A General Approach to Modeling Latent Variable Interactions and Nonlinear Effects

Brian T. Keller, Ph.D.

The University of Missouri

This manuscript has been submitted for peer review.

Keller (2024). *A General Approach to Modeling Latent Variable Interactions and Nonlinear Effects*. Manuscript submitted for publication.

1 Introduction

Interaction or moderation effects are ubiquitous in the behavioral sciences. Within psychology, these interactions stand as a captivating phenomena because they illuminate the intricacies inherent in human behavior, shedding light on how a relationship is influenced by additional factors. Without surprise, interactions play a pivotal role in regression modeling, and substantial methodological work has been devoted to extending moderation effects to latent variable models.

Earlier methods for estimating latent variable interactions originated from the pioneering work by Kenny and Judd (1984). These so-called 'product indicator' methods involve a step-by-step process: first, the methods generate manifest variable products by multiplying specific indicators from different latent variables. These products load onto a third latent variable, effectively representing the interaction between the two latent variables. While advancements have been made to the product indicator approach, their formulation is often complex and challenging. For instance, these methods frequently necessitate intricate specifications of the measurement error covariance structure, which can be prone to misspecification as the model complexity increases. Another drawback of product indicator approaches lies in their lack of parsimony, often demanding additional loadings that necessitate constraint for estimation. Moreover, these methods typically assume continuous data, limiting their applicability to binary, ordinal, and categorical data sets. Furthermore, despite their ability to incorporate missing data in the indicators, it is unclear whether they can provide unbiased estimates because the products are treated as just another variable (Seaman, Bartlett, & White, 2012). I suggest interested readers see Cortina, Markell-Goldstein, Green, and Chang (2021) and Kelava and Brandt (2023) for a more detailed discussion of product indicator methods.

In contrast, alternative analytic methods have emerged to explicitly tackle the distributional assumptions arising from incorporating the product of latent variables. These approaches have come in both frequentist (Klein & Moosbrugger, 2000; Lee & Zhu, 2002) and Bayesian (Arminger & Muthén, 1998; Zhu & Lee, 1999) flavors. The specific modeling of the distributions can differ, but one of the most popular approaches is the general interaction model. The general interaction model extends structural equation modeling (SEM) to allow for an interaction between any two latent variables, and several estimators have been proposed to estimate the model. Two popular estimators are the Latent Moderated Structural Equations (LMS; Klein & Moosbrugger, 2000) and Quasi-Maximum Likelihood (QML; Klein & Muthén, 2007). The general interaction model was further extended to two-level models by Muthén and Asparouhov (2009), and more recently extended by Asparouhov and Muthén (2021b) to accommodate multilevel structural equation models with latent interactions in a Bayesian framework.

Despite these advances, the general interaction methods still have limitations and sometimes require modeling "tricks" to incorporate certain specifications. For example, the general interaction model can only accommodate incomplete manifest-by-latent and incomplete manifest-by-manifest interactions by creating proxy latent variables. These proxy variables are identified by fixing a loading to one and setting the manifest variable's variance near zero. While this can provide unbiased results, it unnecessarily complicates the specification and makes the estimation more computationally demanding. Moreover, similar proxy variable approaches are required for specification for products greater than two variables (e.g., three-way interactions; Asparouhov & Muthén, 2021b) and to perform latent centering in multilevel contexts (Preacher, Zhang, & Zyphur, 2016). The additional difficulties of using the proxy variable approach may pose a significant barrier to entry for some, slowing the method's adaptation by applied researchers.

Bayesian approaches to latent interactions are a third thread of methodological work. Work by Lee and colleagues (Lee & Zhu, 2000; Lee, Song, & Tang, 2007; Lee, 2007) provides discussions of general nonlinear structural equation modeling frameworks focusing on extending LISREL model (Jöreskog & Sörbom, 1996) to allow for arbitrary functions (e.g., interactions, quadratic effects). However, this approach was limited to nonlinearity in the structural model only. Such nonlinearity has become important with moderated nonlinear factor analysis (Bauer, 2017) and investigating latent by manifest interactions in the measurement model. Moreover, the implementations of Lee and colleagues' approach require the usage of general Markov chain Monte Carlo packages such as JAGS (Plummer, 2017) or Stan (Stan Development Team, 2023). While these general packages offer great flexibility, they are computationally slower and require more programming knowledge to implement a model. This programming ability is a huge barrier for many researchers and prevents the widespread adoption of existing Bayesian approaches.

The purpose of this paper is to extend and generalize Bayesian estimation for latent variable interactions, and I discuss the application of a general framework to estimate these effects. A major issue with interaction and nonlinear effects is that they require distributional assumptions incompatible with a multivariate normal distribution. The factored regression framework avoids this problem by specifying multivariate distributions as a collection of simpler (often univariate) distributions that allow the specification of complex data structures required by interactive effects. In the manifest missing data literature, the use of factorization strategies was first introduced by Ibrahim and colleagues (Ibrahim, 1990; Ibrahim, Chen, & Lipsitz, 2002, 1999; Lipsitz & Ibrahim, 1996) and have been applied to missing data problems with Bayesian, multiple imputation, and maximum likelihood methods. More recently, factorization has been used to handle incomplete interactive and polynomial effects (Du, Enders, Keller, Bradbury, & Karney, 2022; Enders, Du, & Keller, 2020; Erler et al., 2016; Zhang & Wang, 2017; Goldstein, Carpenter, & Browne, 2014; Lüdtke, Robitzsch, & West, 2020b; Keller & Enders, 2023). Several software packages offer factored specifications (Keller & Enders, 2022; Robitzsch & Lüdtke, 2021; Bartlett, Keogh, & Bonneville, 2021; Erler, Rizopoulos, & Lesaffre, 2021), although they differ in user-friendliness and the breadth of models they can fit.

By leveraging the previous work in the missing data literature, the factored regression approach to interactive and nonlinear latent variable effects provides important methodological contributions. This approach allows a flexible framework to specify complex structural equation models with interactive and nonlinear effects for latent variables, including three-way or higher interactions and quadratic effects. These nonlinear effects can be specified with any variable, including a mixture of manifest and latent variables. Furthermore, the nonlinear effects are not strictly limited to the structural models, allowing the specification of moderation effects on the indicators in a measurement model (e.g., squared loadings, moderated loadings). The factored regression approach accommodates a diverse set of manifest variable types (e.g., binary, ordinal, multicategorical, continuous normal, skewed continuous, count indicators), which can serve as indicators of latent variables or as manifest predictors with nonlinear relationships. In addition, it easily incorporates missing data on any manifest variable in the model. As an example, a researcher can use factored regression to specify an incomplete multicategorical variable moderating the influence of a latent variable. Finally, the factored regression approach to latent interactions readily extends to multilevel models.

I provide an implementation of the factored regression modeling framework in Blimp (Keller & Enders, 2022), a free standalone software with syntax similarities to many modern-day SEM software packages. From the researcher's perspective, applying the factored regression approach to latent interactions is no different from standard moderated regression. Researchers can use familiar pick-a-point approaches or a Johnson-Neyman plot (P. O. Johnson & Neyman, 1936) to probe the interaction, as illustrated in a substantive data example. In the software syntax, the latent variables function just like other variables; there are no special programming tricks or opaque syntax keywords to specify models. Furthermore, because the latent variable scores are considered missing data to be imputed, I demonstrate that familiar regression diagnostics and graphical methods are available to investigate the functional form of the latent regressions. Specifying these models in Blimp is substantially easier than other software packages, making these models accessible to a much broader audience.

The structure of this article is as follows. First, I begin by discussing the factored regression approach to modeling and how to use it to incorporate measurement models. Second, I illustrate the extension to incorporate latent interactions—including a latent moderation model and a latent by manifest interaction. Third, I provide an overview of the estimation procedure used to estimate the models in a Bayesian paradigm. Fourth, I use a substantive example to demonstrate fitting a latent moderation model.

2 Factored Regression Modeling

A key feature of the factored regression specification is that it avoids working with off-the-shelf multivariate distributions like the multivariate normal distribution. Instead, the factored regression framework aims to chain several regression models (i.e., conditional densities) to characterize the hypothesized joint distribution. Throughout the paper, I represent models via a so-called "functional" notation. To illustrate, suppose I am modeling the joint distribution of three variables, X, Y, and Z. The functional notation representation of the joint distribution is as follows.

$$f\left(X, Y, Z \right) \tag{1}$$

where $f(\ldots)$ represents a generic probability distribution for X, Y, and Z. This unspecified distribution could have any form. For example, I may assume that the three variables follow a multivariate normal distribution with a mean vector and covariance matrix

$$f(X, Y, Z) = \mathcal{N}_3(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \tag{2}$$

where $\boldsymbol{\mu}$ is a 3×1 mean vector and $\boldsymbol{\Sigma}$ is a 3×3 unstructured covariance matrix for a total of nine parameters. Traditionally, SEM approaches work from a multivariate distribution like Equation (2). Ultimately, this is a limiting factor because interactions and nonlinear terms induce characteristics often at odds with the multivariate normal distribution (Liu, Gelman, Hill, Su, & Kropko, 2014; Seaman et al., 2012).

Alternatively, factored regression models the joint density as a sequential product of conditional and marginal densities. By applying the chain rule for probability, I can factor the joint density into a product of univariate distributions. Reconsider the joint density of the three variables, factored regression framework models this joint density as a sequential product of conditional and marginal densities.

$$f(X, Y, Z) = f\big(Y \mid X, Z\big) \times f(X, Z) \tag{3}$$

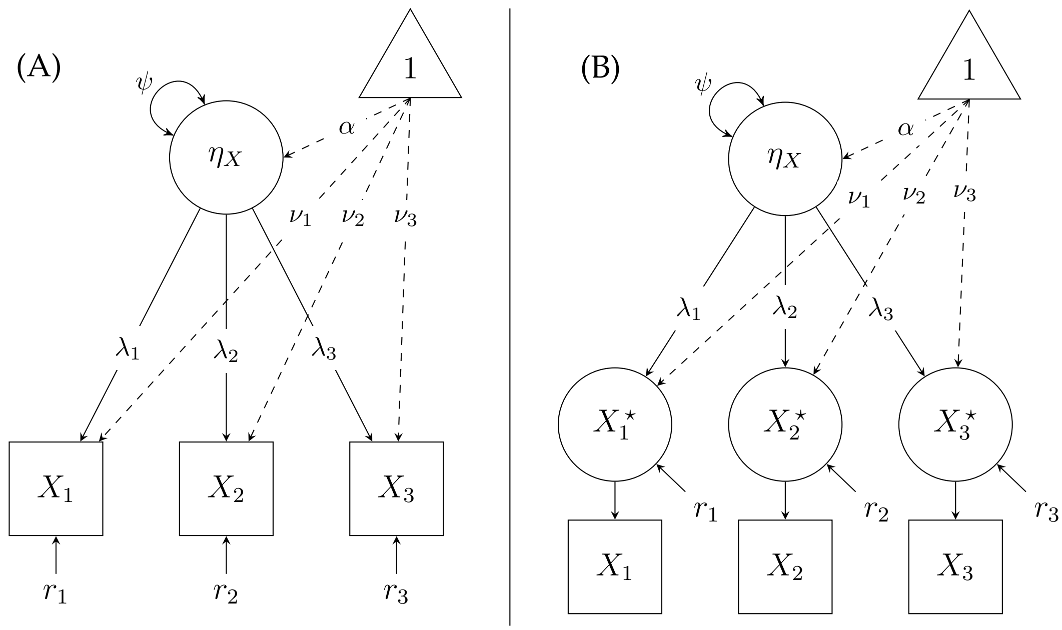
Note, I refer to this as a partially factored specification because the bivariate distribution $f(X, Z)$ can further be factored into two univariate densities.

$$f(X, Y, Z) = f\big(Y \mid X, Z\big) \times f\big(Z \mid X\big) \times f(X) \tag{4}$$

This factorization is sometimes called the "sequential" specification in the literature (Lüdtke et al., 2020b; Lüdtke, Robitzsch, & West, 2020a). I will refer to this as the fully factored specification because I have wholly decomposed the joint density into a product of univariate models. In contrast to a traditional SEM approach, the fully factored specification models

Figure 1

Path Diagrams for One Factor Models



Note: Panel (A) is the path diagram for a single factor model with continuous indicators. Panel (B) is the path diagram for a single factor model with binary indicators. The double-headed curved arrow on η represents the variance of the latent factor. The residuals, r_X, r_Y, and r_Z, also have estimated variances associated with them in (A) but are fixed to 1 in (B) for identification.

the multivariate distribution as a collection of univariate regression models, each potentially with its unique distribution features. This allows for greater flexibility to accurately represent special features induced by nonlinear models.

2.1 *Measurement Models as Factored Regressions*

To illustrate how factored regression applies to a measurement model, consider the model depicted in Figure 1A, where three variables (X_1, X_2, and X_3) load onto a single factor (η_X). To express this model in terms of a factored regression, I fully factor out the joint distribution of all variables into a product of four univariate distributions.

$$f\left(X_1, X_2, X_3, \eta_X\right) = f\left(X_1 \mid X_2, X_3, \eta_X\right) \times f\left(X_2 \mid X_3, \eta_X\right) \times f\left(X_3 \mid \eta_X\right) \times f\left(\eta_X\right) \quad (5)$$

$$= f\left(X_1 \mid \eta_X\right) \times f\left(X_2 \mid \eta_X\right) \times f\left(X_3 \mid \eta_X\right) \times f\left(\eta_X\right) \quad (6)$$

Equation (5) characterizes the full factorization of the measurement model. Based on the model in Figure 1A, I simplify the factorization by removing indicators from the right side of all vertical bars, giving Equation (6). This simplification is possible because each indicator is conditionally independent of the other indicator given the latent factor η_X. Note that the joint distribution can be factorized in different orders. For example, the first term to the right of the equals sign could be $f\left(X_3 \mid X_2, X_1, \eta_X\right)$.

