

The Impact of MNAR Attrition on Estimation of Latent Growth Curve Models with Binary Observed Variables

Jason T. Newsom^a, Brian T. Keller^b, Mallory R. Kroeck^a, and Nicholas A. Smith^c

^aPortland State University; ^bUniversity of Missouri; ^cUniversity of Texas

ABSTRACT

A Monte Carlo simulation was used to evaluate the performance of latent growth curve (LGC) models with binary observed variables when the attrition pattern was missing not at random (MNAR) for five time points. Parameter and standard error biases for three estimation methods were compared: weighted least squares with mean and variance adjustment (WLSMV), categorical robust marginal maximum likelihood (categorical MLR), and Bayes. The results indicated that robust diagonal weighted least squares paired with multiple imputation (MI) performed best when values were missing at random. When data were missing not at random, Bayesian estimates performed best.

KEYWORDS

Attrition; bayesian estimation; missing data; Monte Carlo; weighted least squares

1. Introduction

Growth curve models have become the dominant analysis approach to studying change in the behavioral and social sciences. One important approach to growth curve analysis is the latent growth curve model (LGCM), estimated with structural equation modeling (SEM). Considerable research has been conducted documenting the performance of LGCMs with continuous observed variables under various conditions. Although there is increasing interest in applying these methods to noncontinuous variables (e.g., Masyn et al., 2014; Newsom, 2024), more research is needed about the performance of LGCMs with binary observed variables. Several simulation studies have compared different estimators for binary growth curve models (Finch, 2017; Kim & Whittaker, 2025; Muthén, 1996; Newsom & Smith, 2020), but each has investigated performance with complete data, a rare circumstance in applied longitudinal studies. This is an important omission, as SEM estimators used with binary variables make different assumptions about missing data and might perform differently when there is attrition. Thus, the present study furthers our knowledge by examining bias in the estimation of growth curve models with binary variables missing due to attrition.

Nearly all longitudinal studies involve substantial missing data because of attrition, when participants have missing data for one or more waves through the end of the study. Attrition is a missing data pattern that may be more likely to violate the missing at random (MAR) assumption than intermittently missing data. Ignorable methods for handling missing data require the missing at random (MAR)

assumption (Rubin, 1976). Informally, this means that, after accounting for observed data, the probability that a value is missing does not depend on that variable's unobserved true value. When this assumption is violated, the missingness mechanism is considered non-ignorable, or missing not at random (MNAR). Because those who leave a longitudinal study may be less motivated or have poorer mental or physical health, when factors that may be the focus of the hypotheses about change, data are less likely to meet the MAR assumption.

1.1. Estimation Approaches for LGC Models with Binary Outcomes

SEMs with binary outcomes, including LGCMs, can be estimated using several approaches. The most common is diagonal weighted least squares (DWLS; see Muthén & Muthén, 1998–2017), which uses polychoric correlations as input instead of a raw covariance matrix and is more computationally efficient than general weighted least squares because it inverts only the diagonal matrix. Less frequently used is marginal maximum likelihood (or maximum marginal likelihood), a full-information estimation (see Bock & Aitkin, 1981). Notably, this approach is typically estimated using the expectation maximization (EM) algorithm with numeric integration, which can be computationally intensive. Standard errors for both DWLS and marginal maximum likelihood may be biased when response variables are asymmetric, so robust standard error adjustments (Satorra & Bentler, 1994) are recommended (Bandalos, 2014).



Throughout, we refer to robust DWLS as weighted least squares with mean and variance adjustment (WLSMV) and to robust marginal maximum likelihood as categorical MLR.

Bayes estimation is an approach that can also be used for categorical data (Albert & Chib, 1993), including for LGCMs with binary indicators. Bayesian SEM often uses a Markov Chain Monte Carlo (MCMC) iterative process that draws parameters from a posterior distribution and then summarizes them to provide point estimates. Rubin's (1976) seminal work provides proofs that Bayesian methods work under MAR. In contrast to WLSMV and categorical MLR, Bayesian estimation does not assume an explicit sampling distribution, as its use of the posterior distribution through the MCMC random draw process allows for skewed parameter distributions.

1.2. Estimation Approaches for LGC Models with Incomplete Binary Outcomes

Given that attrition is ubiquitous in longitudinal studies, it is important to consider how missing data are handled differently across estimation approaches for binary outcomes. Because WLSMV, a limited-information estimator, relies on deletion methods to compute polychoric correlations, its missing data handling generally assumes missing completely at random (MCAR) rather than MAR (Enders & Bandalos, 2001; Jia & Wu, 2019). However, multiple imputation (MI) techniques (Rubin, 1987) can be used to satisfy MAR. In contrast, categorical MLR and Bayesian estimation methods are full-information approaches that handle missing data under MAR.

