STRUCTURAL EQUATION MODELING
2025, VOL. 32, NO. 5
ISSN: 1070-5511 | E-ISSN: 1532-8007

Using Small-Variance Priors to Detect Covariate Misspecifications in Latent Class Analysis Models

Sarah Depaoli, Fan Jia, and Marieke Visser

University of California

ABSTRACT

This study examines the detection of direct effects in conditional latent class analysis (LCA) models using Bayesian structural equation modeling (BSEM). By leveraging approximate-zero priors—small-variance priors centered on zero—BSEM relaxes model assumptions, enhancing fit and identifying non-zero parameters. While previous research to relax local independence assumptions in LCA, this study extends their use to detecting direct effects between continuous covariates and latent class indicators, broadening their application within Bayesian modeling. Multiple population models, sample sizes, and class proportion conditions were examined in an extensive simulation study. While priors successfully identified non-zero direct effects, they also produced false positives, particularly under local independence violations. Class-specific direct effects required larger samples for reliable detection. Future research should refine prior specifications and explore alternative methods for detecting direct effects, emphasizing the importance of prior sensitivity analyses in applied research.

KEYWORDS

Bayesian SEM; covariate misspecification; direct effects; latent class analysis; simulation study

Latent class analysis (LCA) is a modeling technique that allows for the identification of an underlying categorical latent variable from a set of observed latent class indicators. The LCA model is considered a measurement model because the categorical latent variable represents the unobservable subgroups within the population. Recent methodological advances have primarily focused on expanding the LCA measurement model to include external variables (e.g., predictors, distal outcomes) as part of a larger structural model. Traditionally, LCA models with external variables were estimated using a one-step approach with the ML estimator (McCutcheon, 1987; Vermunt, 2010). The primary issue with the one-step approach is that each alternation to the external variables in the model results in the reidentification of the model, which can change the interpretation of the latent classes (Asparouhov & Muthén, 2014; Bakk & Kuha, 2018; Bolck, et al., 2004; Vermunt, 2010). An alternative strategy to the one-step approach is to use a stepwise approach to estimation, where the LCA measurement model is established independently of the structural model (Asparouhov & Muthén, 2014; Vermunt, 2010).

Regardless of the estimation strategy being pursued (i.e., one-step vs. stepwise approach), the inclusion of external variables in the LCA model presents an opportunity for model misspecifications. For example, one of the most common external variables to include in LCA is a latent class-predicting covariate. By including a latent class-predicting covariate, researchers can explore variables that explain the clustering in the LCA measurement model. One way of adding a covariate to an LCA model is to regress the latent class variable on the covariate (Nylund-Gibson & Masyn, 2016). By specifying the conditional LCA model this way, an assumption is being made that the relationship between the covariate and observed latent class indicators is fully explained by the indirect effect of the covariate on the latent class variable. In other words, the direct effect of the covariate on each class indicator would be fixed to zero during model estimation. Unfortunately, this assumption does not always hold true in practice. Often, there are direct covariate effects on the observed class indicator variables in applied settings. Misspecification occurs when the true direct covariate effects are fixed to zero in practice.

Several previous methodological studies have explored the consequences of misspecifying covariate effects in LCA models (Collins & Lanza, 2010; Masyn, 2013; Nylund-Gibson & Masyn, 2016; Petras & Masyn, 2010). For example, one simulation study found that the misspecification of covariate effects leads to the over-extraction of the number of latent classes during class enumeration (Nylund-Gibson & Masyn, 2016). For this reason, a growing body of methodological research that suggests class enumeration should take place before modeling external variables (Collins & Lanza, 2010; Masyn, 2013; Nylund-Gibson & Masyn, 2016; Petras & Masyn, 2010). Although covariates may provide additional information that aids class enumeration (Li & Hser, 2011; Lubke & Muthén, 2007; Muthén, 2002), covariates should only be included during class enumeration when the covariate relationships are known *a priori* (Petras & Masyn, 2010). Researchers are unlikely to know the covariate relationships in most applied settings in advance.

A seemingly well-fitting unconditional LCA measurement model can become a poorly fitting conditional LCA model when the covariate is misspecified, regardless of the estimation approach used. When using the one-step approach, ignoring direct effects between the covariate and latent class indicators can bias structural parameters (Janssen et al., 2019). Covariate misspecifications bias structural parameters even more under poor class separation (Janssen et al., 2019). In theory, the stepwise approach to estimation should safeguard the measurement model against covariate misspecifications. However, excluding direct effects still biases structural parameters even when using a stepwise approach to analysis (Asparouhov & Muthén, 2014; Janssen et al., 2019). In other words, failure to model direct effects between the covariate and latent class indicators can compromise the LCA model. Notably, methodological advances have been made on modeling direct covariate effects when using a stepwise approach. Vermunt and Magidson (2021) demonstrate how the ML three-step approach can be modified to include direct effects during the first estimation step. The direct effects must first be detected to be accommodated in the procedure described in Vermunt and Magidson (2021).

This article focuses on detecting direct effects, with particular emphasis on a strategy for identifying non-zero direct effects in conditional LCA models within Bayesian structural equation modeling (BSEM) framework. BSEM allows us to relax model assumptions by employing *approximate-zero priors*, which are near-zero priors centered on zero with very small variances. These priors could help researchers improve model fit and identify non-zero parameters. While previous applications of approximate-zero priors in LCA models have primarily focused on relaxing assumptions about local independence between latent class indicators (Asparouhov & Muthén, 2011; Lee et al., 2020), this study aims to broaden their application to include direct effects between a continuous covariate and latent class indicators.

This article is organized as follows: We begin with an overview of LCA, providing essential details about its fundamental concepts. Next, we address LCA with covariates, examining various estimation approaches (one-step vs. three-step) for incorporating covariates into LCA models. Building upon this, the section on Bayesian LCA with Covariates explores the application of Bayesian methods to LCA models that include covariates, highlighting how Bayesian techniques can address model assumption violations and enhance parameter estimation. We then review the limitations and challenges associated with current methodologies for detecting direct effects and discuss the concept of small-variance priors, elaborating on their role and applications in Bayesian SEM. Following this introduction, we present a simulation study and conclude by evaluating the effectiveness of small-variance priors in detecting and managing non-zero effects in Bayesian LCA models with covariates.

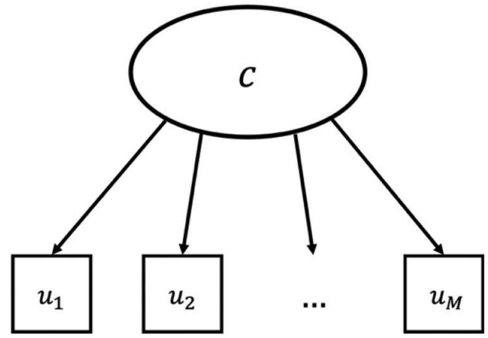

Figure 1. The unconditional LCA model with M binary latent class indicators. The latent class indicators are represented with $u_1, u_2, \ldots, u_M$, and c represents the underlying multinomial latent class variable.

1. Latent Class Analysis

In the LCA model, there are two types of parameters of interest: measurement and structural parameters. The measurement parameters describe the relationship between the observed latent class indicators and the latent class variables (i.e., the class-specific item endorsements probabilities, which are the distribution of the binary class indicators conditional on the latent class variable). In contrast, the structural parameters describe the multinomial distribution of the latent class variable (i.e., the proportion of cases in each latent class). We use the parameterization of the LCA model with binary indicators following notation presented in Nylund-Gibson and Masyn (2016) and Masyn (2017). Figure 1 provides a visual representation of an unconditional LCA model with M binary indicators. Each latent class indicator is observed on n individuals with u_{mi} representing individual i's response to class indicator m The latent class variable has K classes where $c_i = k$ when individual i belongs to Class k. The latent classes are mutually exclusive; therefore, individual i can only be assigned to one of K classes. The relationship between the observed class indicator variables and the latent class variable can be formulated with:

$$\Pr(u_{1i}, u_{2i}, \ldots, u_{Mi}) = \sum_{k=1}^{K} [\pi_k \cdot \Pr(u_{1i}, u_{2i}, \ldots, u_{Mi} | c_i = k)], \tag{1}$$

where π_k is a structural parameter representing the prevalence of individuals in Class k (i.e., class proportions). Considering the latent classes are mutually exclusive, $\sum \pi_k = 1$.

The measurement model for the latent class variable can be parameterized as the relationship between the observed class indicator u_m ($m = 1, 2, \ldots M$) and the latent class variable c, which can be formulated with:

$$\Pr(u_m = 1 | c = k) = \frac{1}{1 + \exp(\tau_{mk})}, \tag{2}$$

where τ_{mk} is the negative log odds of endorsing class indicator u_m given membership to latent class k. In other words, τ_{mk} is equal to -logit(E[$u_m | c = k$]). The class-specific item response probabilities suggest how likely an individual is to endorse a particular item given latent class membership.

For the structural model, the unconditional distribution of the multinomial latent class variable, c, can be parameterized with a multinomial logistic regression formulation. Specifically, the π_k parameters can be defined as intercepts on the inverse multinomial logit scale, such that:

$$\pi_k = \Pr(c = k) = \frac{\exp(\gamma_k)}{\sum_{j=1}^{K}\exp(\gamma_j)}. \quad (3)$$

The γ_k represents the log odds of membership in class k, given membership in either class k or K. For identification purposes, γ_{0K} is constrained to 0.

The LCA model assumes local independence, which suggests the M binary latent class indicators are uncorrelated conditional on class membership. In other words, latent class membership fully explains any correlations between observed class indicators. Software capable of LCA imposes the local independence assumption by default. By making the local independence assumption, Equation (1) is further simplified to:

$$\Pr(u_{1i}, u_{2i}, \ldots, u_{Mi}) = \sum_{k=1}^{K}\left[\pi_k \cdot \left(\prod_{m=1}^{M}\Pr(u_{mi}|c_i = k)\right)\right]. \quad (4)$$

Violating the local independence assumption can impact parameter estimates and model fit indices (Albert & Dodd, 2004; Asparouhov & Muthén, 2011; Lee et al., 2020; Torrance-Rynard & Walter, 1997; Vacek, 1985). In applied settings, the local independence assumption should be evaluated because it is possible to relax the assumption (i.e., allowing residual correlations between two or more latent class indicators in one or more latent classes), if necessary. For a detailed explanation of how to evaluate and relax the local independence assumption in popular mixture modeling software, see Visser and Depaoli (2022).

When estimating an LCA model, users must select the number of latent classes in the population. Applied users often lack prior knowledge about the number of latent classes. A statistical analysis procedure (i.e., class enumeration) can aid applied users in selecting the number of classes. During class enumeration, an iterative procedure is used to estimate several LCA models with a different number of specified classes. The best-fitting LCA model is then selected based on model fit and comparison indices, see Nylund et al. (2007) for a detailed explanation of class enumeration.

2. Latent Class Analysis with Covariates

In many practical applications of LCA, the LCA measurement model is used as part of a larger structural equation model (SEM). These models often include observed explanatory variables (i.e., covariates, predictors, independent variables, external variables, or concomitant variables[1]) that predict the latent class variable. For example, Quirk et al. (2013) extended their LCA model for kindergarten readiness to include several predictors (e.g., student's prior preschool experiences, age, language skills, and gender). The addition of these covariates allows researchers to explore research questions about why an individual was assigned to a particular latent class. A visual example of the latent class model with a covariate (i.e., conditional LCA model) can be seen in Figure 2. The covariate, x_1, can be categorical (Clogg & Goodman, 1985; Haberman, 1979; Hagenaars, 1990; Hagenaars, 1993; Vermunt, 1997) or continuous (Bandeen-Roche et al., 1997; Dayton & Macready, 1988; Kamakura et al., 1994; Yamaguchi, 2000).

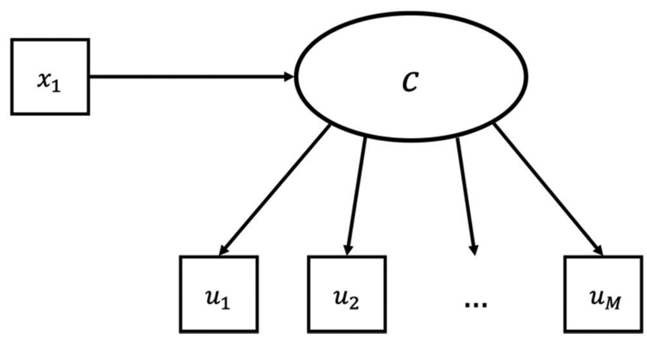

Figure 2. The conditional LCA model with M binary latent class indicators and a single covariate, x_1. The latent class indicators are represented with $u_1, u_2, \ldots, u_M$, and c represents the underlying multinomial latent class variable.

To include a covariate, the latent class model is combined with the latent class regression model into a joint model, which is typically estimated with the maximum-likelihood (ML) estimator. This approach is often referred to as the *one-step approach* in the methodological literature because the measurement model and structural model (i.e., the logistic regression in which the latent classes are related to the covariates) are simultaneously estimated in a single step (Asparouhov & Muthén, 2014; Bandeen-Roche et al., 1997; Dayton & Macready, 1988; Vermunt, 2010). More specifically, the latent class variable is regressed on the covariate using multinomial logistic regression parametrization (Nylund-Gibson & Masyn, 2016). Using notation first presented in Nylund-Gibson and Masyn (2016), the relationship between the LCA model and covariate x_i can be expressed as a multinomial logistic regression model:

$$\Pr(c_i = k|x_i) = \frac{\exp(\gamma_{0k} + \gamma_{1k}\ x_i)}{\sum_{j=1}^{K}\exp(\gamma_{0j} + \gamma_{1j}\ x_i)}, \quad (5)$$

where $\gamma_{0K} = \gamma_{1K} = 0$ for model identification. In Equation (5), the latent class indicator variables are considered independent of the covariate conditional on class membership. Therefore, Equation (4) can be adapted to include covariate x_i such that

$$\Pr(u_{1i}, u_{2i}, \ldots, u_{Mi}|x_i) = \sum_{k=1}^{K}\left[\Pr(c_i = k|x_i) \cdot \left(\prod_{m=1}^{M}\Pr(u_{mi}|c_i = k)\right)\right]. \quad (6)$$

When K number of classes have correctly been identified, the exclusion of x_i from the model has no impact on the

[1]A concomitant variable is a variable that is not the focus of the study, but the variable may influence variables of interest to the study (e.g., the dependent variable).

point estimates (i.e.,τ_{mk}) for each class indicator (i.e., $u_1, u_2, \ldots, u_M$). In other words, latent class membership will depend on x_i, but the class indicator responses should only depend on class membership. Thus, the covariate only has an indirect effect on the latent class indicators via the latent class variable.

The one-step approach may appear straightforward enough, but several drawbacks have been noted in the methodological literature. Specifically, Vermunt (2010) notes that the one-step approach is impractical when using many covariates, as is typical for exploratory studies. With each additional covariate, the LCA measurement model and the structural model must be estimated again. In addition, Vermunt (2010) highlights the model-building issues surrounding the inclusion of covariates. Applied users must decide whether to pick the number of latent classes before or after including covariates. Although covariates can help aid class enumeration when properly specified (Li & Hser, 2011; Lubke & Muthén, 2007; Muthén, 2002), applied researchers are unlikely to properly specify the covariates in the model without prior knowledge. Misspecifying the covariate relationships with the LCA measurement model can impact the class enumeration procedure, resulting in an over-extraction of the number of classes (Nylund-Gibson & Masyn, 2016). Therefore, several methodological studies suggest the number of latent classes should be established prior to including covariates (Collins & Lanza, 2010; Masyn, 2013; Petras & Masyn, 2010). Vermunt (2010) also notes that applied researchers do not find the joint model to be intuitive because they often wish to introduce covariates after classifying individuals. In addition, the applied researcher who establishes the latent class measurement model may not be the same researcher who is building the structural model.