The factorization in Equation (6) makes sense because the latent variable acts as an exogenous predictor, leaving it as the last variable in the chain and including its marginal distribution as the final distribution. The four distributions in Equation (6) map one-to-one onto a specific regression model

$$\begin{aligned} f\left(X_1 \mid \eta_X\right) \rightarrow &\quad x_{1i} = \nu_1 + \lambda_1 \eta_{Xi} + r_{1i} \\ f\left(X_2 \mid \eta_X\right) \rightarrow &\quad x_{2i} = \nu_2 + \lambda_2 \eta_{Xi} + r_{2i} \\ f\left(X_3 \mid \eta_X\right) \rightarrow &\quad x_{3i} = \nu_3 + \lambda_3 \eta_{Xi} + r_{3i} \\ f\left(\eta_X\right) \rightarrow &\quad \eta_{Xi} = \alpha + \zeta_i \end{aligned} \quad (7)$$

These equations represent the common factor model but are expressed in univariate regression models and distributions instead of the usual matrix specification and conditional multivariate normal invoked by SEM . For now, I have assumed that each indicator is continuous. However, as discussed later, factored regression readily accommodates non-normal continuous, binary, ordinal, categorical, or count indicators by specifying the appropriate regression model for the data type. Finally, just like a standard confirmatory factor analysis, I must identify the model via constraints (e.g., $\lambda_1 = 1$, $\alpha = 0$) or narrow priors in a Bayesian context (Muthén & Asparouhov, 2012).

In addition to continuously normally distributed indicators, factored regression readily extends to non-normal indicators and exogenous predictors. For simplicity, suppose the indicators for η_X are all binary indicators. I use a probit model to specify the regression of each item onto η_X (Albert & Chib, 1993; V. E. Johnson & Albert, 2006; Agresti, 2012). The probit model conceptualizes the discrete responses (X_1) as originating from a latent normally distributed propensity (i.e., $X_1^\star$ for X_1) with a threshold dividing the latent variable into the discrete observations via a link function. I have depicted the path model in Figure 1B, with circles denoting the latent response variables. Notice that each indicator now has a latent propensity that is predicted by η_X; therefore, each indicator must incorporate this unobserved latent propensity within the factorization.

For example, reconsider the factorization for X_1 given η_X in Equation (21). This density must now incorporate the latent propensity and the manifest indicator.

$$f\left(X_1 \mid \eta_X\right) = f\left(X_1 \mid X_1^\star\right) \times f\left(X_1^\star \mid \eta_X\right) \tag{8}$$

In words, the distribution of X_1 regressed on η_X is factored into two separate models, a conditional density of X_1 given the latent propensity $X_1^\star$ and a conditional density of $X_1^\star$ conditional on η_X. These two densities result in the following models.

$$\begin{aligned} f\left(X_1 \mid X_1^\star\right) \rightarrow &\quad x_{1i} = \mathcal{I}\left(x_{1i}^\star > 0\right) \\ f\left(X_1^\star \mid \eta_X\right) \rightarrow &\quad x_{1i}^\star = \nu_1 + \lambda_1 \eta_{Xi} + r_{1i} \end{aligned} \tag{9}$$

The first line of Equation (9) is an additional "link" function that relates the observed binary variable to the latent propensity. In words, $\mathcal{I}\left(x_{1i}^\star > 0\right)$ returns one if $x_{1i}^\star > 0$ is true; otherwise, it returns zero. In Figure 1B, the link function is the arrow connecting the $X_1^\star$ to X_1. The second line of the equation characterizes the endogenous normally distributed latent propensity with a mean that relates to the proportion of ones to zeros. Returning to Figure 1B, this is the measurement model where the latent factor influences the latent response. To identify this model, the variance of this propensity is constrained to one and the threshold parameter is fixed at zero. Ordinal indicators share the same model and factorization but simply require additional threshold parameters (e.g., a five-point indicator requires four thresholds with the first threshold fixed for identification). Finally, I have opted to model $f\left(X_1 \mid \eta_X \right)$ via the probit specification but nothing precludes using a logistic regression specification.

3 Factored Regression with Latent Interactions

The factored regression specification readily accommodates interactions and other types of nonlinearities involving latent variables. As an example, consider the latent moderation model depicted in Figure 2. The latent factor, η_Y, is regressed on η_X, η_Z, and their product (represented by the conjoined arrow in the path diagram). Later in the paper I show an application mapping onto this exact example, where η_X and η_Z are latent variables measuring conscientiousness and organizational constraints, and η_Y is a latent measure of counterproductive work behavior.

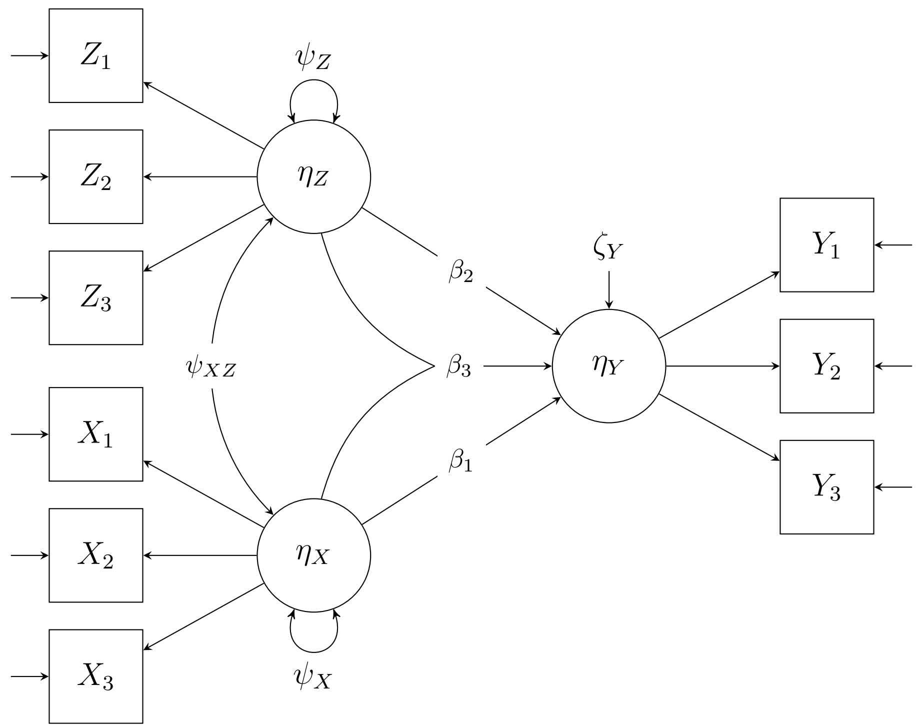
To obtain the factorization for the model in Figure 2, I first separate the joint density into the manifest variables conditional on the unobserved latent factors.

$$\begin{aligned} & f\left(X_1, X_2, X_3, Z_1, Z_2, Z_3, Y_1, Y_2, Y_3, \eta_X, \eta_Z, \eta_Y \right) = \\ & \quad f\left(X_1, X_2, X_3, Z_1, Z_2, Z_3, Y_1, Y_2, Y_3, \mid \eta_X, \eta_Z, \eta_Y \right) \times f\left(\eta_X, \eta_Z, \eta_Y \right) \end{aligned} \tag{10}$$

In words, Equation (10) says that the multivariate distribution of the manifest and latent variables can be expressed as the product of a conditional distribution that links the indicators to their respective latent variables and a multivariate distribution for the latent variables. This factorization directly maps onto the Figure 2 because the manifest indicators are all conditionally independent given their respective latent variables. Said differently, the

Figure 2

Path Diagram for Latent × Latent Interaction Model



Note: The mean structure is not illustrated in the diagram. The conjoined arrow for the β_3 path represents the product between η_X and η_Z.

indicators are all predicted by latent variables only. Notice that this factorization maps directly onto how SEM models are conceptualized, with the first distribution corresponding to a *measurement density* and the second distribution corresponding to a *structural density*.

Because the model implies each indicator is conditionally independent given their respective latent variable and there are no cross-loadings, I factor and simplify the measurement density into three separate distributions.

$$\begin{aligned} &f\left(X_1, X_2, X_3, Z_1, Z_2, Z_3, Y_1, Y_2, Y_3, \mid \eta_X, \eta_Z, \eta_Y \right) = \\ &\quad f\left(X_1, X_2, X_3 \mid \eta_X \right) \times f\left(Z_1, Z_2, Z_3 \mid \eta_Z \right) \times f\left(Y_1, Y_2, Y_3 \mid \eta_Y \right) \end{aligned} \tag{11}$$

As before, the above factorization maps directly onto the path diagram in Figure 2. Notice how each distribution maps directly onto the setup of the single factor model. That is, three indicators are conditional on their respective latent variable. I further factor each of these densities just like I did in Equation (6), but with each respective grouping of variables—e.g., $f\left(X_1, X_2, X_3 \mid \eta_X\right) = f\left(X_1 \mid \eta_X\right) \times f\left(X_2 \mid \eta_X\right) \times f\left(X_3 \mid \eta_X\right)$. Moreover, each distribution is equivalent to the first three lines in Equation (7). Note that the above specification readily extends to more than three indicators, and as discussed in the previous section, these indicators are not required to be continuous. As previously mentioned, the Blimp software also accommodates binary, ordinal, multicategorical, count, and skewed continuous indicators.

Turning to the structural density, I apply the partial factorization in the same manner as in Equation (3), resulting in a conditional and joint distribution for η_X and η_Z.

$$f\left(\eta_X, \eta_Z, \eta_Y\right) = f\left(\eta_Y \mid \eta_X, \eta_Z\right) \times f\left(\eta_X, \eta_Z\right) \tag{12}$$

In words, Equation (12) says that the multivariate distribution of the latent variables can be expressed as the product of a conditional distribution for the endogenous latent variable, η_Y, and a bivariate distribution for the exogenous latent variables, η_X and η_Z. Similar to before, this factorization stems from the path diagram. In Figure 2, η_Y is predicted by the two latent variables and their product. Therefore, I model $f\left(\eta_Y \mid \eta_X, \eta_Z\right)$ via a moderated linear regression model.

$$f\left(\eta_Y \mid \eta_X, \eta_Z\right) \rightarrow \quad \eta_{Yi} = \alpha_Y + \beta_1 \eta_{Xi} + \beta_2 \eta_{Zi} + \beta_3\left(\eta_{Xi} \times \eta_{Zi}\right) + \zeta_{Yi} \tag{13}$$

Notably, the regression in Equation (13) includes the product of the two latent variables. This product is not a unique variable but a function of η_X and η_Z. Finally, the bivariate distribution for η_X and η_Z could be further factored into a conditional and marginal distribution, as I do in the next section.

Although I focus on interaction effects in this paper, it is important to note that the factorization procedure readily accommodates other types of nonlinear associations. For example, suppose that the moderated regression from Equation (13) also included a quadratic effect of η_X. The regression model would be as follows.

$$f\left(\eta_Y \mid \eta_X, \eta_Z\right) \rightarrow \quad \eta_{Yi} = \alpha_Y + \beta_1 \eta_{Xi} + \beta_2 \eta_{Xi}^2 + \beta_3 \eta_{Zi} + \beta_4 \left(\eta_{Xi} \times \eta_{Zi}\right) + \zeta_{Yi}$$

To reiterate, the squared term is not a unique variable but a function of η_X. Similarly, a latent factor could exert a quadratic effect on any of its indicators. By allowing any latent variable to be expressed as a nonlinear function, the factored regression specification affords much greater flexibility to specify complex functions while accounting for the measurement models.

3.1 Latent $\times$ Manifest Interaction with Factored Regression

In addition to accommodating latent by latent interactions, factored regressions readily extend to include manifest predictors (Keller, 2022). Reconsider the latent moderation example but with a single manifest predictor, denoted as X. Figure 3 depicts the path diagram for this model. Note I have opted to specify the relationship between X and η_Z via directed path (as opposed to correlation) because X is manifest, so this specification would be analogous to a multiple indicator multiple cause (MIMIC) model. [1]

To illustrate, I first provide the factorization for the model in Figure 3. I apply the chain rule and simplify the densities to obtain the factorization of the joint density.

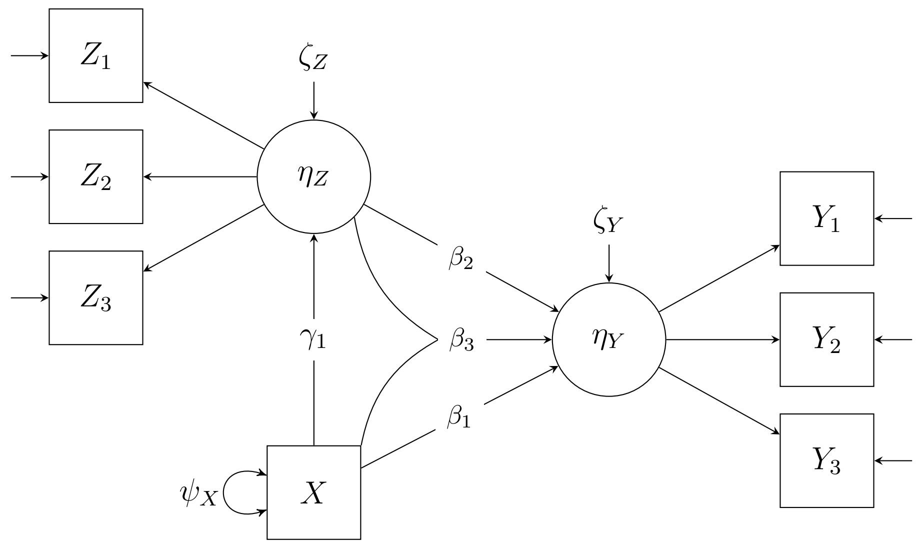
$$\begin{aligned} f\left(X, Z_1, Z_2, Z_3, Y_1, Y_2, Y_3, \eta_Z, \eta_Y\right) = f\left(\eta_Y \mid \eta_Z, X\right) \times f\left(\eta_Z \mid X\right) \times f\left(X\right) \times \\ f\left(Z_1, Z_2, Z_3 \mid \eta_Z\right) \times f\left(Y_1, Y_2, Y_3 \mid \eta_Y\right) \end{aligned} \tag{14}$$

The above factorization maps onto the previous section with only minor modifications.