Typically, auxiliary variables are included in missing data analyses to improve the plausibility of the MAR assumption by helping account for relations between missingness and variables with incomplete data (Collins et al., 2001). Traditionally, researchers have incorporated auxiliary variables using the saturated correlates approach (Graham et al., 1997), which is often implemented through specialized software commands. The saturated correlates approach, however, is not available for categorical outcomes. In these situations, the extra DV approach, which specifies a predictive path to the auxiliary variable, may be used instead. For categorical outcomes, the effectiveness of auxiliary variables may depend on how well they account for missingness. In turn, the relative performance of MLR, WLSMV, Bayesian estimation, and MI may vary because these approaches differ in how they handle incomplete data. Yet the relative performance of WLSMV, categorical MLR, and Bayesian estimators when auxiliary variables are included has not been examined in prior research.

1.3. Prior Simulation Research on LGCMs with Binary Outcomes

To date, nearly all simulation studies evaluating estimators for LGCMs with binary outcomes have only examined data sets with complete cases. Muthén (1996) first conducted a small simulation study to examine the performance of

LGCMs with binary indicators using just DWLS without robust standard errors and manipulating only sample size across four time points and one set of values for the intercept, slope, and variances. Although bias or relative efficiency were not examined, the results indicated that rejection rates for the intercept and slope variances were close to the nominal level. Ye (2016) compared WLSMV estimation in SEM with estimates from multilevel regression and item response software programs, showing that Laplace estimation had fewer parameters and less standard error bias than WLSMV and the other multilevel regression estimation methods examined. Finch (2017) conducted a more elaborate simulation study of categorical LGCMs with complete data, examining models with binary, four ordinal-category, and six ordinal-category indicators. Linear and quadratic models with moderate and large effects were compared across four, six, and ten time points. Results indicated that WLSMV performed well in estimating fixed effects with 200 cases.

In another extensive simulation, Newsom and Smith (2020) compared the performance of WLSMV and categorical MLR estimation in estimating LGCMs with binary variables. They reported that the number of time points was critical for convergence and bias in estimates and standard errors. Among the several γ^* scaling options, the specification using delta parameterization, in which all measurement thresholds are set to 0 and all scale factors are set to 1, had the best overall performance. Categorical MLR estimation was clearly superior to WLSMV in terms of convergence, but categorical MLR and the most constrained WLSMV specifications exhibited similar bias levels, coverage, and Type I error rates when there were at least five time points with 500 cases or seven time points with 200 cases.

Kim and Whittaker (2025) compared three estimation approaches: categorical MLR, weighted least squares (WLS; non-diagonal weight matrix approach), and WLSMV for binary and ordinal indicators. They found that the categorical MLR estimation approach had low or acceptable bias in estimates of average slopes for linear models with binary indicators, outperforming the other two estimation approaches. WLSMV and WLS estimates of average slopes improved with larger sample sizes.

None of these studies, however, has focused on data with missing values. The only simulation study to date comparing estimation approaches for testing LGCMs with binary indicators when missing values are present was conducted by Asparouhov and Muthén (2021), who examined growth parameter values for WLSMV and Bayes estimators in the presence of missing values that were MAR. Two different Bayesian priors were investigated, with results suggesting that improper priors, which are based on a uniform distribution of possible variance values, may have estimation advantages because of fewer restrictions. Their simulation was small, examining only 100 replications and a single sample size of 1,000. To address the biases that can occur with the WLSMV when data are not MCAR, they combined WLSMV estimates with multiple imputation, which only assumes MAR. Findings indicated that WLSMV estimates

without MI were biased, and the Bayes estimates and the WLSMV estimates with MI showed low levels of bias.

1.4. Purpose of the Present Study

We evaluated the performance of estimation choice on latent growth curve parameter estimates under various missing data conditions. The reliance of WLSMV on pairwise deletion suggests that combining this estimator with MI may improve estimates when data are MAR (Asparouhov & Muthén, 2021), making its performance more comparable to categorical MLR or Bayesian estimation when attrition is present. This study, therefore, compared WLSMV, categorical MLR, and Bayesian estimation for missing data conditions when an auxiliary variable partially or fully accounted for the relationship of values to the probability of missingness. Two Bayesian prior settings were used to test whether fewer restrictions on the prior variance distribution might provide an advantage when data are MNAR. Researchers also have provided detailed guidance on the suggested minimum sample sizes (e.g., 200) and the sufficient sample sizes for adequate coverage (e.g., 1,000) for WLSMV and MLR estimators (Kim & Whittaker, 2025; Newsom & Smith, 2020); therefore, determination of sufficient sample sizes was not a focus of this study. Five time points were used throughout based on recommendations in prior research on LGCMs with binary variables (Newsom & Smith, 2020) and to provide more optimal conditions for examining attrition effects (e.g., compared with three or four time points in which only one or two time points could be missing). We go beyond prior work by considering growth parameter bias that occurs with attrition, examining a wider array of SEM estimators for binary variables, and investigating performance under various missing data mechanisms.

