Structural Equation Modeling, 18:465–491, 2011

ISSN: 1070-5511 print/1532-8007 online
DOI: 10.1080/10705511.2011.582408

Psychology Press
Taylor & Francis Group

TEACHER'S CORNER

Advanced Nonlinear Latent Variable Modeling: Distribution Analytic LMS and QML Estimators of Interaction and Quadratic Effects

Augustin Kelava
Department of Psychology, Technical University Darmstadt, Germany

Christina S. Werner, Karin Schermelleh-Engel, Helfried Moosbrugger, and Dieter Zapf
Department of Psychology, Goethe University, Germany

Yue Ma, Heining Cham, Leona S. Aiken, and Stephen G. West
Department of Psychology, Arizona State University

Interaction and quadratic effects in latent variable models have to date only rarely been tested in practice. Traditional product indicator approaches need to create product indicators (e.g., x_1^2, x_1x_4) to serve as indicators of each nonlinear latent construct. These approaches require the use of complex nonlinear constraints and additional model specifications and do not directly address the nonnormal distribution of the product terms. In contrast, recently developed, easy-to-use distribution analytic approaches do not use product indicators, but rather directly model the nonlinear multivariate distribution of the measured indicators. This article outlines the theoretical properties of the distribution analytic Latent Moderated Structural Equations (LMS; Klein & Moosbrugger, 2000) and Quasi-Maximum Likelihood (QML; Klein & Muthén, 2007) estimators. It compares the properties of LMS and QML to those of the product indicator approaches. A small simulation study compares the two approaches and illustrates the advantages of the distribution analytic approaches as multicollinearity increases, particularly in complex models with multiple nonlinear terms. An empirical example from the field of work stress applies LMS and QML to a model with an interaction and 2 quadratic effects. Example syntax for the analyses with both approaches is provided.

Keywords: estimators, interaction, multicollinearity, nonlinear structural equation models, quadratic

Correspondence should be addressed to Augustin Kelava, Technical University Darmstadt, Faculty of Human Sciences, Institute of Psychology, Alexanderstr. 10, 64283 Darmstadt, Germany. E-mail: tino@augustin-kelava.de

Within the behavioral sciences numerous substantive theories hypothesize interaction, quadratic effects, or both between multiple independent and dependent variables (Ajzen, 1987; Cronbach & Snow, 1977; Karasek, 1979; Lusch & Brown, 1996; Snyder & Tanke, 1976). As one example, in studying the relationship between parents' educational level and child's educational expectations, Ganzach (1997) hypothesized and found results consistent with a model with complex interactive and quadratic relationships: When the level of education of one parent is high, the educational expectations of the child will also be high, even if the level of education of the other parent is quite low. For each parent separately, the strength of the relationship between parent's education and child's educational expectations accelerated as parent's educational level increased. This hypothesis was represented by one negative (compensatory) interaction effect and two positively accelerating quadratic effects (one for each parent's educational level). Within the measured variable framework, such hypotheses can be tested using multiple regression (see Aiken & West, 1991):

$$CEE = \beta_0 + \beta_1 ME + \beta_2 FE + \omega_{12} ME \cdot FE + \omega_{11} ME^2 + \omega_{22} FE^2 + \epsilon \qquad (1)$$

In Equation 1, *CEE* is the child's educational expectation, *ME* is the mother's level of education, *FE* is the father's level of education, and ϵ is a residual. The βs are the coefficients of the linear effects. Following Klein and Moosbrugger's (2000) and Klein and Muthén's (2007) notation, the ωs are the coefficients of the nonlinear effects.

Many variables in the behavioral sciences are measured with less than perfect reliability, resulting in biased estimates of the regression coefficients for the nonlinear effects (Bohrnstedt & Marwell, 1978; MacCallum & Mar, 1995). Structural equation modeling (SEM) produces theoretically error-free estimates of the effects of latent variables, overcoming this problem (Marsh, Wen, & Hau, 2006; Schumacker & Marcoulides, 1998). However, SEM has only rarely been used in practice, in part because of the difficulty of model specification within the traditional product indicator (PI) approach. In contrast, newer distribution analytic approaches (Klein & Moosbrugger, 2000; Klein & Muthén, 2007) are easy to use and provide parameter estimates that can be more efficient, yielding greater statistical power, particularly with more complex models. The goals of this article are to provide an introduction to the distribution analytic approaches and to compare the properties of the distribution analytic and PI approaches both on a theoretical level and in a simulation study. We also illustrate the use of the approaches with an empirical example.

COMPLEX NONLINEAR MODELS: LATENT VARIABLE INTERACTIONS AND QUADRATIC EFFECTS

The early literature focused primarily on models with a single latent variable interaction or quadratic effect (e.g., Jöreskog & Yang, 1996; Kenny & Judd, 1984). More recently the literature (e.g., Kelava, Moosbrugger, Dimitruk, & Schermelleh-Engel, 2008; Lee, Song, & Tang, 2007) has begun to consider more complex models that involve simultaneous interaction and quadratic effects like Ganzach's (1997) model of children's educational expectations. Equation 2 expresses a latent model with one interaction and two quadratic effects analogously to Equation 1, but transferred from the manifest variables framework to the latent variables framework:

$$\eta = \alpha + \gamma_1\xi_1 + \gamma_2\xi_2 + \omega_{12}\xi_1\xi_2 + \omega_{11}\xi_1^2 + \omega_{22}\xi_2^2 + \zeta \tag{2}$$

In Equation 2, η denotes the latent criterion, ξ_1 and ξ_2 are latent predictors, the product $\xi_1\xi_2$ represents the interaction term, ξ_1^2 and ξ_2^2 are quadratic terms, α is the intercept, γ_1 and γ_2 are linear effects of the predictors, ω_{12} is the nonlinear effect of the interaction term, ω_{11} and ω_{22} are the nonlinear effects of the quadratic terms, and finally ζ is the latent disturbance. The more general matrix expression is given in Equation 3:

$$\begin{aligned}\eta &= \alpha + \boldsymbol{\Gamma\xi} + \boldsymbol{\xi}'\boldsymbol{\Omega\xi} + \zeta \qquad (3)\\ &= \alpha + \begin{pmatrix}\gamma_1 & \gamma_2\end{pmatrix}\cdot\begin{pmatrix}\xi_1\\ \xi_2\end{pmatrix} + \begin{pmatrix}\xi_1 & \xi_2\end{pmatrix}\cdot\begin{pmatrix}\omega_{11} & \omega_{12}\\ 0 & \omega_{22}\end{pmatrix}\cdot\begin{pmatrix}\xi_1\\ \xi_2\end{pmatrix} + \zeta\end{aligned}$$

In Equation 3, η denotes the latent criterion, α is the latent intercept, $\boldsymbol{\Gamma}$ is the coefficient vector for the linear effects of n latent predictors (summarized in the $\boldsymbol{\xi}$ vector), $\boldsymbol{\Omega}$ is the upper triangular coefficient matrix of the nonlinear effects (with the quadratic effects on the diagonal and the interactions effects off-diagonal), and finally ζ is the latent disturbance. Figure 1 depicts this nonlinear structural equation model with one interaction effect and two quadratic effects.

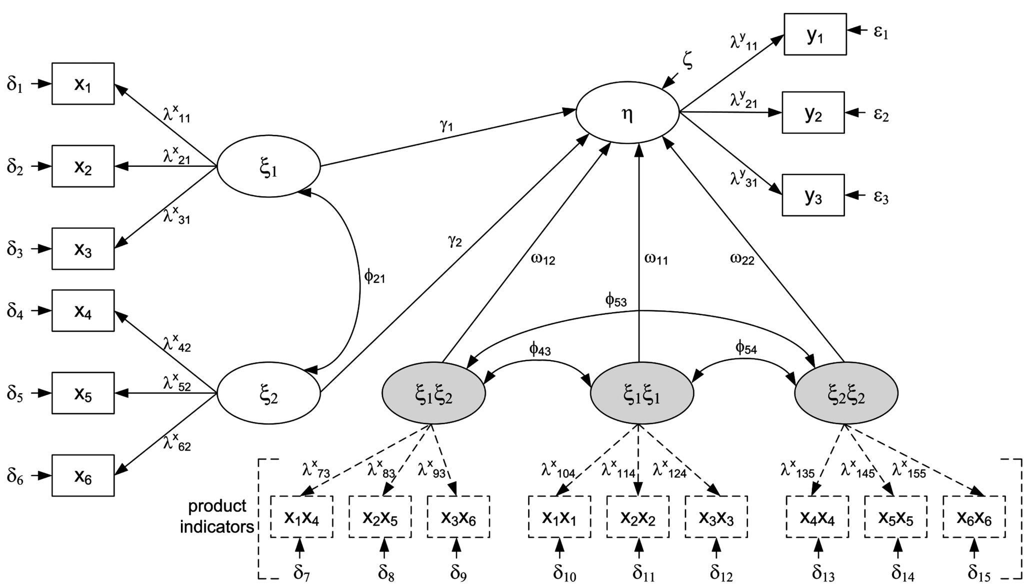

FIGURE 1 Nonlinear structural equation model with one latent interaction effect and two latent quadratic effects. Each linear latent variable (ξ_1, ξ_2, and η) has three indicator variables ($x_1, \ldots, x_3; x_4, \ldots, x_6$; and $y_1, \ldots, y_3$, respectively) as a measurement model. Note that product indicators (e.g., x_1x_4, x_2x_5) are only needed in product indicator approaches as a measurement model for the latent nonlinear terms ($\xi_1\xi_2, \xi_1^2$, and ξ_2^2). Thus, nonlinear measurement models are given in dashed lines. Distribution analytic approaches do not need measurement models for the latent nonlinear terms.

Product Indicator (PI) Approaches

Kenny and Judd (1984) initially developed the basic PI approach for nonlinear SEM. Their approach used multiple PIs for the specification of each nonlinear term's measurement model. Suppose that latent variables ξ_1 and ξ_2 are measured by centered and normally distributed indicators x_1, x_2, x_3 and x_4, x_5, x_6, respectively (Equation 4):

$$\mathbf{x} = \mathbf{\Lambda}^{\mathbf{x}} \cdot \boldsymbol{\xi} + \boldsymbol{\delta}$$

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \lambda^x_{21} & 0 \\ \lambda^x_{31} & 0 \\ 0 & 1 \\ 0 & \lambda^x_{52} \\ 0 & \lambda^x_{62} \end{pmatrix} \cdot \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} + \begin{pmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \\ \delta_5 \\ \delta_6 \end{pmatrix} \tag{4}$$

The interaction term $\xi_1\xi_2$ is measured by products of each latent variable's indicators, for example x_1x_4, x_2x_5, x_3x_6, referred to as PIs (see Figure 1). This PI approach has received subsequent development, particularly as reflected in contributions by Jöreskog and Yang (1996), Algina and Moulder (2001), Wall and Amemiya (2001), and Marsh, Wen, and Hau (2004).