[1]In Blimp, the partial factorization from Equation (12) is also available for binary, ordinal, and normal manifest variables; however, linking exogenous variables with directed pathways accommodates a broader range of predictor metrics

Figure 3

Path Diagram for Latent × Manifest Interaction Model



Note: The mean structure is not illustrated in the diagram. The conjoined arrow for the β_3 path represents the product between X and η_Z.

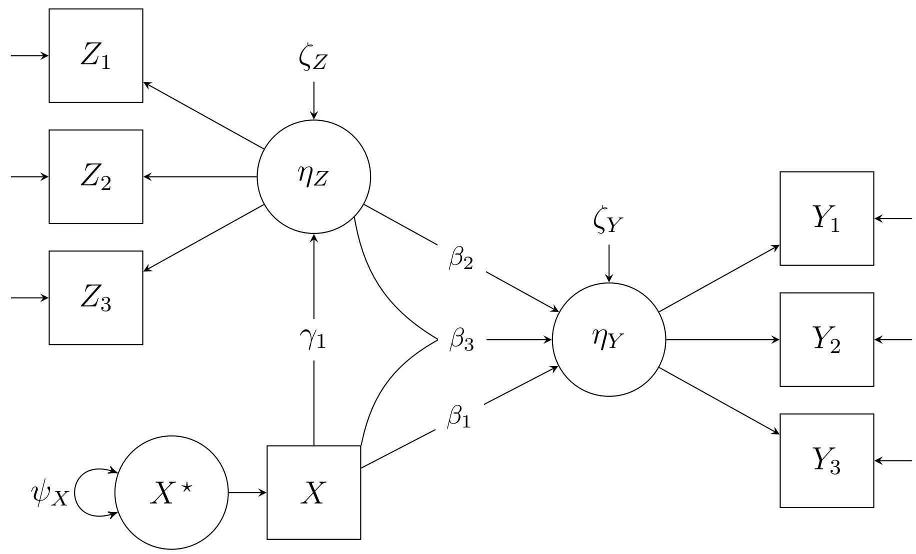
In words, Equation (14) says that there are three structural densities (the first three distributions) and two measurement densities (the last two distributions). Focusing on the two measurement densities, the two sets of indicators can all be expressed in the same way as the previous section because they are conditionally independent given their respective latent variable, and there are no cross-loadings.

Turning to the structural densities, I use the full factorization where the bivariate distribution of the predictors is represented as a model with η_Z conditional on X and a marginal model for X. This specification maps directly onto the path diagram, where X predicts η_Z. These structural densities result in the following regression models.

$$
\begin{aligned}
f\left(\eta_Y \mid \eta_Z, X\right) \rightarrow \quad & \eta_{Yi} = \alpha_Y + \beta_1 \eta_{Zi} + \beta_2 x_i + \beta_3 \left(\eta_{Zi} \times x_i\right) + \zeta_{Yi} \\
f\left(\eta_Z \mid X\right) \rightarrow \quad & \eta_{Zi} = \alpha_Z + \gamma x_i + \zeta_{Zi} \\
f\left(X\right) \rightarrow \quad & x_i = \alpha_X + r_{Xi}
\end{aligned} \tag{15}
$$

Figure 4

Path Diagram for Latent × Binary Interaction Model



Note: The mean structure is not illustrated in the diagram. The conjoined arrow for the β_3 path represents the product between X and η_Z. For identification, ψ_X is generally fixed to one.

Notably, the first regression in Equation (15) includes the product of the latent η_Z and manifest X. Again, this product is not a unique variable but a function of the two variables. The second regression above includes the effect of X predicting η_Z because I used the fully factored specification. Finally, the marginal model for X is an intercept-only model. If X is complete, this model may not be necessary because X could be treated as fixed (i.e., only show up right of the vertical bars). However, if X is incomplete, this model is required to handle the missing data on X appropriately.

The manifest by latent factorization readily incorporates binary, ordinal, multicategorical, and even skewed manifest variables. To illustrate one such extension, now suppose that X is a binary variable. The pathways where X predicts η_Y and η_Z representing a mean shift between the two categories, and X also moderates the effect of η_Z on η_Y. I depict this model in Figure 4. As I did previously, I use the probit specification to include the underlying distributed latent propensity $X^\star$. Notice that the manifest X variable now has a latent propensity that predicts the binary X, and the binary X appears in the structural model.

By using this probit model, the factorization in Equation (14) only requires a slight modification of the marginal density for X to incorporate the latent propensity.

$$f\left(X\right) = f\left(X \mid X^\star\right) \times f\left(X^\star\right) \tag{16}$$

In words, the distribution of X is factored into two separate models, a conditional density of X given the latent propensity $X^\star$ and a marginal density for $X^\star$. These two densities result in the following models.

$$\begin{aligned} f\left(X \mid X^\star\right) \rightarrow \quad x_i &= \mathcal{I}\left(x_i^\star > 0\right) \\ f\left(X^\star\right) \rightarrow \quad x_i^\star &= \alpha_{X^\star} + r_{X^\star i} \end{aligned} \tag{17}$$

The first line of Equation (17) is the additional link function that relates the observed binary variable to the latent propensity. In Figure 4, the link function is the arrow connecting the $X^\star$ to X. The second line of the equation characterizes the exogenous normally distributed latent propensity with a mean that relates to the proportion of ones to zeros. Returning to Figure 4, notice that the binary X remains in the model and still predicts η_Z and η_Y. Therefore, the first two lines of Equation (15) remain unchanged, providing a similar interpretation but with a binary predictor. For example, the interpretation of β_2 and β_3 provide the latent group difference in the intercept and slope of the regression.

As previously mentioned, the factored regression specification accommodates other types of latent by manifest variable interactions. The Blimp software described later allows for continuous, binary, ordinal, multicategorical, and skewed continuous manifest predictor variables. The factorizations for these variables mimic the one from Equation (14) but change depending on the metric of the predictor. For example, if X was a multicategorical nominal variable, the $f(X^\star)$ term in the second line of Equation (17) corresponds to an empty multinomial logistic model, and the link function on the first line is a logit link.A set of dummy codes would appear on the right side of any equation where the discrete X is a predictor.

4 Bayesian Estimation of Factored Regression

Estimating factored regression models with Markov chain Monte Carlo (MCMC) methods allows for sampling from the posterior distribution of the model conditioned on the data. Accumulating many samples from the posterior distribution provides useful summaries corresponding to quantities analogous to point estimates (i.e., mean/median) and 95% interval estimates (i.e., 2.5% and 97.5% quantiles of the distribution). These MCMC methods are iterative by nature, so each iteration produces one sample from each unknown parameter's posterior distribution. This section describes the MCMC sampling steps for factored regression to provide the technical details of Blimp's MCMC steps. Readers not interested in these details can skip to the section titled "Illustrative Analysis Examples via Blimp," which provides a data analysis example using Blimp.

Although one can use different MCMC methods to estimate models, I have found Gibbs sampling (Geman & Geman, 1984) to provide efficient and reliable convergence. Thus, this section provides a conceptual overview of the Gibbs sampler used to estimate the models. The Gibbs sampler divides unknown quantities into blocks. Then, each block is updated in a sequence, holding all other quantities constant at their current values. The updating process for each unknown quantity involves sampling estimates from a distribution of plausible values (a conditional posterior, or a posterior predictive distribution in the case of missing data). For the factored regression models, MCMC estimation can be reduced down to two fundamental blocks: (1) latent variable block and (2) manifest variable block.

To illustrate, consider the entire factorization from the latent moderation model in Figure 2.

$$\begin{aligned} f\left(\mathbf{X},\mathbf{Z},\mathbf{Y},\eta_X,\eta_Z,\eta_Y\right) &= f\left(\mathbf{X},\mathbf{Z},\mathbf{Y},\mid \eta_X,\eta_Z,\eta_Y\right)\times f\left(\eta_X,\eta_Z,\eta_Y\right) \\ &= f\left(\mathbf{X}\mid \eta_X\right)\times f\left(\mathbf{Z}\mid \eta_Z\right)\times f\left(\mathbf{Y}\mid \eta_Y\right)\times \\ &\quad f\left(\eta_Y\mid \eta_X,\eta_Z\right)\times f\left(\eta_X,\eta_Z\right) \end{aligned} \tag{18}$$

Note I use the boldface variable to denote the set of all items that load onto a single factor—e.g., $\mathbf{X} = \{X_1, X_2, X_3\}$. These sets are simply a shorthand notation for the collection of univariate conditional distributions described earlier—e.g., $f\left(\mathbf{X}\mid \eta_X\right) = f\left(X_1\mid \eta_X\right)\times f\left(X_2\mid \eta_X\right)\times f\left(X_3\mid \eta_X\right)$. The five densities above are divided into the following blocks.

1. Latent variable block:
   (a) $f\left(\eta_Y\mid \eta_X,\eta_Z\right)$
   (b) $f\left(\eta_X,\eta_Z\right)$
2. Manifest variable block:
   (a) $f\left(\mathbf{X}\mid \eta_X\right) = f\left(X_1\mid \eta_X\right)\times f\left(X_2\mid \eta_X\right)\times f\left(X_3\mid \eta_X\right)$
   (b) $f\left(\mathbf{Y}\mid \eta_Y\right) = f\left(Y_1\mid \eta_Y\right)\times f\left(Y_2\mid \eta_Y\right)\times f\left(Y_3\mid \eta_Y\right)$
   (c) $f\left(\mathbf{Z}\mid \eta_Z\right) = f\left(Z_1\mid \eta_Z\right)\times f\left(Z_2\mid \eta_Z\right)\times f\left(Z_3\mid \eta_Z\right)$

There are two fundamental steps within each density, regardless of whether the variable is latent or manifest: (1) sampling the estimated parameters for the corresponding regression model and (2) imputing the missing observations.

For step (1), the specific parameters depend on the type of model employed. For example, the $f\left(X_1\mid \eta_X\right)$ density could be a linear regression with a normally distributed residual (e.g., see Equation 7), a probit regression for categorical data, or a regression on the transformed metric if nonnormal (Yeo & Johnson, 2000; Lüdtke et al., 2020b; Keller

& Enders, 2023). To illustrate, suppose X_1 is a continuous variable that is conditionally normal given the latent variable, η_X.

$$\begin{aligned} x_{1i} &= \alpha_1 + \lambda_1 \eta_{Xi} + e_{1i} \\ e_{1i} &\sim \mathcal{N}\left(0, \sigma^2_{e1}\right) \end{aligned} \tag{19}$$

Conceptually, constructing a Gibbs sampler involves drawing parameters conditionally on each other: (a) $f\left(\alpha_1, \lambda_1 \mid \sigma^2_{e1}, \eta_X, X_1\right)$ and (b) $f\left(\sigma^2_{e1} \mid \alpha_1, \lambda_1, \eta_X, X_1\right)$, where the two densities correspond to a multivariate normal and inverse gamma distribution. The sampling of the parameters (i.e., α_1, λ_1, and σ^2_{e1}) conditional on the manifest and latent data follow standard procedures for linear regression models with conjugate priors (e.g., see Lynch, 2007; Gelman et al., 2013; Keller & Enders, 2023). This simplification follows because the imputed values of η_X function as observed/known quantities.

As previously mentioned, the sampling steps for regression models hold true regardless if the variable is latent or manifest. To illustrate, consider drawing a structural regression corresponding to the model in Block (1a) above.

$$\begin{aligned} \eta_{Yi} &= \alpha_Y + \beta_1 \eta_{Zi} + \beta_2 x_i + \beta_3 \left(\eta_{Zi} \times x_i\right) + \zeta_{Yi} \\ \zeta_{Yi} &\sim \mathcal{N}\left(0, \psi^2_Y\right) \end{aligned} \tag{20}$$

Again, treating the current latent imputations as known data allows one to leverage standard Gibbs sampling steps for multiple regression models because the latent data have been imputed and function as observed/known quantities. For example, the conditional posterior for the structural β's with a conjugate prior is a multivariate normal distribution with the OLS point estimates and sampling covariance matrix as its mean vector and covariance matrix, respectively.

4.1 *Imputation of Variables*

As previously discussed, conceptually, there is no difference between a single observation of a latent variable and an unobserved observation of a manifest variable. Both are estimated via imputation—i.e., data augmentation (Tanner & Wong, 1987). The Gibbs sampler obtains imputations from the conditional distribution of the variable to be imputed given all other variables and parameters in the model.

To illustrate, reconsider the latent moderation model in Figure 2. Suppose the manifest variable X_1 has incomplete observations that must be imputed. I determine the conditional distribution by simplifying Equation (18) to only densities that contain X_1.

$$f\left(X_1 \mid X_2, X_3, \mathbf{Y}, \mathbf{Z}, \eta_X, \eta_Y, \eta_Z\right) = f\left(X_1 \mid \eta_X\right) \tag{21}$$

In this instance, there is only one density, and the problem simplifies nicely. If I know what η_X is, I can impute X_1 conditional on this value. I previously characterized $f\left(X_1 \mid \eta_X\right)$ in Equation (19) as a linear regression in the example; therefore, the conditional distribution of X_1 is as follows.

$$x_{1i} \sim \mathcal{N}\left(\alpha_1 + \lambda_1 \eta_{Xi}, \sigma^2_{e1}\right) \tag{22}$$

The mean and variance of the above distribution are defined by the predicted score and residual variance of Equation (19), respectively. Like the latent scores, the parameter values that define the distribution's center and spread are held at their current values.