To address some of the drawbacks of the one-step approach, methodologists have proposed a stepwise approach, a.k.a., *maximum likelihood (ML) three-step approach,* to estimation (Vermunt, 2010; Asparouhov & Muthén, 2014), where the latent class model and the relationship between the latent class variable and covariate are independently evaluated. By using a stepwise approach, the measurement model and structural models are decoupled, which can resolve many of the issues with the one-step approach. We follow Asparouhov and Muthén (2014) in describing the procedure for the ML three-step approach. First, the unconditional LCA model is estimated using only the observed class indicators. In the second step, the most likely class variable, MLC, is created using the latent class posterior distribution, which was produced during the estimation of the unconditional LCA model. The MLC is a nominal variable and for each observation, MLC is set to the class for which $\Pr(c_i = k|\boldsymbol{u}_i)$ is the largest (Asparouhov & Muthén, 2014); c_i is the latent class variable for the ith observation, and $\boldsymbol{u}_i$ represents the class indicator variables for the ith observation. The classification uncertainty rate for a true class given an MLC can be computed. For example, the classification uncertainty rate for Class 2 (C2) given MLC is Class 1 (C1), is expressed as follows:

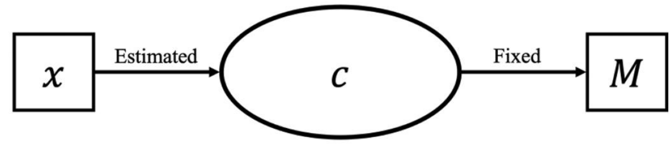

Figure 3. A visual of the ML three-step approach.
The latent class variable *c* is regressed on covariate *x*. The most likely class variable, *M*, is used as a single class indicator of the latent class variable *c*. The relationship between *M* and *c* is fixed, and the relationship between *x* and *c* is freely estimated. Note that this is the depiction only for the third step of estimation

$$p_{C1,C2} = \Pr(c = C2|MLC = C1) = \frac{1}{N_{C1}} \sum_{MLC_i = C1} \Pr(c_i = C2|\boldsymbol{u}_i), \quad (7)$$

where N_{C1} is the number of cases assigned to class C1 by the most likely class variable MLC, and MLC_i is the most likely class variable for the ith observation. The probability $\Pr(c_i = C2|\boldsymbol{u}_i)$ can be computed with the estimated unconditional LCA model from the first step.[2] For a K-class model, K × K classification uncertainty rates should be computed. After calculating the classification uncertainty with Equation (7), it is possible to calculate the classification measurement error. When the true class is C2 while MLC is C1, the classification measurement error is computed as

$$q_{C2,C1} = P(MLC = C1|c = C2) = \frac{p_{C1,C2}N_{C1}}{\sum_c p_{c,C2}N_c}, \quad (8)$$

where N_c is the number of observations classified in class c by the most likely class variable MLC. In this way, the most likely class variable can be treated as an imperfect measurement of c with measurement error $q_{C2,C1}$. The measurement error can then be transformed into logits using $\log(q_{C1,C2}/\ q_{K,C2})$, where Class K is used as a reference class. In the third step, the latent class variable is predicted by x through a multinomial regression while taking into account the K × K measurement errors from the previous step. Specifically, the most likely class variable MLC is used as a single, nominal class indicator of the latent class variable c. The logits are used as fixed parameter values that describe the direct relationship between the latent class variable and the most likely class variable. The multinomial regression of c on predictor x is freely estimated. A visual representation of the three-step approach can be seen in Figure 3, which was adapted from Asparouhov and Muthén (2014).

Ample simulation research has explored the performance of the ML three-step approach under different modeling conditions (e.g., Asparouhov & Muthén, 2014; Bakk et al., 2013; Nylund-Gibson et al., 2019; Vermunt, 2010). Results from these simulation studies suggest the three-step approach can produce unbiased parameter estimates if the

[2]The conditional probabilities for the class assignment given true latent class membership are automatically computed by M*plus* when estimating an LCA model. See Asparouhov and Muthén (2014) and Vermunt (2010) for more details on how to compute the conditional probabilities for the ML three-step approach. In the Vermunt (2010) article, the ML three-step approach is referred to as Modal ML.

LCA measurement model has sufficient class separation. In mixture modeling, class separation refers to how distinct the latent classes are from one another. When class separation is poor, it can be trickier to properly assign cases to latent classes, resulting in increased measurement error in the latent class variable.

Statistical software capable of mixture modeling, such as M*plus* (Muthén & Muthén, 1998-2017)], and Latent GOLD (Vermunt & Magidson, 2025), has largely automated the three-step approach, allowing applied researchers to implement the procedure much more easily. Although this automation can be helpful, the automation limits the user's ability to adjust how the model is estimated. Specifically, the automation limits the user's ability to address missing data, and the estimator is limited to ML. When "manually" implementing the three-step approach instead, the user has a greater ability to adjust how the model is estimated. For example, the direct covariate effects can be detected and modeled in the third step.

3. Bayesian Latent Class Analysis

The previous section discussed estimation strategies available for the conditional LCA models in the frequentist framework (e.g., one-step approach, three-step approach). An alternative method for estimating LCA models is to use Bayesian estimation. In recent years, the Bayesian estimation framework has become increasingly popular as statistical software has made it more accessible to applied researchers (van de Schoot et al., 2017). The primary distinction between the frequentist and Bayesian estimation is the addition of prior distributions in the model. The prior distributions (or priors) represent what a parameter in the model should look like based on a prior belief about the relationship. For every parameter estimated in the model, it is possible to specify a prior distribution that describes these prior beliefs. These prior distributions are incorporated into the estimation process and can provide information about the parameters in the model.

Depending on the certainty of a researcher's prior information, it is possible to specify a prior with varying degrees of informativeness. For example, if a researcher has very specific knowledge about the parameter, the researcher can specify a narrower prior. However, if the researcher is uncertain about what a parameter looks like, the researcher can set a less informative prior. The degree of informativeness about a prior can be set with hyperparameters, the parameters that compromise a probability distribution. When a prior is narrow and contains specific information about the parameter, it is considered an informative prior. A wider, less informative prior is called a noninformative (or diffuse) prior. The information (or lack of information) specified within the prior is incorporated with the data during the estimation process. In mixture models, the ability to incorporate accurate prior knowledge about class-specific parameters and class proportions can greatly improve estimation (Depaoli, 2013, 2014; Lu et al., 2011). Bayesian estimation offers several advantages when addressing model estimation issues, such as model assumption violations (Asparouhov & Muthén, 2011; Bauer, 2007), convergence to local maxima (Hipp & Bauer, 2006), and inaccurate parameter estimates (Depaoli, 2013).

One important concept in latent class modeling is how separated the latent classes are from one another at the population level. When the latent classes are difficult to distinguish from one another (i.e., poor class separation), it can be less clear which latent class a particular case belongs to. In addition, it may not be obvious how many latent classes are present in the population. Estimating the class-specific parameters can be much more difficult when class separation is poor. When using Bayesian estimation, it is possible to incorporate prior knowledge about the latent classes, which can be a helpful tool for accurately estimating class-specific parameters. In contrast, in the frequentist estimation framework, one of the only viable options for overcoming these estimation challenges is to collect a much a larger sample (Depaoli, 2013, 2014; Lu et al., 2011).

Another source of estimation issues in latent class models is the relative size of the latent classes. When a latent class is small relative to the other latent classes (e.g., Class 1 = 18% vs. Class 2 = 82%), parameters specific to the minority class are much more difficult to estimate (Depaoli, 2013, 2014; Lu et al., 2011; Tueller & Lubke, 2010). Bayesian estimation can be a useful tool for incorporating prior knowledge about the relative size of the latent classes in the model. By incorporating accurate prior distributions about the relative size of the latent classes, the model is better able to identify and accurately estimate small latent classes (Depaoli et al., 2017).

4. Existing Techniques for Detecting Direct Covariate Effects in LCA

A key challenge in conditional LCA is correctly specifying the effects of covariates. Covariate misspecification, including the exclusion of true direct effects between the covariate and latent class indicators, can lead to biased structural parameters estimates (Asparouhov & Muthén, 2014; Janssen et al., 2019). The existing methodological research on detecting direct effects in conditional LCA models is limited to one previous simulation study (Janssen et al., 2019). The authors illustrated how both residual and fit statistics could be used to identify direct effects in the ML estimation framework. Residual statistics can test for potentially problematic restrictions in the model. For example, residual statistics could be used to test whether the correlation between a pair of class indicators should be freed. In addition, residual statistics can be used to determine whether the path between the covariate and a class indicator should be freed. For a detailed explanation of implementing residual statistics in M*plus* and Latent GOLD, see Visser and Depaoli (2022). Alternatively, an inferential method (e.g., Wald Test) could also be used to test whether the inclusion of a direct effect improves model fit. According to Janssen et al. (2019), the effectiveness of the residual and fit statistics at detecting direct effects largely depends on the number of direct effects present, the size of the direct effects, and whether the direct effects were class-specific. Although the detection methods in Janssen et al. (2019) were promising, the methods struggled to detect multiple direct effects.

Another study proposed the use of multiple indicator multiple cause (MIMIC) modeling procedures that are commonly used in the item-response theory (IRT) framework to detect differential item functioning (DIF) in conditional LCA models (Masyn, 2017). In the context of LCA models with a class-predicting covariate, DIF occurs in two main ways (Nylund-Gibson & Masyn, 2016). First, the latent class indicator could function differently for individuals with different covariate values. Second, the probability of endorsing the latent class indicator could correspond to a particular difference in the covariate for an individual in a specific latent class. The second type of DIF would occur when there is a class-varying direct effect. An abundance of past methodological work suggests the MIMIC modeling procedure to detecting DIF in factor analysis models is effective (Finch, 2021; Muthén, 1985; Wang et al., 2009; Willse & Goodman, 2008; Woods, 2009; Woods & Grimm, 2011). Masyn (2017) expanded on this methodological work to develop a similar MIMIC modeling procedure for LCA. The proposed LCA MIMIC modeling procedure is an iterative procedure that could be used in conjunction with a stepwise approach to estimation; see Masyn (2017) for a detailed explanation of how to implement the proposed LCA MIMIC modeling procedure. This new LCA MIMIC modeling procedure's performance has yet to be evaluated in a simulation study. However, Masyn (2017) points out that there may be Type I error inflation due to the iterative testing procedure used to test each possible direct effect.

5. Small-Variance Priors

An alternative strategy for detecting non-zero direct effects in conditional LCA models that may be effective for a broader range of situations is Bayesian structural equation modeling (BSEM). BSEM allows for a very flexible modeling experience, where it is possible to relax model assumptions in restrictive models. Ample methodological research has demonstrated the value of using BSEM to relax model assumptions for confirmatory factor analysis (CFA) models (Muthén & Asparouhov, 2012; Stromeyer et al., 2015; Xiao et al., 2019). For example, BSEM has been used to relax model assumptions about cross-loadings in CFA, residual correlations in CFA, and measurement non-invariance in MIMIC modeling (Muthén & Asparouhov, 2012). In addition, BSEM has been used to relax assumptions about measurement invariance in CFA models (Hoijtink & van de Schoot, 2018; Sedding & Leitgöb, 2018; Winter & Depaoli, 2020). The Bayesian methodology allows us to relax model assumptions by implementing *approximate-zero priors*, which are near-zero priors centered on zero with very small variances. By applying approximate-zero priors to parameters that are typically constrained to zero in the ML estimation framework, a less restrictive version of the model can be estimated, improving model fit and interpretation (Depaoli, 2021).

Approximate-zero priors can also help detect non-zero direct effects. For example, BSEM has handled direct effects between a covariate and factor indicator variables in CFA models (Muthén & Asparouhov, 2012). In traditional CFA models, the direct effects between covariates and the factor indicators are typically fixed to zero, introducing problematic model misspecifications. Using small-variance, normal priors centered at zero for all direct effects between the covariate and the factor indicators, it is possible to relax this strict assumption and say the direct effects are approximately zero. The near-zero priors will allow some "wiggle" room surrounding the cross-loadings, but the substantive meaning remains (e.g., the cross-loadings are minor and unimportant). If a cross-loading is truly non-zero, the data should conflict with the near-zero priors, resulting in a non-zero estimate for the cross-loading. The approximate-zero strategy can improve model fit and aid researchers in detecting non-zero cross-loadings (Muthén & Asparouhov, 2012).

Previous applications of approximate-zero priors in LCA models have been limited to relaxing assumptions about local independence between latent class indicators (Asparouhov & Muthén, 2011; Lee et al., 2020). More specifically, near-zero priors can help detect non-zero residual correlations between latent class indicators in LCA models. For a detailed guide for detecting and modeling conditional dependence in LCA, see Visser and Depaoli (2022). The current study seeks to expand approximate-zero priors' application to include direct effects between a continuous covariate and latent class indicators. Considering the approximate-zero strategy is effective for direct effects in CFA models (Muthén & Asparouhov, 2012), we anticipate similar benefits will be seen in LCA models.

6. Simulation Study

6.1. Population Models

This study aimed to illustrate how small-variance priors can effectively identify and model non-zero direct effects that commonly occur in applied settings. Random samples were generated from six different population models with known covariate relationships that commonly occur in applied research to investigate the utility of the approximate-zero strategy. All population models were generated with two latent classes ($K = 2$), five binary class indicator variables ($u_1 - u_5$), and one standard normal covariate (x). As seen in Figure 4, LCA population models with six known covariate effect specifications for x were considered (population models 1-6, labeled P1-P6). Across the population models, there were two pathways through which the covariate x could influence an observed latent class indicator. The first pathway is indirect via the latent class variable c, as seen in the P1-P4 population models. The second pathway is a direct pathway that bypasses c, as seen in P2-P6 population models. Some population models have indirect and direct effects, whereas others have only indirect or direct effects. All data were generated using M*plus*.

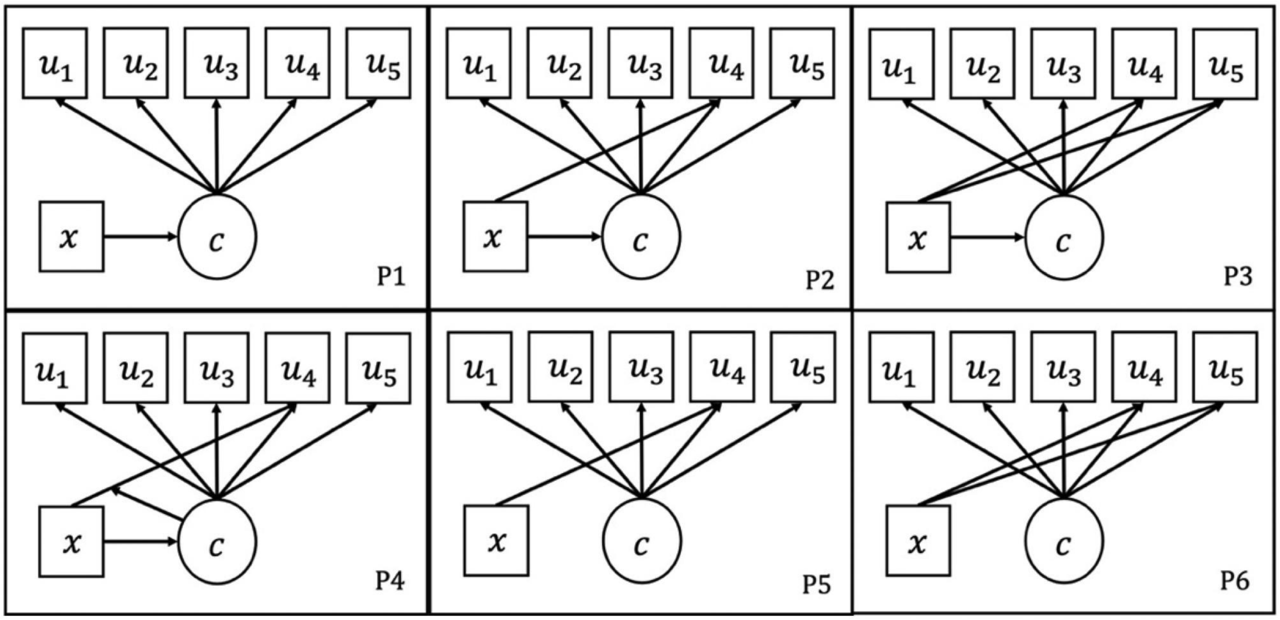

Figure 4. Conditional LCA population models in the current study (P1-P6). The paths from x to c represent the indirect effects of the covariate on the class indicators (i.e., u_1-u_5) via the latent class variable, c. Each path from x to an observed class indicator (e.g., u_4) represents a direct effect. In P4, the arrow pointing from c to the direct effect path indicates a class-varying direct effect.