Nonlinear constraints. Unfortunately, this approach has been rarely used by applied researchers. One reason is that the PI approach involves the specification of nonlinear parameter constraints that are difficult for researchers to implement. As depicted in Figure 1, x_2 and x_5 are indicators of the centered and normally distributed latent predictor variables ξ_1 and ξ_2; $x_2 = \lambda^x_{21}\xi_1 + \delta_2$ and $x_5 = \lambda^x_{52}\xi_2 + \delta_5$, where λ^x_{21} and λ^x_{52} are factor loadings and δ_2 and δ_5 are measurement errors, respectively. The product indicator x_2x_5 of the interaction term $\xi_1\xi_2$ can be expressed as:

$$\begin{aligned} x_2x_5 &= \lambda^x_{21}\lambda^x_{52}\xi_1\xi_2 + \lambda^x_{52}\xi_2\delta_2 + \lambda^x_{21}\xi_1\delta_5 + \delta_2\delta_5 \\ &= \lambda^x_{83}\xi_1\xi_2 + \delta_8 \end{aligned} \tag{5}$$

For the factor loading λ^x_{83}, the first subscript 8 refers to the eighth indicator (x_2x_5), and the second subscript 3 refers to the third latent variable $\xi_1\xi_2$ in the model (see Figure 1). The variance decomposition of the product indicator x_2x_5, which is required for model specification, is given by:

$$Var(x_2x_5) = {\lambda^x_{83}}^2\phi_{33} + \theta^\delta_{88}\text{, where:} \tag{6}$$

$$\lambda^x_{83} = \lambda^x_{21}\lambda^x_{52}$$

$$\phi_{33} = \phi_{11}\phi_{22} + \phi^2_{21}$$

$$\theta^\delta_{88} = {\lambda^x_{21}}^2\phi_{11}\theta^\delta_{55} + {\lambda^x_{52}}^2\phi_{22}\theta^\delta_{22} + \theta^\delta_{22}\theta^\delta_{55}$$

$$\phi_{11} = Var(\xi_1),\ \phi_{21} = Cov(\xi_1, \xi_2),\ \phi_{22} = Var(\xi_2),\ \theta^\delta_{22} = Var(\delta_2),\ \theta^\delta_{55} = Var(\delta_5)$$

Because loadings and variances of the indicator products are functions of the loadings and variances of the linear indicators, this estimation approach demands the specification of non-

TABLE 1
Specification of Nonlinear Constraints in the Product Indicator Approach

Nonlinear Model	*Interaction Model*	*Quadratic Model*
Product indicators	$\eta = \alpha + \gamma_1\xi_1 + \gamma_2\xi_2 + \omega_{12}\xi_1\xi_2 + \zeta$	$\eta = \alpha + \gamma_1\xi_1 + \gamma_2\xi_2 + \omega_{11}\xi_1^2 + \zeta$
Nonlinear constraints	x_1x_4, x_2x_5, x_3x_6	x_1^2, x_2^2, x_3^2
Factor loadings	$\lambda_{73}^x = 1, \lambda_{83}^x = \lambda_{21}^x\lambda_{52}^x, \lambda_{93}^x = \lambda_{31}^x\lambda_{62}^x$	$\lambda_{73}^x = 1, \lambda_{83}^x = \left(\lambda_{21}^x\right)^2, \lambda_{93}^x = \left(\lambda_{31}^x\right)^2$
Mean of latent nonlinear effect	$\underline{\boldsymbol{E(\xi_1\xi_2)=\phi_{21}}}$	$\underline{\boldsymbol{E(\xi_1^2)=\phi_{11}}}$
Variance and covariances of latent nonlinear effect	$Var(\xi_1\xi_2) = \phi_{11}\phi_{22} + \phi_{21}^2$ $Cov(\xi_1\xi_2, \xi_1) = Cov(\xi_1\xi_2, \xi_2) = 0$	$Var(\xi_1^2) = 2\phi_{11}^2$ $Cov(\xi_1^2, \xi_1) = Cov(\xi_1^2, \xi_2) = 0$
Error variances and covariances	$\theta_{77}^\delta = \phi_{11}\theta_{44}^\delta + \phi_{22}\theta_{11}^\delta + \theta_{11}^\delta\theta_{44}^\delta$ $\theta_{88}^\delta = \left(\lambda_{21}^x\right)^2\phi_{11}\theta_{55}^\delta + \left(\lambda_{52}^x\right)^2\phi_{22}\theta_{22}^\delta + \theta_{22}^\delta\theta_{55}^\delta$ $\theta_{99}^\delta = \left(\lambda_{31}^x\right)^2\phi_{11}\theta_{66}^\delta + \left(\lambda_{62}^x\right)^2\phi_{22}\theta_{33}^\delta + \theta_{33}^\delta\theta_{66}^\delta$ $\theta_{87}^\delta = \theta_{97}^\delta = \theta_{98}^\delta = 0$	$\theta_{77}^\delta = 4\phi_{11}\theta_{11}^\delta + 2\left(\theta_{11}^\delta\right)^2$ $\theta_{88}^\delta = 4\left(\lambda_{21}^x\right)^2\phi_{11}\theta_{22}^\delta + 2\left(\theta_{22}^\delta\right)^2$ $\theta_{99}^\delta = 4\left(\lambda_{31}^x\right)^2\phi_{11}\theta_{33}^\delta + 2\left(\theta_{33}^\delta\right)^2$ $\theta_{87}^\delta = \theta_{97}^\delta = \theta_{98}^\delta = 0$

Note. For both nonlinear models, the linear measurement model is $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \lambda_{21}^x & 0 \\ \lambda_{31}^x & 0 \\ 0 & 1 \\ 0 & \lambda_{52}^x \\ 0 & \lambda_{62}^x \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} + \begin{pmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \\ \delta_5 \\ \delta_6 \end{pmatrix}$ with covariance matrix $\begin{matrix} \xi_1 \\ \xi_2 \\ \zeta \\ \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \\ \delta_5 \\ \delta_6 \end{matrix} \begin{pmatrix} \phi_{11} & & & & & & & & \\ \phi_{21} & \phi_{22} & & & & & & & \\ 0 & 0 & \psi & & & & & & \\ 0 & 0 & 0 & \theta_{11}^\delta & & & & & \\ 0 & 0 & 0 & 0 & \theta_{22}^\delta & & & & \\ 0 & 0 & 0 & 0 & 0 & \theta_{33}^\delta & & & \\ 0 & 0 & 0 & 0 & 0 & 0 & \theta_{44}^\delta & & \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \theta_{55}^\delta & \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \theta_{66}^\delta \end{pmatrix}$. All linear indicators and linear latent variables of the measurement model are assumed to be centered and normally distributed. This table lists the required nonlinear constraints for the constrained product indicator (PI) approach. For partially constrained and unconstrained approaches of PI, some nonlinear constraints are released. The required constraints for the unconstrained approach (Marsh et al., 2004, 2006) are in boldface font and underlined. When linear latent variables are nonnormally distributed, the covariances between higher order terms and their linear terms (e.g., $Cov(\xi_1\xi_2, \xi_1)$) are nonzero in general. In the unconstrained approach, these covariances are estimated freely (Marsh et al., 2004, 2006). The specification of nonlinear constraints will become more complex if additional product indicators are used. This table summary is made based on earlier work by Algina and Moulder (2001), Jöreskog and Yang (1996), Kenny and Judd (1984), Lee et al. (2004), Marsh et al. (2004, 2006), and Wall and Amemiya (2001).

linear parameter constraints, which is very error prone. Table 1 presents the full set of nonlinear constraints needed for a model containing a single interaction $\xi_1\xi_2$ between two latent predictors or a single latent quadratic term ξ_1^2 for the *constrained PI approach* when the observed variables have been centered.

Marsh et al. (2004, 2006) proposed a so-called *unconstrained model* that relaxes most of the constraints in the Jöreskog and Yang (1996) model for a single latent variable interaction. In Marsh et al.'s approach, the factor loadings, the variance of the latent nonlinear effect, and the measurement error variances are all freely estimated[1] (see Table 1). Only the mean of the latent

[1]Note that the specification of the parameters in the unconstrained approach is conditional on the distribution of the variables. For example, when the latent exogenous variables $(\xi_1, \xi_2)'$ are nonsymmetrically distributed, the covariance between higher order terms and their linear terms (e.g., $Cov(\xi_1\xi_2, \xi_1)$ and $Cov(\xi_1\xi_2, \xi_2)$) also has to be specified (Marsh et al., 2004). In that case, the covariance matrix Φ is estimated freely.

nonlinear term is constrained to ϕ_{21} for a model with a single interaction effect and to ϕ_{11} for a model with a single quadratic effect (in boldface in Table 1). The unconstrained approach has shown generally good performance in simulation studies in terms of high convergence rates and small bias in the estimate of the nonlinear latent term under conditions of nonnormally distributed variables. However, even when the observed x variables have a multivariate normal distribution, the unconstrained approach shows a modest loss of statistical power in the test of the latent variable interaction relative to the Jöreskog and Yang (1996) fully constrained approach (Marsh et al., 2004).

Unfortunately, in more complex models involving multiple nonlinear terms, such as the Ganzach model depicted in Figure 1, the constraints and specifications become much more complex. Additional specifications are required when each measured indicator variable contributes to more than one product indicator. Recently, Kelava (2009), Kelava and Brandt (2009), and Moosbrugger, Schermelleh-Engel, Kelava, and Klein (2009) proposed an *extended unconstrained approach* that identifies the additional specifications that are needed for proper estimation in more complex nonlinear models. Table 2 presents the specifications that are required for Ganzach's model with one latent interaction and two quadratic effects (see Figure 1). With other complex models, the required specifications must be developed following guidelines presented in Kelava and Brandt (2009). Additional complexity in this approach occurs if unequal numbers of indicators are available for each latent exogenous variable or if indicators are nonnormally distributed.

Nonnormality of product terms. A second reason for the lack of use of this approach is that latent variable interactions and quadratic terms do not have a normal distribution, even when the measured and latent variables have normal distributions (Aroian, 1944; Ma, 2010; Moosbrugger, Schermelleh-Engel, & Klein, 1997). Therefore, normal theory-based standard errors and hence significance tests and confidence intervals for the effects of interest will be incorrect (cf. Jöreskog & Yang, 1996).

Distribution Analytic Approaches: LMS and QML

More recently Klein and Moosbrugger (2000) developed a Latent Moderated Structural Equations (LMS) approach that employs a unique model specification that does not involve PIs. LMS produces asymptotically correct standard errors for nonlinear effects. Because this approach becomes computationally (numerically) intensive as the number of nonlinear effects increases, Klein and Muthén (2007) subsequently developed a Quasi-Maximum Likelihood (QML) approach. QML permits the estimation of multiple nonlinear effects with a smaller increase of computational burden by taking a small loss of precision because a "quasi" likelihood (described later) is maximized. Figure 2 provides estimates of our empirical example from work stress (discussed in detail in a later section) analyzing a model containing one latent interaction and two quadratic effects, as specified for LMS or QML; note that there are no PIs.