Generally speaking, obtaining the imputations of the latent variables is more complicated because they appear in more than one density. The latent moderation example has three latent variables that need to be imputed. The conditional density for any variable can be expressed up to proportionality by returning to Equation (18) and removing any density that does not contain the specific variable. For example, η_X's conditional density is expressed up to proportionality as follows.

$$f\left(\eta_X \mid \mathbf{X}, \mathbf{Z}, \mathbf{Y}, \eta_Y, \eta_Z\right) \propto f\left(\eta_Y \mid \eta_X, \eta_Z\right) \times f\left(\eta_X, \eta_Z\right) \times f\left(\mathbf{X} \mid \eta_X\right) \tag{23}$$

Importantly, the density $f\left(\eta_Y \mid \eta_X, \eta_Z\right)$ maintains the nonlinear relationship between η_X and η_Y by including the $\eta_X \times \eta_Z$ product (i.e., β_3 in Equation 13). Similarly, the imputations of η_Z and η_Y can be obtained by sampling from their conditional densities.

$$f\left(\eta_Z \mid \mathbf{X}, \mathbf{Z}, \mathbf{Y}, \eta_X, \eta_Y\right) \propto f\left(\eta_Y \mid \eta_X, \eta_Z\right) \times f\left(\eta_X, \eta_Z\right) \times f\left(\mathbf{Z} \mid \eta_Z\right) \tag{24}$$

$$f\left(\eta_Y \mid \mathbf{X}, \mathbf{Z}, \mathbf{Y}, \eta_X, \eta_Z\right) \propto f\left(\eta_Y \mid \eta_X, \eta_Z\right) \times f\left(\mathbf{Y} \mid \eta_Y\right) \tag{25}$$

Under the assumption of normally distributed residuals, all three conditional latent densities above have closed-form solutions. Each latent variable follows a normal distribution, but η_X and η_Z are heteroscedastic conditional on η_Y.

The technical Appendix A describes the recursive algorithmic approach used by Blimp to obtain the solutions for the mean and variance of the latent variable distributions. In addition, I include an R implementation at `https://github.com/blimp-stats/Latent-Interactions` to demonstrate how the close-form solutions are obtained. Although beyond the scope of this paper, this approach readily extends to a wide range of models, including interactions, multivariate models, and multilevel models.

Alternatively, when the conditional distribution cannot be solved, I can leverage various general MCMC sampling techniques that require knowing the density only up to proportionality and embed these within the Gibbs sampler. For example, I can employ a random walk Metropolis step to sample from the density (Lynch, 2007). Such techniques are necessary when an incomplete or latent variable is squared or a categorical mediator is present.

4.2 Prior Distributions

In this section, I briefly discuss the default priors used by Blimp. For more specific technical details about Blimp's default prior distributions in the online supplemental documents from Enders et al. (2020) and Keller and Enders (2023)[2], as well as the user manual (Keller & Enders, 2022). Generally speaking, Blimp provides uninformative or weakly informative priors by default (when possible), and the priors for regression coefficients (i.e., structural coefficients or factor loadings), thresholds, and variances (latent or manifest) are all considered independent. Note, informative priors are possible via the `PARAMETERS` command.

Returning to Figure 2, each path coefficient (i.e., loading or regression slope) and intercept would invoke an improper uniform prior; however, Blimp can incorporate other priors that are user specified. For path coefficients, if a univariate normal distribution prior is requested, then Blimp will sample the coefficient from a full conditional distribution. For factor loadings, if the variance of a latent variable is fixed at one for identification, then a truncated prior distribution over the positive range is usually desirable. Otherwise, the posterior distributions of the loadings may be bimodal (Levy & Mislevy, 2017).

For unstructured covariances matrices, they are given a weakly informative inverse Wishart prior. For structured covariance matrices, the prior largely depends on the nature of the matrix. For example, if only a variance is fixed, then I maintain the inverse Wishart prior and use a transformation to maintain the fixed variance (Lynch, 2007; Asparouhov & Muthén, 2010a). Alternatively, if covariances are fixed, I place a uniform prior on the elements and use a metropolis-hasting algorithm to draw each element conditionally on all the others. Separation type specifications that place distinct priors on variances and correlations are also possible in Blimp (Keller & Enders, 2023). Finally, residual variances are given inverse gamma priors that map onto the univariate special case of the inverse Wishart priors.

Turning to regressions of categorical outcomes, Blimp imposes a default normal prior

[2]Both online supplemental documents are available at `www.appliedmissingdata.com/blimp-papers`.

with a mean of zero and a variance of five. In my experience, this weakly informative prior is useful for stabilizing the estimates where the regression coefficient ranges are more limited due to scaling. Finally, for ordinal probit models, the thresholds are drawn according to Cowles (1996).

5 Illustrative Analysis Examples via Blimp

To introduce estimating models with Blimp (Version 3.2; Keller & Enders, 2022), I use an illustrative data example. Throughout this section, I supplement the discussion with code excerpts that show how to specify the model. All syntax files, output files, and the raw data are provided at `https://github.com/blimp-stats/Latent-Interactions`.

The substantive example is based on Zhou, Meier, and Spector (2014), who investigated the interactions between personality traits and organizational constraints when predicting counterproductive work behavior. Based on the reported summary statistics, effects, and discussion of scales, I generated data to mimic theirs. I generated all items to follow a 5-point Likert scale, just as they were measured in the original article.

Initially, I focus on conscientiousness (`consci`) and organizational constraints (`orgcon`) scales consisting of 10 items each, and how they predict counterproductive work behavior directed at people (`cwb`). The conscientiousness scale measures the "Big Five" conceptualization of the personality trait, which relates to the degree of organization and self-discipline in a participant (Costa & McCrae, 1992). The organizational constraints scale captures how difficult it is for a participant to perform their job. An example item might ask if a participant feels the organization provides poor equipment or supplies. Finally, the counterproductive work behavior scale consisted of 23 items, asking participants how often they had performed counterproductive behaviors at their jobs. The concept of counterproductive work behavior refers to intentional behaviors that harm organizations or people in organizations (Spector & Fox, 2005). An example item might ask how often they might yell at someone at work or damage an organization's property.

The initial models map directly onto the discussion of the latent moderation model (i.e., Figure 2), with latent factors of conscientiousness and constraints as η_X and η_Z, respectively, and counterproductive work behavior as the endogenous latent factor, η_Y. Because each item follows a Likert scale, I use an ordered probit model to characterize each item (Albert & Chib, 1993; V. E. Johnson & Albert, 2006). The metrics of the indicators are specified by listing the indicator names after the `ORDINAL` command. Blimp will automatically use the data to determine the number of categories and fit the appropriate ordered probit model.

5.1 *Fitting a Latent Regression Model*

When fitting these models, I advocate for a modeling-building approach driven by substantive theory. I begin by fitting the structural model without the interaction.

$$\texttt{cwb}_i = \beta_0 + \beta_1(\texttt{consci}_i) + \beta_2(\texttt{orgcon}_i) + \zeta_i \tag{26}$$

First, I invoke the modeling section using the `MODEL` command. Each subsequent line under this command specifies a regression equation from the broader model. The following code excerpt specifies the regression of the two latent exogenous variables onto the endogenous latent counterproductive work behavior variable. Below, the first code excerpt includes two commands (denoted as capitalized names followed by a colon).

```
## Define latent variables
LATENT: consci orgcon cwb;

## Begin Modeling Syntax
MODEL:
    ## Latent Regression Model (no moderation)
    cwb ~ consci orgcon;

    # Fix residual variance to 1 for identification
    cwb ~~ cwb@1;
```

Note that Blimp's syntax uses a hashtag or pound symbol (`#`) to turn the remaining line into comments and a semicolon (`;`) to terminate any statement. The `LATENT` command defines three new latent variables. The names themselves are arbitrary, but I use them to represent the three latent variables for conscientiousness (`consci`), organizational constraints (`orgcon`), and counterproductive work behavior (`cwb`). The `MODEL` command begins the specification of the model. The first non-comment line in the `MODEL` section specifies the regression without the interaction (Equation 26). Blimp syntax uses a tilde ($\sim$) to denote "regressed on" or "predicted by". Thus, the first equation in the `MODEL` section reads as `cwb` is predicted by `consci` and `orgcon`. Finally, the double tidle syntax ($\sim\sim$) is used to specify correlations or variances. In this case, I use the syntax to fix the residual variance of `cwb` to be equal to one for identification to specify a fixed factor scaling (Klopp & Klößner, 2021). I also use the double tilde syntax to specify the models for the two exogenous factors.

```
    # Correlate two exogenous variables
    consci ~~ orgcon;

    # Fix variance to 1 for identification
    consci ~~ consci@1;
    orgcon ~~ orgcon@1;
```

Mapping onto a standard multiple regression, I correlate the two exogenous latent variables. I also set their variances to one for identification of the latent constructs.

The final section of the `MODEL` command is devoted to specifying the measurement models (i.e., Equation 11). I allow all slopes (i.e., loadings) to be freely estimated because I have identified the latent variables by constraining their variances. Below, I regress each item onto its specific latent variable.

```
    # Measurement Models for conscientiousness items
    consci -> consci1@l_con consci2:consci10;

    # Measurement Models for org constraints items
    orgcon -> orgcon1@l_org orgcon2:orgcon10;

    # Measurement Models for CWB items
    cwb -> cwb1@l_cwb cwb2:cwb23;

# Specify first loading as always positive
PARAMETERS:
  l_con ~ truncate(0, Inf);
  l_org ~ truncate(0, Inf);
  l_cwb ~ truncate(0, Inf);
```

Blimp uses a right-pointing arrow (->) to specify that the variable to the left (a latent factor) predicts every variable to the right (its indicators). As a shortcut, Blimp allows for the specification of multiple variables using a colon. For example, above the '`consci2:consci10`' denotes a consescutive range of `consci2` to `consci10`. As discussed in the previous section, I must identify the latent variable by setting the first loading to always be positive. I accomplish this by first specifying a parameter name for each loading using the ampersand ('`@`') followed by the name of my choosing (e.g., `consci1@l_con`). I then used the `PARAMETERS` command to specify a truncated prior distribution that constrains the first loading of each factor to always be positive (Levy & Mislevy, 2017).

The Blimp script also requests 50,000 iterations after the initial warm-up period (see Appendix B for a full syntax example). Thus, each parameter is characterized by an em-

Table 1: Results for Latent Regression Models

Predictors	Coefficient	SD	Credible Intervals 2.5%	97.5%
Model 1				
`consci`	-0.288	0.039	-0.365	-0.213
`orgcon`	0.611	0.043	0.527	0.697
R^2	0.335	0.026	0.285	0.386
Model 2				
`consci`	-0.300	0.040	-0.379	-0.223
`orgcon`	0.632	0.045	0.546	0.721
`consci` × `orgcon`	-0.223	0.042	-0.307	-0.142
R^2	0.362	0.026	0.310	0.412
Model 3				
`consci`	-0.311	0.042	-0.396	-0.230
`orgcon`	0.653	0.048	0.562	0.749
`emosta`	-0.017	0.039	-0.094	0.060
`consci` × `orgcon`	-0.251	0.045	-0.341	-0.165
`consci` × `emosta`	-0.072	0.039	-0.148	0.004
`orgcon` × `emosta`	0.084	0.043	0.001	0.169
`consci` × `orgcon` × `emosta`	-0.093	0.044	-0.181	-0.008
R^2	0.377	0.026	0.325	0.427

Note: All variables are unobserved latent constructs. The "Coefficient" column is the posterior median.

pirical posterior distribution of 50,000 samples. In Table 1, the "Coefficient" column is the median and is analogous to a frequentist point estimate. The posterior standard deviation values in the second column give the standard deviation of the 50,000 parameters, which are numerically similar to frequentist standard errors. The "Credible Intervals" columns report the 95% interval, which is similar to a frequentist confidence interval, but they make no reference to repeated sampling. Said differently, the credible interval gives the 95% probability that the true estimate would lie within the interval given the observed data and model. For researchers who prefer to work with null hypothesis significance testing logic, the credible interval can be used to test such hypotheses. Specifically, suppose the null value of zero falls outside this interval. In that case, one can conclude that the effect is "significant" because the probability that a parameter value of zero produced the data is less than 0.05. For researchers who prefer a frequentist-like p-value, the Bayesian Wald chi-square tests (Asparouhov & Muthén, 2021a) and can be obtained for every model parameter (see Chapter 2, `OUTPUT` section; Keller & Enders, 2022).