6.1.1. Indirect Effects

P1, the first covariate effect specification in Figure 4, represents the most common conditional LCA model, with only an indirect effect on the latent class indicators via the latent class variable c (Nylund-Gibson & Masyn, 2016). The relationship between x and c can be represented with a single regression path (i.e., γ_1). In the P2, P3, and P4 population models, x was also specified to have an indirect effect on the latent class indicators via the latent class variable c. For population models P1-P4, the effect of x on c (i.e., specified as "c on x" in M*plus* language) was fixed to $\gamma_1 = 1$, which corresponds to an odds ratio of 2.72 for membership in Class 1 compared to Class 2 for a 1 standard deviation (SD) increase in x. The fifth and sixth covariate effect specifications in Figure 4, P5 and P6, differ from the other population models because there was no relationship between x and c (i.e., $\gamma_1 = 0$).

6.1.2. Direct Effects

The P2-P6 population models also have direct effects between x and one or more latent class indicators. Expressly, P2 and P5 were specified to have a single direct effect between x and u_4, whereas P3 and P6 were specified to have direct effects between x and u_4 as well as x and u_5. P4 was very similar to the P2 population model with a single direct effect between x and u_4. However, in the P4 covariate specification, the path between x and u_4 was class-varying. The class-varying effect is represented in Figure 4 with the arrow pointing from c to the path between x and u_4. All the population models (excluding P1) have a direct effect of x on u_4, which can be represented with the regression coefficient β_4. Across population models, $\beta_4 = 1$, which corresponds to an odds ratio of 2.72 for item endorsement of u_4 for a 1 SD increase in x, given latent class membership. In population model P4, the class-varying direct effect of x on u_4 was specified such that the Class 1 direct effect was fixed to $\beta_{41} = 0.5$, and the Class 2 direct effect was fixed to $\beta_{42} = 1.5$. Population models P3 and P6 also had a direct effect of x on u_5, which was set to $\beta_5 = 1$. Notably, both P3 and P6 represent a violation of the local independence assumption because covariate x was a shared antecedent of u_4 and u_5 above and beyond the latent class variable.

6.2. Design Factors

For each population model, class proportions (equal and unequal) and sample size ($N = 500$ and $N = 1{,}000$) were varied. Each generated dataset was analyzed with several Bayesian conditional LCA models with different prior specifications. To explore the impact of prior specifications (or misspecifications), in the analysis model we included different prior conditions on γ_1 (i.e., x predicting c) and the direct effects (i.e., x predicting $u_1 - u_5$). The prior specifications for each population are discussed in detail below. There were 324 cells in this simulation study, and each cell had 500 replications, as previous studies using conditional LCA models have found 500 replications to be sufficient (Janssen et al., 2019; Nylund-Gibson & Masyn, 2016).

6.2.1. Class Proportions

Data were generated according to two different class proportion conditions (equal vs. unequal). For the equal class proportion condition, the classes had a 50%-50% split ($\pi_1 = \pi_2 = 0.5$). For the unequal class proportion condition, the classes had an 82%-18% split ($\pi_1 = 0.82, \pi_2 = 0.18$). Considering unequal class sizes pose a greater estimation challenge in mixture models (Kim, 2014; Lubke & Tueller,

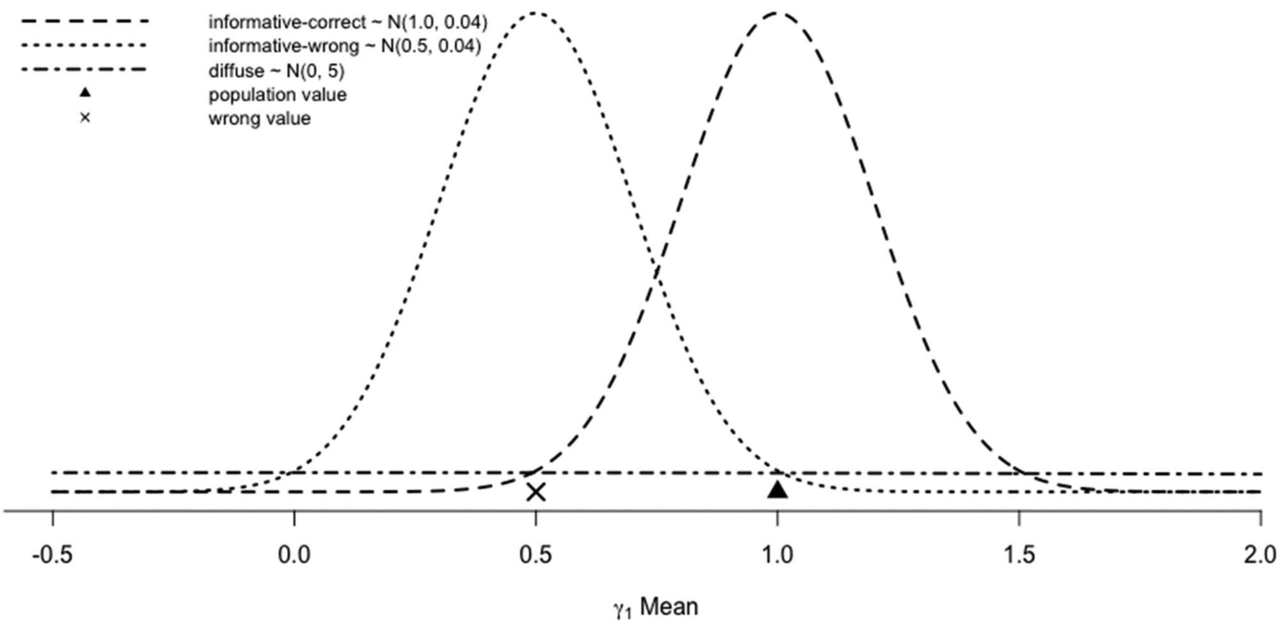

Figure 5. The prior specification levels (informative-correct, informative-wrong, and diffuse) for γ_1 in population models P1-P4.

2010; Nylund et al., 2007), this factor could be important for illustrating the value of informative prior specifications for the binomial logistic regression slope. The population values selected were equivalent to a previous simulation study that used conditional LCA models (Nylund-Gibson & Masyn, 2016).

6.2.2. Sample Size

Two levels of sample size were considered in the current study ($N = 500$ vs. $N = 1{,}000$). The $N = 500$ is a common sample size for mixture models in applied settings (Sterba, 2014), whereas $N = 1{,}000$ represents an ideal scenario. These sample sizes have previously been used in a simulation study using conditional LCA models (Nylund-Gibson & Masyn, 2016).

6.2.3. Prior Specifications

This study assumes that the 2-class unconditional LCA model has already been correctly selected during class enumeration. Our primary aim is to accurately model the covariate relations with the measurement model (e.g., latent class variable and class indicators). For each replication from each population model in Figure 4, we estimated a series of LCA analysis models with different covariate effects and prior specifications. We varied the prior specifications for γ_1 and the direct effects (e.g., β_1-β_5). The levels for the γ_1 prior specifications and the direct effect prior specifications depended on the population model used for data generation. All analyses were performed in M*plus* using the Bayesian estimator with a single MCMC chain per parameter. Each analysis model used 30,000 iterations, and the first 15,000 iterations were discarded as burn-in. Convergence was assessed by carefully examining trace plots and autocorrelation plots and monitoring the potential scale reduction factor (PSRF). The PSRF is a diagnostic tool to assess the convergence of MCMC, with a value less than 1.01 indicating that the chains have converged, and the estimates are reliable. The M*plus* default priors were utilized for all other parameters in the analysis models. More specifically, the default prior specifications implement a normal prior $\sim N(0, 5)$ for the individual item thresholds and the intercept of the latent class variable.

6.2.3.1. Prior Specifications of γ_1. Depending on the population model used for data generation, there may or may not be an indirect effect. For population models with an indirect effect (e.g., P1-P4), three levels of normal priors (informative-correct, informative-wrong, and diffuse) were considered for γ_1. Figure 5 provides a visual representation of the three levels of priors used for γ_1 in P1-P4. The informative-correct conditions assigned a normal prior $\sim N(1, 0.04)$ for γ_1, which is an informative prior correctly centered on 1. The informative-wrong conditions assigned a normal prior $\sim N(0.5, 0.04)$ for γ_1, which is an informative prior incorrectly centered on 0.5. The diffuse conditions assigned a $\sim N(0, 5)$ for γ_1, which is the Mplus default prior.

For population models without an indirect effect (i.e., P5 and P6), we explored the impact of misspecifying γ_1 in the analysis model. More specifically, we considered three different specifications for γ_1 (not specified, informative misspecification, and diffuse misspecification). The γ_1 parameter was not included in the analysis model in the *not specified* conditions, and no prior was assigned. For the γ_1 misspecification conditions, the normal prior $\sim N(0.5, 0.04)$ was assigned for the *informative misspecification* condition and the M*plus* default prior was used for the $\sim N(0, 5)$ *diffuse misspecification* condition.

6.2.3.2. Prior Specifications of Direct Effect. All population models were analyzed with two levels of small-variance

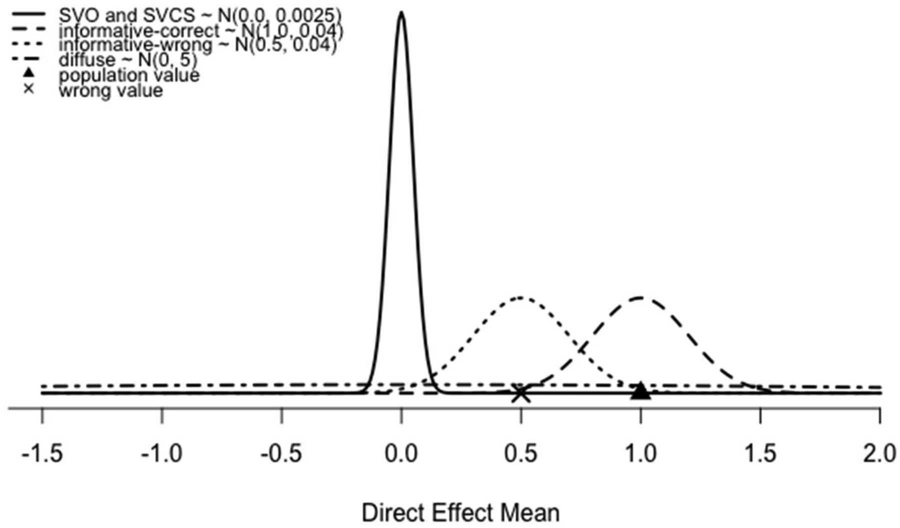

Figure 6. The prior specification levels (SVO, SVCS, informative-correct, informative-wrong, and diffuse) for the direct effects (e.g., β_4) in population models P2, P3, P5, and P6.

priors (overall and class-specific) for the direct effects of $u_1 - u_5$ on x, which can be represented with regression coefficients β_1-β_5. For the small-variance *overall* prior level, β_1-β_5 were assigned a normal prior $\sim N(0.0, 0.0025)$. For the small-variance class-specific prior level, the direct effects were estimated as class-varying effects (i.e., Class 1 = β_{11}-β_{51}; Class 2 = β_{12}-β_{52}), and each regression coefficient was assigned a normal prior $\sim N(0.0, 0.0025)$. Conditions utilizing the small-variance overall (SVO) priors will likely have fewer estimation problems (Nylund-Gibson & Masyn, 2016); however, the small-variance class-specific (SVCS) priors are essential for detecting class-varying effects such as β_{41} and β_{42} in P4. The small-variance prior conditions aim to determine if the approximate-zero strategy effectively detects non-zero direct effects. We would expect the regression coefficients that are truly zero in the population model to be estimated close to zero and the regression coefficients that are non-zero in each population model to "escape" the restrictive small-variance prior (especially when sample sizes are relatively higher).

In addition to the two levels of small-variance priors, the P2-P6 population models had three additional levels of direct effect priors (informative-correct, informative-wrong, and diffuse). These three additional levels illustrate how a more parsimonious model can be estimated after detecting the non-zero direct effects with the small-variance priors. For these three levels of priors, only truly non-zero direct effects were included in the analysis models. Population models P2, P3, P5, and P6 included the direct effect β_4, and population models P3 and P6 also had direct effect β_5. The *informative-correct* prior level assigned $\sim N(1.0, 0.04)$ to each non-zero direct effect. In contrast, *informative-wrong* prior level assigned $\sim N(0.5, 0.04)$ to each non-zero direct effect. The informative-correct prior is likely to be effective because it will be centered on the population value; however, this level of accuracy in the prior is unlikely to occur in practice. Therefore, we included an informative-wrong prior condition where the prior is centered on a wrong value (i.e., 0.5). The wrong value represents a weaker direct effect, which may be tempting because the estimated direct effect will be pulled towards zero under a small-variance prior. It is important to understand the impact of wrong informative prior specifications because, in most applications, the researcher is likely to be off from the "truth" (i.e., population value). The proposed wrong prior level mimics a situation that is likely to occur in applied research. In cases where the researcher does not want to include informative priors (e.g., no previous knowledge available, exploratory analysis), a common practice is to use the default prior in the software. To assess the viability of this estimation strategy, we included the direct effect *diffuse* prior level, which is the M*plus* default prior $\sim N(0, 5)$. Figure 6 provides a visual representation of the five levels of direct effect priors (SVO, SVCS, informative-correct, informative-wrong, and diffuse) for population models P2, P3, P5, and P6.

The P4 population model is unique because it has a class-varying direct effect ($\beta_{41} = 0.5$ and $\beta_{42} = 1.5$). To adjust for the class-varying direct effect in the analysis models, the *informative-correct* prior level was adapted to estimate the class-specific direct effect of x on u_4. The Class 1 (C1) regression coefficient β_{41} was assigned $\sim N(0.5, 0.04)$, and the Class 2 (C2) regression coefficient β_{42} was assigned $\sim N(1.5, 0.04)$. Figure 7 illustrates the adapted informative-correct prior condition.