The distribution analytic approaches make the same standard assumptions of latent variable models as the PI approaches (except for normally distributed y variables). On the predictor side, ξ and δ are assumed to be multivariate normally distributed with means equal to 0. Each of the measured x variables $(x_1, x_2, \ldots, x_q)'$ is normally distributed and centered internally by the program. On the criterion side, the ϵ and ζ variables are assumed to be multivariate

TABLE 2
Specification of Ganzach's Model in the Extended Unconstrained Approach

Nonlinear structural model

$$\eta = \alpha + \gamma_1\xi_1 + \gamma_2\xi_2 + \omega_{12}\xi_1\xi_2 + \omega_{11}\xi_1^2 + \omega_{22}\xi_2^2 + \zeta$$

Linear measurement models

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \lambda_{21}^x & 0 \\ \lambda_{31}^x & 0 \\ 0 & 1 \\ 0 & \lambda_{52}^x \\ 0 & \lambda_{62}^x \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} + \begin{pmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \\ \delta_5 \\ \delta_6 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \lambda_{21}^y & 0 \\ \lambda_{31}^y & 0 \end{pmatrix} \eta + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \end{pmatrix}$$

Nonlinear measurement model

$$\begin{pmatrix} x_1x_4 \\ x_2x_5 \\ x_3x_6 \\ x_1^2 \\ x_2^2 \\ x_3^2 \\ x_4^2 \\ x_5^2 \\ x_6^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ \lambda_{83}^x & 0 & 0 \\ \lambda_{93}^x & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \lambda_{11,4}^x & 0 \\ 0 & \lambda_{12,4}^x & 0 \\ 0 & 0 & 1 \\ 0 & 0 & \lambda_{14,5}^x \\ 0 & 0 & \lambda_{15,5}^x \end{pmatrix} \begin{pmatrix} \xi_1\xi_2 \\ \xi_1^2 \\ \xi_2^2 \end{pmatrix} + \begin{pmatrix} \delta_7 \\ \delta_8 \\ \delta_9 \\ \delta_{10} \\ \delta_{11} \\ \delta_{12} \\ \delta_{13} \\ \delta_{14} \\ \delta_{15} \end{pmatrix}$$

Covariance matrices

$$\begin{matrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \\ \delta_5 \\ \delta_6 \end{matrix} \begin{pmatrix} \theta_{11}^\delta & & & & & \\ 0 & \theta_{22}^\delta & & & & \\ 0 & 0 & \theta_{33}^\delta & & & \\ 0 & 0 & 0 & \theta_{44}^\delta & & \\ 0 & 0 & 0 & 0 & \theta_{55}^\delta & \\ 0 & 0 & 0 & 0 & 0 & \theta_{66}^\delta \end{pmatrix},$$

$$\begin{matrix} \delta_7 \\ \delta_8 \\ \delta_9 \\ \delta_{10} \\ \delta_{11} \\ \delta_{12} \\ \delta_{13} \\ \delta_{14} \\ \delta_{15} \end{matrix} \begin{pmatrix} \theta_{77}^\delta & & & & & & & & \\ 0 & \theta_{88}^\delta & & & & & & & \\ 0 & 0 & \theta_{99}^\delta & & & & & & \\ \boldsymbol{\theta_{10,7}^\delta} & 0 & 0 & \theta_{10,10}^\delta & & & & & \\ 0 & \boldsymbol{\theta_{11,8}^\delta} & 0 & 0 & \theta_{11,11}^\delta & & & & \\ 0 & 0 & \boldsymbol{\theta_{12,9}^\delta} & 0 & 0 & \theta_{12,12}^\delta & & & \\ \boldsymbol{\theta_{13,7}^\delta} & 0 & 0 & 0 & 0 & 0 & \theta_{13,13}^\delta & & \\ 0 & \boldsymbol{\theta_{14,8}^\delta} & 0 & 0 & 0 & 0 & 0 & \theta_{14,14}^\delta & \\ 0 & 0 & \boldsymbol{\theta_{15,9}^\delta} & 0 & 0 & 0 & 0 & 0 & \theta_{15,15}^\delta \end{pmatrix}$$

$$\begin{matrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \end{matrix} \begin{pmatrix} \theta_{11}^\epsilon & 0 & 0 \\ 0 & \theta_{22}^\epsilon & 0 \\ 0 & 0 & \theta_{33}^\epsilon \end{pmatrix}, \begin{matrix} \xi_1 \\ \xi_2 \\ \xi_1\xi_2 \\ \xi_1^2 \\ \xi_2^2 \end{matrix} \begin{pmatrix} \phi_{11} & & & & \\ \phi_{21} & \phi_{22} & & & \\ 0 & 0 & \phi_{33} & & \\ 0 & 0 & \phi_{43} & \phi_{44} & \\ 0 & 0 & \phi_{53} & \phi_{54} & \phi_{55} \end{pmatrix}, \quad \zeta \; \begin{pmatrix} \psi \end{pmatrix}$$

Latent expectations

$$E(\xi_1\xi_2) = \kappa_3, E(\xi_1^2) = \kappa_4, E(\xi_2^2) = \kappa_5$$

Note. All linear indicators $(x_1, \ldots, x_6)'$, exogenous latent variables $(\xi_1, \xi_2)'$, measurement errors $(\delta_1, \ldots, \delta_6, \epsilon_1, \ldots, \epsilon_3)'$ of the linear measurement models, and the latent disturbance ζ are assumed to be centered and normally distributed. This table lists the required specifications for the extended unconstrained approach (for details, see Kelava & Brandt, 2009). All parameters are estimated freely. Necessary measurement error covariances are in boldface font (e.g., $\theta_{10,7}^\delta$). Additional measurement error covariances (e.g., θ_{71}^δ) and additional covariances of the latent predictors (e.g., ϕ_{31}) need to be specified if linear indicator variables $(x_1, \ldots, x_6)'$ and latent variables $(\xi_1, \xi_2)'$ are non-normally distributed (Kelava & Brandt, 2009).

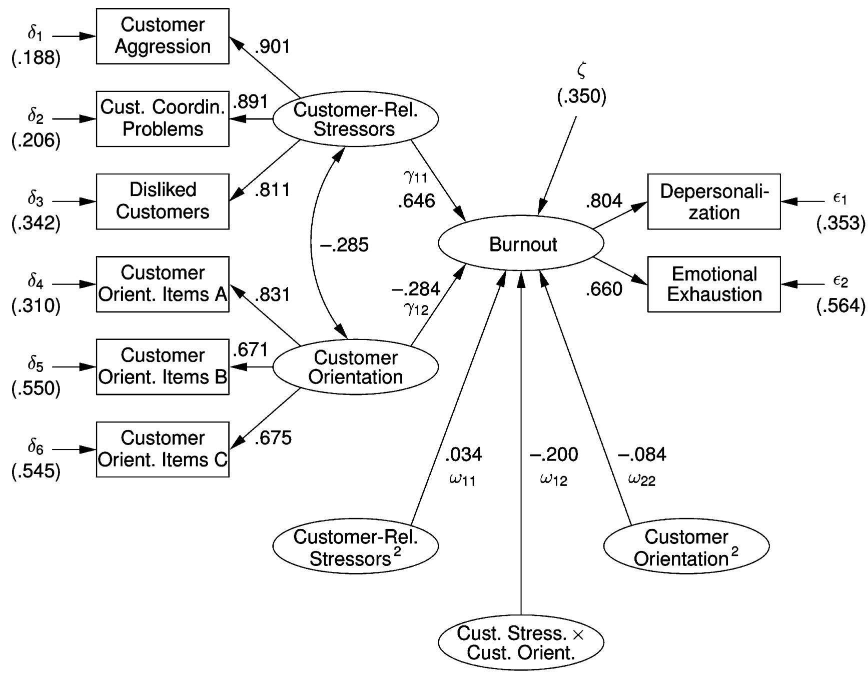

FIGURE 2 Complete nonlinear structural equation model including linear effects of customer-related stressors and customer orientation, interaction (Customer-Related Stressors × Customer Orientation), and quadratic effects (customer-related stressors^2 Customer Orientation^2) on burnout. Coefficients are standardized Quasi-Maximum Likelihood (QML) parameter estimates (variance estimates given in parentheses). With the exception of the quadratic effect of customer-related stressors, all model coefficients are significant. No product indicators are required in QML or Latent Moderated Structural Equations.

normally distributed with means equal to 0. As a consequence of the nonlinear effects, the latent η variable and the measured y variables $(y_1, y_2, \ldots, y_p)'$ will be nonnormal (Kenny & Judd, 1984). This nonnormality can be problematic for SEM procedures that use PIs. This problem can be solved by using distribution analytic approaches. Distribution analytic approaches use alternative procedures to maximize the (transformed) likelihood function, which takes the nonnormality of the nonlinear effects into account. These approaches yield more reliable estimates of standard errors of the nonlinear effects, but leave the estimates and interpretation of the nonlinear effects $(\omega_{12}, \omega_{11}, \omega_{22})$ unchanged.[2] We provide a brief overview of the LMS and QML estimation methods for a general audience. Readers wishing more detailed technical presentations should consult the original Klein and Moosbrugger (2000) and Klein and Muthén (2007) articles.

[2]All parameters associated with the two original linear latent predictor variables ξ_1, ξ_2 and the latent criterion variable η are identical in the distributional analytic and PI approaches. The distribution analytic approaches do not rely on PIs. Parameters associated with the PIs and with relationships involving latent exogenous interaction and quadratic variables and other exogenous variables are not estimated.

LMS estimator. The LMS procedure builds on two key statistical concepts. The first is the concept of mixture distributions—the observed (nonnormal) distribution of a variable can be represented by a combination of normal distributions having different means and variances. The second is the concept of conditional distribution, which is the distribution of a variable holding one or more other variables constant, each at a particular value.

To illustrate these ideas, consider the distribution of the height of adults in the United States which is nonnormal. Males have a distribution that is approximately normal ($\mu = 69.41$, $\sigma = 4.48$ inches) and females have a distribution that is approximately normal ($\mu = 63.86$, $\sigma = 4.39$ inches; McDowell, Fryar, Ogden, & Flegal, 2008) as well. In other words, there are two conditional normal distributions, one for *gender* = *male* and one for *gender* = *female*. Combining these two conditional distributions into one distribution represents the nonnormal distribution in the entire population. Statisticians often use this idea and combine several normal distributions to represent complex nonnormal distributions. The challenge is to find a conditioning variable like gender in the preceding example that identifies the mean and variance of the specific conditional normal distributions to be combined.

The LMS procedure builds on these two central ideas. First, although overall interaction ($\xi_1\xi_2$) and quadratic (ξ_1^2, ξ_2^2) effects are nonlinear, the conditional effects are linear when a variable is controlled that causes the nonlinearity. Second, the multivariate distribution of the observed indicator variables $(x_1, x_2, \ldots, x_q, y_1, y_2, \ldots, y_p)'$ can be approximated by a weighted combination of conditionally normal distributions. For both parts, the challenge is in finding the proper variable on which to condition.

LMS uses a matrix operation known as a Cholesky decomposition. Like principal components analysis (over which it has mathematical advantages), the Cholesky decomposition permits the analyst to replace the original variables (here, the latent ξ variables) with another set of orthogonal variables. In LMS the Cholesky decomposition is applied to the positive definite $(m \times m)$ covariance matrix $\mathbf{\Phi}$ of the m latent exogenous variables $(\xi_1, \ldots, \xi_m)'$, not including higher order nonlinear terms. Regardless of the number of nonlinear effects in the structural model, all models considered in this article contain two latent exogenous variables $(\xi_1, \xi_2)'$, therefore, m is equal to 2. More formally, the Cholesky decomposition can be expressed as:

$$\mathbf{\Phi} = \boldsymbol{\xi}\boldsymbol{\xi}' = \mathbf{AA}' = \mathbf{AIA}' = \mathbf{Azz}'\mathbf{A}' = (\mathbf{Az})(\mathbf{Az})' \tag{7}$$

where $\mathbf{I}$ is an $(m \times m)$ identity matrix. $\mathbf{I}$ is replaced by the vector product of a $(m \times 1)$ vector $\mathbf{z} = (z_1, \ldots, z_m)'$ with itself. Each z variable from the z vector is standardized and normally distributed ($z \sim N(0, 1)$) and is orthogonal to the remaining z variables. As can be seen from Equation 7, the decomposition of $\mathbf{\Phi}$ replaces the correlated ξ variables by an $\mathbf{A}$ matrix and by a $\mathbf{z}$ vector of m independent z variables. $\mathbf{z}$ can be separated into two subvectors:

$$\mathbf{z} = (z_1, \ldots, z_m)' = [\mathbf{z}_1', \mathbf{z}_2']' \tag{8}$$

where $\mathbf{z_1} = (z_1, \ldots, z_k)'$ and $\mathbf{z_2} = (z_{k+1}, \ldots, z_m)'$. The first k elements in $\mathbf{z_1}$ are the z variables that correspond to ξ variables involved in nonlinear terms. Here, k is equal to 2 because ξ_1 and ξ_2 were the only latent exogenous variables and both were involved in the nonlinear terms (e.g., $\xi_1\xi_2$). The remaining elements ($k + 1$ to m) in $\mathbf{z_2}$ (here, no elements) are those that are *only* involved in linear terms, but not in nonlinear terms. This procedure creates orthogonal components that allow us to partition the distribution of the y variables into linear and nonlinear parts.