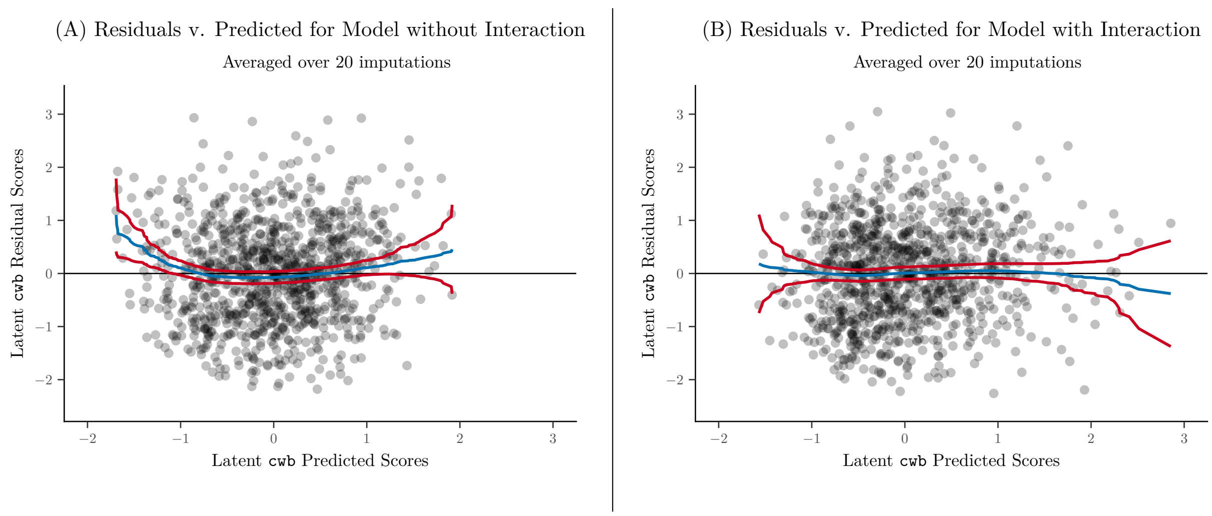
Under the Model 1 heading, Table 1 gives the posterior summaries for the latent regression model with no interactions. I can see that zero is not included within the 95% credible interval for both effects. Looking at the coefficient for latent conscientiousness, with the fixed factor scaling, the coefficient is interpreted as a one standard deviation unit increase in latent conscientiousness, I would expect the latent counterproductive work behavior to decrease by 0.288 standard deviations of its residual (Klopp & Klößner, 2021). Similarly, turning to latent organizational constraints, with a one standard deviation unit increase in the latent organizational constraints, I would expect the latent counterproductive work behavior to increase by 0.611 standard deviations of its residual. Finally, the results suggest that the inclusion of the two latent predictors explains 33.5% of the variance of the latent counterproductive work behavior.

5.2 Graphical Diagnostics on Latent Imputations

One of important practical advantages of the factored regression approach is that it yields latent variable imputations as a byproduct of estimation. These latent variable imputations can be leveraged to produce familiar graphical diagnostics for multiple regression models. To illustrate, I used the follow code block to save 20 imputed data sets that included imputations, predicted values, and residuals for every variable in the model (latent or manifest).

Figure 5

Regression Diagnostic Plots Based on Imputed Scores



Note: Panel (A) is the plot of the residuals versus predicted for `cwb` regressed on `consci` and `orgcon`. Panel (B) is the plot of the residuals versus predicted for `cwb` regressed on `consci`, `orgcon`, and the product of `consci` $\times$ `orgcon`. Each point represents the posterior mean of an observation's predicted and residual scores across twenty imputations. The blue loess line represents the average fitted value of a loess line across twenty imputations. The red lines represent ± 2 standard error of the fitted value pooled across twenty imputations.

```
NIMPS: 20;                  # Number of imputed data sets
OPTIONS: SaveLatent;        # Request for saving latent imputations
         SavePredicted;     # Request for saving predicted scores
         SaveResiduals;     # Request for saving residuals
SAVE: stacked = m0.txt;     # File to save imputations into 'm0.txt'
```

Each imputed data set will contain different imputed values. The imputations will be equally spaced across the requested post burn-in iterations and chains. Because these are imputations by nature, nothing precludes analyzing them as complete data sets and using multiple imputation pooling rules (Asparouhov & Muthén, 2010b).

Frequentist inference using Rubin's (1987) pooling rules is a typical application for multiply imputed data sets. One could fit the structural regression using the latent variable imputations if desired. Instead, I use the residual and predicted scores for `cwb` from the model to construct diagnostic regression plots (Cook & Weisberg, 1999). Panel A in Figure 5 plots the residuals versus predicted based on the model without the interaction. Each point represents the average values across the twenty imputations, and the loess line's fitted values and ± 2 standard error bars are based upon the pooled results using Rubin's rules (Rubin, 1987). Although looking at regression plots involves some subjectivity, the curvilinear form of the loess line in Panel A indicates that an unmodeled nonlinear relationship may remain in the latent regression. Paired with a substantive theory about a stressor-personality interaction (Bowling & Eschleman, 2010; Fox, Spector, & Miles, 2001; Penney & Spector, 2005), it would make sense to model the interaction between conscientiousness (`consci`) and organizational constraints (`orgcon`).

5.3 *Fitting a Latent Moderation Model*

Fitting a latent moderation model like the one in Figure 2 requires only a minor change to the syntax from Section 5.1. Below I provide the equation for the focal structural model.

$$\texttt{cwb}_i = \beta_0 + \beta_1(\texttt{consci}_i) + \beta_2(\texttt{orgcon}_i) + \beta_3(\texttt{consci}_i \times \texttt{orgcon}_i) + \zeta_i \tag{27}$$

Equation (27) illustrates that I must include the interaction between the two latent variables, `consci` and `orgcon`. In Blimp's syntax, I explicitly specify the interaction within the regression model. The modified `MODEL` excerpt is as follows.

```
## Begin Modeling Syntax
MODEL:
    ## Latent Moderation Model
    # Explicitly write out the product with '*'
    cwb ~ consci orgcon consci*orgcon;
```

All other parts of the model specification exactly follow Section 5.1, and Appendix B gives the full Blimp script for the example.

Under the Model 2 heading, Table 1 gives the posterior summaries for the latent moderation model. The negative interaction coefficient implies that this negative association between conscientiousness (`consci`) and counterproductive work behavior (`cwb`) strengthens (becomes more negative) as organizational constraints increase (`orgcon`). Because the 95% credible interval does not contain zero, one could conclude that the effect is "statistically significant" at $p < 0.05$. Turning to the lower-order effects, they have clear interpretations because the latent predictors have zero means. For example, at the average latent organizational constraints, the conditional effect of conscientiousness is negative, and zero is not within the 95% credible intervals. Thus, higher levels of conscientiousness are associated with lower levels of counterproductive work behavior. Finally, note that including the product term has explained approximately 3% more variance in the latent outcome.

Like the model in Section 5.1, I evaluated the residual diagnostics for the latent moderation model. Returning to Figure 5, Panel B plots the latent `cwb` residuals against its predicted values, both of which were now generated under the (correct) assumption that there is a latent interaction. Focusing on both panels, the diagnostic plot illustrates that including the latent interaction has most of the systematic relationship between the predicted and residual scores. Visually, the loess line now falls closer to the horizontal zero bar. Furthermore, the ± 2 standard error bands consistently incorporate the zero within the predictive range, indicating that I have adequately modeled the functional form.

5.4 Conditional Effects Analysis for the Latent Moderation Model

In addition to estimating the latent moderation model, Blimp can perform conditional effects analysis to probe the interaction at different values of a moderator (Aiken & West, 1991; Bauer & Curran, 2005). Using the `SIMPLE` command, I requested the simple intercept and slope slope of the focal predictor (`consci`) at three levels of the moderator (`orgcon`): one standard deviation below the latent `orgcon` mean, at the latent `orgcon` mean, and one standard deviation above the latent `orgcon` mean.

Table 2: Results for Model 2 Conditional Effects

Parameter	Coefficient	SD	Credible Intervals 2.5%	97.5%
consci $\vert$ **orgcon** @ -1				
Intercept	-0.632	0.045	-0.721	-0.546
Slope	-0.076	0.055	-0.184	0.031
consci $\vert$ **orgcon** @ 0				
Intercept	0.000	0.000	0.000	0.000
Slope	-0.300	0.040	-0.379	-0.223
consci $\vert$ **orgcon** @ $+1$				
Intercept	0.632	0.045	0.546	0.721
Slope	-0.523	0.061	-0.646	-0.406

Note: All variables are unobserved latent constructs. The "Coefficient" column is the posterior median.

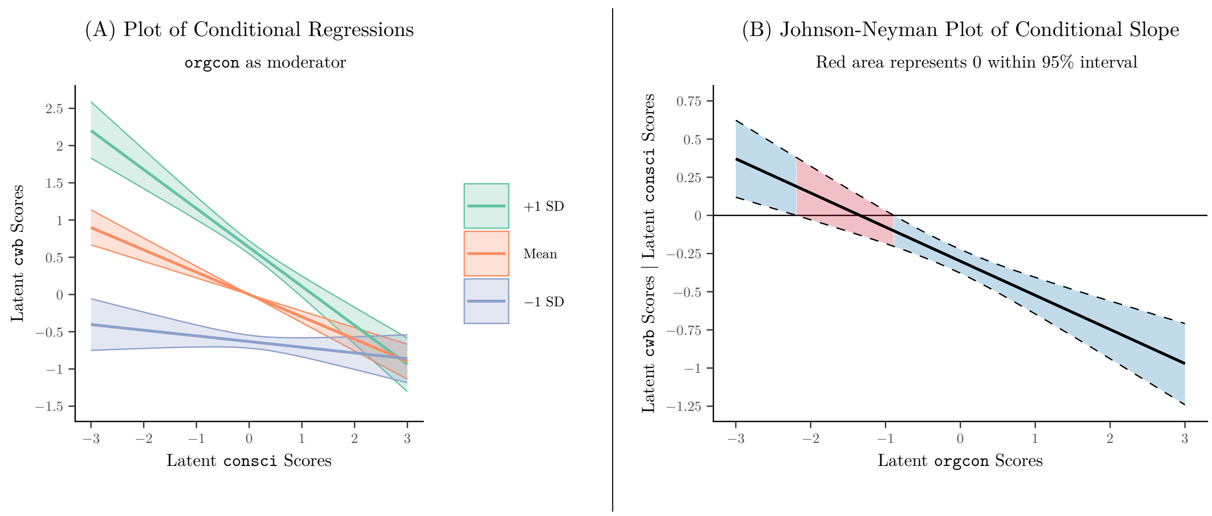
```
## Conditional Effects Analysis with +/- 1 SD of orgcon
# mean = 0 by definition
# SD   = 1 by definition
SIMPLE:
    consci | orgcon @ -1.0;
    consci | orgcon @  0.0;
    consci | orgcon @  1.0;
```

The variable to the left of the vertical bar is the focal variable, and the variable to the right is the moderator. The numeric value after the ampersand (`@`) is the moderator value being conditioned on. Recall that I chose identification constraints that standardized the latent predictor variables. Thus, the numeric value of the moderator listed after the ampersand reflects standard deviation units from the mean.

Because the conditional effects are simple functions of the estimated parameters, Blimp provides summary statistics of the posterior distribution for each simple intercept and simple slope. Table 2 gives the posterior summaries for the conditional effect's analysis at

Figure 6

Conditional Effects Plots for `cwb` Regressed on `consci` Moderated by `orgcon`



Note: Panel A is the regression of latent counterproductive work behavior (`cwb`) on latent conscientiousness (`consci`) moderated at different levels of latent organizational constraint (`orgcon`). Panel B is the Johnson-Neyman plot to evaluate when the credible intervals for the conditional slope. The ribbons around each line depict the corresponding 95% credible intervals.

one standard deviation above `orgcon`'s mean, at average `orgcon`, and one standard deviation below `orgcon`'s mean. The conditional effect's analysis shows that the negative relationship between counterproductive work behavior and an employee's conscientiousness gets stronger as organizational constraints increase.

To further illustrate the interaction, I computed the conditional effect estimates and used them to construct plots. Figure 6 illustrates two established approaches for probing interactions. I include the commented R code to produce these plots at `https://github.com/blimp-stats/Latent-Interactions`. To begin, Panel A plots the conditional regression equations, illustrating the relationship between the latent counterproductive work behavior as a function of the latent conscientiousness factor evaluated at the different levels of organizational constraints. The bands around each regression line represent the 95% credible intervals of prediction. The flatter purple line near the bottom of the chart shows that having lower organizational constraints predicts little to no relationship between counterproductive work behavior and an employee's conscientiousness. Only as employee's perceive more constraints does lower conscientiousness predict higher counterproductive work behavior.

Panel B of Figure 6 illustrates a plot for the Johnson-Neyman approach (P. O. Johnson & Neyman, 1936) to probing interactions. On the x-axis, I have different moderator values—i.e., organizational constraints. On the y-axis, I have the values of the conditional slope of latent counterproductive work behavior regressed on the latent conscientiousness factor. The red-shaded region represents the values of the moderator that produce a conditional effect where zero falls within the 95% credible intervals (i.e., a range of moderator scores that produce a non-significant conditional slope). The plot graphically illustrates that zero is a plausible value for the conditional relationship at about one standard deviation below the mean of organizational constraints. Furthermore, zero remains within the interval until a bit past two standard deviations below the mean of organizational constraints.

5.5 Adding a Three-Way Latent Interaction

As previously noted, the general interaction model used by LMS and other procedures is restricted to the product of two latent variables. Fitting a model with a three-way interaction requires a complex model specification where two-way effects are defined by a proxy latent variable (Asparouhov & Muthén, 2021b). In contrast, the factored regression approach readily accommodates three-way effects. Extending the example, Zhou et al. (2014) also investigated how emotional stability (`emosta`) might moderate the two-way effects of organizational constraints and conscientiousness on counterproductive work behaviors. The emotional stability contained ten 5-point Likert items that loaded onto a common factor. Thus, the primary analysis model with a three-way latent variable interaction is as follows.

$$
\begin{aligned}
\texttt{cwb}_i = \beta_0 &+ \beta_1(\texttt{consci}_i) + \beta_2(\texttt{orgcon}_i) + \beta_3(\texttt{emosta}_i) \\
&+ \beta_4(\texttt{consci}_i \times \texttt{orgcon}_i) + \beta_5(\texttt{consci}_i \times \texttt{emosta}_i) \\
&+ \beta_6(\texttt{orgcon}_i \times \texttt{emosta}_i) + \beta_7(\texttt{consci}_i \times \texttt{orgcon}_i \times \texttt{emosta}_i) + \zeta_i
\end{aligned}
\tag{28}
$$

Fitting the three-way latent interaction model requires only a minor change to the code. In Blimp's syntax, I explicitly specify all interactions within the regression model. The modified `MODEL` excerpt is as follows.

```
## Begin Modeling Syntax
MODEL:
    ## Threeway Latent Moderation Model
    cwb ~ consci@b1 orgcon@b2 emosta@b3
          consci*orgcon@b4 consci*emosta@b5 orgcon*emosta@b6
          consci*orgcon*emosta@b7;
```

In the code above, I have included parameter labels on each coefficient using the ampersand (`@`) followed by the name based on Equation (28). As demonstrated below, I use these parameter labels to probe the three-way interaction. Note that the code also specifies a measurement model (not shown here) for `emosta`, and it correlates the new latent factor with the other two exogenous latent variables. The entire script can be accessed at `https://github.com/blimp-stats/Latent-Interactions`.