7. Results

7.1. Convergence

To prevent within-chain label switching across replications of the simulation study, a model constraint was included on the latent class indicator u_1 such that the values adhered to the following order: Class $2 >$ Class 1. A single MCMC chain was utilized for parameter estimation to prevent between-chain label switching. The number of iterations was set to 30,000 for all analyses, and the first 15,000 iterations were discarded as burn-in. Each cell in the simulation study

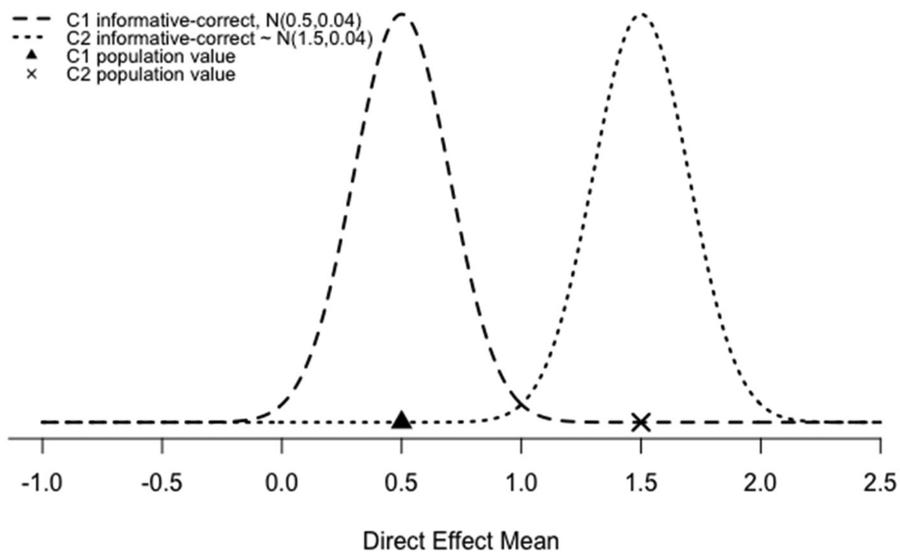

Figure 7. The direct effect prior specifications for the informative-correct conditions (C1 informative-correct, C2 informative-correct) in population model P4.

converged without issue and a set of stable estimates for the model parameters was obtained. Convergence for each replication was assessed using PSRF. According to the "less than 1.01" criterion, all replications converged as expected.

7.2. Detecting Direct Effects with Small-Variance Priors

In this section, we examine the power of small-variance priors to detect non-zero direct effects between the covariate and the latent class indicators. If a non-zero direct effect escapes the small-variance prior that is centered on zero, there is evidence to suggest a direct effect may need to be included when estimating the relationship between x and c. If the direct effect is truly zero, it should be held to "approximately zero" by the small-variance prior. Small-variance priors were specified on the $\beta_1 - \beta_5$ parameters for each of the six population models (P1-P6). P1 results were included to illustrate whether an inflated Type I error is likely to occur when there is no direct effect in the population model. P2 and P3 results explored whether the performance of the small-variance priors was impacted by the number of direct effects (one vs. two). P4 results were included to determine whether a class-varying direct effect can be detected with small-variance priors. P5 and P6 were included to examine the impact of misspecifying γ_1 when detecting direct effects with small-variance priors. Factors such as sample size and class sizes could impact the power to detect the direct effects when using small-variance prior.

Results from P1-P6 are presented in Tables 1–6, respectively; see Figure 4 for a visual representation of the population models. For each population model, two types of small-variance priors (SVO vs. SVCS) were applied to the direct effects of $u_3 - u_5$ on x, which can be denoted as $\beta_3 - \beta_5$. The SVO prior represents a small-variance prior applied to the overall direct effect, whereas the SVCS prior is a class-specific small-variance prior. In addition to the small-variance priors for the direct effect, three different levels of priors (informative-correct, informative-wrong, and diffuse) were applied to γ_1.

Results presented in Tables 1–6 include the average parameter estimates for three of the direct effects for each latent class (i.e., $\beta_3 - \beta_5$) and γ_1. In addition, the tables display the percentage of replications in which the null hypothesis (e.g., $\beta_5 = 0$) is rejected. When the population value is non-zero, this percentage represents an estimate of power for a single parameter (i.e., the probability of rejecting the null hypothesis when it is false). A cut-off of 80% was used as the standard for power. When the population value is truly zero, this percentage represents an estimate of the Type I error (i.e., the probability of rejecting the null hypothesis when it is true). We would expect a Type I error (i.e., false positive) rate of 5% or less when a population parameter equals zero.

7.2.1. P1 – No Direct Effects

Table 1 displays the mean coefficient estimate and percentage of replications with a non-zero coefficient in each condition in population model P1, which has an indirect effect and no direct effects. We define a covariate as having a non-zero (i.e., detectable) direct effect if the 95% Bayesian credible interval for its estimated coefficient does not include zero. The left side of the table contains the results for SVO priors, and the right side of the table has results for SVCS. Table 1 includes two different sample size conditions ($N = 500$ vs. $N = 1,000$) and two different class sizes (equal vs. unequal). The aim of including the P1 population model was to demonstrate that small-variance priors do not result in an inflated Type I error rate for the truly zero direct effects. As expected, the small-variance priors (SVO and SVCS) consistently produced approximately zero estimates for the truly zero direct effects, regardless of the sample size, class sizes, and prior specification. The number of false positives was acceptable (<5%) for the truly zero direct effects, regardless of condition. The γ_1 regression coefficient was unbiased in conditions with an informative-correct

Table 1. The average parameter estimates and % of non-zero (detectable) coefficients for P1 parameters when using two levels of small variance priors (SVO and SVCS) on β_3-β_5 and three levels of priors (diffuse, informative-correct, and informative-wrong) on γ_1.

Sample Size		$N = 500$				$N = 1,000$				$N = 500$				$N = 1,000$			
Class Prop.		50%/50%		82%/18%		50%/50%		82%/18%		50%/50%		82%/18%		50%/50%		82%/18%	
Parameter	Pop. Value	Est.	% non-zero	Est.	% non-zero	Est.	% non-zero	Est.	% non-zero	Est.	% non-zero	Est.	% non-zero	Est.	% non-zero	Est.	% non-zero
		(SVO) β_1-β_5~N(0,0.0025) & γ_1~N(0,5)								(SVCS) β_1-β_5~N(0,0.0025) & γ_1~N(0,5)							
$C1 : \beta_3$	0	.000	0	.001	0	.000	0	.000	0	.001	0	−.005	0	.001	0	−.009	0
β_4	0	.000	0	.001	0	−.001	0	−.001	0	.000	0	−.005	0	−.001	0	−.010	0
β_5	0	.001	0	.001	0	.001	0	.000	0	.000	0	−.005	0	−.001	0	−.010	0
$C2 : \beta_3$	0	.000	0	.001	0	.000	0	.000	0	−.002	0	.006	0	−.001	0	.013	0
β_4	0	.000	0	.001	0	−.001	0	−.001	0	−.001	0	.008	0	.000	0	.014	0
β_5	0	.001	0	.001	0	.001	0	.000	0	−.001	0	.007	0	.002	0	.016	0
γ_1	1	1.020	100	1.011	100	1.008	100	1.003	100	1.012	100	1.011	100	1.008	100	1.002	100
		(SVO) β_1-β_5~N(0,0.0025) & γ_1~N(1,0.04)								(SVCS) β_1-β_5~N(0,0.0025) & γ_1~N(1,0.04)							
$C1 : \beta_3$	0	.000	0	.001	0	.000	0	.000	0	.001	0	−.005	0	.001	0	−.009	0
β_4	0	.000	0	.001	0	−.001	0	−.001	0	.000	0	−.005	0	−.001	0	−.010	1
β_5	0	.001	0	.001	0	.001	0	.000	0	.000	0	−.005	0	.000	0	−.010	0
$C2 : \beta_3$	0	.000	0	.001	0	.000	0	.000	0	−.002	0	.007	0	−.001	0	.013	0
β_4	0	.000	0	.001	0	−.001	0	−.001	0	−.001	0	.008	0	.000	0	.014	0
β_5	0	.001	0	.001	0	.001	0	.000	0	.001	0	.007	0	.002	0	.014	0
γ_1	1	1.012	100	1.005	100	1.006	100	1.003	100	1.013	100	1.005	100	1.006	100	1.001	100
		(SVO) β_1-β_5~N(0,0.0025) & γ_1~N(0.5,0.04)								(SVCS) β_1-β_5~N(0,0.0025) & γ_1~N(0.5,0.04)							
$C1 : \beta_3$	0	.003	0	.004	0	.003	0	.003	1	.004	0	−.003	0	.003	0	−.007	0
β_4	0	.003	0	.004	0	.002	0	.003	0	.002	0	−.003	0	.001	0	−.008	0
β_5	0	.005	0	.004	0	.004	1	.003	0	.002	0	−.003	0	.001	0	−.008	0
$C2 : \beta_3$	0	.003	0	.004	0	.003	0	.003	1	.000	0	.008	0	.001	0	.014	0
β_4	0	.003	0	.004	0	.002	0	.003	0	.001	0	.009	0	.002	0	.015	0
β_5	0	.005	0	.004	0	.004	1	.003	0	.003	0	.009	0	.004	0	.016	0
γ_1	1	.872	100	.832	100	.922	100	.895	100	.873	100	.833	100	.923	100	.895	100

prior or diffuse prior on γ_1. In contrast, the informative-wrong prior on γ_1 resulted in underestimation on the direct effect.

7.2.2. P2 – One Direct Effect

Table 2 provides the mean coefficient estimate and percentage of replications with a non-zero (detectable) coefficient in each condition in population model P2, which has an indirect effect and one direct effect β_4. Again, we define a covariate as having a non-zero (i.e., detectable) direct effect if the 95% Bayesian credible interval for its estimated coefficient does not include zero. The left side of the table contains the results for SVO priors, and the right side of the table has results for SVCS. Table 2 includes two different sample size conditions ($N = 500$ vs. $N = 1,000$) and two different class sizes (equal vs. unequal). The P2 population model represents an ideal modeling situation with a single direct effect.

When using SVO priors, the non-zero direct effect β_4 was consistently detected as non-zero, regardless of the sample size, class size, and the γ_1 prior. Although β_4 was underestimated, there was enough power to overcome the restrictive SVO prior, as evidenced by 100% of the replications having a non-zero (detectable) coefficient (credible interval does not cover zero). The truly zero direct effects, β_3 and β_5, had an acceptable level of false positives (<5%). When using the SVCS priors, the situation was more complicated. In conditions with equal class sizes, both latent classes had an acceptable level of power to detect β_4. In conditions with unequal class sizes, the non-zero direct effect was consistently detected in the majority class (C1), but there was inadequate power (<80%) to detect the direct effect in the minority class (C2). Across small-variance prior conditions (SVO vs. SVCS), there was adequate power to detect γ_1.

7.2.3. P3 – Two Direct Effects

Table 3 provides the mean coefficient estimate and percentage of replications with a non-zero (detectable) coefficient in each condition in population model P3, which has an indirect effect and two direct effects, β_4 and β_5. Again, we define a covariate as having a non-zero (i.e., detectable) direct effect if the 95% Bayesian credible interval for its estimated coefficient does not include zero. The left side of the table contains the results for SVO priors, and the right side of the table has results for SVCS. Table 3 includes two different sample size conditions ($N = 500$ vs. $N = 1,000$) and two different class sizes (equal vs. unequal). The P3 population model includes a local independence assumption violation because the latent class indicators u_4 and u_5 are related via the covariate x.

The local independence assumption violation introduces complications when using small-variance priors. When using the SVO priors, there was enough power to detect the non-zero direct effects (β_4 and β_5), regardless of sample size, class sizes, and γ_1 prior specification. However, in conditions with $N = 1,000$, there was an inflated Type I error rate (>5%) for the truly zero direct effects. As the sample size increased, the SVO prior struggled to hold truly zero direct effects to approximately zero when there was a local independence assumption violation. According to the right side of Table 3, the SVCS priors further complicated results.

Table 2. The average parameter estimates and % of non-zero (detectable) coefficients for P2 parameters when using two levels of small variance priors (SVO and SVCS) on β_3-β_5 and three levels of priors (diffuse, informative-correct, and informative-wrong) on γ_1.

Sample Size		$N = 500$				$N = 1,000$				$N = 500$				$N = 1,000$			
Class Prop.		50%/50%		82%/18%		50%/50%		82%/18%		50%/50%		82%/18%		50%/50%		82%/18%	
Parameter	Pop. Value	Est.	% non-zero	Est.	% non-zero	Est.	% non-zero	Est.	% non-zero	Est.	% non-zero	Est.	% non-zero	Est.	% non-zero	Est.	% non-zero
		(SVO) β_1-β_5~N(0,0.0025) & γ_1~N(0,5)								(SVCS) β_1-β_5~N(0,0.0025) & γ_1~N(0,5)							
$C1:\beta_3$	0	−.020	0	−.016	0	−.026	2	−.020	1	−.013	0	−.022	0	−.020	0	−.032	3
β_4	1	.196	100	.219	100	.320	100	.351	100	.110	80	.185	100	.197	100	.303	100
β_5	0	−.020	0	−.014	0	−.024	3	−.017	1	−.012	0	−.020	0	−.018	0	−.029	3
$C2:\beta_3$	0	−.020	0	−.016	0	−.026	2	−.020	1	−.013	0	.004	0	−.020	0	.009	0
β_4	1	.196	100	.219	100	.320	100	.351	100	.111	82	.050	1	.197	100	.098	54
β_5	0	−.020	0	−.014	0	−.024	3	−.017	1	−.013	0	.004	0	−.019	0	.010	0
γ_1	1	1.183	100	1.192	100	1.142	100	1.148	100	1.209	100	1.228	100	1.174	100	1.190	100
		(SVO) β_1-β_5~N(0,0.0025) & γ_1~N(1,0.04)								(SVCS) β_1-β_5~N(0,0.0025) & γ_1~N(1,0.04)							
$C1:\beta_3$	0	−.019	0	−.014	0	−.025	2	−.019	1	−.012	0	−.021	0	−.019	0	−.031	3
β_4	1	.198	100	.222	100	.322	100	.353	100	.112	82	.188	100	.198	100	.305	100
β_5	0	−.018	0	−.013	0	−.023	3	−.016	2	−.011	0	−.019	0	−.018	0	−.028	3
$C2:\beta_3$	0	−.019	0	−.014	0	−.025	2	−.019	1	−.013	0	.005	0	−.020	0	.009	0
β_4	1	.198	100	.222	100	.322	100	.353	100	.113	86	.052	1	.199	100	.099	56
β_5	0	−.018	0	−.013	0	−.023	3	−.016	2	−.012	0	.005	0	−.019	0	.010	0
γ_1	1	1.120	100	1.109	100	1.113	100	1.110	100	1.136	100	1.129	100	1.139	100	1.141	100
		(SVO) β_1-β_5~N(0,0.0025) & γ_1~N(0.5,0.04)								(SVCS) β_1-β_5~N(0,0.0025) & γ_1~N(0.5,0.04)							
$C1:\beta_3$	0	−.015	0	−.011	0	−.021	2	−.015	1	−.010	0	−.018	0	−.017	0	−.028	2
β_4	1	.203	100	.227	100	.326	100	.357	100	.112	88	.192	100	.202	100	.308	100
β_5	0	−.014	0	−.009	0	−.020	2	−.013	1	−.009	0	−.016	0	−.016	0	−.026	2
$C2:\beta_3$	0	−.015	0	−.011	0	−.021	2	−.015	1	−.011	0	.006	0	−.017	0	.011	0
β_4	1	.203	100	.227	100	.326	100	.357	100	.116	89	.054	1	.202	100	.102	63
β_5	0	−.014	0	−.009	0	−.020	2	−.013	1	−.011	0	.006	0	−.017	0	.011	0
γ_1	1	.965	100	.923	100	1.019	100	.993	100	.980	100	.941	100	1.043	100	1.022	100

Table 3. The average parameter estimates and % of non-zero (detectable) coefficients for P3 parameters when using two levels of small variance priors (SVO and SVCS) on β_3-β_5 and three levels of priors (diffuse, informative-correct, and informative-wrong) on γ_1.