The $\mathbf{z_1}$ vector is used as the conditioning variable. Klein and Moosbrugger (2000) showed that the joint distribution of the original measured x and y variables is conditionally multivariate normal when vector $\mathbf{z_1}$ is used as the basis for conditioning, $[\left((x_1, \ldots, x_p, y_1, \ldots, y_q)\right)' | \mathbf{z_1} \sim N\left(\mu(\mathbf{z_1}), \Sigma(\mathbf{z_1})\right)$, where p and q are the number of observed x and y variables, respectively]. Based on this result, they suggested using a mixture distribution to represent the multivariate distribution of the x and y variables in which $\mathbf{z_1}$ is used to determine the means, variances, and covariances of the set of normal distributions used in the mixture. These multivariate normal distributions are weighted and summed to represent the multivariate distribution of the observed variables. A numeric approximation procedure known as Hermite–Gaussian quadrature (see Freund & Hoppe, 2007) is used to approximate the mixture distribution.[3] The weights used by the quadrature process are those that produce the best approximation of the multivariate surface. Because LMS represents the nonnormal distribution as a mixture of conditionally normal distributions, no separate indicators of the product terms are needed. Figure 3 presents an illustration showing in the univariate case how different normal distributions can be combined to approximate the nonnormal distribution of y_1.

As with many difficult estimation problems, particularly mixture models, the expectation-maximization algorithm (EM; Dempster, Laird, & Rubin, 1977) is used to produce maximum likelihood estimates and standard errors for each of the parameters. Unlike standard SEM, LMS uses the full information contained in the raw data, not just the means and covariances. Wald z tests can be used to evaluate each parameter estimate compared to its standard error. Alternatively, likelihood ratio tests for nested models can be used to compare the full model to one in which each key parameter in turn is restricted to 0. Although asymptotically equivalent, likelihood ratio tests could be more accurate than Wald tests given realistic sample sizes. Klein and Moosbrugger (2000) provide the full technical details of the LMS procedure.

QML estimator. The QML procedure takes a different approach to solving the same problem of the nonnormality of the y variables.

The first key idea is that all but one of the measured y variables are corrected for the nonnormality that is caused by the nonlinear effects on η. For ease of presentation, we consider the special case in which the measured y variables have a τ-equivalent measurement structure[4] in which each of the indicators has the same unstandardized loading ($\lambda = 1$) on the underlying latent variable. For each measured y variable, we have $y_i = \eta + \epsilon_i$. Under standard assumptions, the most reliable y variable is designated as y_1, termed the scaling variable, which is taken as a proxy for η. As long as all variables are sufficiently reliable, another variable could be chosen as the scaling variable with little change in the results. However, the choice of an unreliable scaling variable could lead to substantially poorer estimates. Because the structural model for η contains latent nonlinear terms (e.g., $\xi_1\xi_2$), η will be nonnormally distributed. In contrast, the ϵ_is are assumed to have independent normal distributions. If we create a set of difference scores $\mathbf{y}^* = (y_2^* = y_2 - y_1, y_3^* = y_3 - y_1, \ldots, y_p^* = y_p - y_1)'$, then each difference score

[3]In the unidimensional case, quadrature approximation proceeds by using a series of rectangles to approximate the area under the curve. By using smaller widths, the rectangles provide better and better approximation of the area under the curve, but at a cost of increased computational burden. In the multidimensional case, the computational burden increases rapidly as the number of dimensions increases.

[4]The QML procedure assumes only the standard congeneric measurement structure so in practice the factor loadings can differ.

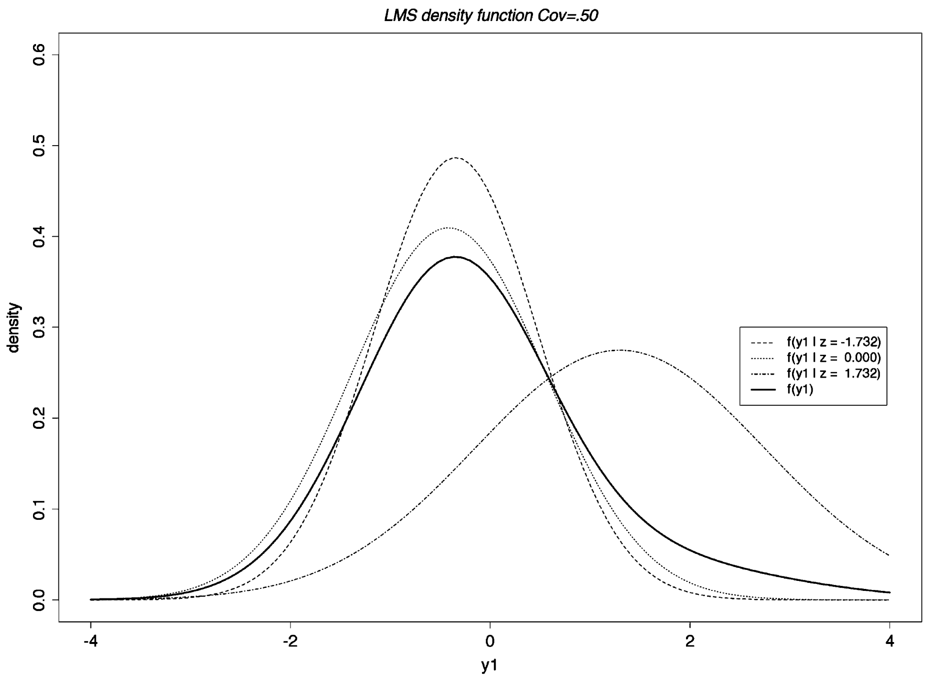

FIGURE 3 Latent Moderated Structural Equations (LMS) mixture density $f(y_1)$ and single mixture components: $Cov(\xi_1, \xi_2) = .50$. The three normal distributions, each conditioned on a different value of $\mathbf{z_1}$, that are components of the mixture distribution are depicted with thin lines using different line styles. The resulting nonnormal LMS mixture distribution representing a weighted sum of the component distributions is depicted with a thick black line.

will reflect the difference in the normally distributed measurement errors ($y_i^* = y_i - y_1 = [(\eta+\epsilon_i)-(\eta+\epsilon_1)] = (\epsilon_i - \epsilon_1)$). The resulting y variables vector $(y_1, y_2^*, \ldots, y_p^*)'$ contains one nonnormally distributed variable y_1 and $p-1$ difference score variables $\mathbf{y}^*$ that are normally distributed. In this formulation y_1 contains all the nonnormality resulting from the nonlinear effects on η.

The second key idea uses the idea of conditioning to remove normal parts from the y_1 distribution. A distribution $f(y_1|\mathbf{x}, \mathbf{y}^*)$ is created by conditioning y_1 on the x variables and the $p-1$ difference scores variables $\mathbf{y}^*$, both of which are assumed to be normally distributed. In effect, this conditional distribution partials out the normal parts of $(\mathbf{x}, \mathbf{y}^*)$ from the y_1 distribution, leaving the nonnormal parts.

The third key idea is that the joint multivariate nonnormal distribution $f(\mathbf{x}, \mathbf{y})$ of the x and y variables can be represented as the product of the conditional distribution $f(y_1|\mathbf{x}, \mathbf{y}*)$ and an unconditional distribution $f(\mathbf{x}, \mathbf{y}^*)$. The unconditional distribution $f(\mathbf{x}, \mathbf{y}^*)$ is a multivariate normal distribution of the x variables and the $\mathbf{y}^*$ difference score variables. This idea is expressed in Equation 9.

$$f(\mathbf{x}, \mathbf{y}) = f(y_1|\mathbf{x}, \mathbf{y}^*) f(\mathbf{x}, \mathbf{y}^*) \tag{9}$$

Estimating the density of the original multivariate distribution $f(\mathbf{x}, \mathbf{y})$ is difficult because the conditional distribution $f(y_1|\mathbf{x}, \mathbf{y}^*)$ is nonnormal and complex. QML solves this problem by replacing the nonnormal distribution $f(y_1|\mathbf{x}, \mathbf{y}^*)$ with $f^*(y_1|\mathbf{x}, \mathbf{y}^*)$ which is a normal distribution having the same mean and variance. The product of $f^*(y_1|\mathbf{x}, \mathbf{y}^*)$ and $f(\mathbf{x}, \mathbf{y}^*)$ provides an *approximation* $f^*(\mathbf{x}, \mathbf{y})$ of the original multivariate nonnormal distribution $f(\mathbf{x}, \mathbf{y})$ of the x and y variables as expressed in Equation 10.

$$f(\mathbf{x}, \mathbf{y}) \approx f^*(y_1|\mathbf{x}, \mathbf{y}^*) f(\mathbf{x}, \mathbf{y}^*) = f^*(\mathbf{x}, \mathbf{y}) \tag{10}$$

Using standard numerical procedures commonly used in maximum likelihood estimation (e.g., Newton-type procedures), the likelihood of this approximation expressed in Equation 10 is maximized and estimates of each of the linear and nonlinear effects and their standard errors are obtained. The procedure of replacing the nonnormal distribution with a normal one and maximizing the approximation of the likelihood function is termed quasi-maximum likelihood. The cost of this procedure is a small loss of efficiency compared to LMS. The information about the nonlinear effects is provided by the nonnormality of the y_1 indicator; no separate indicators of the product terms are needed.

Similarities and differences between the LMS and QML estimators. LMS uses a proper maximum likelihood function and represents the nonnormal distribution by a mixture of normal distributions. In contrast, QML uses a quasi-maximum likelihood estimation procedure that only approximates the true likelihood function. We focus here on theoretical implications associated with the different estimators that might be important for users. These implications will be most apparent in models with a larger magnitude and number of nonlinear effects.