Under the Model 3 heading, Table 1 gives the posterior summaries for the three-way latent interaction model. Because all latent variables have a mean of zero, the two-way interactions in the table are conditional effects at the mean of the third variable. For example, consider the latent conscientiousness by latent organizational constraints interaction from Section 5.3. The two-way interaction and lower-order terms are similar in magnitude and sign, so the substantive interpretations only slightly change by referencing an employee with average emotional stability. The three-way interaction captures the degree to which the two-way effect changes as a function of emotional stability. The negative slope means that the two-way product of conscientiousness and organizational constraints becomes more negative as emotional stability increases. Because zero does not fall inside the 95% credible interval, one could conclude that the parameter differs from zero. For researchers who prefer a frequentist-like test statistic, a Bayesian Wald chi-square test (Asparouhov & Muthén, 2021a) of the three-way interaction (β_7) was statistically significant, $\chi^2(1) = 4.502$, $p = 0.034$.

To further investigate the three-way interaction, I probe the three-way effect by computing the two-way interaction between `consci` and `orgcon` conditional on different levels of `emosta`. In Blimp, I computed these auxiliary conditional effect parameters using the `PARAMETERS` command and the parameter labels specified earlier.

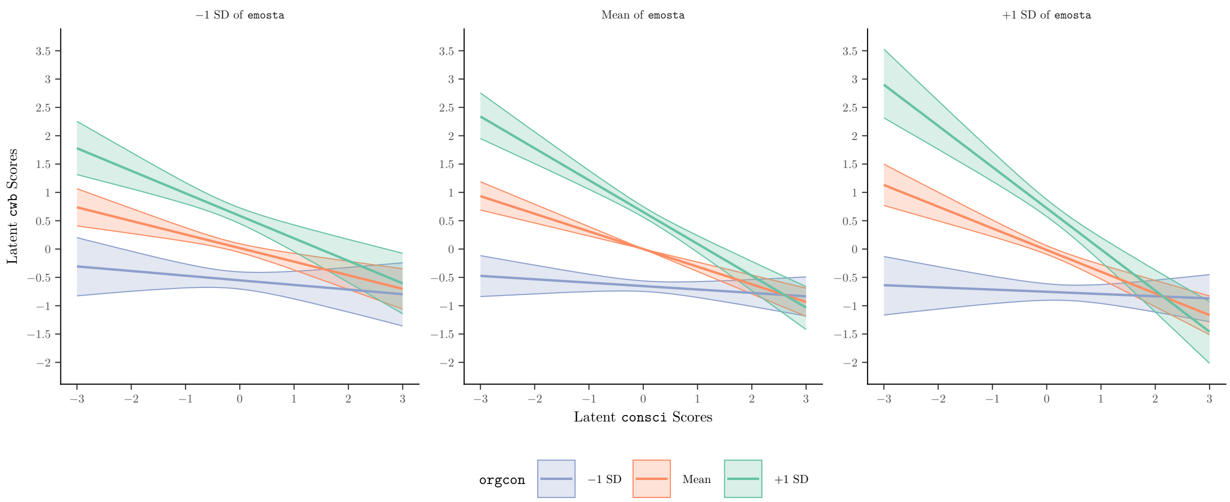
```
# Obtain two-way conditional effects
PARAMETERS:
  inter.emo_m1 = b4 + b7 * (-1);
  inter.emo_0  = b4 + b7 * ( 0);
  inter.emo_p1 = b4 + b7 * (+1);
```

The excerpt above creates three new generate quantities that represent the two-way interaction above, below and at the mean of `emosta`. Because these three quantities are functions of the estimated parameters, Blimp provides summary statistics of the posterior distribution.

To further illustrate the interaction, I compute the conditional effect estimates and used them to construct plots to probe the interactions. Figure 7 illustrates the relationship between the latent counterproductive work behavior as a function of the latent conscientiousness evaluated at the different levels of organizational constraints and employee emotional stability. The bands around each regression line represent the 95% credible intervals of prediction. Moving from the left to the right in Figure 7 illustrates that the interaction becomes strong (i.e., it's coefficient more negative) as emotional stability increases.

Figure 7

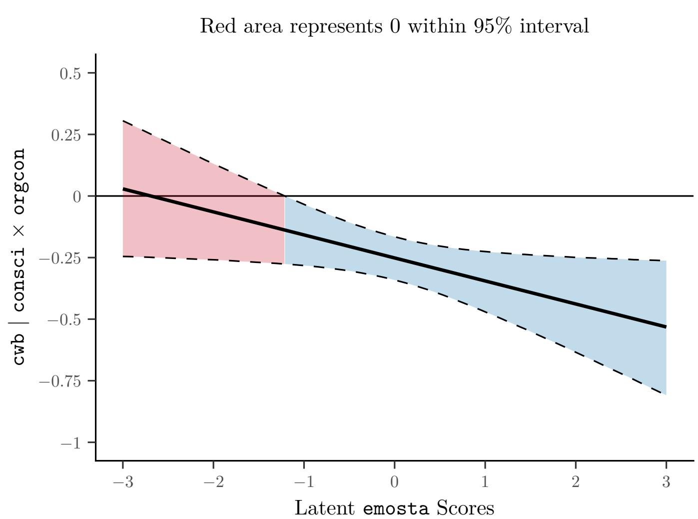
Conditional Effects Plots for `cwb` Regressed on `consci` Moderated by `orgcon` and `emosta`



Note: Plots of the regression of latent counterproductive work behavior (`cwb`) on latent conscientiousness (`consci`) moderated at different levels of latent organizational constraint (`orgcon`) and latent emotional stability (`emotsa`). The ribbons around each line depict the corresponding 95% credible intervals.

Figure 8

Johnson-Neyman Plot for `consci` × `orgcon` Moderated by `emotsa`



Note: A Johnson-Neyman plot to evaluate when the credible intervals for the two-way interaction between latent conscientiousness (`consci`) and organizational constraint (`orgcon`) at different levels of latent emotional stability (`emotsa`). The ribbons around each line depict the corresponding 95% credible intervals.

In addition to the conditional effects plot, Figure 8 illustrates the Johnson-Neyman plot for the two-way interaction between conscientiousness and organizational constraints moderated by emotional stability. On the x-axis, I have different values of emotional stability. On the y-axis, I have the values of the conditional slope for the two-way interaction between conscientiousness and organizational constraints. The red-shaded region represents the values of the moderator that produce a conditional effect where zero falls within the 95% credible intervals (i.e., two-way effects that would not be deemed statistically significant). The plot graphically illustrates that a zero interaction is plausible value for the conditional relationship at about one standard deviation below the mean of emotional stability.

5.6 Accommodating Latent $\times$ Discrete Interaction

One of the great strengths of factored regression's framework is that it readily accommodates a range of manifest predictor variables. As previously noted, Blimp supports binary, ordinal, multicategorical, skewed continuous, and normal predictors. The model specification is largely the same across variable types. For discrete predictors, one defines the variable type using the `ORDINAL` or `NOMINAL` command. The manifest by latent interaction is then specified in the same manner as a latent by latent interaction. To illustrate, I created a binary version of the organizational constraints variable by performing a median split on the simulated latent variable scores. The structural regression equation is the same as Equation (27), except that `orgcon` would now be a dummy variable characterizing low and high organizational constraints.

Returning to the path diagram in Figure 4, the manifest by latent interaction requires two modeling changes. First, the observed binary organizational constraints (`bin_org`) would need a model to accommodate any potential missing data. Second, in Figure 4, the predictors link via a regression equation rather than a covariance. The moderated regression itself is left unchanged. The key changes to the Blimp script are shown in the following syntax excerpt, and the full script is available at `https://github.com/blimp-stats/Latent-Interactions`.

```
ORDINAL: bin_org;    # Define binary variable
LATENT:  cwb consci; # Define latent variables

## Begin Modeling Syntax
MODEL:
  bin_org ~ 1;
  consci ~ bin_org;
  cwb ~ consci bin_org consci*bin_org;
  # Specify cwb and consci measurement models here

## Simple effects for binary orgcon
SIMPLE: consci | bin_org;
```

Listing `bin_org` on the `ORDINAL` line invokes a probit formulation like Equation (17) for the first regression after the `MODEL` command. Alternatively, listing `bin_org` after the `NOMINAL` command would invoke a logistic regression. The second regression after the `MODEL` command is the aforementioned directed pathway between `consci` and `bin_org`. In this example, the `SIMPLE` command produces conditional effects for each level of the discrete moderator (i.e., the effect of conscientiousness on counterproductive work behavior in the low and high organizational constraint groups).

The parameter interpretations parallel any standard multiple regression model with an interaction between a binary and continuous predictor. Returning to Figure 4, the γ coefficient would represent the group mean difference on latent `consci` between the low and high organizational constraint groups. The lower-order effects have familiar interpretations. For example, the β_1 slope is the counterproductive work behavior group mean difference at the average latent `consci`. The β_2 slope would convey the conditional effect of latent `consci` on latent `cwb` in the low organizational constraints group. Finally, the interaction coefficient (β_3) would represent the `consci` slope difference for the high organizational constraints group.

Next, suppose that organizational constraints is a multicategorical variable (`cat_org`) with low, medium, and high groups coded 1, 2, and 3. The structural model would now include two dummy codes representing the difference relative to a reference category. Each dummy code (D_2 and D_3) would then interact with the latent conscientiousness factor. The regression equation is as follows.

$$\begin{aligned}\texttt{cwb}_i = \beta_0 + \beta_1(\texttt{consci}_i) + \beta_2(D_{2i}) + \beta_3(D_{3i}) \\ + \beta_4(\texttt{consci}_i \times D_{2i}) + \beta_5(\texttt{consci}_i \times D_{3i}) + \zeta_i\end{aligned} \tag{29}$$

The Blimp syntax excerpt below lists `cat_org` on the `NOMINAL` line, but it is otherwise identical to the binary model.

```
NOMINAL: cat_org;    # Define nominal variables
LATENT:  cwb consci; # Define latent variables

## Begin Modeling Syntax
MODEL:
  cat_org ~ 1;
  consci ~ cat_org;
  cwb ~ consci cat_org consci*cat_org;
  # Specify cwb and consci measurement models here

## Simple effects for categorical orgcon
SIMPLE: consci | cat_org;
```

In this case, Blimp invokes a multinomial logistic model for `cat_org`. In both models where it is a predictor, Blimp automatically creates a pair of dummy codes, treating the lowest numeric code as the reference group. The following syntax excerpt is an equivalent specification of the focal structural model that makes the automatic dummy coding more explicit.

```
## Begin Modeling Syntax
MODEL:
  cwb ~ consci cat_org.2 cat_org.3
        consci*cat_org.2 consci*cat_org.3;
```

Above, adding a period followed by the category's numeric code references the dummy code for that category. For example, `cat_org.2` is equivalent to D_2 in Equation (29). The substantive interpretations parallel any multiple regression with an interaction between multicategorical and continuous predictors. Where the dummy codes represent the differences between the reference group and the category. Finally, alternative or custom coding schemes are possible by embedding logical statements as predictors in a regression equation (see Chapter 2 in the Blimp User Guide for details; Keller & Enders, 2022).

6 Discussion

Moderated regression models are ubiquitous across the social sciences, and much methodological interest has gone into extending moderated regression models with latent variables. Unsurprisingly, several different approaches have been proposed to estimate these models (see Kelava & Brandt, 2023, for more details), with the two prominent approaches being the product indicator method (Kenny & Judd, 1984) and the general interaction model (Klein & Moosbrugger, 2000). The product indicator has a long history of slight modifications to improve estimation but still remains severely limited. Among other things, these methods frequently require intricate specifications of covariance structures, lack parsimony, assume continuous data, and potentially provide biased results with incomplete indicators. The general interaction model rectifies this by modeling the latent product directly. However, as I discussed, it is not general enough. This modeling approach requires complex specifications to approximate models with proxy latent variables (e.g., three-way interactions, latent centering in multilevel moderation models, handling of latent by manifest interactions). Furthermore, in certain circumstances, it does not extend to handling variables of different metrics (e.g., binary, ordinal, nominal). This article provides a fully generalized Bayesian approach to latent variable interactions and other nonlinear effects. I provide a free software implementation to estimate these models that does not require researchers to program their models in difficult-to-use languages such as JAGS (Plummer, 2017) or Stan (Stan Development Team, 2023).

The factored regression framework expresses the multivariate distribution of variables as a collection of simpler distributions. These simpler distributions can be configured to accommodate complex distributional features (e.g., heteroscedastic variation of a latent predictor) that are incompatible with the multivariate normality assumption invoked by traditional SEMs. Overall, the factored regression approach builds on previous work both in the nonlinear SEM (Lee, 2007) and missing data (Ibrahim et al., 2002). I leverage this previous work to layout a general approach to latent variable modeling. The framework treats latent variables as missing data that should be estimated via imputation. Once the latent missing data are imputed, the estimation steps parallel those for any regression model.

