabilities. Hence, as shown in Figure 7, as we move across the range of positive emotions, from low to high, the weight (conditional probability) assigned to the Class 2 regression line in computing the global regression function diminishes and the weight assigned to the Class 1 regression line increases. Thus the global regression function computed by Equation 5 and shown in Figure 6 smoothly shifts from the Class 2 line to the Class 1 line.

More substantively, Figure 6 clearly shows that the relationship between positive emotions and heuristic processing is positive and monotonic but not strictly linear; beyond a score of about 3, the effect of positive emotions on heuristic processing becomes more acute. This pattern indicates that use of heuristic processing strategies increases more sharply under the influence of a strongly positive mood. At the extreme, this may contribute to the inability to attend to details and impaired judgment shown by individuals with bipolar disorder during a manic state (American Psychiatric Association, 1994).

The effect of negative emotions on heuristic processing.   The results of fitting a standard linear SEM to the negative emotions and heuristic processing data are shown in Table 1 and Figure 8. Again, there is little basis for making the assumption that the relationship between the latent variables is strictly linear. Fitting SEMMs to the data, we once again found that two classes were optimal for the data. In this case, however, the best model fit was obtained by allowing the variance of the negative emotions factor to differ over classes. In support of nonlinearity, Conditions A and B were both rejected, $\Delta\chi^2(2) = 244.486$, $p < .0001$ and $\Delta\chi^2(2) = 52.152$, $p < .0001$. Paralleling Figures 6 and 7 for positive emotions, Figures 9 and 10 depict the regression of cognitive heuristic processing on negative emotions. We again consider the information presented in each plot, followed by its substantive interpretation.

The model-implied bivariate contour plot for the two latent variables is depicted in the primary panel of Figure 9, with the bold line tracing out the global regression function derived by Equation 5. The contour plot does not provide any contraindications for interpreting the global regression function (e.g., discontinuities, strong multimodality). It does, however, indicate that the vast majority of participants reported a low level of negative emotions and that some care should be taken in interpreting the regression function within the sparse lower right quadrant. This characteristic of the data is similarly seen in the marginal distribution for negative emotions, which shows considerable positive skew. In contrast, the distribution for heuristic processing is much more normal in shape (consistent with Figure 6).
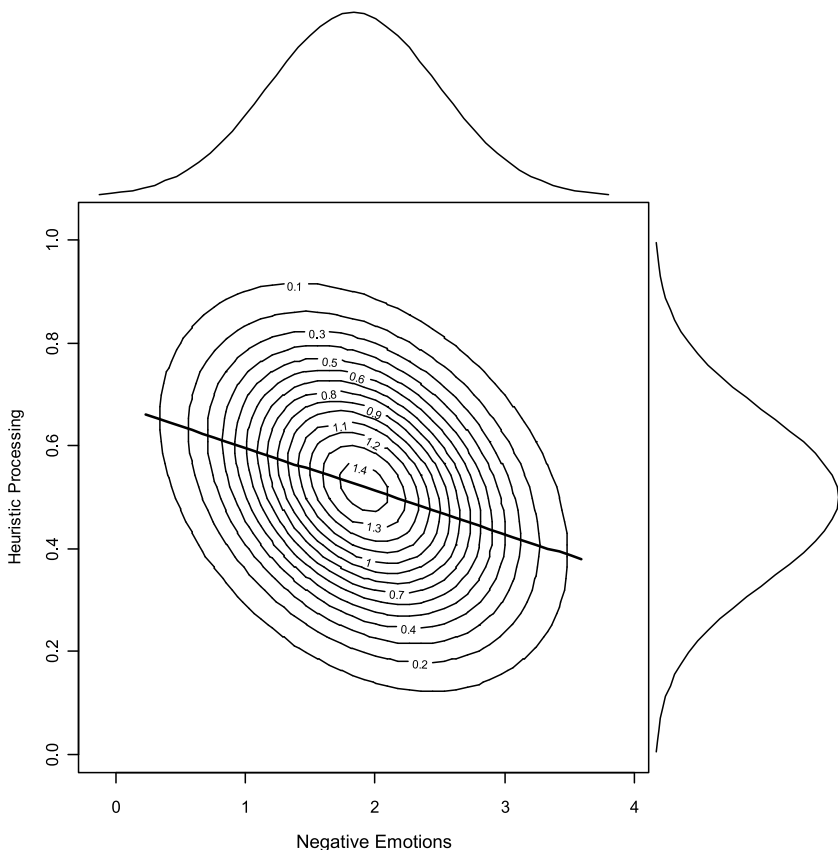
FIGURE 8   Plots of the marginal densities for heuristic processing and negative emotions, the bivariate contour plot, and the linear regression of heuristic processing on negative emotions.