1. When predictor variables and measurement error variables are normally distributed, LMS and QML should provide nearly identical estimates. LMS should have a small advantage in precision of estimates because it utilizes the true maximum likelihood function, instead of an approximation of it.
2. As the correlation between the latent predictors (here ξ_1 and ξ_2) that form the higher order terms in the structural equation increases (multicollinearity), a slight advantage should be found for LMS. With increasing multicollinearity, the distribution of the measured y variables becomes more nonnormal. The quality of the approximation of the conditional distribution of y_1 used in QML will decrease, leading to a slightly higher bias of the estimates relative to LMS.
3. When predictor variables and measurement error variables are normally distributed, QML should provide also slightly more biased estimates than LMS as the number of nonlinear terms increases. Once again, the distribution of the conditional y_1 variable will become more nonnormal and therefore the approximation will be less precise in QML.
4. When the assumption of normality of the predictor variables and measurement errors $(\delta, \epsilon, \xi, \zeta)$ is violated, QML is likely to produce less biased estimates than LMS, unless the distribution of the latent exogenous variables (ξs), the measurement errors (δs), or both are substantially skewed (Klein & Muthén, 2007). The ability of the mixture model representation in LMS to represent these more extreme forms of multivariate nonnormality decays more quickly than the approximation of the conditional y_1 distribution in

QML. For example, LMS assumes that its mixture components are normally distributed after conditioning for the mixing variable $\mathbf{z_1}$. If this assumption is violated, LMS applies a likelihood function for parameter estimation that is a (finite) sum of misspecified models. LMS is only able to approximate the nonnormality that is due to the nonlinear effects (ω_{12}, etc.). Conditioning on $\mathbf{z_1}$, the joint distribution of the indicator variables $(\mathbf{x}, \mathbf{y})$ becomes a (weighted) sum of normal distributions only when assumptions are met.

5. For more complex models with many nonlinear terms, the computational burden will increase. Of importance, this burden increases exponentially faster in LMS than in QML and can exceed the capacity of current personal computers when several nonlinear terms are involved. Complex models might also require far more computer time than is typical for problems not having nonlinear effects.
6. LMS is currently implemented in M*plus* (Muthén & Muthén, 1998–2007), a standard latent variable analysis software package in which one simply states the equations to be estimated. QML is currently a freestanding program that uses a matrix-based format similar to LISREL (Jöreskog & Sörbom, 1996). QML is available from Andreas Klein (aklein25@uwo.ca).

Differences Between the LMS/QML and PI Approaches

Recall that the PI approaches use products of observed variables to serve as indicators of latent variable interactions and quadratic effects. LMS and QML use conditional distributions to represent the nonlinear effects. The result is that the nonnormality of the y variables that is due to nonlinearity is directly addressed in LMS and QML, but not in PI approaches. Maximum likelihood estimation in PI approaches assumes multivariate normality, which will be violated by the product variables (Aroian, 1944) and the y variables.

Again, we focus next on theoretical implications associated with the different estimators for users.

1. In PI approaches, the distribution of the terms representing interactions and quadratic effects will always be nonnormal with the degree of nonnormality increasing as the latent predictor correlation and the magnitude of the interaction and quadratic effects increase (Aroian, 1944; Dimitruk, Schermelleh-Engel, Kelava, & Moosbrugger, 2007; Klein & Moosbrugger, 2000; Ma, 2010).[5] When maximum likelihood estimation rests on normal theory, the standard errors will be underestimated. This result will lead to an increased Type I error rate and confidence intervals that are too narrow. Consequently, estimates of standard errors should be corrected for nonnormality. Simulation studies using the constrained PI approach (e.g., Moulder & Algina, 2002) have found that common methods of correcting standard errors for nonnormality (e.g., Satorra & Bentler, 1994) fail to yield improved standard errors. Bootstrapping procedures (Brandt, 2009) might be able to improve estimates of standard errors, leading to acceptable Type I error rates at a modest cost in statistical power. In contrast, the LMS and QML estimators address the issue of nonnormality directly and provide more accurate standard errors and test

[5] The product term $\xi_1\xi_2$ has a skewness of 2.33 and a kurtosis of 9.84 when the latent predictors ξ_1 and ξ_2 are standard normally distributed and correlated at $\phi_{21} = .50$ (Ma, 2010).

statistics without correction. More accurate standard errors offer an important theoretical benefit: Adequate statistical power (e.g., .80) can be achieved with a smaller sample size than with PI approaches.

2. In PI approaches, the nonlinear effects are represented by second-order product terms (e.g., x_1x_2, x_1^2). The variances of these product terms are based on fourth-order moments. Estimation of these higher order moments is highly unstable. LMS and QML do not utilize product terms and do not require the estimation of higher moments, so their estimates should be more reliable.
3. The calculation of the χ^2 test of goodness of fit requires specification of a saturated model. The standard saturated model implemented in current SEM software that is used by PI approaches is *not* correct for nonlinear latent variable models, because these models include restrictions on the mean and covariance structures that are not appropriate for a saturated model (Klein & Schermelleh-Engel, 2010). Conclusions based on the χ^2 test of fit and practical fit indexes based on the χ^2 statistic are suspect. LMS appropriately does not provide measures of fit. QML includes a new measure of fit based on a theoretically appropriate saturated model; however, its performance has not yet been investigated. χ^2 difference (likelihood ratio) tests can be implemented in all approaches, permitting appropriate comparison of nested models. The Akaike's Information Criterion (AIC) and Bayesian Information Criterion (BIC) can also be calculated, permitting comparison of nonnested models.
4. PI approaches provide modification indexes (Lagrange multiplier tests) that can be used for exploratory model improvement, although modification indexes will often be misleading. Even under the theoretically ideal conditions of a model involving no nonlinear terms with data that are multivariately normal, these tests must be used cautiously (MacCallum, 1986) as they can potentially capitalize on chance, identifying incorrect aspects of the model for modification. These tests will be further compromised with latent variable models involving nonlinear terms because the χ^2 values associated with the nonlinear latent terms and their indicators will be inflated by nonnormality, exacerbating the problem of identifying inappropriate model modifications. LMS and QML do not offer misleading model modification indexes.
5. In PI approaches, measurement models using the products of measured x variables have to be constructed for latent variable interactions and quadratic effects. The correlations between higher order terms (e.g., $\xi_1\xi_2$ and ξ_1^2) have to be estimated. In contrast, LMS and QML bypass the construction of indicators of latent nonlinear variables so that these correlations are not estimated.

SIMULATION STUDY

To illustrate some finite sample differences between the distribution analytic approaches and the PI approach, we specified a nonlinear model with one interaction and two quadratic effects corresponding to Ganzach's model (Equation 2; see also Figure 1), using the population parameter values given in Table 3. We investigated the performance of LMS, QML, and the extended unconstrained approach (Kelava, 2009; Kelava & Brandt, 2009; Moosbrugger et al., 2009). We varied the correlation between the two latent first-order variables ξ_1 and ξ_2 to study

TABLE 3
Six True Models: Variation of the Multicollinearity Conditions and Nonlinear Effects

	Parameter	*Type I Error Rate Simulations True Value*	*Power simulations True Value*
	ω_{11}	0.000	0.101
$\phi_{21} = .000$	ω_{12}	0.000	0.148
	ω_{22}	0.000	0.101
	ω_{11}	0.000	0.101
$\phi_{21} = .375$	ω_{12}	0.000	0.139
	ω_{22}	0.000	0.101
	ω_{11}	0.000	0.101
$\phi_{21} = .625$	ω_{12}	0.000	0.126
	ω_{22}	0.000	0.101

Note. The following parameters were constant across conditions: $Var(\xi_1) = Var(\xi_2) = 1$; $\gamma_1 = \gamma_2 = .316$; $\lambda^x_{11} = \lambda^x_{42} = \lambda^y_{11} = 1$; $\lambda^x_{21} = \lambda^x_{31} = \lambda^x_{52} = \lambda^x_{62} = \lambda^y_{21} = \lambda^y_{31} = .894$; $\theta^\delta_{11} = \theta^\delta_{44} = \theta^\epsilon_{11} = .25$; $\theta^\delta_{22} = \theta^\delta_{33} = \theta^\delta_{55} = \theta^\delta_{66} = \theta^\epsilon_{22} = \theta^\epsilon_{33} = .20$. See Figure 1 for a path diagram defining the parameters.

the effects of increasing multicollinearity. We also examined the Type I error rates for test statistics when there are no true nonlinear effects in the population ($\omega_{11} = \omega_{12} = \omega_{22} = 0$). The key nonlinear effects ω_{11}, ω_{12}, and ω_{22} and the linear effects γ_1 and γ_2 that are of central interest to researchers are theoretically expected to be identical in the extended PI and distribution analytic approaches so they can be directly compared. More accurate estimation of the nonlinear effects associated with the distribution analytic approaches effects was expected to be reflected in smaller standard errors and higher statistical power for the tests of the nonlinear effects without an increase in the Type I error rate.

Design of the Simulation Study

Data were generated specifying the latent predictor correlations to be .000, .375, and .625. These latent predictor correlations correspond to correlations between measured indicators of .000, .300, and .500 when indicators have a reliability of .800. In terms of Cohen's 1988 norms, correlations between measured variables of .300 and .500 are described as moderate and strong, respectively. Linear effects explained a total of 10% of the latent criterion's variance in each structural model. First, we examined the Type I error rates, setting all nonlinear effects to zero. Second, we specified three nonlinear effects that each explained 2.2% of the latent criterion's variance. Champoux and Peters (1987) and Chaplin (1991, 2007) have documented the small effect size of continuous variable interactions that typify personality and industrial-organizational literatures. Theoretically, larger magnitudes of nonlinear effects would be expected to produce greater nonlinearity of the observed y variables, further improving the performance of LMS and QML (Klein & Moosbrugger, 2000; Klein & Muthén, 2007) relative to the extended unconstrained approach.

The sample size was $N = 400$, a relatively large but realistic value (see MacCallum & Austin, 2000). The reliability of each indicator of ξ_1 and ξ_2 was set to be .800. Indicator variances were homogenous. The latent linear predictor variables, latent disturbances, and measurement error variables were normally distributed and centered. These specifications imply that the observed predictor variables x_1 to x_6 were also normally distributed in the population. A total of 500 replications were generated from each of the three true population models. A solution was considered proper and selected when there were no negative variances or standard error estimates (cf. Paxton, Curran, Bollen, Kirby, & Chen, 2001). Table 3 provides a summary of the true parameter values of the three true models.

The 500 replication samples from each of the three true nonlinear models were analyzed using LMS, QML, and the extended unconstrained approach, always specifying a nonlinear model with one interaction and two quadratic effects. The parameter estimates and standard errors of the nonlinear latent effects were examined.

Software and Implementations

Data were generated using PRELIS version 2.7 (Jöreskog & Sörbom, 1999). Analyses with the extended unconstrained approach were conducted using LISREL version 8.72 (Jöreskog & Sörbom, 1996). Analyses with Klein and Moosbrugger's LMS approach were carried out using M*plus* version 5.21 (Muthén & Muthén, 1998–2007). Analyses with Klein and Muthén's QML approach were conducted using the QML stand-alone software version 3.11 (Klein, 2007). Examples of syntax files for these approaches are included in Appendices A and B available at http://www.augustin.kelava.de/pubs. In this simulation study, we always used the default settings (e.g., default start values, default number of iterations, and default numerical algorithms) supplied by the programs, reflecting the typical practice of most users.

Results of the Simulation Study

We focus next on the results for the central nonlinear effects of interest ($\omega_{11}, \omega_{12}, \omega_{22}$). No appreciable differences among LMS, QML, and the extended unconstrained approach for the linear effects γ_1, and γ_2 or the factor loadings on ξ_1 and ξ_2 were expected or observed.

Table 4 shows the results of the simulations when no nonlinear effects are present in the true model. In other words, ω_{11}, ω_{12}, and ω_{22} were set to 0 in Equation 2 in the data generation process. In each multicollinearity condition, mean parameter estimates of the nonlinear effects produced by LMS, QML, and the extended unconstrained approach were close to zero. The standard error estimates were good approximations of the observed standard deviations of the parameter estimates. None of the standard errors was severely underestimated, using 10% as a criterion for severe bias. Across all multicollinearity conditions, Type I error rates were within acceptable ranges and not systematically inflated or deflated. The convergence rate for LMS and the extended unconstrained approach was always 100%. For QML, the convergence rate was slightly lower (96.2%–98.2%).