The factored regression approach to interactive and nonlinear latent variable effects provides a flexible framework to specify complex structural equation models with interactive and nonlinear effects for latent variables, including three-way or higher interactions and quadratic effects. These nonlinear effects can be specified with any variable, including a mixture of manifest and latent variable, and they are not strictly limited to the structural models. The factored regression approach accommodates a diverse set of manifest variable types (e.g., binary, ordinal, multicategorical, continuous normal, skewed continuous, count indicators), which can serve as indicators of latent variables or as manifest predictors with nonlinear relationships. In addition, it easily incorporates missing data on any manifest variable in the model. While the framework solves interesting estimation issues, the practical advantage of the framework is that it is plug-and-play with familiar procedures with moderated regression models—e.g., centering, conditional effects analysis, regression diagnostics, Johnson-Neyman plots.

To link to existing research, I provided a few examples of the types of latent variable interactions this framework can handle. However, nothing precludes factored regression from incorporating more complex models. For example, the framework allows for moderated nonlinear factor analysis (MNFLA Bauer, 2017), where manifest predictor variables moderate loadings and variances without the complications of imposing nonlinear constraints (Enders, Vera, Keller, Lenartowicz, & Loo, 2023). The framework can further extend this to allow latent variables to moderate the loadings and variances. Another example is intensive longitudinal data, which has recently shown an interest in heterogeneous variances. In these models, the magnitude of the intra-individual variation is represented as a level-2 latent variable (McNeish, 2021; Lester, Cullen-Lester, & Walters, 2021; Feng & Hancock, 2022; Ernst, Albers, & Timmerman, 2023). In the factored regression framework, the latent variable representing the heterogeneous variance can be a moderator for any other variable in the model. Other examples include interaction effects in dynamic structural equation models discussed in Asparouhov, Hamaker, and Muthén (2018) and latent moderated mediation models with single-level or multilevel data (Keller, 2022). More broadly, general nonlinear structural equation models (see Harring & Zou, 2023) can be estimated, allowing for nonlinear effects, and general nonlinear structural equation models with latent and manifest variables (e.g., quadratic effects, sine-cosine models, Jenss–Bayley functions; Jenss & Bayley, 1937).

Although not discussed in this paper, factored regression readily extends to multilevel latent variable modeling, where sampling units are clustered within another unit. My approach easily accommodates the decomposition of a multilevel interaction effect into its within, between, and cross-level interactions via latent cluster mean centering discussed by Preacher et al. (2016). Current methods using the general interaction model are limited, often relying on unintuitive proxy variables. By requiring these proxy variables, the general interaction model cannot handle categorical moderators and mediators in the context of multilevel models (Asparouhov & Muthén, 2021b). In contrast, the factored regression framework provides a straightforward path for incorporating a categorical moderator or mediator. Factored regression also readily incorporates within-cluster centering at level-2 and level-3 latent group means, and latent centering is available for continuous, binary, ordinal, and multicategorical predictors. The Blimp User's Guide (Keller & Enders, 2022) provides illustrative syntax examples for multilevel modeling.

All of the modeling possibilities described above provide avenues for future methodological work. Research is still needed to determine the sample size and data requirements needed to reliably estimate these complex models. Similarly, power studies are needed to provide researchers with guidelines linked to presumed effect sizes. Finally, future studies should investigate the robustness of factored regression to violations of model-specific distributions.

In summary, this paper provides researchers with a general approach to fitting latent variable interactions of all types. It fully generalizes past work and offers interesting extensions that were not previously possible. All of this is provided in a user-friendly and free software package.

References

Agresti, A. (2012). *Analysis of ordinal categorical data* (3rd ed.). Hoboken, NJ: Wiley.

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

Albert, J. H., & Chib, S. (1993). Bayesian analysis of binary and polychotomous response data. *Journal of the American Statistical Association*, *88*, 669–679.

Arminger, G., & Muthén, B. O. (1998). A bayesian approach to nonlinear latent variable models using the gibbs sampler and the metropolis-hastings algorithm. *Psychometrika*, *63*, 271–300.

Asparouhov, T., Hamaker, E. L., & Muthén, B. (2018). Dynamic structural equation models. *Structural equation modeling: a multidisciplinary journal*, *25*(3), 359–388.

Asparouhov, T., & Muthén, B. (2010a). *Bayesian analysis using mplus: Technical implementation* [Electronic Article]. Retrieved from `https://www.statmodel.com/download/Bayes3.pdf`

Asparouhov, T., & Muthén, B. (2010b). *Plausible values for latent variables using mplus* [Electronic Article]. Retrieved from `www.statmodel.com/download/Plausible.pdf`

Asparouhov, T., & Muthén, B. (2021a). Advances in bayesian model fit evaluation for structural equation models. *Structural equation modeling: a multidisciplinary journal*, *28*(1), 1–14.

Asparouhov, T., & Muthén, B. (2021b). Bayesian estimation of single and multilevel models with latent variable interactions. *Structural Equation Modeling: A Multidisciplinary Journal*, *28*(2), 314–328.

Bartlett, J., Keogh, R., & Bonneville, E. F. (2021). smcfcs: Multiple imputation of covariates by substantive model compatible fully conditional specification. Retrieved from `https://CRAN.R-project.org/package=smcfcs`

Bauer, D. J. (2017). A more general model for testing measurement invariance and differential item functioning. *Psychological methods*, *22*(3), 507.

Bauer, D. J., & Curran, P. J. (2005). Probing interactions in fixed and multilevel regression: Inferential and graphical techniques. *Multivariate behavioral research*, *40*(3), 373–400.

Bowling, N. A., & Eschleman, K. J. (2010). Employee personality as a moderator of the relationships between work stressors and counterproductive work behavior. *Journal of occupational health psychology*, *15*(1), 91.

Cook, R., & Weisberg, S. (1999). Applied regression including computing and graphics.

Cortina, J. M., Markell-Goldstein, H. M., Green, J. P., & Chang, Y. (2021). How are we testing interactions in latent variable models? surging forward or fighting shy? *Organizational Research Methods*, *24*(1), 26–54.

Costa, P. T., & McCrae, R. R. (1992). Normal personality assessment in clinical practice: The neo personality inventory. *Psychological assessment*, *4*(1), 5.

Cowles, M. K. (1996). Accelerating monte carlo markov chain convergence for cumulative-link generalized linear models. *Statistics and Computing*, *6*, 101–111.

Du, H., Enders, C., Keller, B. T., Bradbury, T. N., & Karney, B. R. (2022). A bayesian latent variable selection model for nonignorable missingness. *Multivariate Behavioral Research*, *57*(2-3), 478-512. (PMID: 33529056) doi: 10.1080/00273171.2021.1874259

Enders, C. K., Du, H., & Keller, B. T. (2020). A model-based imputation procedure for multilevel regression models with random coefficients, interaction effects, and other nonlinear terms. *Psychological Methods*, *25*, 88–112.

Enders, C. K., Vera, J. D., Keller, B. T., Lenartowicz, A., & Loo, S. K. (2023). Building a simpler moderated nonlinear factor analysis model with bayesian estimation. Manuscript submitted for publication.

Erler, N. S., Rizopoulos, D., & Lesaffre, E. M. E. H. (2021). JointAI: Joint analysis and imputation of incomplete data in R. *Journal of Statistical Software*, *100*(20), 1–56. doi: 10.18637/jss.v100.i20

Erler, N. S., Rizopoulos, D., Rosmalen, J. v., Jaddoe, V. W., Franco, O. H., & Lesaffre, E. M. (2016). Dealing with missing covariates in epidemiologic studies: a comparison between multiple imputation and a full bayesian approach. *Statistics in medicine*, *35*(17), 2955–2974.

Ernst, A. F., Albers, C. J., & Timmerman, M. E. (2023). A comprehensive model framework for between-individual differences in longitudinal data. *Psychological Methods*.

Feng, Y., & Hancock, G. R. (2022). A structural equation modeling approach for modeling variability as a latent variable. *Psychological Methods*.

Fox, S., Spector, P. E., & Miles, D. (2001). Counterproductive work behavior (cwb) in response to job stressors and organizational justice: Some mediator and moderator tests for autonomy and emotions. *Journal of vocational behavior*, *59*(3), 291–309.

Gelman, A., Stern, H. S., Carlin, J. B., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian data analysis*. Chapman and Hall/CRC.

Geman, S., & Geman, D. (1984). Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Transactions on pattern analysis and machine intelligence*(6), 721–741.

Goldstein, H., Carpenter, J. R., & Browne, W. J. (2014). Fitting multilevel multivariate models with missing data in responses and covariates that may include interactions and non-linear terms. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *177*(2), 553–564.

Harring, J. R., & Zou, J. (2023). Nonlinear structural equation models: Advanced methods and applications. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 681–700). Guilford Publications.

Ibrahim, J. G. (1990). Incomplete data in generalized linear models. *Journal of the American Statistical Association*, *85*(411), 765–769.

Ibrahim, J. G., Chen, M.-H., & Lipsitz, S. R. (1999). Monte carlo em for missing covariates in parametric regression models. *Biometrics*, *55*(2), 591–596.

Ibrahim, J. G., Chen, M.-H., & Lipsitz, S. R. (2002). Bayesian methods for generalized linear models with covariates missing at random. *Canadian Journal of Statistics*, *30*(1), 55–78.

Jenss, R. M., & Bayley, N. (1937). A mathematical method for studying the growth of a child. *Human Biology*, *9*(4), 556.

Johnson, P. O., & Neyman, J. (1936). Tests of certain linear hypotheses and their application to some educational problems. *Statistical research memoirs*.

Johnson, V. E., & Albert, J. H. (2006). *Ordinal data modeling*. Springer Science & Business Media.

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

Kelava, A., & Brandt, H. (2023). Latent interaction effects. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 427–446). Guilford Publications.

Keller, B. T. (2022). An introduction to factored regression models with blimp. *Psych*, *4*(1), 10.

Keller, B. T. (2024). Automatically derived full conditionals for factored regression models. Manuscript in preparation.

Keller, B. T., & Enders, C. K. (2022). *Blimp user's guide (version 3.0)*. Los Angeles, CA.

Keller, B. T., & Enders, C. K. (2023). An investigation of factored regression missing data methods for multilevel models with cross-level interactions. *Multivariate Behavioral Research*, *58*(5), 938–963.

Kenny, D. A., & Judd, C. M. (1984). Estimating the nonlinear and interactive effects of latent variables. *Psychological bulletin*, *96*(1), 201.

Klein, A., & Moosbrugger, H. (2000). Maximum likelihood estimation of latent interaction effects with the lms method. *Psychometrika*, *65*(4), 457–474.

Klein, A., & Muthén, B. O. (2007). Quasi-maximum likelihood estimation of structural equation models with multiple interaction and quadratic effects. *Multivariate behavioral research*, *42*(4), 647–673.

Klopp, E., & Klößner, S. (2021). The impact of scaling methods on the properties and interpretation of parameter estimates in structural equation models with latent variables. *Structural Equation Modeling: A Multidisciplinary Journal*, *28*(2), 182–206.

Lee, S.-Y. (2007). *Structural equation modeling: A bayesian approach*. John Wiley & Sons.

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: A Multidisciplinary Journal*, *14*(3), 404–434.

Lee, S.-Y., & Zhu, H.-T. (2000). Statistical analysis of nonlinear structural equation models with continuous and polytomous data. *British Journal of Mathematical and Statistical Psychology*, *53*(2), 209–232.

Lee, S.-Y., & Zhu, H.-T. (2002). Maximum likelihood estimation of nonlinear structural equation models. *Psychometrika*, *67*(2), 189–210.

Lester, H. F., Cullen-Lester, K. L., & Walters, R. W. (2021). From nuisance to novel research questions: Using multilevel models to predict heterogeneous variances. *Organizational Research Methods*, *24*(2), 342–388.

Levy, R., & Mislevy, R. J. (2017). *Bayesian psychometric modeling*. CRC Press.

Lipsitz, S. R., & Ibrahim, J. G. (1996). A conditional model for incomplete covariates in parametric regression models. *Biometrika*, *83*(4), 916–922.

Liu, J., Gelman, A., Hill, J., Su, Y.-S., & Kropko, J. (2014). On the stationary distribution of iterative imputations. *Biometrika*, *101*(1), 155–173.

Lüdtke, O., Robitzsch, A., & West, S. G. (2020a). Analysis of interactions and nonlinear effects with missing data: a factored regression modeling approach using maximum likelihood estimation. *Multivariate Behavioral Research*, *55*(3), 361–381.

Lüdtke, O., Robitzsch, A., & West, S. G. (2020b). Regression models involving nonlinear effects with missing data: A sequential modeling approach using bayesian estimation. *Psychological methods*, *25*, 157–181.

Lynch, S. M. (2007). *Introduction to applied bayesian statistics and estimation for social scientists*. Springer Science & Business Media.

McNeish, D. (2021). Specifying location-scale models for heterogeneous variances as multilevel sems. *Organizational Research Methods*, *24*(3), 630–653.

Muthén, B., & Asparouhov, T. (2009). Growth mixture modeling: Analysis with non-gaussian random effects. In G. Fitzmaurice, M. Davidian, G. Verbeke, & G. Molenberghs (Eds.), *Longitudinal data analysis* (pp. 143–165). Chapman & Hall/CRC.

Muthén, B., & Asparouhov, T. (2012). Bayesian structural equation modeling: a more flexible representation of substantive theory. *Psychological methods*, *17*(3), 313.