2. Method

2.1. Design

2.1.1. Overview

The study used a 2 (sample size: $N = 200$ vs. $N = 1000$) $\times$ 4 (missingness: MCAR vs. MAR vs. partial MNAR vs. pure MNAR) $\times$ 10 (estimation: WLSMV-LD vs. WLSMV-D vs. WLSMV-MI vs. MLR-LD vs. MLR-D vs. MLR-MI vs. Bayes $IW[I, p + 1]$ -LD vs. Bayes $IW[I, p + 1]$ -D vs. Bayes $IW[0, p - 1]$ -LD vs. Bayes $IW[0, p - 1]$ -D) design. LD stands for listwise deletion, D for direct missing data estimation, and MI for multiple imputation. Each cell of the design used 1,000 replications and five time points with binary observed variables. The design was not fully crossed, however, because direct estimation and MI are not needed with listwise deletion and MI is not needed with Bayesian estimation. Only four estimation approaches were compared for complete data (WLSMV, MLR, Bayes $IW[I, p + 1]$, and Bayes $IW[0, p - 1]$).

2.1.2. Data Generation

Data were generated in SAS version 9.4 and the RandomMVBinary macro (Wicklin, 2013) to create 1,000

samples for every condition, each with five variables representing five time points (y_1 , y_2 , y_3 , y_4 , and y_5). The first time point, y_1 , was created in the sample to have a mostly-symmetric distribution for the baseline variable, y_1 with the proportion of 1s to 0s equal to .45 using Bernoulli random number generation. Missing values were created to mimic three attrition patterns (missing on y_3 , y_4 , and y_5 ; missing on y_4 and y_5 , or missing on y_5), which were combined in each data set with complete cases. The proportion of cases with one of the three attrition patterns was .55 and the average rate of attrition per wave was equal to .2, consistent with the attrition rate in an ongoing national population-based study (Heeringa & Connor, 1995). Finally, a normally distributed continuous variable was created for the auxiliary variable with varying levels of association with missingness and proportion change over time, as described below.

2.1.3. Slope

The average initial proportion (.45) was the same for the first variable, y_1 . The slope was generated to have a linear increase with an average proportional increment of .05 at each wave, which corresponds with a moderate standardized slope of approximately .4. Between-person variation was incorporated so that intercepts and slopes would have significant variability of moderate magnitude. An additional slope condition with an average value of 0 is discussed in the Supplemental Material.

2.1.4. Missingness

There were four missingness conditions: MCAR, MAR, partial MNAR, and pure MNAR. The MCAR condition had attrition patterns as described above, but with the probability of missingness unrelated to proportion increments over time. The three remaining missing data conditions were produced by conditioning the attrition pattern on the magnitude of proportion change ($r = .30$), thus reflecting a violation of the MAR assumption. The three conditions differed, however, in the magnitude of the relationship of missingness to an auxiliary variable. For MAR, the auxiliary variable was associated with the rate of change ($r = .40$) and missingness ($r = .75$) such that the relationship between the slope and missingness would be eliminated when these associations were accounted for in the model. For partial MNAR (called indirect MNAR by Gomer & Yuan, 2021), the association between the auxiliary variable and the rate of change ($r = .40$) and the association between the auxiliary variable and missingness ($r = .40$) only accounted for part of the relationship between the rate of change and missingness. For pure MNAR (called direct MNAR by Gomer & Yuan, 2021), the auxiliary variable was uncorrelated with either the rate of change or with missingness ($r = .00$), leaving a moderate association between the slope and missingness ($r = .30$).



2.1.5. Model Specification

Each model specified a latent intercept and slope with five indicators, with means and variances of both factors estimated (see [Supplemental Material](#) for example Mplus code). Because of limitations in the program for inclusion of auxiliary variables with binary indicators, the “extra DV” approach suggested by Graham et al. (1997) was used, which is comparable to the “saturated correlates” approach that is implemented in the auxiliary feature for software such as Mplus and lavaan. A visual representation of the design is available in the [Supplemental Material](#). To implement the “extra DV” approach with WLSMV, theta parameterization is required, so a theta parameterized model specification equivalent to the best performing delta specification reported by Newsom and Smith (2020) was used.