Sample Size		$N = 500$				$N = 1,000$				$N = 500$				$N = 1,000$			
Class Prop.		50%/50%		82%/18%		50%/50%		82%/18%		50%/50%		82%/18%		50%/50%		82%/18%	
Parameter	Pop. Value	Est.	% non-zero	Est.	% non-zero	Est.	% non-zero	Est.	% non-zero	Est.	% non-zero	Est.	% non-zero	Est.	% non-zero	Est.	% non-zero
		(SVO) β_1-β_5~N(0,0.0025) & γ_1~N(0,5)								(SVCS) β_1-β_5~N(0,0.0025) & γ_1~N(0,5)							
$C1:\beta_3$	0	−.042	3	−.038	2	−.059	27	−.052	18	−.025	0	−.033	0	−.042	3	−.057	17
β_4	1	.145	99	.160	98	.255	100	.292	100	.077	18	.116	78	.143	99	.222	100
β_5	1	.146	99	.159	98	.257	100	.293	100	.076	17	.114	71	.143	99	.222	100
$C2:\beta_3$	0	−.042	3	−.038	2	−.059	27	−.052	18	−.026	0	−.010	0	−.042	3	−.009	0
β_4	1	.145	99	.160	98	.255	100	.292	100	.078	20	.049	0	.144	98	.096	52
β_5	1	.146	99	.159	98	.257	100	.293	100	.080	25	.050	0	.146	99	.097	56
γ_1	1	1.511	100	1.673	100	1.421	100	1.472	100	1.570	100	1.764	100	1.52	100	1.593	100
		(SVO) β_1-β_5~N(0,0.0025) & γ_1~N(1,0.04)								(SVCS) β_1-β_5~N(0,0.0025) & γ_1~N(1,0.04)							
$C1:\beta_3$	0	−.040	2	−.035	1	−.057	23	−.049	16	−.024	0	−.034	0	−.041	3	−.056	17
β_4	1	.153	100	.176	100	.263	100	.302	100	.081	26	.131	90	.148	99	.237	100
β_5	1	.154	100	.175	100	.264	100	.303	100	.081	25	.129	89	.148	99	.237	100
$C2:\beta_3$	0	−.040	2	−.035	1	−.057	23	−.049	16	−.025	0	−.006	0	−.041	3	−.004	0
β_4	1	.153	100	.176	100	.263	100	.302	100	.082	26	.052	0	.149	100	.098	55
β_5	1	.154	100	.175	100	.264	100	.303	100	.084	34	.053	0	.152	100	.099	59
γ_1	1	1.303	100	1.327	100	1.313	100	1.316	100	1.338	100	1.367	100	1.384	100	1.382	100
		(SVO) β_1-β_5~N(0,0.0025) & γ_1~N(0.5,0.04)								(SVCS) β_1-β_5~N(0,0.0025) & γ_1~N(0.5,0.04)							
$C1:\beta_3$	0	−.037	1	−.032	1	−.053	21	−.045	11	−.023	0	−.033	0	−.039	2	−.055	16
β_4	1	.161	100	.186	100	.270	100	.311	100	.085	35	.142	94	.153	99	.248	100
β_5	1	.162	100	.185	100	.272	100	.312	100	.085	32	.140	93	.153	100	.248	100
$C2:\beta_3$	0	−.037	1	−.032	1	−.053	21	−.045	11	−.023	0	−.004	0	−.039	2	.000	0
β_4	1	.161	100	.186	100	.270	100	.311	100	.086	36	.053	0	.154	100	.098	54
β_5	1	.162	100	.185	100	.272	100	.312	100	.088	41	.054	0	.156	100	.099	59
γ_1	1	1.125	100	1.111	100	1.197	100	1.174	100	1.161	100	1.147	100	1.266	100	1.233	100

When $N = 1,000$ with equal class sizes, there was enough power to detect the non-zero direct effects, and there was an acceptable number of false positives (<5%) for the truly zero direct effects. When class sizes are unequal, there was enough power (>80%) to detect the non-zero direct effects in the majority class (C1). The only exception to this trend was when $N = 500$, and a diffuse prior was used for the γ_1. A small sample size $N = 500$ with equal class sizes had limited power to detect the non-zero direct effects when using SVCS priors.

Table 4. The average parameter estimates and % of non-zero (detectable) coefficients for P4 parameters when using two levels of small variance priors (SVO and SVCS) on β_3-β_5 and three levels of priors (diffuse, informative-correct, and informative-wrong) on γ_1.

Sample Size		$N = 500$				$N = 1,000$				$N = 500$				$N = 1,000$			
Class Prop.		50%/50%		82%/18%		50%/50%		82%/18%		50%/50%		82%/18%		50%/50%		82%/18%	
Parameter	Pop. Value	Est.	% non-zero	Est.	% non-zero	Est.	% non-zero	Est.	% non-zero	Est.	% non-zero	Est.	% non-zero	Est.	% non-zero	Est.	% non-zero
		(SVO) β_1-β_5~N(0,0.0025) & γ_1~N(0,5)								(SVCS) β_1-β_5~N(0,0.0025) & γ_1~N(0,5)							
$C1 : \beta_3$	0	−.019	0	−.013	0	−.024	2	−.018	1	−.009	0	−.018	0	−.013	0	−.026	2
β_4	0.5	.184	100	.149	100	.299	100	.237	100	.075	17	.111	86	.133	99	.179	100
β_5	0	−.018	0	−.012	0	−.022	2	−.015	1	−.009	0	−.016	0	−.012	0	−.024	2
$C2 : \beta_3$	0	−.019	0	−.013	0	−.024	2	−.018	1	−.014	0	.003	0	−.022	0	.005	0
β_4	1.5	.184	100	.149	100	.299	100	.237	100	.134	94	.052	1	.238	100	.102	59
β_5	0	−.018	0	−.012	0	−.022	2	−.015	1	−.014	0	.003	0	−.022	0	.006	0
γ_1	1	1.172	100	1.183	100	1.131	100	1.150	100	1.191	100	1.210	100	1.154	100	1.183	100
		(SVO) β_1-β_5~N(0,0.0025) & γ_1~N(1,0.04)								(SVCS) β_1-β_5~N(0,0.0025) & γ_1~N(1,0.04)							
$C1 : \beta_3$	0	−.017	0	−.012	0	−.023	2	−.017	1	−.009	0	−.017	0	−.013	0	−.025	2
β_4	0.5	.186	100	.151	100	.300	100	.239	100	.076	19	.112	88	.134	99	.181	100
β_5	0	−.016	0	−.011	0	−.021	1	−.014	1	−.008	0	−.015	0	−.011	0	−.023	2
$C2 : \beta_3$	0	−.017	0	−.012	0	−.023	2	−.017	1	−.014	0	.003	0	−.022	0	.005	0
β_4	1.5	.186	100	.151	100	.300	100	.239	100	.136	95	.054	1	.240	100	.103	63
β_5	0	−.016	0	−.011	0	−.021	1	−.014	1	−.014	0	.003	0	−.021	0	.006	0
γ_1	1	1.111	100	1.105	100	1.104	100	1.112	100	1.124	100	1.120	100	1.123	100	1.135	100
		(SVO) β_1-β_5~N(0,0.0025) & γ_1~N(0.5,0.04)								(SVCS) β_1-β_5~N(0,0.0025) & γ_1~N(0.5,0.04)							
$C1 : \beta_3$	0	−.014	0	−.009	0	−.019	1	−.014	1	−.007	0	−.015	0	−.011	0	−.023	1
β_4	0.5	.191	100	.156	100	.305	100	.243	100	.078	23	.115	90	.136	99	.183	100
β_5	0	−.013	0	−.008	0	−.017	1	−.011	1	−.006	0	−.013	0	−.009	0	−.021	1
$C2 : \beta_3$	0	−.014	0	−.009	0	−.019	1	−.014	1	−.012	0	.004	0	−.020	0	.006	0
β_4	1.5	.191	100	.156	100	.305	100	.243	100	.141	97	.058	1	.244	100	.107	69
β_5	0	−.013	0	−.008	0	−.017	1	−.011	1	−.012	0	.004	0	−.019	0	.007	0
γ_1	1	.957	100	.919	100	1.010	100	.995	100	.969	100	.933	100	1.028	100	1.016	100

Across SVO and SVCS priors, there was enough power to detect γ_1, regardless of condition and prior specification. However, γ_1 was grossly overestimated in conditions with diffuse and informative-correct priors on γ_1, whereas the informative-wrong prior tended to produce a less biased estimate. Overall, the γ_1 parameter estimates illustrate the impact a local independence assumption violation can have on conditional LCA models. When latent class indicators are related to one another via a covariate, the relationship between the covariate and the latent class variable can be inflated.

7.2.4. P4 – Class-Varying Direct Effect

Table 4 provides the mean coefficient estimate and percentage of replications with a non-zero (detectable) coefficient in each condition in population model P4, which has an indirect effect and a class-varying direct effect, β_4. Again, we define a covariate as having a non-zero (i.e., detectable) direct effect if the 95% Bayesian credible interval for its estimated coefficient does not include zero. The left side of the table contains the results for SVO priors, and the right side of the table contains results for SVCS. Table 4 includes two different sample size conditions ($N = 500$ vs. $N = 1,000$) and two different class sizes (equal vs. unequal). The P4 population model represents a tricky modeling situation in which the direct effect in stronger in C2 ($\beta_4 = 1.5$) than in C1 ($\beta_4 = 0.5$).

The SVO priors were effective at detecting the presence of the non-zero direct effect β_4, regardless of sample size, class size, and γ_1 priors. Although the SVO prior is not capable, of detecting the class-varying aspect of the direct effect, it was effective at flagging the non-zero direct effect β_4 while constraining truly zero direct effects (i.e., β_3 and β_5) to be approximately zero. There truly zero direct effects rarely lead to false positives when using the SVO prior. The overall pattern of results for the SVO prior was akin to what was seen in Table 2 with the P2 population model, which also had a single direct effect.

The SVCS priors are important for identifying the class-varying aspect of the direct effect because the direct effect parameters are no longer held constant across classes. As seen on the right side of Table 4, the SVCS priors complicated the situation. The only condition in which the SVCS prior had enough power to detect the class-varying direct effect in each latent class was when $N = 1,000$ and the classes were equal in size. In contrast, when $N = 1,000$ with unequal class sizes, there was only enough power to detect the direct effect in the majority class (C1) but not the minority class (C2). When $N = 500$ with equal class sizes, there was enough power to detect the direct effect in C2 but not C1 because the direct effect is stronger in C2 than C1. When $N = 500$ with unequal class sizes, there was enough power to detect the direct effect in the majority class (C1) but not the minority class (C2).

7.2.5. P5 – No Indirect Effect with One Direct Effect

Table 5 provides the mean coefficient estimate and percentage of replications with a non-zero (detectable) coefficient in each condition in population model P5, which has a single direct effect β_4 and no indirect effect. Again, we define a covariate as having a non-zero (i.e., detectable) direct effect if the 95% Bayesian credible interval for its estimated coefficient does not include zero. The left side of the table

Table 5. The average parameter estimates and % of non-zero (detectable) coefficients for P5 parameters when using two levels of small variance priors (SVO and SVCS) on β_3-β_5 with three levels of prior specifications (i.e., no prior, informative misspecification, and diffuse misspecification) for γ_1.

Sample Size		$N = 500$				$N = 1,000$				$N = 500$				$N = 1,000$			
Class Prop.		50%/50%		82%/18%		50%/50%		82%/18%		50%/50%		82%/18%		50%/50%		82%/18%	
Parameter	Pop. Value	Est.	% non-zero	Est.	% non-zero	Est.	% non-zero	Est.	% non-zero	Est.	% non-zero	Est.	% non-zero	Est.	% non-zero	Est.	% non-zero
		(SVO) β_1-β_5~$N(0,0.0025)$								(SVCS) β_1-β_5~$N(0,0.0025)$							
$C1:\beta_3$	0	−.016	0	−.012	0	−.020	2	−.014	1	−.011	0	−.011	0	−.016	0	−.015	1
β_4	1	.275	100	.278	100	.412	100	.416	100	.167	100	.243	100	.275	100	.374	100
β_5	0	−.015	1	−.011	0	−.018	3	−.011	2	−.009	0	−.009	0	−.014	0	−.012	1
$C2:\beta_3$	0	−.016	0	−.012	0	−.020	2	−.014	1	−.011	0	−.004	0	−.016	0	−.006	0
β_4	1	.275	100	.278	100	.412	100	.416	100	.169	100	.071	8	.276	100	.130	99
β_5	0	−.015	1	−.011	0	−.018	3	−.011	2	−.012	0	−.005	0	−.016	0	−.006	0
		(SVO) β_1-β_5~$N(0,0.0025)$ & γ_1~$N(1,0.04)$								(SVCS) β_1-β_5~$N(0,0.0025)$ & γ_1~$N(1,0.04)$							
$C1:\beta_3$	0	−.024	0	−.019	0	−.028	4	−.022	2	−.016	0	−.017	0	−.021	0	−.022	2
β_4	1	.270	100	.273	100	.407	100	.411	100	.164	100	.241	100	.272	100	.371	100
β_5	0	−.022	1	−.017	1	−.027	5	−.019	4	−.013	0	−.015	0	−.019	0	−.019	2
$C2:\beta_3$	0	−.024	0	−.019	0	−.028	4	−.022	2	−.015	0	−.006	0	−.022	0	−.009	0
β_4	1	.270	100	.273	100	.407	100	.411	100	.166	100	.066	2	.273	100	.124	97
β_5	0	−.022	1	−.017	1	−.027	5	−.019	4	−.016	0	−.007	0	−.022	0	−.010	0
γ_1	0	.166	47	.227	52	.118	42	.159	45	.172	50	.238	56	.127	48	.176	52
		(SVO) β_1-β_5~$N(0,0.0025)$ & γ_1~$N(0,5)$								(SVCS) β_1-β_5~$N(0,0.0025)$ & γ_1~$N(0,5)$							
$C1:\beta_3$	0	−.020	0	−.015	0	−.025	3	−.018	1	−.013	0	−.014	0	−.019	0	−.019	1
β_4	1	.273	100	.276	100	.409	100	.413	100	.165	100	.243	100	.273	100	.373	100
β_5	0	−.019	1	−.014	1	−.023	3	−.015	2	−.011	0	−.012	0	−.017	0	−.016	2
$C2:\beta_3$	0	−.020	0	−.015	0	−.025	3	−.018	1	−.013	0	−.005	0	−.019	0	−.008	0
β_4	1	.273	100	.276	100	.409	100	.413	100	.168	100	.067	3	.274	100	.125	97
β_5	0	−.019	1	−.014	1	−.023	3	−.015	2	−.014	0	−.006	0	−.019	0	−.008	0
γ_1	0	.081	14	.013	15	.069	17	.082	15	−.090	17	.118	16	.080	22	.102	21

contains the results for SVO priors, and the right side of the table contains results for SVCS. In Table 5, the γ_1 regression coefficient is either correctly not specified (top panel), misspecified with an informative prior (middle panel), or misspecified with a diffuse prior (bottom panel). Table 5 includes two sample size conditions ($N = 500$ vs. $N = 1,000$) and two class sizes (equal vs. unequal). The P5 population model is a situation where a single latent class indicator u_4 is related to the covariate x, but the latent class variable c is unrelated to the covariate. Often researchers are unaware of the covariate relationships a priori; therefore, the P5 population model represents a realistic situation in which the indirect effect does not exist.

The SVO priors were effective at detecting the non-zero direct effect β_4, regardless of the γ_1 specification, sample size, and class proportions. The SVCS priors were also able to effectively detect the non-zero direct effect β_4 in most conditions. The only exception was in conditions with $N = 500$ and unequal class sizes, where there was not enough power to detect the direct effect in the minority class (C2). Across small-variance prior conditions (SVO and SVCS), the truly zero direct effects were held to approximately zero and had an acceptable level of false positive (<5%). Notably, in conditions with a misspecified γ_1 there was an inflated Type I error rate (>5%) for γ_1. The misspecification with an informative prior on γ_1 resulted in a higher number of false positives compared to the misspecification with a diffuse prior.

7.2.6. P6 – No Indirect Effect with Two Direct Effects

Table 6 provides the mean coefficient estimate and percentage of replications with a non-zero (detectable) coefficient in each condition in population model P6, which has two direct effects (β_4 and β_5) and no indirect effect. Again, we define a covariate as having a non-zero (i.e., detectable) direct effect if the 95% Bayesian credible interval for its estimated coefficient does not include zero. The left side of the table contains the results for SVO priors, and the right side of the table contains results for SVCS. In Table 6, γ_1 is either correctly not specified (top panel), misspecified with an informative prior (middle panel), or misspecified with a diffuse prior (bottom panel). Table 6 includes two different sample size conditions ($N = 500$ vs. $N = 1,000$) and two different class sizes (equal vs. unequal). The P6 population model represents a very challenging modeling situation in which a local independence assumption violation is present without an indirect effect. Specifically, the covariate x is related to two latent class indicators, u_4 and u_5, but is unrelated to the latent class variable c.