Penney, L. M., & Spector, P. E. (2005). Job stress, incivility, and counterproductive work behavior (cwb): The moderating role of negative affectivity. *Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior*, *26*(7), 777–796.

Plummer, M. (2017). *Jags version 4.3.0 user manual* [Computer Software]. Lyon, France.

Preacher, K. J., Zhang, Z., & Zyphur, M. J. (2016). Multilevel structural equation models for assessing moderation within and across levels of analysis. *Psychological methods*, *21*(2), 189.

Robitzsch, A., & Lüdtke, O. (2021). mdmb: Model based treatment of missing data. Retrieved from `https://CRAN.R-project.org/package=mdmb`

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. New York: Wiley.

Seaman, S. R., Bartlett, J. W., & White, I. R. (2012). Multiple imputation of missing covariates with non-linear effects and interactions: an evaluation of statistical methods. *BMC medical research methodology*, *12*(1), 46.

Spector, P. E., & Fox, S. (2005). The stressor-emotion model of counterproductive work behavior.

Stan Development Team. (2023). *The Stan Core Library* [Computer Software]. Retrieved from `http://mc-stan.org`

Tanner, M. A., & Wong, W. H. (1987). The calculation of posterior distributions by data augmentation. *Journal of the American statistical Association*, *82*(398), 528–540.

Yeo, I.-K., & Johnson, R. A. (2000). A new family of power transformations to improve normality or symmetry. *Biometrika*, *87*(4), 954–959.

Zhang, Q., & Wang, L. (2017). Moderation analysis with missing data in the predictors. *Psychological Methods*, *22*(4), 649.

Zhou, Z. E., Meier, L. L., & Spector, P. E. (2014). The role of personality and job stressors in predicting counterproductive work behavior: A three-way interaction. *International Journal of Selection and Assessment*, *22*(3), 286–296.

Zhu, H.-T., & Lee, S.-Y. (1999). Statistical analysis of nonlinear factor analysis models. *British Journal of Mathematical and Statistical Psychology*, *52*(2), 225–242.

A Algorithmically Derived Conditional Distributions

This discussion provides an overview of how Blimp derives the analytical distribution for general products and centering. Further details, including the extension to multilevel and multivariate models, are given in Keller (2024). Finally, note that if a model does not map onto the models discussed in this section, Blimp can still estimate them using a less efficient sampler such as a Metropolis-Hastings algorithm or elliptical slice sampler.

Derivation of General Underlying Model

Suppose I am interested in deriving the conditional density $f\left(\, X \mid Y, \ldots \right)$, where the ellipsis ('...') represent any other variables in the model. I express this density up to proportionality as follows.

$$f\left(\, Y \mid X, \ldots \right) \times f\left(\, X \mid \ldots \right)$$

Under normality assumptions for the two densities in the product, if I rearrange the two densities into the following forms

$$f\left(\, Y \mid X, \ldots \right) \rightarrow \quad y_i \sim \mathcal{N}\left(\mathrm{A}_i + \mathrm{B}_i x_i, \sigma_e^2\,\right) \tag{A1}$$

$$f\left(\, X \mid \ldots \right) \rightarrow \quad x_i \sim \mathcal{N}\left(\mathrm{C}_i, \sigma_r^2\,\right) \tag{A2}$$

where A_i, B_i, and C_i are any computation that does not include X or Y, then I can solve for $f\left(\, X \mid Y, \ldots \right)$ as follows.

$$\begin{aligned} f\left(\, X \mid Y, \ldots \right) &\propto \exp\left(-\frac{\left(y_i - [\mathrm{A}_i + \mathrm{B}_i x_i]\right)^2}{2\sigma_e^2}\right) \times \exp\left(-\frac{\left(x_i - \mathrm{C}_i\right)^2}{2\sigma_r^2}\right) \\ &\propto \exp\left(-\left[\frac{\left([y_i - \mathrm{A}_i] - \mathrm{B}_i x_i\right)^2}{2\sigma_e^2} + \frac{\left(x_i - \mathrm{C}_i\right)^2}{2\sigma_r^2}\right]\right) \end{aligned}$$

$$
\begin{aligned}
&\propto \exp\left(-\left[\frac{x_i^2\mathrm{B}_i^2 - 2x_i\mathrm{B}_i\left(y_i - \mathrm{A}_i\right) + \text{const}}{2\sigma_e^2} + \frac{x_i^2 - 2x_i\mathrm{C}_i + \text{const}}{2\sigma_r^2}\right]\right)\\
&\propto \exp\left(-\left[\frac{x_i^2\mathrm{B}_i^2 - 2x_i\mathrm{B}_i\left(y_i - \mathrm{A}_i\right)}{2\sigma_e^2} + \frac{x_i^2 - 2x_i\mathrm{C}_i}{2\sigma_r^2}\right]\right)\\
&\propto \exp\left(-\frac{x_i^2\mathrm{B}_i^2\sigma_r^2 - 2x_i\mathrm{B}_i\sigma_r^2\left(y_i - \mathrm{A}_i\right) + x_i^2\sigma_e^2 - 2x_i\sigma_e^2\mathrm{C}_i}{2\sigma_e^2\sigma_r^2}\right)\\
&\propto \exp\left(-\frac{x_i^2\left[\mathrm{B}_i^2\sigma_r^2 + \sigma_e^2\right] - 2x_i\left[\mathrm{B}_i\sigma_r^2\left(y_i - \mathrm{A}_i\right) + \sigma_e^2\mathrm{C}_i\right]}{2\sigma_e^2\sigma_r^2}\right)\\
&\propto \exp\left(-\frac{x_i^2 - \frac{2x_i\left[\mathrm{B}_i\sigma_r^2(y_i - \mathrm{A}_i) + \sigma_e^2\mathrm{C}_i\right]}{\mathrm{B}_i^2\sigma_r^2 + \sigma_e^2}}{\frac{2\sigma_e^2\sigma_r^2}{\mathrm{B}_i^2\sigma_r^2 + \sigma_e^2}}\right)\\
&\propto \exp\left(-\frac{x_i^2 - \frac{2x_i\left[\mathrm{B}_i\sigma_r^2(y_i - \mathrm{A}_i) + \sigma_e^2\mathrm{C}_i\right]}{\mathrm{B}_i^2\sigma_r^2 + \sigma_e^2} + \text{const}}{\frac{2\sigma_e^2\sigma_r^2}{\mathrm{B}_i^2\sigma_r^2 + \sigma_e^2}}\right)\\
&\propto \exp\left(-\frac{\left[x_i - \frac{\sigma_r^2\mathrm{B}_i(y_i - \mathrm{A}_i) + \sigma_e^2\mathrm{C}_i}{\sigma_r^2\mathrm{B}_i^2 + \sigma_e^2}\right]^2}{2\left(\frac{\sigma_e^2\sigma_r^2}{\mathrm{B}_i^2\sigma_r^2 + \sigma_e^2}\right)}\right)
\end{aligned}
$$

$$
\therefore \quad x_i \sim \mathcal{N}\left(\frac{\sigma_r^2\mathrm{B}_i\left(y_i - \mathrm{A}_i\right) + \sigma_e^2\mathrm{C}_i}{\sigma_r^2\mathrm{B}_i^2 + \sigma_e^2}, \frac{\sigma_e^2\sigma_r^2}{\sigma_r^2\mathrm{B}_i^2 + \sigma_e^2}\right) \tag{A3}
$$

The above result illustrates that if I can rearrange the two densities into some form that maps onto equations (A1) and (A2), then the product will result in a distribution that is also normally distributed. Note that the i subscript is both in the mean and the variance; therefore, the conditional distribution in Equation (A3) is computed on every observation. However, this maintains a very general result that easily extends to common situations. For example, any moderator variables for the regression of Y on X can be included in the B term. Similarly, anything additive to X can be in A. As another example, if X is centered, then A would include the subtraction of the mean of X times the regression coefficient.

Chaining Products of Multiple Densities

Suppose now I have three variables, X, Y, and Z, with a fully factored distribution as follows

$$f\left(Y \mid X, Z\right) \times f\left(Z \mid X\right) \times f\left(X\right) \tag{A4}$$

with the following three regressions as the models for each density.

$$\begin{aligned} f\left(Y \mid X, Z\right) \rightarrow &\quad y_i \sim \mathcal{N}\left(\beta_0 + \beta_1 x_i + \beta_2 z_i + \beta_3 (x_i \times z_i), \sigma^2_Y\right) \\ f\left(Z \mid X\right) \rightarrow &\quad z_i \sim \mathcal{N}\left(\alpha_0 + \alpha_1 x_i, \sigma^2_Z\right) \\ f\left(X\right) \rightarrow &\quad x_i \sim \mathcal{N}\left(\gamma_0, \sigma^2_X\right) \end{aligned} \tag{A5}$$

If the goal is to obtain the distribution for $f\left(X \mid Y, Z\right)$—e.g., as required for a Gibbs sampler to impute X—we first break down this product into two separate steps.

1. Obtain $f\left(X \mid Y, Z\right) \propto f\left(Y \mid X, Z\right) \times f\left(X \mid Z\right)$
2. Obtain $f\left(X \mid Z\right) \propto f\left(Z \mid X\right) \times f\left(X\right)$

By breaking the problem into products of two distributions, I simplify the problem and perform the steps backwards to obtain the desired distribution. Starting with Step 2, I apply the result from Equation (A3) to obtain the conditional distribution.

$$f\left(X \mid Z\right) \rightarrow \quad x_i \sim \mathcal{N}\left(\frac{\sigma^2_X \alpha_1 \left(z_i - \alpha_0\right) + \sigma^2_Z \gamma_0}{\sigma^2_X \alpha^2_1 + \sigma^2_Z}, \frac{\sigma^2_Z \sigma^2_X}{\sigma^2_X \alpha^2_1 + \sigma^2_Z}\right) \tag{A6}$$

Next, I take the above distribution and apply Equation (A3) again in Step 1 to obtain a normal distribution

$$f\left(X \mid Y, Z\right) \rightarrow \quad x_i \sim \mathcal{N}\left(\mu_{X|Y,Z}, \sigma^2_{X|Y,Z}\right) \tag{A7}$$

where the mean and variances are as follows.

$$
\begin{aligned}
\mu_{X|Y,Z} &= \frac{\left(\frac{\sigma^2_Z \sigma^2_X}{\sigma^2_X \alpha^2_1 + \sigma^2_Z}\right)(\beta_1 + \beta_3 z_i)\left(y_i - [\beta_0 + \beta_2 z_i]\right) + \sigma^2_Y \left(\frac{\sigma^2_X \alpha_1 (z_i - \alpha_0) + \sigma^2_Z \gamma_0}{\sigma^2_X \alpha^2_1 + \sigma^2_Z}\right)}{\left(\frac{\sigma^2_Z \sigma^2_X}{\sigma^2_X \alpha^2_1 + \sigma^2_Z}\right)(\beta_1 + \beta_3 z_i)^2 + \sigma^2_Y} \\
\sigma^2_{X|Y,Z} &= \frac{\sigma^2_Y \left(\frac{\sigma^2_Z \sigma^2_X}{\sigma^2_X \alpha^2_1 + \sigma^2_Z}\right)}{\left(\frac{\sigma^2_Z \sigma^2_X}{\sigma^2_X \alpha^2_1 + \sigma^2_Z}\right)(\beta_1 + \beta_3 z_i)^2 + \sigma^2_Y}
\end{aligned} \tag{A8}
$$

The above distribution may not have a nice compact form; however, using element-wise calculations, it is straightforward to compute for every observation, resulting in mean and variance vectors. These vectors can then be used to sample from the derived conditional distribution for every observation. More importantly, the described algorithm is straightforward to implement in code, thus allowing a wide range of models to have the conditional posterior predictive distributions automatically derived. I include an R implementation of this algorithm and demonstrate how to use functional programming techniques at `https://github.com/blimp-stats/Latent-Interactions`.

B Full Blimp Syntax for Latent Moderation Example

```
# Data set's file name
DATA: data_cat.txt;

# Variable names in text data set
VARIABLES:
    consci1:consci10 emosta1:emosta10
    orgcon1:orgcon10 cwb1:cwb23
    bin_orgcon cat_orgcon;

## Specify that the items are ordinal.
ORDINAL:
    consci1:consci10 orgcon1:orgcon10
    cwb1:cwb23;

## Define latent variables
LATENT: consci orgcon cwb;

## Begin Modeling Syntax
MODEL:
    ## Latent Moderation Model
    cwb_model:
        # Explicitly write out the product between
      # the two latent variables
        cwb ~ consci orgcon consci*orgcon;

        # Fix residual variance to 1 for identification
    cwb ~~ cwb@1;

    ## Correlating predictor latent variables
    predictor_model:
        consci ~~ orgcon;
        consci ~~ consci@1;
        orgcon ~~ orgcon@1;

    ## Measurement Models for conscientiousness items
    measurement_consci:
      consci -> consci1@l_con consci2:consci10;

    ## Measurement Models for org constraints items
    measurement_orgcon:
      orgcon -> orgcon1@l_org orgcon2:orgcon10;

    ## Measurement Models for CWB items
    measurement_cwb:
      cwb -> cwb1@l_cwb cwb2:cwb23;

# Specify first loading as always positive
PARAMETERS:
  l_con ~ truncate(0, Inf);
  l_org ~ truncate(0, Inf);
  l_cwb ~ truncate(0, Inf);

## Conditional Effects Analysis with +/- 1 SD of orgcon
# mean = 0 by definition
# SD   = 1 by definition
SIMPLE:
    consci | orgcon @ -1.0;
    consci | orgcon @  0.0;
    consci | orgcon @  1.0;

# Algorithmic Options
SEED: 298722;   # PRNG Seed
CHAINS: 10;     # Number of chains
BURN: 20000;    # Burn-in iterations
ITER: 50000;    # Post burn-in iterations
```