2.1.6. Estimation

Models were tested using Mplus v8.9 (Muthén & Muthén, 1998–2017) with either WLSMV, MLR, or Bayes estimation (two prior settings and probit link). WLSMV is the default estimator in Mplus and the R lavaan package when categorical variables are declared. MLR, which is only available for general structural models in Mplus, used a Monte Carlo integration with eight integration points and a probit link for parameters of comparable scale to WLSMV. Bayesian estimation depends on the initial priors for the parameters, and two different prior distributions for latent variances were compared, referred to throughout as Bayes $IW(I, p+1)$ and Bayes $IW(0, -p-1)$ for their hyperparameter values. This notation is in reference to latent variance parameter priors, where IW refers to the inverse Wishart distribution (Asparouhov & Muthén, 2021) with p equal to the number of latent variable variances (e.g., 3 for a typical unconditional LGC), that place constraints on the latent variable variances and covariance. The Bayes $IW(I, p+1)$ notation uses I to refer to identity, implying a uniform distribution for the factor correlation bounded by -1 and $+1$, which we refer to as a “proper prior.” The alternative prior, Bayes $IW(0, -p-1)$, a generalization of a Jeffreys prior, is equivalent to a uniform distribution with $-\infty$ to $+\infty$ bounds on the factor correlation and, therefore, can be described as an “improper prior.” Defaults were used for the loading and threshold values.

For WLSMV and MLR estimation, models were tested comparing listwise deletion (WLSMV-LD and MLR-LD), direct missing data estimation (i.e., estimator allows missing data; WLSMV-D and MLR-D), and multiple imputation (WLSMV-MI and MLR-MI). For Bayesian estimation, models were tested comparing listwise deletion and direct missing data estimation but not MI: Bayes $IW(I, p+1)$ -LD, Bayes $IW(I, p+1)$ -D, Bayes $IW(0, -p-1)$ -LD, and Bayes $IW(0, -p-1)$ -D. MI was not used with Bayesian estimation due to redundancy, as most MI procedures inherently depend on an underlying Bayesian model to generate imputations (Enders, 2022).

MI for WLSMV and MLR estimators was implemented separately for each sample in Mplus for 20 imputed data sets. The number of imputed data sets was chosen after ensuring that samples had acceptable potential scale reduction (PSR) factors close to 1 (Gelman & Rubin, 1992).

MplusAutomation package in R (version 1.1.1; Hallquist & Wiley, 2018) was used to repeat the MI and model estimation process for the 1,000 replications, and then evaluation criteria were computed in R. The imputation method employed by Mplus is an unrestricted model with sequential regression (Asparouhov & Muthén, 2021).

2.2. Evaluation Criteria

To evaluate performance in each condition, we computed percent bias for average intercept and slope estimates, along with their standard errors, mean square error (MSE), and coverage. Relative bias, expressed as a percent, was computed for all parameter estimates. The percent relative bias computation is the same as what is commonly used, in which the percent bias is a function of the deviation of the average sample value from the population value relative to the absolute value of the population value and, then, multiplied by 100. Bias for standard errors was computed by subtracting the empirical standard deviation for the parameter from the average standard error, dividing by the empirical standard deviation, and then multiplying this quantity by 100. The comparison population values were based on complete data and the same estimator for $N=30$ million. We considered percent bias greater than approximately 5% to be unacceptable (Hoogland & Boomsma, 1998). Coverage rates, expressed as a proportion of the replications with a 95% confidence interval (for WLSMV and MLR) and credible interval (for Bayes) that contained the population value, were computed for the sample estimates of the average slope value. ([Supplemental Material](#) also includes Type I error rates for the zero-slope condition).

3. Results

Tables 1–4 contain results for MCAR, MAR, Partial MNAR, and Pure MNAR conditions.

3.1. MCAR

3.1.1. WLSMV

Overall, when missing values were MCAR, listwise deletion led to unacceptable levels of bias for most parameters and standard errors regardless of the estimation approach. When using direct estimation or MI, WLSMV showed positive bias in variance estimates in the 200 sample size condition but not the 1,000 sample size condition. Both WLSMV-LD and WLSMV-D had unbiased or only modestly biased estimates (approximately 7–9%) for all parameters and standard errors for $N=1,000$. Regardless of sample size, coverage values were close to nominal values for each of the WLSMV missing data handling approaches. MI reduced parameter bias modestly in the smaller sample size condition, but added little to no advantage in the larger sample size condition except intercept variance standard errors with WLSMV-MI (27.77%). Coverage with MI was too high for the covariance for $N=200$, however.

Table 1. Medium slope, MCAR.