Examining the within-class estimates, we can see that the positive skew for negative emotions is accommodated through the mixture of distributions for Class 1, with a higher mean and large variance, and Class 2, with a lower mean and small variance. In turn, this implies that the within-class regression line for Class 1 will be given greater weight (as determined by Equation 6) when negative emotions are high and less weight for lower levels of negative emotions (for which Class 2 will be given greater weight). This pattern is clearly evident in Figure 10 when examining negative emotion scores above 1. Note that given the wide variance of Class 1, the mixing probability for this class is also higher than Class 2 when negative emotions are very low
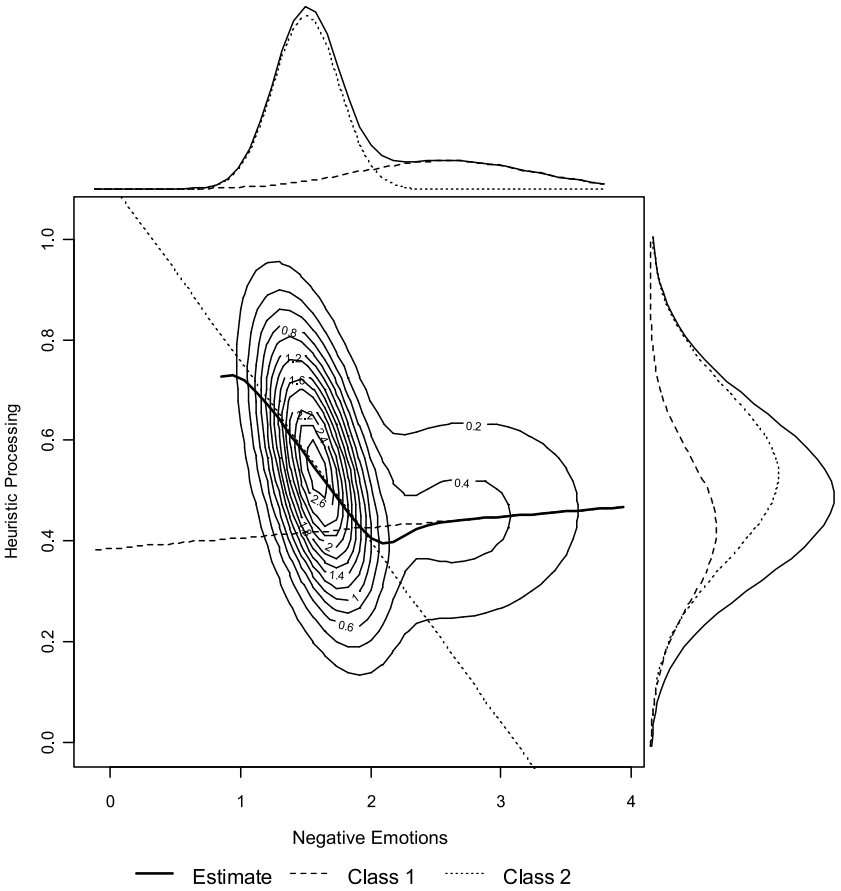
FIGURE 9   Plots of the marginal mixture densities for heuristic processing and negative emotions, the two locally linear regression estimates, the bivariate contour plot, and the nonlinear regression of heuristic processing on negative emotions.