There was enough power to detect the non-zero direct effects with SVO priors, regardless of the γ_1 specification (or misspecification), sample size, and class proportions. However, the SVO prior did not always hold the truly zero direct effects to be approximately zero. Specifically, there was inflated number of false positives in most conditions. The only condition with an acceptable level of false positives (<5%) for the truly zero direct effect had a sample size of $N = 500$, unequal class sizes, and γ_1 was not specified.

When using SVCS priors, there was enough power to detect non-zero direct effects when there were equal class sizes. When the class sizes are unequal, there was not enough power to detect the direct effect in the minority class (C2). The only condition with unequal class sizes that could detect the direct effect in the minority class had a

Table 6. The average parameter estimates and % of non-zero (detectable) coefficients for P6 parameters when using two levels of small variance priors (SVO and SVCS) on β_3-β_5 with three levels of prior specifications (i.e., no prior, informative misspecification, and diffuse misspecification) for γ_1.

Sample Size		$N = 500$				$N = 1,000$				$N = 500$				$N = 1,000$			
Class Prop.		50%/50%		82%/18%		50%/50%		82%/18%		50%/50%		82%/18%		50%/50%		82%/18%	
Parameter	Pop. Value	Est.	% non-zero	Est.	% non-zero	Est.	% non-zero	Est.	% non-zero	Est.	% non-zero	Est.	% non-zero	Est.	% non-zero	Est.	% non-zero
		(SVO) β_1-β_5~N(0,0.0025)								(SVCS) β_1-β_5~N(0,0.0025)							
$C1 : \beta_3$	0	−.047	6	−.033	2	−.058	28	−.038	10	−.033	0	−.033	2	−.047	5	−.044	12
β_4	1	.257	100	.266	100	.389	100	.402	100	.155	100	.234	100	.257	100	.362	100
β_5	1	.257	100	.266	100	.391	100	.403	100	.154	99	.233	100	.258	100	.363	100
$C2 : \beta_3$	0	−.047	6	−.033	2	−.058	28	−.038	10	−.031	0	−.009	0	−.046	5	−.012	0
β_4	1	.257	100	.266	100	.389	100	.402	100	.153	100	.061	1	.256	100	.111	80
β_5	1	.257	100	.266	100	.391	100	.403	100	.155	100	.062	3	.258	100	.111	79
		(SVO) β_1-β_5~N(0,0.0025) & γ_1~N(1,0.04)								(SVCS) β_1-β_5~N(0,0.0025) & γ_1~N(1,0.04)							
$C1 : \beta_3$	0	−.071	32	−.064	27	−.087	70	−.066	40	−.049	3	−.050	6	−.069	27	−.071	44
β_4	1	.237	100	.226	99	.369	100	.382	100	.138	94	.172	75	.240	100	.329	99
β_5	1	.238	100	.225	99	.371	100	.383	100	.138	94	.172	77	.240	100	.330	99
$C2 : \beta_3$	0	−.071	32	−.064	27	−.087	70	−.066	40	−.049	4	−.032	4	−.069	25	−.025	4
β_4	1	.237	100	.226	99	.369	100	.382	100	.137	95	.060	3	.239	100	.104	62
β_5	1	.238	100	.225	99	.371	100	.383	100	.138	95	.061	3	.241	100	.105	62
γ_1	0	.345	90	.533	92	.272	96	.342	93	.395	93	.625	94	.320	99	.439	97
		(SVO) β_1-β_5~N(0,0.0025) & γ_1~N(0,5)								(SVCS) β_1-β_5~N(0,0.0025) & γ_1~N(0,5)							
$C1 : \beta_3$	0	−.068	25	−.067	31	−.083	66	−.063	35	−.048	2	−.047	5	−.067	23	−.069	39
β_4	1	.234	100	.206	90	.372	100	.381	100	.139	93	.142	56	.241	100	.317	96
β_5	1	.240	100	.207	91	.374	100	.382	100	.129	93	.142	55	.242	100	.318	96
$C2 : \beta_3$	0	−.068	25	−.067	31	−.083	66	−.063	35	−.047	3	−.041	6	−.067	23	−.030	8
β_4	1	.234	100	.206	90	.372	100	.381	100	.137	95	.058	1	.240	100	.104	64
β_5	1	.240	100	.207	91	.374	100	.382	100	.139	94	.058	2	.242	100	.105	64
γ_1	0	.295	71	.683	70	.235	85	.311	71	.364	79	.891	77	.290	93	.474	86

sample size of $N = 1,000$ and γ_1 was not specified. Notably, there was an unacceptable level of false positives (>5%) for the truly zero direct effects when the sample size was $N = 1,000$. Additionally, when the sample size was $N = 500$ with unequal class sizes, conditions with a misspecified γ_1 also resulted in a higher level of false positives.

When using small-variance priors, the local independence assumption violation results in a dramatic inflation in the Type I error rate for γ_1. Misspecifying γ_1 resulted in a high number of false positives for γ_1, regardless of whether an informative or diffuse prior was utilized. The P6 population model results highlight how unreliable the γ_1 parameter estimate is when using small-variance priors in the presence of a local independence assumption violation.

7.3. Modeling Direct Effects with Bayesian Estimation

After the non-zero direct effect(s) have been detected, a more parsimonious conditional LCA model can be specified that only includes the non-zero direct effects (instead of all possible direct effects) and γ_1. In this section of results, we explore how robust the direct effect and γ_1 parameter estimates are to different prior specifications, using the datasets generated from the P2 population model.[3] Specifically, three levels of priors were specified on the direct effect (informative-correct, informative-wrong, and diffuse) and three levels of priors were used on γ_1 (informative-correct, informative-wrong, and diffuse) for each condition. Figure 8 provides the relative bias in the β_4 direct effect, whereas Figure 9 displays the relative bias in γ_1.

Figure 8 reveals the direct effect parameter estimate was robust to different prior specification on the direct effect and γ_1. Regardless of sample size, class size, and prior specifications, minimal bias was produced for the β_4 parameter estimate. In contrast, Figure 9 reveals the γ_1 parameter estimates were affected by γ_1 prior specification, but not the direct prior specification. When using an informative-wrong prior on the γ_1 regression coefficient, the γ_1 parameter tended to be underestimated. However, the informative-correct prior and the diffuse prior resulted in an unbiased γ_1, regardless of sample size, class size, and the prior specification on the direct effect. Overall, Figures 8 and 9 demonstrate that the conditional LCA model results are relatively robust to different prior specifications.

7.4. Modeling Class-Varying Direct Effects

In some conditional LCA modeling situations, the relationship between the covariate and a latent class indicator can be class-varying. To assess the impact of prior specification on model parameter estimates, we analyzed the data generated from the P4 population model with a variety of priors on the direct effect (informative-correct, informative-wrong, and diffuse) and γ_1 (informative-correct, informative-wrong, and diffuse). The informative-correct prior on the direct effect was the only condition that allowed for the estimation of a class-varying direct effect, whereas the diffuse and informative-wrong prior constrained the direct effect to be equal across classes. Considering γ_1 is typically of greater substantive interest, it is important to understand the

[3]The pattern of results for the P3 population model were very similar to the pattern of results for the P2 population model; therefore, the P3 population model results are not presented here but are available upon request.

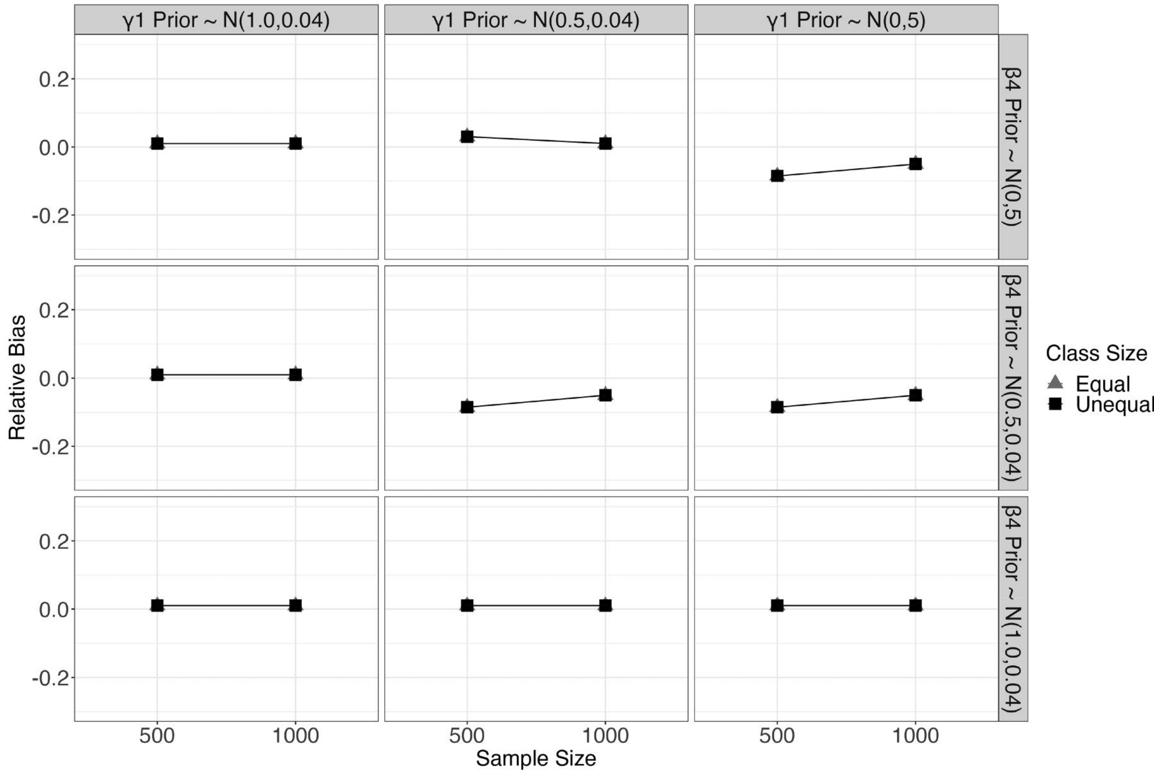

Figure 8. The relative bias in β_4 from P2 when using three levels of priors on γ_1 and three levels of priors on the direct effect.

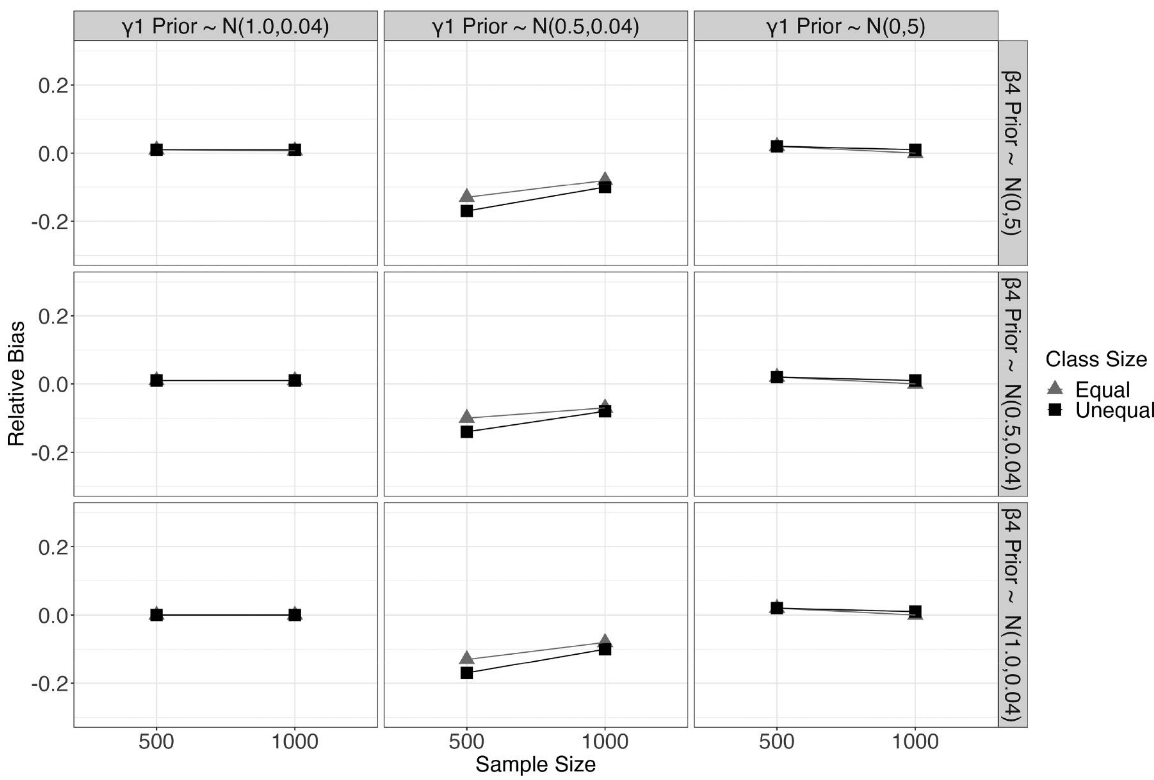

Figure 9. The relative bias in γ_1 from P2 when using three levels of priors on γ_1 and three levels of priors on the direct effect.

impact of direct effect prior specification on the γ_1 parameter estimate.

Figures 10 and 11 display the relative bias in the β_{41} and β_{42} parameters, respectively. Unsurprisingly, specifying informative-wrong and diffuse priors biased the β_{41} and β_{42} parameters. The β_{41} tended to be overestimated, especially in conditions with equal class sizes. In contrast, the β_{42} parameter was underestimated, regardless of class sizes. The

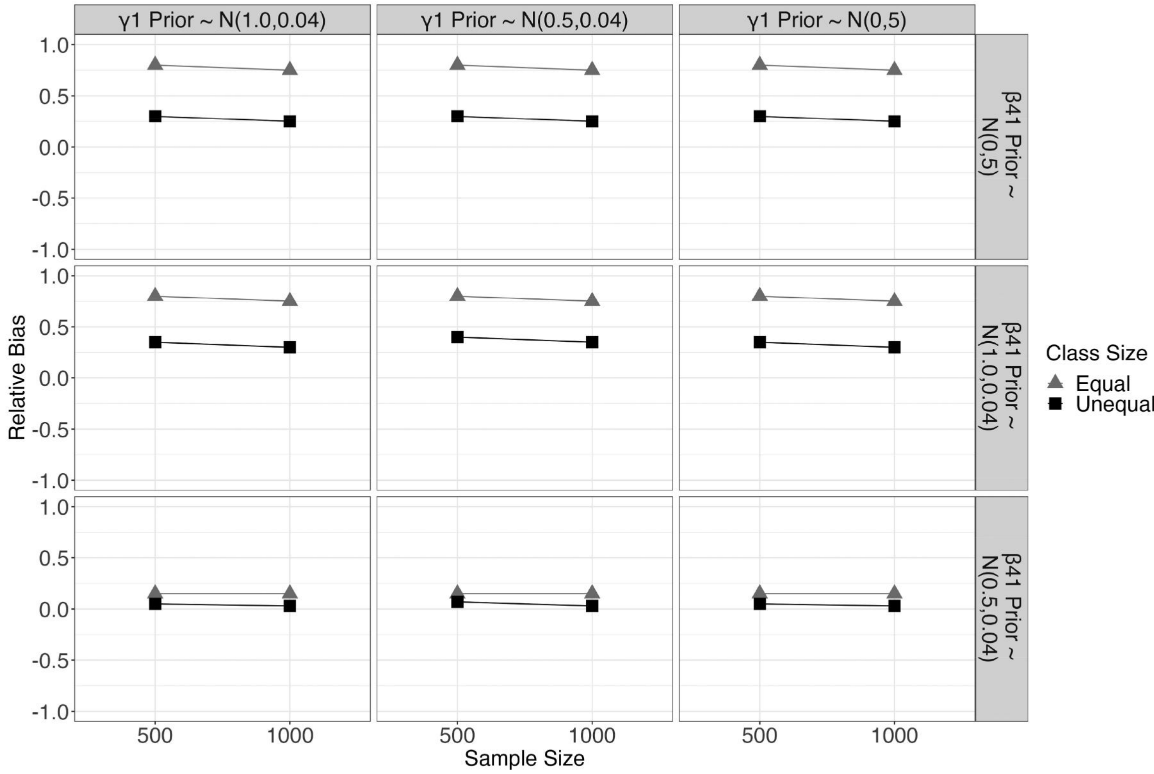

Figure 10. The relative bias in β_{41} from P4 when using three levels of priors on γ_1 and three levels of priors on the direct effect.