N	WLSMV			MLR			Bayes $IW(I, p + 1)$		Bayes $IW(0, -p - 1)$	
	LD	D	MI	LD	D	MI	LD	D	LD	D
200										
cov i with s										
% bias	-72.290	-28.626	-11.804	-36.000	12.960	-4.238	20.960	20.480	-142.800	-62.080
SE % bias	-14.332	-4.630	11.664	-28.942	-32.639	-17.135	9.096	1.598	-3.837	-6.453
MSE	0.086	0.037	0.083	0.065	0.032	0.022	0.028	0.021	0.119	0.044
Coverage	0.965	0.964	1.000	0.803	0.744	0.980	0.945	0.927	0.862	0.881
means i										
% bias	-13.695	-8.966	-6.687	-37.241	-2.217	0.422	-7.488	-2.315	-17.586	-7.438
SE % bias	2.803	3.265	11.555	-32.801	-19.143	-49.314	4.550	4.204	9.117	6.179
MSE	0.041	0.018	0.017	0.096	0.029	0.019	0.033	0.015	0.044	0.018
coverage	0.966	0.958	0.963	0.818	0.895	0.996	0.946	0.955	0.950	0.957
means s										
% bias	9.216	4.510	3.347	5.941	-3.168	-1.694	12.376	11.188	15.248	12.277
SE % bias	8.442	1.083	4.599	-11.125	-6.724	-15.686	10.292	4.562	9.351	2.407
MSE	0.004	0.003	0.009	0.006	0.003	0.003	0.004	0.003	0.005	0.003
coverage	0.963	0.960	0.954	0.933	0.916	0.977	0.957	0.951	0.952	0.943
var i										
% bias	30.615	13.682	13.254	-1.900	-19.926	-7.652	-6.790	-4.438	61.123	26.778
SE % bias	-18.744	-8.884	41.820	-26.277	-21.750	-120.510	0.142	-5.021	-8.612	-8.386
MSE	1.806	0.582	1.860	0.993	0.423	0.311	0.612	0.341	2.796	0.757
coverage	0.946	0.955	0.948	0.789	0.673	0.949	0.920	0.912	0.841	0.862
var s										
% bias	62.319	42.464	34.878	77.727	40.303	55.489	148.788	130.152	142.576	96.212
SE % bias	2.081	10.961	4.507	-42.934	-35.419	-13.065	53.394	46.234	2.948	-1.966
MSE	0.008	0.004	0.012	0.015	0.007	0.005	0.012	0.009	0.017	0.009
coverage	0.999	0.996	0.934	0.769	0.728	0.980	0.238	0.311	0.767	0.797
1000										
cov i with s										
% bias	-8.244	-0.305	-2.773	-13.680	10.080	3.138	0.480	-3.600	-18.000	-11.280
SE % bias	-4.241	-5.767	6.290	-61.967	-62.904	11.145	-1.139	-6.222	-7.533	0.113
MSE	0.008	0.006	0.007	0.031	0.021	0.006	0.006	0.005	0.009	0.006
coverage	0.937	0.939	0.934	0.548	0.506	0.994	0.943	0.915	0.916	0.893
means i										
% bias	-1.281	-0.197	-0.740	-27.044	-2.660	12.414	-1.724	-1.133	-2.167	-1.133
SE % bias	-3.288	-0.680	5.537	-61.375	-49.221	14.295	-3.802	-1.188	-3.378	-1.176
MSE	0.008	0.004	0.004	0.054	0.013	0.007	0.008	0.004	0.008	0.004
coverage	0.947	0.954	0.953	0.524	0.684	1.000	0.938	0.953	0.937	0.953
means s										
% bias	1.471	0.294	1.120	2.376	-6.535	-3.888	3.960	3.168	2.178	1.683
SE % bias	0.368	3.493	2.112	-33.638	-35.014	2.944	2.158	3.057	-0.364	1.333
MSE	0.001	0.001	0.001	0.002	0.001	0.001	0.001	0.001	0.001	0.001
coverage	0.953	0.958	0.944	0.831	0.814	0.997	0.953	0.952	0.942	0.954
var i										
% bias	4.041	0.931	2.160	-14.849	-20.543	-14.390	-1.006	0.981	8.080	4.864
SE % bias	-1.959	-4.062	23.766	-53.462	-50.925	37.777	-1.162	-5.045	-92.966	-7.802
MSE	0.147	0.078	0.090	0.424	0.271	0.124	0.119	0.066	0.001	0.081
coverage	0.944	0.945	0.947	0.521	0.454	0.959	0.933	0.922	0.885	0.911
var s										
% bias	9.275	7.391	7.664	51.212	23.636	38.314	50.303	48.485	20.455	20.455
SE % bias	-3.472	-2.299	2.193	-71.429	-71.674	4.670	25.000	16.304	-7.483	-12.360
MSE	0.001	0.001	0.001	0.008	0.005	0.002	0.002	0.001	0.001	0.001
coverage	0.955	0.947	0.934	0.454	0.410	1.000	0.703	0.655	0.885	0.849

Note. LD = listwise deletion; D = direct missing data estimation; Bayes $IW(I, p + 1)$ = proper prior setting; Bayes $IW(0, -p - 1)$ = improper prior setting.

3.1.2. MLR

MI substantially reduced bias in all parameter estimations using MLR estimation, often to acceptable levels. Standard errors, however, were not improved for MLR-MI for $N = 200$, and, in some cases, produced substantially poorer results (e.g., % bias = -120.510 for intercept variance standard errors). Standard errors for intercept and slope were considerably better estimated with MLR-MI for $N = 1,000$ but variance parameters were not (intercept variance %bias = -14.39%; slope variance %bias = 38.31%). MLR-D generally produced more biased estimates and standard errors than MLR-MI and had coverage values that were too low.