(at the left side of Figure 10). Viewing these probabilities in relation to the marginal distribution depicted at the top of Figure 9, however, clarifies that this reversal essentially takes place outside of the range of the data. Thus, the global regression function in Figure 9 closely resembles the within-class regression line for Class 2 for negative emotion scores from 1 to 2, then smoothly shifts to the within-class regression line for Class 1 when negative emotion scores exceed about 2.
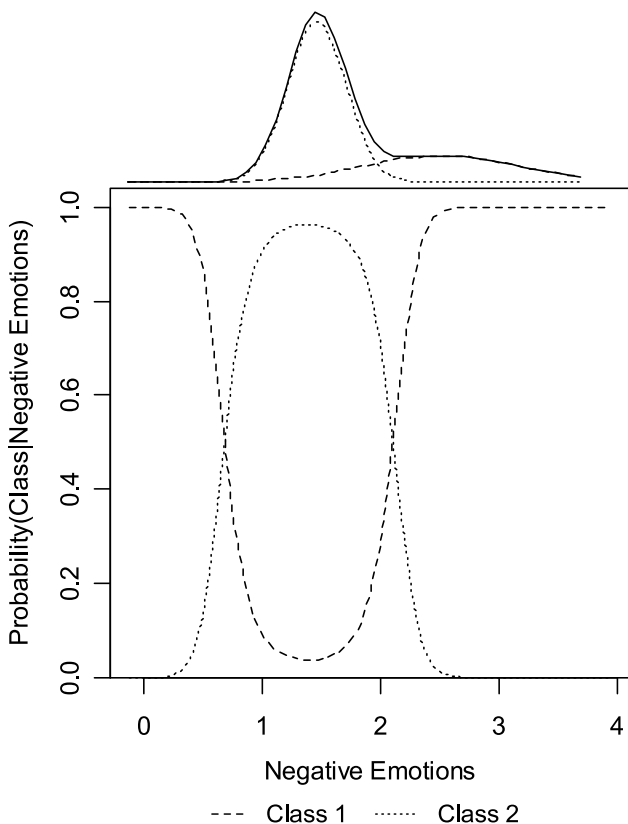
FIGURE 10   Plots of the marginal densities of negative emotions and the class probabilities across the range of negative emotions.

Overall, these results show that the regression of heuristic processing on negative emotions is also not linear. With increasing levels of negative emotions, the tendency to rely on heuristics in decision making decreases sharply, then essentially asymptotes when negative emotions exceed about two. This pattern shows that the influence of negative emotions on heuristic processing is somewhat distinct from the absence of positive emotion and further suggests a limit to "depressive realism" in that individuals appear to retain some heuristic processing even under the influence of strong negative emotion (Alloy & Abramson, 1979). Although suggestive, we must be mindful that these results are most stable for low levels of negative emotions, given the skew of the negative emotions distribution. To make stronger conclusions, it would be necessary to

recruit a sample including a larger proportion of individuals with high levels of negative emotions.

DISCUSSION

The assumption that latent variables are linearly related is often made for convenience, not on the basis of theory or data. Here we have shown that this assumption can be easily evaluated empirically by using a technique that is analogous to the LOWESS regression procedure routinely used to evaluate observed-variable relationships. Although the methods illustrated here were proposed previously by Bauer (2005), they have not been widely implemented. We have sought to remove several impediments to the use of this approach by providing a more accessible description of the approach, numerous example analyses, a standardized display format for the results, and software utilities that automate the post-processing of SEMM results to generate the displays. Although we focused on models including just one latent outcome, these utilities may also be applied without modification to visualize relationships between a latent predictor and multiple latent outcomes. For example, if the effect of negative affect on both cognitive processing and memory was examined, a single SEMM would be fit with emotion predicting the two outcomes. Two separate sets of graphs are then generated—those for negative affect on heuristic processing and those for negative affect on memory. Regardless of the number of outcomes, the relationship of the predictor to each outcome is examined in turn.

A limitation of our article is that we did not address the modeling of nonlinear interactions between two or more latent predictors via the semiparametric approach. There are two primary reasons for this omission. First, although Bauer (2005) noted that applying this procedure for SEMMs with multiple latent predictors could accommodate potentially nonlinear interaction effects, no empirical research has yet been carried out to evaluate the performance of this approach. Second, optimally depicting multidimensional relationships among latent variables will require different types of graphs than those that suffice for the bivariate case. For example, in order to portray interactions between two latent variables, one might use wire frame plots to depict the three-dimensional regression surface, and such plots would require wholly new graphing utilities. A second limitation of the graphing utilities is that they do not provide confidence intervals for the nonlinear latent regression estimate, although we are currently developing methods to do so. These tasks thus provide an important challenge to future research on the SEMM approach for modeling nonlinear effects. Despite these limitations, however, we hope that this article will facilitate the use of SEMMs to semiparametrically model bivariate latent variable relationships of unknown functional form under minimal distributional assumptions. The glob-

ally nonlinear regression estimates obtained by this approach will often be of interest in their own right and may also be useful for motivating the use of more standard polynomial regression models for latent variables.