Table 5 shows the results when the three nonlinear effects are present in the true model. Again, the convergence rate of LMS and the extended unconstrained approach was always 100%. For QML, the convergence rate ranged between 98.0% and 99.4% with higher convergence rates as multicollinearity increased.

When the latent predictors were uncorrelated, unbiased mean parameter estimates of the nonlinear effects were obtained for each approach (Table 5, Panel A). LMS, QML, and the extended unconstrained approach provided accurate standard error estimates. The largest bias

TABLE 4
Type I Error Rates for the Three Approaches

	True Value	*Approach*	*N (Converged)*	*M Est.*	*SD*	*SE*	*SE/SD*	*Type I Error Rate*
Panel A		LMS	500	.002	.038	.036	.958	.048
	$\omega_{11} = .000$	QML	481	.002	.038	.037	.980	.058
		Unconstrained	500	.002	.039	.038	.976	.056
		LMS	500	−.003	.051	.052	1.002	.052
$\phi_{21} = .000$	$\omega_{12} = .000$	QML	481	−.003	.052	.051	.991	.050
		Unconstrained	500	−.003	.052	.053	1.023	.048
		LMS	500	.003	.039	.036	.933	.062
	$\omega_{22} = .000$	QML	481	.003	.039	.037	.948	.067
		Unconstrained	500	.003	.040	.038	.938	.052
Panel B		LMS	500	.004	.043	.042	.963	.064
	$\omega_{11} = .000$	QML	491	.003	.043	.042	.969	.063
		Unconstrained	500	.004	.045	.044	.976	.058
		LMS	500	−.005	.065	.064	.974	.074
$\phi_{21} = .375$	$\omega_{12} = .000$	QML	491	−.006	.065	.066	1.014	.055
		Unconstrained	500	−.005	.067	.066	.992	.060
		LMS	500	.003	.044	.041	.947	.060
	$\omega_{22} = .000$	QML	491	.003	.044	.043	.968	.061
		Unconstrained	500	.003	.046	.044	.957	.042
Panel C		LMS	500	.006	.064	.061	.955	.064
	$\omega_{11} = .000$	QML	491	.006	.064	.061	.957	.067
		Unconstrained	500	.006	.068	.066	.973	.054
		LMS	500	−.009	.109	.104	.956	.064
$\phi_{21} = .625$	$\omega_{12} = .000$	QML	491	−.011	.108	.104	.959	.059
		Unconstrained	500	−.009	.114	.112	.977	.048
		LMS	500	.005	.064	.061	.957	.060
	$\omega_{22} = .000$	QML	491	.005	.064	.061	.957	.059
		Unconstrained	500	.005	.068	.066	.974	.028

Note. LMS = Latent Moderated Structural Equations; QML = Quasi-Maximum Likelihood.

of the nonlinear parameter estimates was 4.5%. Relative to the actual standard deviation of the parameter estimates, the bias in the mean standard error estimate never exceeded 10%. But there were differences among approaches with respect to power for detecting interaction and quadratic effects. For each nonlinear effect, the power for detecting the effect was lowest with the extended unconstrained approach.

With a latent predictor correlation of .375, acceptable parameter estimates were obtained using all three approaches (Table 5, Panel B). All approaches slightly overestimated the two quadratic effects and underestimated the interaction effect. But, the highest overestimation did not exceed 6.0%. The standard error estimates were acceptable for all three approaches and did not deviate substantially from the standard deviations of the parameters. The highest empirical standard deviations of the parameter estimates were shown by the extended unconstrained approach. Accordingly the power for detecting a nonlinear effect was lower with the extended unconstrained approach than with LMS or QML.

Given a correlation between latent predictors of .625, the extended unconstrained approach produced a slightly larger bias in the estimates of the nonlinear effects than LMS and QML (Table 5, Panel C). For example, the extended unconstrained approach overestimated the first

TABLE 5
Results for Three Models with One Interaction Effect and Two Quadratic Effects in the Context of Increasing Multicollinearity

	True Value	*Approach*	*N (Converged)*	*M Est.*	*Difference*	*Bias %*	*SD*	*SE*	*SE/SD*	*Power*
Panel A		LMS	500	.104	.003	3.10	.037	.036	.970	.826
	$\omega_{11} = .101$	QML	490	.104	.003	3.00	.037	.036	.975	.824
		Unconstrained	500	.106	.005	4.50	.038	.037	.965	.812
		LMS	500	.145	−.003	−1.80	.051	.051	.994	.816
$\phi_{21} = .000$	$\omega_{12} = .148$	QML	490	.145	−.003	−2.00	.051	.051	1.004	.818
		Unconstrained	500	.147	−.001	−1.00	.052	.052	.998	.800
		LMS	500	.104	.003	2.50	.038	.036	.935	.832
	$\omega_{22} = .101$	QML	490	.104	.003	2.60	.039	.036	.921	.816
		Unconstrained	500	.105	.004	4.00	.040	.037	.923	.792
Panel B		LMS	500	.106	.005	4.70	.041	.040	.975	.760
	$\omega_{11} = .101$	QML	496	.106	.005	4.50	.041	.040	.970	.762
		Unconstrained	500	.107	.006	6.00	.042	.041	.975	.744
		LMS	500	.134	−.005	−3.30	.062	.060	.963	.598
$\phi_{21} = .375$	$\omega_{12} = .139$	QML	496	.134	−.005	−3.60	.062	.060	.972	.615
		Unconstrained	500	.135	−.004	−3.00	.064	.063	.976	.602
		LMS	500	.104	.003	2.80	.042	.039	.945	.752
	$\omega_{22} = .101$	QML	496	.104	.003	3.00	.042	.039	.941	.758
		Unconstrained	500	.106	.005	4.50	.044	.041	.939	.718
Panel C		LMS	500	.108	.007	6.60	.059	.057	.958	.464
	$\omega_{11} = .101$	QML	497	.108	.007	6.70	.059	.057	.960	.483
		Unconstrained	500	.110	.009	8.80	.063	.061	.973	.440
		LMS	500	.119	−.007	−6.00	.100	.095	.943	.276
$\phi_{21} = .625$	$\omega_{12} = .126$	QML	497	.118	−.008	−6.40	.099	.095	.952	.274
		Unconstrained	500	.118	−.008	−6.20	.105	.102	.970	.228
		LMS	500	.105	.004	3.90	.059	.056	.952	.474
	$\omega_{22} = .101$	QML	497	.105	.004	3.80	.059	.056	.952	.477
		Unconstrained	500	.107	.006	6.30	.063	.061	.960	.446

Note. LMS = Latent Moderated Structural Equations; QML = Quasi-Maximum Likelihood.

quadratic effect by 8.8%. The extended unconstrained approach showed the highest empirical standard deviations of its parameter estimates. For each nonlinear effect, the power for detecting the effect was again lower with this approach. LMS and QML showed a power of approximately 46.4% to 48.3% to detect a quadratic effect, whereas the power of the extended unconstrained approach was approximately 44.0% to 44.6%. The power for detecting the interaction effect was low for LMS (27.6%) and QML (27.4%), but even lower for the extended unconstrained approach (22.8%).

EMPIRICAL EXAMPLE

We illustrate the analysis of nonlinear models using an empirical example. In certain service jobs, employees frequently encounter customer-related social stressors (Dormann & Zapf, 2004), ranging from annoyances (e.g., customers disturbing an employee's workflow) to harassment (e.g., customers insulting an employee). Regular exposure to customer-related stressors can seriously affect employees' well-being, leading to burnout (Dormann & Zapf, 2004). For coping with customer-related stressors, an employee's personal resources can be crucial: Confronted with an angry customer, an employee possessing a high personal customer orientation might try to empathize and remain friendly. This strategy ideally leads to both customer satisfaction and a sense of professional efficacy on the employee's part. In contrast, an employee with a low personal customer orientation might interpret the customer's anger as a personal attack and escalate the conflict, resulting in customer dissatisfaction and emotional exhaustion on the employee's part. Thus, customer orientation can be expected to moderate the effect of customer-related stressors on employee burnout.

In addition to an interaction effect, nonlinear effects of stressors and resources could be expected as well. Severe stressor levels might have a particularly detrimental effect on health, corresponding with a positive quadratic effect of the stressor. On the other hand, if customer orientation serves as a resource in coping with customer-related stressors, high levels of customer orientation could be expected to be particularly effective in changing stressor appraisal and facilitating coping, leading to a negative quadratic effect of customer orientation on burnout.

Based on these predictions, we estimated the following model (Equation 11, cf. Equation 2), where *BO* represents the latent criterion burnout and *CRS* and *CO* represent the latent predictors customer-related stressors and customer orientation (cf. Figure 2):

$$BO = \alpha + \gamma_1\, CRS + \gamma_2\, CO + \omega_{12} CRS \cdot CO + \omega_{11}\, CRS^2 + \omega_{22} CO^2 + \zeta \qquad (11)$$

This model was tested in a sample of 400 employees from a public welfare agency in Germany. Customer-related social stressors were measured by 27 items, mostly taken from Dormann and Zapf (2004). Three correlated subscales, Customer Aggression (9 items), Customer Coordination Problems (15 items), and Disliked Customers (3 items), were used as indicators of the latent predictor, customer-related stressors. Customer orientation was measured by 16 items that were essentially unidimensional; three item parcels of 5 or 6 items each were used as indicators of customer orientation.[6] Burnout was measured with the Maslach Burnout Inventory (MBI; Maslach, Jackson, & Leiter, 1996) in its German version (Büssing & Perrar, 1992). For this example, the Depersonalization and Emotional Exhaustion subscales were used as indicators of the latent burnout variable.[7]

We illustrate the statistical analysis of the hypothesized nonlinear effects using the stand-alone QML program (Klein & Muthén, 2007), the M*plus* implementation of LMS (Klein & Moosbrugger, 2000, cf. Muthén & Muthén, 1998–2007), and the extended unconstrained PI approach (Kelava, 2009; Kelava & Brandt, 2009; Moosbrugger et al., 2009; cf. Marsh et al., 2004, 2006), using default settings for estimation.

The distribution analytic QML and LMS approaches require raw data for analysis, and can use data files in free format or space-delimited fixed format. By default, the latent variables constituting the interaction and quadratic terms are modeled as centered by the software, and intercepts (τ_x, τ_y) for measurement equations are estimated (Equations 12 and 13):

$$x = \tau_x + \Lambda_x \cdot \xi + \delta \qquad (12)$$

$$y = \tau_y + \Lambda_y \cdot \eta + \epsilon \qquad (13)$$

[6]In practice, item parcels have strengths and weaknesses; the effects of parceling and conditions for their proper use are described in Bandalos (2002) and Little, Cunningham, Shahar, and Widaman (2002).

[7]The third subscale, Personal Accomplishment, was not used as it showed relationship patterns distinct from the other subscales (cf. Leiter, 1993, for models of MBI subscales' relationships).

Annotated QML input syntax and LMS input syntax using M*plus* (cf. Muthén & Muthén, 1998–2007, Example 5.13) for our illustrative example are included in Appendices A and B available at http://www.augustin-kelava.de/pubs. Results of the LMS analysis using M*plus* were generally similar to those obtained by QML. To help readers understand the results, we also present annotated output obtained from QML and LMS in Appendices C and D.