3.1.3. Bayes

Bayes estimation with either type of prior had unacceptably high bias for the 200 sample size condition, but all parameters except slope variance and the intercept-slope covariance in the 1,000 sample size condition were unbiased. Bayes $IW(I, p + 1)$ had somewhat lower bias than Bayes $IW(I, p - 1)$. Bias in standard errors with Bayes estimation was similar for the two types of priors and were unbiased for all parameters except the slope variance (16.30% and -12.36%, respectively). Bayesian estimates and standard errors were less biased than with any of the MLR missing data approaches and the WLSMV-MI approach. Coverage rates for Bayes estimates were variable for both types

**Table 2.** Medium slope, MAR.

N	WLSMV			MLR			Bayes $IW(l, p + 1)$		Bayes $IW(0, -p - 1)$	
	LD	D	MI	LD	D	MI	LD	D	LD	D
200										
cov i with s										
% bias	-28.473	28.931	-15.963	60.800	74.800	61.849	60.800	30.160	-73.120	-43.280
SE % bias	-7.785	3.420	10.628	-1.291	-0.525	1.605	22.164	7.182	4.046	-2.700
MSE	0.063	0.026	0.043	0.023	0.022	0.019	0.024	0.018	0.068	0.033
coverage	0.955	0.931	0.936	0.755	0.717	0.852	0.928	0.932	0.939	0.923
means i										
% bias	163.498	22.759	-4.727	160.197	11.872	5.549	160.099	0.591	164.138	-4.286
SE % bias	2.923	2.900	12.143	0.254	12.935	2.023	4.011	4.689	8.107	5.803
MSE	0.150	0.018	0.019	0.144	0.016	0.015	0.137	0.015	0.152	0.018
coverage	0.615	0.929	0.958	0.589	0.960	0.977	0.561	0.957	0.661	0.955
means s										
% bias	-166.373	-60.000	-3.711	-164.059	-77.525	-13.809	-171.089	7.030	-170.495	6.535
SE % bias	0.782	1.670	5.478	1.045	7.391	2.541	9.063	4.545	5.477	-0.168
MSE	0.033	0.007	0.005	0.031	0.008	0.004	0.034	0.004	0.034	0.004
coverage	0.224	0.779	0.942	0.171	0.596	0.945	0.269	0.958	0.253	0.952
var i										
% bias	30.110	2.660	14.489	-22.795	-26.724	-22.445	-12.969	-7.358	48.586	21.302
SE % bias	-9.818	-3.551	37.583	-7.122	-7.091	-53.320	6.350	-2.448	-2.250	-6.430
MSE	1.505	0.407	0.727	0.526	0.392	0.350	0.522	0.309	2.048	0.619
coverage	0.962	0.931	0.948	0.753	0.662	0.816	0.925	0.918	0.890	0.899
var s										
% bias	28.841	15.362	29.367	-45.303	-54.848	-51.123	119.848	117.121	73.788	75.606
SE % bias	4.197	9.926	4.144	-2.500	3.869	-1.402	63.934	50.578	7.977	-1.713
MSE	0.005	0.003	0.005	0.003	0.002	0.002	0.008	0.007	0.007	0.007
coverage	0.997	0.994	0.956	0.634	0.530	0.760	0.403	0.366	0.896	0.840
1000										
cov i with s										
% bias	28.092	47.328	0.420	89.120	95.760	87.629	27.760	6.320	13.600	-0.800
SE % bias	-3.321	-2.806	6.363	0.486	13.943	1.948	0.693	-5.344	-6.794	-10.724
MSE	0.009	0.008	0.007	0.016	0.016	0.015	0.006	0.004	0.007	0.005
coverage	0.879	0.789	0.930	0.476	0.323	0.525	0.913	0.922	0.916	0.899
means i										
% bias	164.384	29.015	-0.612	165.172	15.764	9.778	164.335	0.197	164.631	0.690
SE % bias	0.718	-1.568	5.782	0.591	12.249	2.477	-2.033	-2.349	-1.641	-2.653
MSE	0.118	0.007	0.004	0.120	0.004	0.004	0.118	0.004	0.119	0.004
coverage	0.016	0.799	0.947	0.026	0.923	0.966	0.017	0.945	0.024	0.945
means s										
% bias	-164.804	-63.627	-0.041	-160.495	-78.416	-10.462	-168.812	0.396	-167.327	-2.871
SE % bias	0.388	0.000	2.500	0.422	8.374	2.145	2.642	-2.290	-1.158	-4.981
MSE	0.029	0.005	0.001	0.027	0.007	0.001	0.030	0.001	0.029	0.001
coverage	0.000	0.204	0.933	0.000	0.042	0.929	0.000	0.944	0.000	0.929
var i										
% bias	2.203	-8.253	1.961	-27.101	-35.688	-34.913	-1.142	-1.302	7.006	2.444
SE % bias	-1.562	-3.010	24.064	-13.342	-5.136	10.066	-1.557	-5.615	-5.265	-8.484
MSE	0.139	0.084	0.088	0.289	0.374	0.362	0.112	0.064	0.153	0.075
coverage	0.951	0.876	0.945	0.528	0.168	0.355	0.942	0.921	0.923	0.915
var s										
% bias	-26.957	-20.145	6.272	-84.091	-87.576	-79.512	28.636	42.576	-12.424	14.394
SE % bias	-3.543	-6.557	2.144	5.833	41.176	0.202	30.968	14.045	-10.526	-12.941
MSE	0.001	0.001	0.001	0.003	0.003	0.003	0.001	0.001	0.001	0.001
coverage	0.823	0.837	0.936	0.133	0.113	0.192	0.903	0.690	0.888	0.872