ACKNOWLEDGMENTS

REFERENCES

Alloy, L. B., & Abramson, L. Y. (1979). Judgment of contingency in depressed and nondepressed students: Sadder but wiser? *Journal of Experimental Psychology: General, 108,* 41–485.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author

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.

Arminger, G., Wittenberg, J., & Schepers, A. (1996). *MECOSA 3: User guide*. Friedrichsdorf, Germany: Additive GmbH.

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

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.

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

Cleveland, W. S. (1981). LOWESS: A program for smoothing scatter plots by robust locally weighted regression. *The American Statistician, 35,* 54.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analyses for the behavioral sciences* (3rd. ed.). Mahwah, NJ: Erlbaum.

Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology, 24,* 349–354.

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

Evans, R. G., & Forbach, G. B. (1982). Intellectual ability correlation of the Marlowe-Crowne Social Desirability Scale. *Journal of Personality Assessment, 46,* 59–62.

Fredrickson, B. L., Tugade, M. M., Waugh, C. E., & Larkin, G. R. (2003). What good are positive emotions in crises? A prospective study of resilience and emotions following the terrorist attacks on the United States on September 11th, 2001. *Journal of Personality and Social Psychology, 84,* 365–376.

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., & 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., & 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.

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. G., & Muthén, B. O. (2007). Quasi-maximum likelihood estimation of structural equation models with multiple interaction and quadratic effects. *Multivariate Behavioral Research, 42,* 647–673.

Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York: Guilford.

Lubinski, D., & Humphreys, L. G. (1990). Assessing spurious "moderator effects": Illustrated substantively with the hypothesized ("synergistic") relation between spatial and mathematical ability. *Psychological Bulletin, 107,* 385–393.

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

Mooijaart, A., & Bentler, R. (1986). Random polynomial factor analysis. In E. Diday, M. Jambu, L. Lebart, J. Pages, & R. Tomassone (Eds.), *Data analysis and informatics, IV: Proceedings of the fourth international symposium on data analysis and informatics* (pp. 241–250). North Holland, Amsterdam: Elsevier Science Publishers.

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. (2007). *Mplus user's guide* (5th ed.). Los Angeles: Author.

Neale, M. C., Boker, S. M., Xie, G., & Maes, H. H. (2003). *Mx: Statistical modeling* (6th ed.). Richmond, VA: Department of Psychiatry.

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

Schwarz, N. (2000). Emotion, cognition and decision making. *Cognition Emotion, 14*, 433–440.

Schwarz, N., & Clore, G. L. (1996). Feelings and phenomenal experiences. In E. T. Higgins & A. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 433–465). New York: Guilford.

Shrauger, J. S. (1972). Performance on the Marlowe-Crowne Scale and its perceived purpose. *Journal of Personality Assessment, 36,* 287–290.

Smith, P. (1979). Splines as a useful and convenient statistical tool. *The American Statistician, 33,* 57–62.

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

Wold, S. (1974). Spline functions in data analysis. *Technometrics, 16,* 1–11.

Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology, 18,* 459–482.

APPENDIX A

R Code for Positive Emotions and Heuristic Processing Example

```
#user specified matrices
pi <- c(0.602, 0.398)
alpha1 <- c(3.529, 2.317)
alpha2 <- c(0.02, 0.336)
beta21 <- c(0.152, 0.053)
psi11 <- c(0.265, 0.265)
psi22 <- c(0.023, 0.023)

#calling the setup function to generate data to be used in plots
plotSEMM_setup(pi, alpha1, alpha2, beta21, psi11, psi22)

#functions for the two plots and the legend

plotSEMM_contour(Eta1, Eta2)
plotSEMM_probability(Eta2)
plotSEMM_legend()
```