QML and LMS provide unstandardized parameter estimates and standard error estimates. As in linear structural equation models, unstandardized estimates are based on the specified structural model (Equation 11) and the measurement models (Equations 12 and 13), thus are linked to the manifest variables' empirical variances and the latent variables' variances (irrespective of data transformations employed in estimating the likelihoods used by the QML and LMS algorithms, which remain internal to the program and are not seen by the user). Unstandardized estimates could be preferred for interpretation if manifest variables possess well-defined, empirically meaningful metrics and can thus serve as marker variables for latent variables linked to them. In typical applications, though, standardized estimates might be preferred for interpretation. QML also includes a properly standardized solution. In general, for PI approaches, standardization is complicated by product constructs (see Aiken & West, 1991, chap. 3), so that current SEM software produces improperly standardized coefficients. The distribution analytic QML and LMS approaches, by contrast, estimate interaction and curvilinear effects (i.e., effects involving product terms) without separate product constructs, making proper standardization very straightforward. Special procedures must be taken to produce properly standardized solutions using the PI approach (Wen, Marsh, & Hau, 2010).

We present the results for the complete nonlinear model with all standardized QML parameter estimates, including the interaction and both quadratic effects in Figure 2. The corresponding structural model for the relationships between the latent variables is given in Equation 14:

$$BO = -.007 + .646\,CRS - .284\,CO - .200\,CRS \cdot CO + .034\,CRS^2 - .084\,CO^2 + \zeta \qquad (14)$$

Unstandardized solutions for the LMS, QML, and extended PI approaches are presented in Table 6. A comparison of the standardized QML effect estimates for the structural model with the respective results obtained by the extended unconstrained PI approach is also given in Table 6, including both the improperly standardized solution supplied by standard SEM software for PI models, and a corrected standardization based on the standard deviations of the constructs customer-related stressors, customer orientation, and burnout (analogous to the procedure described in Wen et al., 2010, for the standardization of latent interaction coefficients).

In addition to parameter estimates and standard errors of individual parameter estimates, QML offers likelihood ratio tests for nested structural models. Nested model tests permit easy comparison of the complete nonlinear model represented in Equation 11 to a model with only linear effects, or to test any particular nonlinear (or linear) effect for significance against a nested model without this effect.

Suppose a researcher wishes to compare the complete nonlinear model represented in Equation 11 to a model without any nonlinear effects as represented in Equation 15:

$$BO = \alpha + \gamma_1\,CRS + \gamma_2\,CO + \zeta \qquad (15)$$

TABLE 6
Effect Estimates in the Structural Model for Predicting Burnout (Empirical Example)

	Linear		*Quadratic*		*Interaction*
Estimation	*CRS* γ_1	*CO* γ_2	*CRS*^{2} ω_{11}	*CO*^{2} ω_{22}	*CRS × CO* ω_{12}
QML (distribution analytic)					
Unstandardized	.558	−.256	.033	−.089	−.204
Standardized	.646	−.284	.034	−.084	−.200
LMS/M*plus* (distribution analytic)					
Unstandardized	.556	−.263	.032	−.099	−.220
Extended unconstrained (product indicators)					
Unstandardized[a]	.546	−.251	.053	−.063	−.117
Improperly standardized (software-supplied)[a]	.624	−.285	.091	−.082	−.125
Standardized according to Wen, Marsh, and Hau (2010)[b]	.624	−.285	.058	−.068	−.127

Note. Complete nonlinear model with predictors CRS = customer-related stressors; CO = customer orientation; QML = Quasi-Maximum Likelihood; LMS = Latent Moderated Structural Equations.
[a]LISREL estimation. [b]Described for interaction effects in Wen et al. (2010), analogous procedures applied to quadratic effects here.

Employing the likelihood ratio test, the researcher could compare the fit of the two models. In our example, the complete nonlinear model showed a superior fit to the linear model, as can be seen from the QML test results in Table 7 (for comparison, results from LMS/M*plus* and the extended unconstrained approach are also included). The researcher could also probe whether each term made a significant contribution to the model fit. Each successive nonlinear term can be omitted from the complete nonlinear model and the effect tested. In our example, QML showed that the small quadratic effect of customer-related stressors (ω_{11}) was not statistically significant. The other two nonlinear effects, the interaction effect (ω_{12}) and the quadratic effect of customer orientation (ω_{22}), were both statistically significant. LMS in M*plus* led to identical

TABLE 7
Likelihood Ratio Test Results for Nonlinear Effects on Burnout Using Different Approaches (Empirical Example)

		QML		*LMS/M*plus		*Extended Unconstrained*	
Model Comparison	df_Δ	χ^2_Δ	*p*	χ^2_Δ	*p*	χ^2_Δ	*p*
Complete nonlinear vs. linear	3	18.061	.001	12.653	.005	7.693	.053
Complete nonlinear vs. $\omega_{11} = 0$	1	0.671	.410	0.270	.603	1.365	.243
Complete nonlinear vs. $\omega_{12} = 0$	1	11.171	.001	4.994	.025	4.282	.039
Complete nonlinear vs. $\omega_{22} = 0$	1	4.243	.040	8.693	.003	1.385	.239

Note. The complete nonlinear model included quadratic effects ω_{11} (customer-related stressors^2), ω_{22} (customer orientation^2), and the interaction effect ω_{12} (Customer-Related Stressors × Customer Orientation) on burnout. For the linear model, all three effects were fixed to zero. QML = Quasi-Maximum Likelihood; LMS = Latent Moderated Structural Equations.

results. However, in contrast to the results of both LMS and QML, ω_{22} failed to achieve statistical significance using the extended unconstrained approach, likely reflecting its lower statistical power.

In cases in which the researcher desires to compare nonnested models, QML and LMS in M*plus* provide the AIC (Akaike, 1974) and the BIC (Schwarz, 1978). Researchers should choose the model with the smallest AIC and BIC as the optimal model according to these criteria. The values of the AIC and BIC calculated by the distribution analytic approaches will differ from those calculated by the PI approaches because additional product terms are not estimated in the distribution analytic approaches.

Complex nonlinear effects can be difficult to visualize; two-dimensional and three-dimensional representations can be helpful. Figure 4 shows a two-dimensional representation of the results of our example. This graph illustrates the quadratic relationship between customer orientation and burnout at different levels of customer-related stressors (at 1 *SD* below the mean, at the mean value, and at 1 *SD* above the mean). These figures can be plotted using any graphics package and overlaying lines corresponding to substituting values of -1, 0, and $+1$ into the standardized structural model equation (here, Equation 14).

At customer-related stressors values of $+1$ *SD*, the value of burnout showed continual but declining increases as the value of the latent variable customer orientation decreases. At customer-related stressors values of -1 *SD*, the value of burnout initially increases to a maximum for medium values of customer orientation, and then decreases in value.

Three-dimensional plots can also be helpful in interpreting nonlinear effects. In three-dimensional plots a surface corresponding to the structural model equation (Equation 14) is constructed, potentially using color graphics to highlight changes in the level of the dependent variable. Such plots can be constructed using specialized graphical software (e.g., R, Statgraphics) or increasingly in standard statistical packages (e.g., SAS, SPSS).

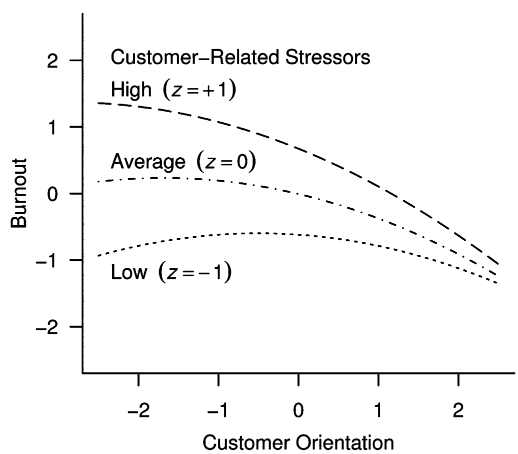

FIGURE 4 Effect of customer orientation on burnout, depicted for high, average, and low levels of customer-related stressors (latent variables, all standardized, Quasi-Maximum Likelihood estimates): The effect of customer orientation on burnout is curvilinear (quadratic effect) and additionally depends on the level of customer-related stressors (interactive effect).

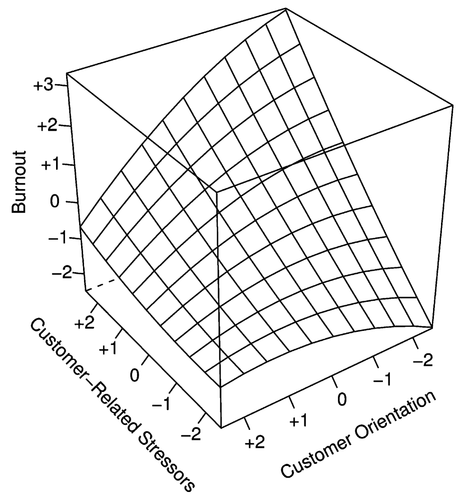

FIGURE 5 Linear effects of customer-related stressors and customer orientation, interaction effect Customer-Related Stressors × Customer Orientation, and quadratic effect of customer orientation on burnout (standardized Quasi-Maximum Likelihood estimates). Given low customer orientation, customer-related stressors predict increased burnout, whereas for high customer orientation, this relation disappears (interaction). Given high customer-related stressors, increasing customer orientation predicts a pronounced decrease in burnout (negative quadratic effect).

As can be seen in Figure 5, given low customer orientation, increasing customer-related stressors predict increased burnout, whereas for high customer orientation, customer-related stressors are essentially unrelated to burnout (interaction effect). Furthermore, given the presence of customer-related stressors, increasing customer orientation leads to a particularly pronounced reduction in burnout (negative quadratic effect).

DISCUSSION

In this article we provided a nontechnical introduction to two distribution analytic approaches, LMS and QML, for the estimation and testing of nonlinear effects in latent variable models. We provided a description of the key ideas underlying each approach and how they differ from PI approaches. LMS and QML do not require the specification of nonlinear constraints. This is a particular advantage in complex models for which the needed constraints and model specifications required in the PI approaches might not be available in the published literature. Nonlinear effects are also easily implemented in the M*plus* package (Muthén & Muthén, 1998–2007) and freestanding QML (Klein, 2007) software. In a small-scale simulation, LMS and QML consistently showed acceptable Type I error rates and a modest advantage (2%–4%) in statistical power compared with the extended unconstrained approach (Kelava, 2009; Kelava & Brandt, 2009; Moosbrugger et al., 2009), which is based on the most commonly used of the PI approaches (Marsh et al., 2004, 2006). This advantage could be important in practice given the small effect sizes (1%–3% of variation accounted for) that have typified interaction effects between measured variables (Champoux & Peters, 1987; Chaplin, 1991, 2007; Donovan & Radosevich, 1998). LMS and QML detected an additional hypothesized nonlinear (quadratic) effect of interest in an actual data set on work stress and job burnout, perhaps due to the greater statistical power of the distribution analytic approaches.

The results of the simulation study were obtained using the default settings provided by the software packages. LMS and the extended unconstrained PI approach achieved 100% convergence; QML convergence rates were only slightly lower, exceeding 96% in all conditions. These results were obtained with a sample size of 400 and measured x variables that were normally distributed. Analyses conducted with smaller sample sizes or with nonnormal x variables would be more likely to have problems in estimation, potentially leading to lower convergence rates.