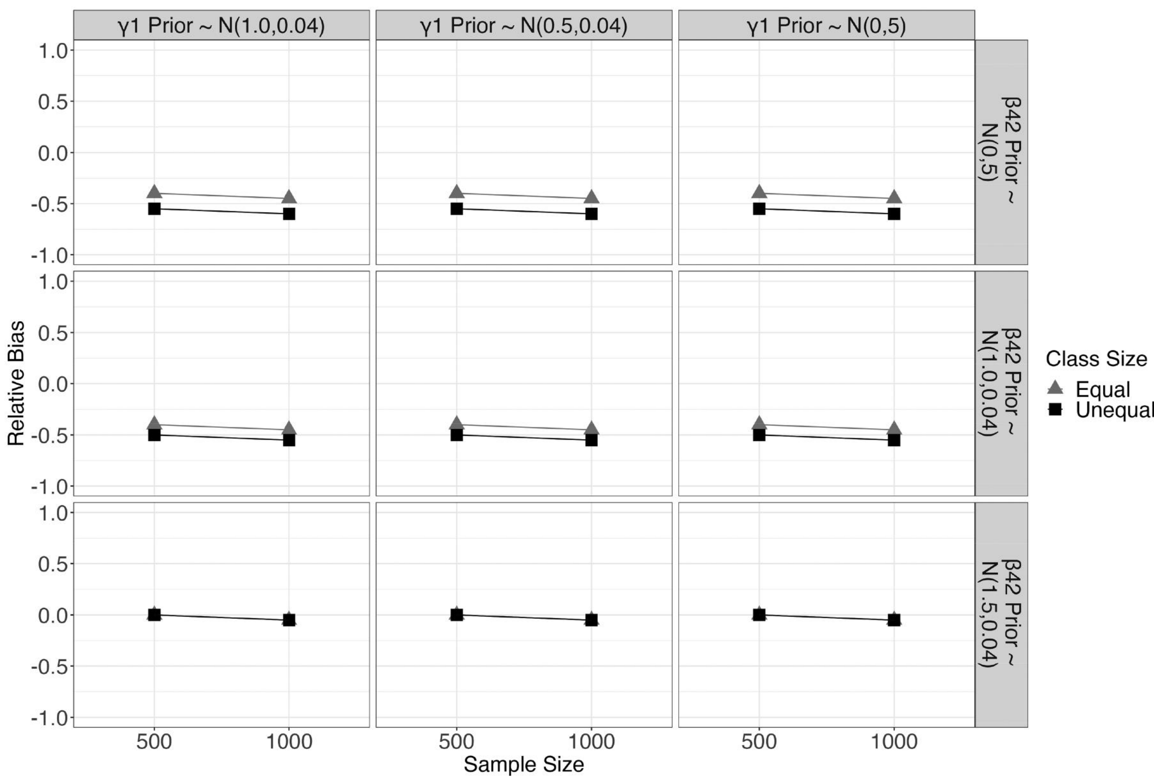

Figure 11. The relative bias in β_{42} from P4 when using three levels of priors on γ_1 and three levels of priors on the direct effect.

direct effect results were robust to the prior specification of γ_1. Figure 12 provides the relative bias in the γ_1 parameter. When an informative-wrong prior is specified on γ_1, the γ_1 was biased in conditions with a sample size of $N = 500$, but not in conditions with $N = 1,000$. The γ_1 parameter had minimal bias when a diffuse or informative-correct prior was specified on γ_1, regardless of sample size and class size. Notably, the direct effect prior specification did not bias γ_1.

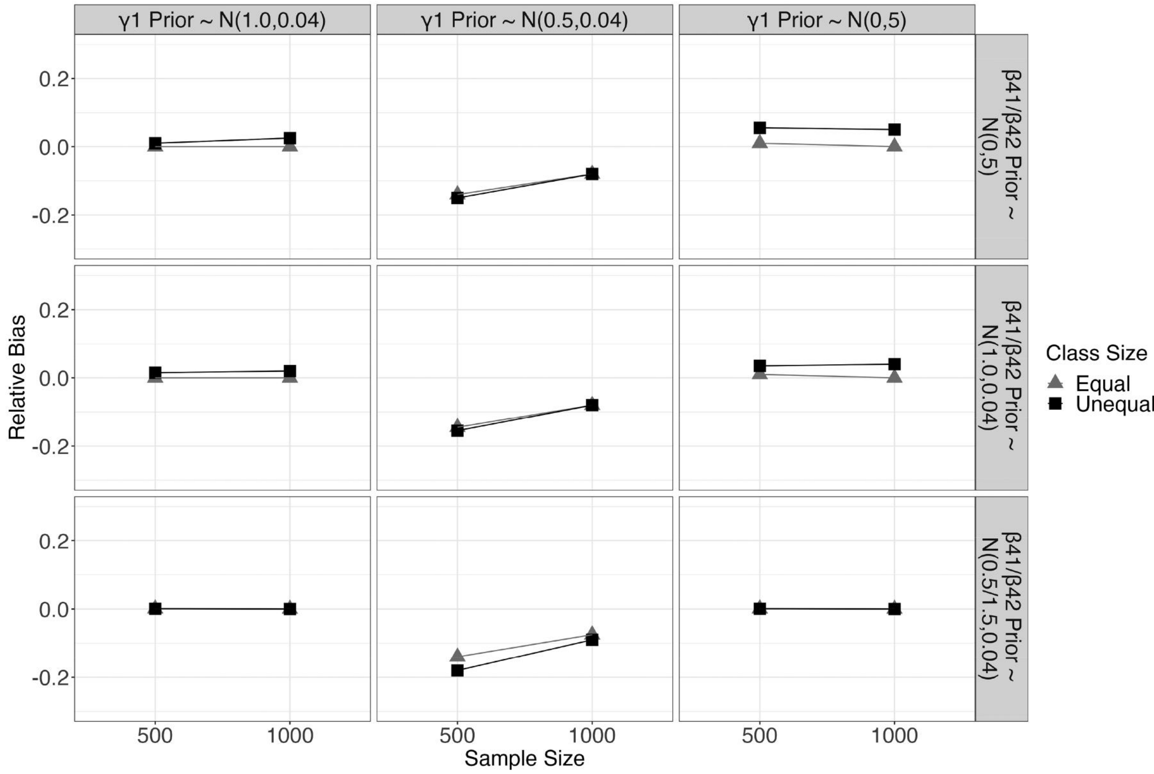

Figure 12. The relative bias in γ_1 from P4 when using three levels of priors on γ_1 and three levels of priors on the direct effect.

Table 7. The relative bias and % of non-zero (detectable) coefficients for P5 parameters when using three levels of priors (diffuse, informative-correct, and informative-wrong) on β_4 and two levels of prior misspecifications (informative misspecification, diffuse misspecification) on γ_1.

Sample Size		$N = 500$				$N = 1,000$				$N = 500$				$N = 1,000$			
Class Prop.		50%/50%		82%/18%		50%/50%		82%/18%		50%/50%		82%/18%		50%/50%		82%/18%	
Parameter	Pop. Value	Bias	% non-zero	Bias	% non-zero	Bias	% non-zero	Bias	% non-zero	Bias	% non-zero	Bias	% non-zero	Bias	% non-zero	Bias	% non-zero
		β_4-~N(1,0.04) & γ_1~N(1,0.04)								β_4-~N(1,0.04) & γ_1~N(0,5)							
β_4	1	.007	100	.002	100	.009	100	.005	100	.012	100	.007	100	.013	100	.009	100
γ_1	0	.098	18	.153	25	.055	13	.089	16	.002	5	.008	7	.002	6	.004	5
		β_4-~N(0.5,0.04) & γ_1~N(1,0.04)								β_4-~N(0.5,0.04) & γ_1~N(0,5)							
β_4	1	−.096	100	−.098	100	−.049	100	−.052	100	−.091	100	−.092	100	.046	100	.048	100
γ_1	0	.103	19	.159	28	.058	13	.093	17	.009	6	.016	7	.006	6	.009	5
		β_4-~N(0,5) & γ_1~N(1,0.04)								β_4-~N(0,5) & γ_1~N(0,5)							
β_4	1	.011	100	.004	100	.011	100	.006	100	.018	100	.012	100	.015	100	.010	100
γ_1	0	.098	18	.097	25	.055	13	.089	15	.002	6	.008	7	.002	6	.004	6

7.5. Misspecification of γ_1

Applied researchers often assume the presence of an indirect effect when including a covariate variable in the LCA model. In the P5 population model, γ_1 was fixed to zero and there was single direct effect. To explore the impact of misspecifying γ_1 on parameter estimates under different prior conditions, we analyzed the data generated from the P5 population model with three levels of priors on the direct effects (informative-correct, informative-wrong, and diffuse) and two levels of priors on the misspecified γ_1 (informative misspecification, diffuse misspecification).[4] Table 7 displays the parameter bias and percentage of replications with a non-zero (detectable) coefficient under two sample size conditions ($N = 500$ vs. $N = 1,000$) two class size conditions (equal vs. unequal), and six combinations of prior specifications.

Table 7 shows the results for the P5 population model, which has a single direct effect and no indirect effect. The direct effect β_4 was unbiased in all conditions, regardless of sample size, class size, and prior specification. When the truly zero γ_1 was misspecified with informative priors, γ_1 was typically overestimated in the $N = 500$ conditions. Despite γ_1 often being unbiased, there was alarming number of false positives. The γ_1 parameter tended to have an inflated Type I error rate (>5%). The most problematic conditions utilized an informative prior on the misspecified γ_1. When comparing sample size ($N = 500$ vs. $N = 1,000$), conditions with a smaller sample size produced more false positives. The

[4]The P6 population model, which has no indirect effect and two direct effects (i.e., local independence assumption violation), had a very similar pattern of results; therefore, the results from the P6 population model are not presented here but are available upon request.

combination of a smaller sample size and unequal class sizes produced the greatest number of false positives.

8. Discussion

The primary goal of this article was to explore the performance of Bayesian SEM when modeling direct effects in conditional LCA models. The use of small-variance priors to detect non-zero direct effects between covariates and latent class indicators is a novel application of Bayesian SEM. In the conditions we investigated, small-variance priors on the overall direct effects had the power to detect non-zero direct effects in all population models, regardless of sample size and class sizes. Notably, this includes conditions with a single direct effect, two direct effects (i.e., local independence assumption violation), a class-varying direct effect, and a misspecified γ_1. However, the small-variance priors tended to produce a high number of false positives for truly zero direct effects in conditions with a local independence assumption violation, especially when the sample size was large. In addition, conditions with local independence assumption violations and no indirect effect (i.e., population model P6) tended to produce a high number of false positives for γ_1 when it was misspecified. These findings illustrate how problematic local independence assumption violations can be in LCA models.

In addition to exploring the performance of small-variance priors on the overall direct effects, we also examined the power of small-variance priors on the class-specific direct effects. The aim of a small-variance prior on the class-specific direct effect is to explore the possibility of a class-varying direct effect. Class-specific direct effects are much more difficult to estimate. The primary finding was that the overall sample size and the class size had to be large to detect a class-varying direct effect with small-variance priors. In situations where the class-varying direct effect is not of substantive interest to applied researchers and the sample size is limited, only estimating the overall direct effect may be a better strategy.

Although the SVCS prior provides a useful regularization strategy when modeling direct effects in the three-step approach, it can struggle in scenarios involving small classes with high classification uncertainty. One potential solution—using empirical class sizes from Step 1 to further inform prior specification—may seem appealing but risks "double dipping" into the data, thereby compromising the independence between class enumeration and covariate modeling. We discourage this practice and instead recommend caution in applying SVCS when small or poorly separated classes are present. Sensitivity analyses or alternative modeling strategies may be more appropriate in such cases.

The small-variance prior simulations were not without limitations. First, the small-variance prior conditions set the variance hyperparameter to 0.0025, across conditions. Simulation results may be different when a wider (or narrower) prior is specified on the direct effects. A second limitation of the study is how we generated the data for each population model. To examine the performance of small-variance priors on direct effects, we generated data from population models with a relatively strong direct effect between the covariate and latent class variable. The strength of the direct effect would impact the power available to detect the non-zero direct effects. In addition, the covariate was normally distributed with no missing data, which could impact results. Future simulation studies should consider a wider variety of small-variance prior specifications and population models.

Another important limitation of the present study is that we assume covariates are uncorrelated. In practice, especially with demographic predictors, some degree of multicollinearity is almost inevitable. The SV and SVCS priors, as implemented here, do not account for correlation among covariates, which may impact their ability to recover true direct effects when only a subset of predictors are associated with the latent classes. Future work should explore the robustness of these priors under varying levels of multicollinearity and evaluate whether alternative approaches—such as multivariate priors or covariate selection techniques—offer improved performance in such contexts.

Another avenue for future methodological research is model fit. Methodologists should investigate the power of posterior predictive p-value (PPP; Scheines et al., 1999) to detect covariate misspecifications in conditional LCA models with small-variance priors. Past research on CFA models suggests PPP lacks power to detect model misspecifications unless the small-variance priors were very restrictive, and the sample size and misspecification are large (Jorgensen et al., 2019). Methodologists should also compare the performance of Bayesian SEM with other methods available for detecting direct effects such as the LCA MIMIC modeling procedure proposed by Masyn (2017) and the residual and fit statistics discussed in Janssen et al. (2019). Each of these procedures for detecting direct effects have their own limitations and should be explored via simulation research.

An additional limitation of the present study is the exclusive use of the modal assignment rule for class membership. While this approach aligns with common practice due to its interpretability, prior work (e.g., Alagöz & Vermunt, 2022) has shown that modal assignment can be more sensitive to model misspecification, particularly when respondents exhibit high classification uncertainty. Future research could explore whether the use of proportional assignment (soft partitioning) improves performance when modeling direct effects with small variance priors, especially in borderline—or fuzzy—classification cases (e.g., posterior probabilities near 0.5 in two-class solutions).

In addition to examining small-variance priors, this study also explored how robust the conditional LCA model results are to different combinations of prior specifications (informative-correct, informative-wrong, and diffuse) on β_4 (direct effect of x on u_4) and γ_1 (effect of x on c). Specifically, we examined the bias in the β_4 and γ_1 parameters. Regardless of the prior specification, the parameter estimates for β_4 and γ_1 were robust. The γ_1 regression coefficient was most impacted by the prior specification on γ_1. Despite these findings, applied researchers should use a prior sensitivity to explore different combinations of priors on the direct effects and γ_1 when modeling a conditional LCA model. For an example of how to implement a prior sensitivity analysis, see Depaoli et al. (2020).