Note. LD = listwise deletion; D = direct missing data estimation; Bayes $IW(l, p + 1)$ = proper prior setting; Bayes $IW(0, -p - 1)$ = improper prior setting.

of priors—too low for $N = 200$ but acceptable for most parameters for $N = 1,000$.

3.2. MAR

3.2.1. WLSMV

Estimates had greater bias for most parameters across all estimation approaches. Compared with direct estimation, when MI was added to WLSMV, bias was reduced for $N = 200$ for some parameters but not others. Most were estimated with unacceptable bias, although the standard errors were frequently unbiased. Parameter estimation bias was eliminated with WLSMV-MI for $N = 1,000$

for all parameters and standard errors except the intercept variance standard error (24.06%). WLSMV-LD and WLSMV-D had unacceptable parameter bias for nearly all parameters, although the standard errors were generally unbiased. For $N = 200$, coverage for WLSMV-D and WLSMV-MI was near the nominal level, but not WLSMV-LD. For $N = 1,000$, however, only WLSMV-MI consistently had coverage near .95.

3.2.2. MLR

Adding MI to MLR substantially reduced bias in parameter estimates compared with MLR-LD and MLR-D; however, bias remained unacceptably high for most parameters regardless of

Table 3. Medium slope, partial MNAR.

N	WLSMV			MLR			Bayes $IW(I, p + 1)$		Bayes $IW(0, p - 1)$	
	LD	D	MI	LD	D	MI	LD	D	LD	D
200										
cov i with s										
% bias	-38.779	7.405	10.952	60.880	69.760	78.054	55.360	32.960	-80.000	-36.880
SE % bias	-4.118	-0.533	11.401	6.804	3.391	-5.824	25.211	3.306	6.041	-5.639
MSE	0.059	0.029	0.039	0.023	0.021	0.022	0.022	0.019	0.068	0.034
coverage	0.963	0.956	0.911	0.757	0.754	0.793	0.937	0.914	0.928	0.913
means i										
% bias	160.394	22.562	18.554	155.419	19.261	25.939	155.517	14.680	158.473	10.542
SE % bias	3.359	3.753	11.505	11.337	11.598	-5.062	4.390	4.344	8.896	5.107
MSE	0.145	0.018	0.018	0.135	0.017	0.017	0.130	0.016	0.144	0.018
coverage	0.648	0.936	0.944	0.602	0.941	0.960	0.588	0.942	0.681	0.952
means s										
% bias	-165.784	-61.275	-51.187	-162.970	-77.426	-57.525	-167.327	-31.287	-166.436	-31.782
SE % bias	4.523	1.848	4.720	3.163	8.547	0.718	10.811	8.255	8.307	3.970
MSE	0.032	0.007	0.006	0.030	0.008	0.006	0.033	0.004	0.032	0.004
coverage	0.247	0.783	0.838	0.187	0.618	0.812	0.283	0.930	0.266	0.904
var i										
% bias	28.740	6.251	8.039	-24.337	-26.878	-25.597	-14.173	-8.259	46.988	19.747
SE % bias	-10.895	-6.123	41.514	-3.558	-1.423	-11.059	8.333	-4.034	-0.465	-7.847
MSE	1.480	0.458	0.599	0.534	0.386	0.376	0.499	0.318	1.929	0.608
coverage	0.959	0.939	0.926	0.734	0.685	0.767	0.915	0.901	0.884	0.888
var s										
% bias	36.667	26.377	24.214	-42.121	-52.121	-53.189	123.333	117.424	83.030	72.576
SE % bias	10.106	8.834	4.479	4.847	11.677	-0.861	73.295	53.314	12.899	0.468
MSE	0.005	0.004	0.005	0.002	0.002	0.002	0.008	0.007	0.008	0.006
coverage	0.999	0.993	0.958	0.656	0.576	0.710	0.368	0.351	0.876	0.838
1000										
cov i with s										
% bias	21.603	31.603	22.188	88.640	91.360	99.264	24.240	11.200	9.760	4.640
SE % bias	-3.669	-4.674	5.960	4.072	17.805	1.706	1.120	-6.555	-5.206	-11.872
MSE	0.008	0.007	0.007	0.016	0.015	0.018	0.006	0.005	0.007	0.005
coverage	0.888	0.870	0.890	0.487	0.370	0.375	0.900	0.899	0.908	0.891