Users wishing to consider LMS or QML might wonder which approach to choose. The answer to this question depends on both practical considerations and the specific research questions of the user. On a practical level, LMS is currently implemented in commercial M*plus* software (Muthén & Muthén, 1998–2007), a general package for latent variable models. QML is currently implemented as a noncommercial stand-alone package, as of this writing at no cost. Users need to consider the trade-off between the cost versus the availability of a technical support infrastructure associated with commercial and noncommercial programs. Users familiar with equation-based programs (e.g., M*plus*, EQS) will find their programming experience readily generalizes to LMS. Users familiar with matrix-based programs (e.g., LISREL) will find that their programming experience readily generalizes to QML. Neither LMS nor QML currently offer point-and-click analyses.

In terms of research questions, the types of models that can be addressed and the content of the output that is produced by LMS and QML largely overlap. However, there are important exceptions that exist in the current implementations of the two programs.

1. LMS allows researchers to build more complicated SEM models involving multiple latent endogenous variables, whereas QML is currently limited to models with only one latent outcome variable. As one example, a latent variable moderated mediation model in which the path from the mediator to the outcome is moderated by another latent variable can be specified in the M*plus* package, but not using QML software.
2. QML might have an advantage in computational speed in models with several nonlinear terms because of the heavier computational demand in LMS.
3. QML will compute a properly standardized solution if requested by the user, whereas the user must compute the proper standardized solution by hand in LMS.

In conclusion, for applied researchers, employing distribution analytic approaches leads to more powerful tests that address the specific hypothesized substantive effects of interest in their data. With the advent of user-friendly SEM software implementing these approaches, employing LMS and QML for empirical analyses becomes feasible. Our simulation study and empirical example illustrated the advantages of these new approaches in terms of ease of specification and by providing slightly more powerful tests of nonlinear effects in latent variable models. The advantages of LMS and QML over the extended PI approaches are theoretically expected to increase as the correlation between the latent exogenous variables and the effect size of the latent nonlinear effects increase. Readers should note that these models demand larger sample sizes than those often seen in practice. For this simulation study and empirical illustration involving a complex nonlinear model with two small quadratic effects and small interaction, a sample size of 400 was adequate for estimation. Even larger sample sizes will be needed to achieve adequate statistical power given high multicollinearity. For a model involving only a single interaction or quadratic effect, a sample size of 200 might be sufficient for estimation. A fuller understanding of the performance of each of the approaches under conditions in which there are severe levels of excess kurtosis, skewness, or both in the distributions of the observed variables will require further study. Our hope is that this article provides applied researchers with the information necessary to use distribution analytic approaches when they wish to include interaction, quadratic, or both types of nonlinear terms in latent variable models.

ACKNOWLEDGMENTS

We would like to thank Associate Editor Deborah Bandalos and the reviewers for providing comments that led to substantial improvements of the article. Stephen G. West was supported by a Forschungspreis (research prize) from the Alexander von Humboldt Foundation.

Appendices A through D including the computer code and excerpts from the output for the LMS and QML analyses are available at http://www.augustin-kelava.de/pubs.

REFERENCES

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions.* Newbury Park, CA: Sage.

Ajzen, I. (1987). Attitudes, traits, and actions: Dispositional prediction of behavior in personality and social psychology. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 20, pp. 1–63). New York, NY: Academic Press.

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

Algina, J., & Moulder, B. C. (2001). A note on estimating the Jöreskog–Yang model for latent variable interaction using LISREL 8.3. *Structural Equation Modeling, 8,* 40–52.

Aroian, L. A. (1944). The probability function of the product of two normally distributed variables. *The Annals of Mathematical Statistics, 18*, 265–271.

Bandalos, D. L. (2002). The effects of item parceling on goodness-of-fit and parameter estimate bias in structural equation modeling. *Structural Equation Modeling, 9*, 78–102.

Bohrnstedt, G., & Marwell, G. (1978). The reliability of the products of two random variables. *Sociological Methodology, 9*, 254–273.

Brandt, H. (2009). *Der verteilungskorrigierte Bootstrap zur Schätzung von Konfidenzintervallen nichtlinearer Strukturgleichungsmodelle* [The distribution-corrected bootstrap for the estimation of confidence intervals in nonlinear structural equation models]. Unpublished master's thesis, Goethe University, Frankfurt, Germany.

Büssing, A., & Perrar, K.-M. (1992). Die Messung von Burnout. Untersuchung einer deutschen Fassung des Maslach Burnout Inventory (MBI-D) [The measurement of burnout: Studies on a German version of the Maslach Burnout Inventory (MBI-D)]. *Diagnostica, 38*, 328–353.

Champoux, J. E., & Peters, W. S. (1987). Form, effect size, and power in moderated multiple regression. *Journal of Occupational Psychology, 17*, 585–605.

Chaplin, W. (1991). The next generation of moderator research in personality psychology. *Journal of Personality, 59*, 143–178.

Chaplin, W. (2007). Moderator and mediator models in personality research: A basic introduction. In R. W. Robins, C. Fraley, & R. F. Krueger (Eds.), *Handbook of research methods in personality psychology* (pp. 602–632). New York, NY: Guilford.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cronbach, L. J., & Snow, R. E. (1977). *Aptitudes and instructional methods: A handbook for research on interactions.* New York, NY: Irvington.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B, 39,* 1–38.

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

Donovan, J. J., & Radosevich, D. J. (1998). The moderating role of goal commitment on the goal difficulty performance relationship: A meta-analytic review and critical reanalysis. *Journal of Applied Psychology, 83,* 308–315.

Dormann, C., & Zapf, D. (2004). Customer-related social stressors and burnout. *Journal of Occupational Health Psychology, 9*, 61–82.

Freund, R., & Hoppe, R. H. W. (2007). *Stoer/Bulirsch: Numerische Mathematik 1* [Numerical mathematics 1]. Heidelberg, Germany: Springer.

Ganzach, Y. (1997). Misleading interaction and curvilinear terms. *Psychological Methods, 2*, 235–247.

Jöreskog, K. G., & Sörbom, D. (1996). *LISREL 8: User's reference guide*. Lincolnwood, IL: Scientific Software International.

Jöreskog, K. G., & Sörbom, D. (1999). *PRELIS 2: User's reference guide.* Lincolnwood, IL: Scientific Software International.

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. 57–87). Mahwah, NJ: Lawrence Erlbaum Associates.

Karasek, R. A. (1979). Job demands, job decision latitude, and mental strain: Implication for job redesign. *Administrative Science Quarterly, 24*, 285–308.

Kelava, A. (2009). *Multikollinearität in nicht-linearen latenten Strukturgleichungsmodellen* [Multicollinearity in nonlinear structural equation models]. Unpublished doctoral dissertation, Goethe University, Frankfurt, Germany. Retrieved from http://publikationen.ub.uni-frankfurt.de/volltexte/2009/6336/

Kelava, A., & Brandt, H. (2009). Estimation of nonlinear latent structural equation models using the extended unconstrained approach. *Review of Psychology, 16*, 123–131.

Kelava, A., Moosbrugger, H., Dimitruk, P., & Schermelleh-Engel, K. (2008). Multicollinearity and missing constraints: A comparison of three approaches for the analysis of latent nonlinear effects. *Methodology, 4*, 51–66.

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

Klein, A. G. (2007). *QuasiML 3.10: Quick reference manual.* Unpublished manuscript, University of Western Ontario, London, ON, Canada.

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

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*, 647–674.

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

Lee, S. Y., Song, X. Y., & Tang, N. S. (2007). Bayesian methods for analyzing structural equation models with covariates, interaction, and quadratic latent variables. *Structural Equation Modeling, 14*, 404–434.

Leiter, M. P. (1993). Burnout as a developmental process: Consideration of models. In W. B. Schaufeli, C. Maslach, & T. Marek (Eds.), *Professional burnout: Recent developments in theory and research* (pp. 237–250). Washington, DC: Taylor & Francis.

Little, T. D., Cunningham, W. A., Shahar, G., & Widaman, K. F. (2002). To parcel or not to parcel: Exploring the question, weighing the merits. *Structural Equation Modeling, 9*, 151–173.

Lusch, R. F., & Brown, J. R. (1996). Interdependency, contracting, and relational behavior in marketing channels. *Journal of Marketing, 60*, 19–38.

Ma, Y. (2010). Comparison of statistical power for testing interactive versus quadratic effects. Unpublished doctoral dissertation, Arizona State University, Tempe, Arizona.

MacCallum, R. (1986). Specification searches in covariance structure modeling. *Psychological Bulletin, 100*, 107–120.

MacCallum, R. C., & Austin, J. T. (2000). Applications of structural equation modeling in psychological research. *Annual Review of Psychology, 51*, 201–226.

MacCallum, R. C., & Mar, C. M. (1995). Distinguishing between moderator and quadratic effects in multiple regression. *Psychological Bulletin, 118*, 405–421.

Marsh, H. W., Wen, Z., & Hau, K.-T. (2004). Structural equation models of latent interactions: Evaluation of alternative estimation strategies and indicator construction. *Psychological Methods, 9*, 275–300.

Marsh, H. W., Wen, Z., & Hau, K.-T. (2006). Structural equation models of latent interaction and quadratic effects. In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: A second course* (pp. 225–265). Greenwich, CT: Information Age.

Maslach, C., Jackson, S. E., & Leiter, M. P. (1996). *The Maslach Burnout Inventory manual*. Palo Alto, CA: Consulting Psychologists Press.

McDowell, M. A., Fryar, C. D., Ogden, C. L., & Flegal, K. M. (2008). Anthropometric reference data for children and adults: United States, 2003–2006. *National Health Statistics Report, 10*, 1–45. Retrieved from http://www.cdc.gov/nchs/data/nhsr/nhsr010.pdf

Moosbrugger, H., Schermelleh-Engel, K., Kelava, A., & Klein, A. G. (2009). Testing multiple nonlinear effects in structural equation modeling: A comparison of alternative estimation approaches. In T. Teo & M. Khine (Eds.), *Structural equation modeling in educational research: Concepts and applications* (pp. 103–136). Rotterdam, The Netherlands: Sense Publishers.

Moosbrugger, H., Schermelleh-Engel, K., & Klein, A. G. (1997). Methodological problems of estimating latent interaction effects. *Methods of Psychological Research Online, 2*, 95–111.

Moulder, B. C., & Algina, J. (2002). Comparison of methods for estimating and testing latent variable interactions. *Structural Equation Modeling, 9*, 1–19.

Muthén, L. K., & Muthén, B. O. (1998–2007). *M*plus *user's guide* (5th ed.). Los Angeles, CA: Muthén & Muthén.

Paxton, P., Curran, P. J., Bollen, K. A., Kirby, J., & Chen, F. (2001). Monte Carlo experiments: Design and implementation. *Structural Equation Modeling, 8*, 287–312.

Satorra, A., & Bentler, P. M. (1994). Corrections to test statistics and standard errors in covariance structure analysis. In A. von Eye & C. C. Clogg (Eds.), *Latent variables analysis: Applications for developmental research* (pp. 399–419). Thousand Oaks, CA: Sage.

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

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

Snyder, M., & Tanke, E. D. (1976). Behavior and attitude: Some people are more consistent than others. *Journal of Personality, 44*, 501–517.

Wall, M. M., & Amemiya, Y. (2001). Generalized appended product indicator procedure for nonlinear structural equation analysis. *Journal of Educational and Behavioral Statistics, 26*, 1–29.

Wen, Z., Marsh, H. W., & Hau, K.-T. (2010). Structural equation models of latent interactions: An appropriate standardized solution and its scale-free properties. *Structural Equation Modeling, 17*, 1–22.