In some modeling situations, the direct effect is not equal in both classes. Based on the small-variance prior results from this study, we know that detecting a class-varying direct effect that is unknown *a priori* is difficult. In addition to a larger sample size requirement, the number of cases assigned to each latent class and the strength of the direct effect in each class impacts our power to detect the direct effect. Considering how challenging class-varying direct effects can be, it was important to understand the impact of misspecifying the class-varying direct effect as an overall direct effect instead. As expected, this misspecification biased the class-specific direct effect parameter estimates. However, the misspecification of the direct effect did not impact the γ_1 parameter estimates. Therefore, it may not be necessary for applied researchers with limited sample sizes or class sizes to consider the class-varying component of the direct effect, if the primary focus of the research is γ_1. The accuracy of the γ_1 parameter estimates was most influenced by the γ_1 prior specification.

The effect of x on c (γ_1) is often the primary interest of applied researchers using conditional LCA models. The effect of x on c allows researchers to answer questions about why a case was assigned to a particular latent class. Often researchers explore the relationship between demographic variables and the latent class assignment. In some situations, applied researchers may inadvertently assume a covariate is related to the latent class variable (c) when the covariate is only related to a latent class indicator (u_m). This would be a misspecification of γ_1. When using an informative prior on γ_1, the Type I error rate for γ_1 is inflated. Smaller sample sizes and unequal class sizes tended to increase the probability of a false positive, but the Type I error remained inflated in all conditions. A more diffuse prior distribution held the Type I error rate to 0.05-0.07, which is still somewhat inflated. A prior sensitivity analysis on γ_1 would be helpful in this situation because it would demonstrate how influential the prior specification is on γ_1.

References

Alagöz, Ö. E. C., & Vermunt, J. K. (2022). Stepwise latent class analysis in the presence of missing values on the class indicators. *Structural Equation Modeling: A Multidisciplinary Journal*, *29*, 784–790. https://doi.org/10.1080/10705511.2022.2030743

Albert, P. S., & Dodd, L. E. (2004). A cautionary note on the robustness of latent class models for estimating diagnostic error without a gold standard. *Biometrics*, *60*, 427–435. https://doi.org/10.1111/j.0006-341X.2004.00187.x

Asparouhov, T., & Muthén, B. (2014). Auxiliary variables in mixture modeling: Three-step approaches using Mplus. *Structural Equation Modeling: A Multidisciplinary Journal*, *21*, 329–341. https://doi.org/10.1080/10705511.2014.915181

Asparouhov, T., & Muthén, B. O. (2011). Using Bayesian priors for more flexible latent class analysis. In *Joint statistical proceedings, government statistics section* (pp. 4979–4993). American Statistical Association.

Bakk, Z., & Kuha, J. (2018). Two-step estimation of models between latent classes and external variables. *Psychometrika*, *83*, 871–892. https://doi.org/10.1007/s11336-017-9592-7

Bakk, Z., Tekle, F. B., & Vermunt, J. K. (2013). Estimating the association between latent class membership and external variables using bias-adjusted three-step approaches. *Sociological Methodology*, *43*, 272–311. https://doi.org/10.1177/0081175012470644

Bandeen-Roche, K., Miglioretti, D. L., Zeger, S. L., & Rathouz, P. J. (1997). Latent variable regression for multiple discrete outcomes. *Journal of the American Statistical Association*, *92*, 1375–1386. https://doi.org/10.1080/01621459.1997.10473658

Bauer, D. J. (2007). Observations on the use of growth mixture models in psychological research. *Multivariate Behavioral Research*, *42*, 757–786. https://doi.org/10.1080/00273170701710338

Bolck, A., Croon, M., & Hagenaars, J. (2004). Estimating latent structure models with categorical variables: One-step versus three-step estimators. *Political Analysis*, *12*, 3–27. https://doi.org/10.1093/pan/mph001

Clogg, C. C., & Goodman, L. A. (1985). Simultaneous latent structure analysis in several groups. In N. B. Tuma (Ed.), *Sociological methodology 1985* (pp. 81–110). Jossey-Bass.

Collins, L. M., & Lanza, S. T. (2010). *Latent class and latent transition analysis: With applications in the social, behavioral, and Health Sciences*. Wiley. https://doi.org/10.1002/9780470567333

Dayton, C. M., & Macready, G. B. (1988). Concomitant-variable latent class models. *Journal of the American Statistical Association*, *83*, 173–178. https://doi.org/10.2307/2288938

Depaoli, S. (2013). Mixture class recovery in GMM under varying degrees of class separation: Frequentist versus Bayesian estimation. *Psychological Methods*, *18*, 186–219. https://doi.org/10.1037/a0031609

Depaoli, S. (2014). The impact of inaccurate "informative" priors for growth parameters in Bayesian growth mixture modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *21*, 239–252. https://doi.org/10.1080/10705511.2014.882686

Depaoli, S. (2021). *Bayesian structural equation modeling*. The Guilford Press.

Depaoli, S., Rus, H., Clifton, J., van de Schoot, R., & Tiemensma, J. (2017). An introduction to Bayesian statistics in health psychology. *Health Psychology Review*, *11*, 248–264. https://doi.org/10.1080/17437199.2017.1343676

Depaoli, S., Winter, S. D., & Visser, M. (2020). The importance of prior sensitivity analysis in bayesian statistics: demonstrations using an interactive shiny app. *Frontiers in Psychology*, *11*(November), 608045. https://doi.org/10.3389/fpsyg.2020.608045

Finch, H. (2021). Comparison of distance measures in cluster analysis with dichotomous data. *Journal of Data Science*, *3*, 85–100. https://doi.org/10.6339/JDS.2005.03(1).192

Haberman, S. J. (1979). *Analysis of qualitative data, Volume 2: New developments*. Academic Press.

Hagenaars, J. A. (1990). *Categorical Longitudinal Data: Log-Linear Panel, Trend, and Cohort Analysis*. Sage. https://doi.org/10.2307/1165391

Hagenaars, J. A. (1993). *Loglinear models with latent variables*. Sage. https://doi.org/10.4135/9781412984850

Hipp, J. R., & Bauer, D. J. (2006). Local solutions in the estimation of growth mixture models. *Psychological Methods*, *11*, 36–53. https://doi.org/10.1037/1082-989X.11.1.36

Hoijtink, H., & van de Schoot, R. (2018). Testing small variance priors using prior-posterior predictive p values. *Psychological Methods*, *23*, 561–569. https://doi.org/10.1037/met0000131

Janssen, J. H. M., van Laar, S., de Rooij, M. J., Kuha, J., & Bakk, S. (2019). The Detection and Modeling of Direct Effects in Latent Class Analysis. *Structural Equation Modeling: A Multidisciplinary Journal*, *26*, 280–290. https://doi.org/10.1080/10705511.2018.1541745

Jorgensen, T. D., Garnier-Villarreal, M., Pornprasermanit, S., & Lee, J. (2019). Small-variance priors can prevent detecting important misspecifications in Bayesian confirmatory factor analysis. In Wiberg M., Culpepper S., Janssen R., González J., Molenaar D. (Eds.), *Quantitative psychology: IMPS 2017. Springer proceedings in mathematics & statistics* (pp. 255–263). Springer. https://doi.org/10.1007/978-3-030-01310-3_23

Kamakura, W. A., Wedel, M., & Agrawal, J. (1994). Concomitant variable latent class models for conjoint analysis. *International Journal of Research in Marketing*, *11*, 451–464. https://doi.org/10.1016/0167-8116(94)00004-2

Kim, S. (2014). Determining the number of latent classes in single- and multiphase growth mixture models. *Structural Equation Modeling: a Multidisciplinary Journal*, *21*, 263–279. https://doi.org/10.1080/10705511.2014.882690

Lee, J., Jung, K., & Park, J. (2020). Detecting conditional dependence using flexible Bayesian latent class analysis. *Frontiers in Psychology*, *11*, 1987. https://doi.org/10.3389/fpsyg.2020.01987

Li, L., & Hser, Y. I. (2011). On inclusion of covariates for class enumeration of growth mixture models. *Multivariate Behavioral Research*, *46*, 266–302. https://doi.org/10.1080/00273171.2011.556549

Lu, Z. L., Zhang, Z., & Lubke, G. (2011). Bayesian inference for growth mixture models with latent class dependent missing data. *Multivariate Behavioral Research*, *46*, 567–597. https://doi.org/10.1080/00273171.2011.589261

Lubke, G., & Muthén, B. (2007). Performance of factor mixture models as a function of model size, covariate effects, and class-specific parameters. *Structural Equation Modeling: A Multidisciplinary Journal*, *14*, 26–47. https://doi.org/10.1080/10705510709336735

Lubke, G., & Tueller, S. (2010). Latent class detection and class assignment: A comparison of the MAXEIG taxometric procedure and factor mixture modeling approaches. *Structural Equation Modeling: a Multidisciplinary Journal*, *17*, 605–628. https://doi.org/10.1080/10705511.2010.510050

Masyn, K. (2013). Latent class analysis and finite mixture modeling. In T. D. Little (Ed.), *The Oxford handbook of quantitative methods in psychology* (Vol. 2, pp. 551–611). Oxford University Press.

Masyn, K. E. (2017). Measurement invariance and differential item functioning in latent class analysis with stepwise multiple indicator multiple cause modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *24*, 180–197. https://doi.org/10.1080/10705511.2016.1254049

McCutcheon, A. L. (1987). *Latent class analysis*. Sage. https://doi.org/10.4135/9781412984713

Muthén, B. (1985). A method for studying the homogeneity of test items with respect to other relevant variables. *Journal of Educational Statistics*, *10*, 121–132. https://doi.org/10.3102/10769986010002121

Muthén, B. O. (2002). Beyond SEM: General latent variable modeling. *Behaviormetrika*, *29*, 81–117. https://doi.org/10.2333/bhmk.29.81

Muthén, B. O., & Asparouhov, T. (2012). Bayesian SEM: A more flexible representation of substantive theory. *Psychological Methods*, *17*, 315–335. https://doi.org/10.1037/a0026802

Muthén, L. K., & Muthén, B. O. (1998–2017). *Mplus user's guide* (8th ed.). Muthén & Muthén.

Nylund, K. L., Asparouhov, T., & Muthén, B. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling: A Multidisciplinary Journal*, *14*, 535–569. https://doi.org/10.1080/10705510701575396

Nylund-Gibson, K., & Masyn, K. E. (2016). Covariates and Mixture Modeling: Results from a Simulation Study Exploring the Impact of Misspecified Effects on Class Enumeration. *Structural Equation Modeling: A Multidisciplinary Journal*, *23*, 782–797. https://doi.org/10.1080/10705511.2016.1221313

Nylund-Gibson, K., Grimm, R. P., & Masyn, K. E. (2019). Prediction from latent classes: A demonstration of different approaches to include distal outcomes in mixture models. *Structural Equation Modeling: A Multidisciplinary Journal*, *26*, 967–985. https://doi.org/10.1080/10705511.2019.1590146

Petras, H., & Masyn, K. (2010). General growth mixture analysis with ante- cedents and consequences of change. In A. R. Piquero & D. Weisburd (Eds.), *Handbooks of quantitative criminology* (pp. 69–100). Springer.

Quirk, M., Nylund-Gibson, K., & Furlong, M. (2013). Exploring patterns of Latino/a children's school readiness at kindergarten entry and their relations with grade 2 achievement. *Early Childhood Research Quarterly*, *28*, 437–449. https://doi.org/10.1016/j.ecresq.2012.11.002

Scheines, R., Hoijtink, H., & Boomsma, A. (1999). Bayesian estimation and testing of structural equation models. *Psychometrika*, *64*, 37–52. https://doi.org/10.1007/BF02294318

Sedding, D., & Leitgöb, H. (2018). Exact and Bayesian approximate measurement invariance. In Davidov, E., Schmidt, P., Billiet, J., & Meuleman, B. *Cross-cultural analysis methods and applications* (2nd ed., pp. 553–579). Routledge. https://doi.org/10.4324/9781315537078-20

Sterba, S. K. (2014). Handling missing covariates in conditional mixture models under missing at random assumptions. *Multivariate Behavioral Research*, *49*, 614–632. https://doi.org/10.1080/00273171.2014.950719

Stromeyer, W. R., Miller, J. W., Sriramachandramurthy, R., & DeMartino, R. (2015). The prowess and pitfalls of Bayesian structural equation modeling: Important considerations for management research. *Journal of Management*, *41*, 491–520. https://doi.org/10.1177/014920631455196

Torrance-Rynard, V. L., & Walter, S. D. (1997). Effects of dependent errors in the assessment of diagnostic test performance. *Statistics in Medicine*, *16*, 2157–2175. https://doi.org/10.1002/(SICI)1097-0258(19971015)16:19<2157::AID-SIM653>3.0.CO;2-X

Tueller, S., & Lubke, G. (2010). Evaluation of structural equation mixture models: Parameter estimates and correct class assignment. *Structural Equation Modeling: a Multidisciplinary Journal*, *17*, 165–192. https://doi.org/10.1080/10705511003659318

Vacek, P. M. (1985). The effect of conditional dependence on the evaluation of diagnostic tests. *Biometrics*, *41*, 959–968. https://doi.org/10.1111/j.1467-985X.2008.00544.x

van de Schoot, R., Winter, S., Zondervan-Zwijnenburg, M., Ryan, O., & Depaoli, S. (2017). A systematic review of Bayesian applications in psychology: The last 25 years. *Psychological Methods*, *22*, 217–239. https://doi.org/10.1037/met0000100

Vermunt, J. K. (1997). *Log-linear models for event histories*. Sage.

Vermunt, J. K. (2010). Latent class modeling with covariates: Two improved three-step approaches. *Political Analysis*, *18*, 450–469. https://doi.org/10.1093/pan/mpq025

Vermunt, J. K., & Magidson, J. (2021). How to perform three-step latent class analysis in the presence of measurement non-invariance or differential item functioning. *Structural Equation Modeling: A Multidisciplinary Journal*, *28*, 356–364. https://doi.org/10.1080/10705511.2020.1818084

Vermunt, J. K., & Magidson, J. (2025). *Upgrade Manual for Latent GOLD Basic, Advanced, Syntax, and Choice Versions 6.0 and 6.1*. Statistical Innovations Europe.

Visser, M., & Depaoli, S. (2022). A guide to detecting and modeling local dependence in latent class analysis models. *Structural Equation Modeling: A Multidisciplinary Journal*. https://doi.org/10.1080/10705511.2022.2033622

Wang, W.-C., Shih, C.-L., & Yang, C.-C. (2009). The MIMIC method with scale purification for detecting differential item functioning. *Educational and Psychological Measurement*, *69*, 713–731. https://doi.org/10.1177/0013164409332228

Willse, J. T., & Goodman, J. T. (2008). Comparison of multiple-indicators, multiple-causes- and item response theory-based analyses of subgroup differences. *Educational and Psychological Measurement*, *68*, 587–602. https://doi.org/10.1177/0013164407312601

Winter, S. D., & Depaoli, S. (2020). An illustration of Bayesian approximate measurement invariance with longitudinal data and a small sample size. *International Journal of Behavioral Development*, *44*, 371–382. https://doi.org/10.1177/0165025419880610

Woods, C. M. (2009). Empirical selection of anchors for tests of differential item functioning. *Applied Psychological Measurement*, *33*, 42–57. https://doi.org/10.1177/0146621607314044

Woods, C. M., & Grimm, K. J. (2011). Testing for nonuniform differential item function with multiple indicator multiple cause models. *Applied Psychological Measurement*, *35*, 339–361. https://doi.org/10.1177/01466216114059

Xiao, Y., Liu, H., & Hau, K. T. (2019). A comparison of CFA, ESEM, and BSEM in test structure analysis. *Structural Equation Modeling: A Multidisciplinary Journal*, *26*, 665–677. https://doi.org/10.1080/10705511.2018.1562928

Yamaguchi, K. (2000). Multinomial logit latent class regression models: An analysis of predictors of gender role attitudes among Japanese women. *American Journal of Sociology*, *105*, 1702–1740. https://doi.org/10.1086/210470