APPENDIX B

M*plus* Code for Positive Emotions and Heuristic Processing Example

```
title:
    Positive emotions -> Heuristic processing
data:
    file is hp.dat;
variable:
    names are sid pe1 pe2 pe3 ne1 ne2 ne3 hp1 hp2 hp3;
    usevariables pe1-pe3 hp1-hp3;
    classes = class(2);
analysis:
    type = mixture;
    estimator = mlr;          !Robust maximum likelihood estimation
    starts = 500 10;          !Estimate 10 best of 500 random starts
  CLASSES = Class(2);          !Number of classes specified as 2
model:
    %overall%   !Parameters declared here are held constant across classes
    PE by pe1@1 pe2* pe3*;      !Measurement model for PE
    [pe1@0];                    !Fixing intercept of first PE variable at 0
    PE;                         !Covariance of PE
    HP by hp1@1 hp2* hp3*;      !Measurement model for HP
```

```
    [hp1@0];                        !Fixing intercept of first HP variable at 0
    HP;                             !Covariance of HP

%class#1%
    [PE*];                          !Mean of PE for class 1
    [HP*];                          !Mean of HP for class 1
    HP on PE;                       !Slope for latent variable in class 1

%class#2%
    [PE*];                          !Mean of PE for class 2
    [HP*];                          !Mean of HP for class 2
    HP on PE;                       !Slope for latent variable in class 2
```

APPENDIX C

Mx Code for Negative Emotions and Heuristic
Processing Example

```
TITLE Negative Emotions -> Heuristic Processing;
DATA NG=2 NI=10 NO=507 NModel=2
REC FILE=HP.dat

LABELS sid pe1 pe2 pe3 ne1 ne2 ne3 hp1 hp2 hp3
SELECT ne1 ne2 ne3 hp1 hp2 hp3;

BEGIN MATRICES;
C DIAG 2 2 Free   !Covariance matrix of latent variables Class1
D DIAG 2 2 Free   !Covariance matrix of latent variables Class2
E DIAG 6 6 Free   !Covariance matrix of measured variables
L FULL 6 2 Free   !Factor loading matrix
S FULL 2 2 Free   !Slopes for latent variables class 1
T FULL 2 2 Free   !Slopes for latent variables class 2
A FULL 2 1 Free   !Intercepts for latent variables class 1
B FULL 2 1 Free   !Intercepts for latent variables class 2
N FULL 6 1 Free   !Intercepts for measured variables
P FULL 2 1 Free   !Mixing probabilities
I IDEN 2 2 Fixed  !Identity matrix
END MATRICES;

EQUATE C 2 2 D 2 2     !Constraining variance of HP equal across classes

BEGIN ALGEBRA;
F = L*(I-S)~*C*((I-S)~)'*L' + E;     !Covariance structure for class 1
G = L*(I-T)~*D*((I-T)~)'*L' + E;     !Covariance structure for class 2
X = N + L*(I-S)~*A;                  !Mean structure for class 1
```

```
Y = N + L*(I-T)~*B;                 !Mean structure for class 2
END ALGEBRA;

Means X'_Y';
Covariance F_G;
Weight P;

!Matrix patterns
PATTERN L        !Specifying matrix elements to be estimated
0 0
1 0
1 0
0 0
0 1
0 1
PATTERN S
0 0
1 0
PATTERN T
0 0
1 0
PATTERN N 0 1 1 0 1 1

!Identification Constraints
VALUE 1 L 1 1 L 4 2    !Fix loadings as 1.0 for first loading on each factor
VALUE 0 N 1 1 N 4 1    !Fix intercepts as 0 for first measured variable.

!START VALUES
MATRIX E .3 .3 .3 .3 .3 .3
MATRIX D .3 .02
MATRIX L
1 0
1 0
1 0
0 1
0 1
0 1
MATRIX N 0 0 0 0 0 0
MATRIX S
0 0
.4 0
MATRIX T
0 0
1.1 0

MATRIX A 3.5 0
MATRIX B 2.5 .5
MATRIX P .3 .7
BO .05 .95 P 1 1 to P 2 1
END
```

```
TITLE mixing proportions sum to unity
CONSTRAINT_GROUP
BEGIN MATRICES;
P     FULL 2 1 = P(1)
I     UNIT 1 2
J     UNIT 1 1
END MATRICES
CON (I*P)-J;

END
```