means i										
% bias	163.941	28.719	25.521	164.483	25.468	33.191	161.921	14.631	161.970	15.025
SE % bias	-1.297	-1.042	5.471	2.071	10.623	0.953	-2.751	-1.712	-2.573	-1.864
MSE	0.118	0.007	0.006	0.118	0.006	0.007	0.115	0.004	0.115	0.004
coverage	0.022	0.816	0.839	0.024	0.847	0.821	0.020	0.909	0.026	0.910
means s										
% bias	-164.510	-63.137	-53.470	-160.297	-79.604	-58.570	-166.040	-37.228	-164.455	-39.604
SE % bias	2.353	2.655	2.129	2.575	8.673	1.079	3.422	5.333	-0.386	2.703
MSE	0.029	0.005	0.004	0.027	0.007	0.004	0.029	0.002	0.028	0.002
coverage	0.000	0.205	0.389	0.000	0.035	0.257	0.000	0.643	0.000	0.583
var i										
% bias	0.688	-5.587	-3.270	-30.012	-35.657	-37.660	-3.407	-2.753	4.772	0.870
SE % bias	-1.932	-3.951	22.740	-9.970	1.161	9.766	-0.738	-5.908	-3.558	-8.923
MSE	0.135	0.078	0.083	0.327	0.373	0.407	0.109	0.066	0.137	0.073
coverage	0.929	0.893	0.920	0.467	0.176	0.274	0.924	0.913	0.928	0.904
var s										
% bias	-17.536	-8.406	-3.409	-82.424	-85.758	-81.689	32.727	42.273	-5.152	12.879
SE % bias	-4.528	-5.578	2.088	-2.878	45.378	0.187	28.049	11.602	-9.728	-16.288
MSE	0.001	0.001	0.001	0.003	0.003	0.003	0.001	0.001	0.001	0.001
coverage	0.862	0.907	0.926	0.162	0.172	0.146	0.871	0.692	0.919	0.863

Note. LD = listwise deletion; D = direct missing data estimation; Bayes $IW(I, p + 1)$ = proper prior setting; Bayes $IW(0, p - 1)$ = improper prior setting.

sample size. Standard errors tended to be relatively unbiased, particularly with a larger sample size. MLR-MI, but not MLR-LD or MLR-D, had standard errors that were unbiased for $N = 1,000$, except standard errors for the intercept variance (10.07%). Acceptable coverage was obtained only for the average intercept and slope with MLR-MI estimation.

3.2.3. Bayes

Both Bayes $IW[I, p + 1]$ -D and Bayes $IW[I, p - 1]$ -D had lower bias levels than MLR. However, bias was worse than WLSMV-MI for $N = 200$, with overestimation of variances, in particular. For $N = 1,000$, bias was eliminated with both Bayesian prior

types except for slope variance estimates. Standard errors were also generally accurate with Bayes estimation in the $N = 1,000$ condition, except that standard errors for slope variances were overestimated with Bayes $IW(I, p + 1)$ and underestimated with Bayes $IW(I, p - 1)$. Coverage for Bayes estimates with direct missing data estimation was unacceptably low for $N = 200$ but close to the expected proportion for $N = 1,000$, except for the slope variance in which values were far too low.

3.3. Partial MNAR

Contrary to results when the MAR assumption could be conditionally met, bias for most parameters and coverage

Table 4. Medium slope, pure MNAR.

N	WLSMV			MLR			Bayes $IW(l,p+1)$		Bayes $IW(0,-p-1)$	
	LD	D	MI	LD	D	MI	LD	D	LD	D
200										
cov i with s										
% bias	-31.603	19.008	43.804	31.040	24.320	29.957	56.640	41.120	-89.760	-33.920
SE % bias	-5.074	1.300	10.988	6.803	-19.369	-22.304	17.734	9.905	3.547	-0.291
MSE	0.058	0.027	0.036	0.050	0.045	0.017	0.025	0.019	0.080	0.031
coverage	0.951	0.934	0.871	0.755	0.727	0.963	0.932	0.939	0.913	0.915
means i										
% bias	161.281	23.547	40.979	140.099	37.340	43.518	159.606	27.438	164.680	23.103
SE % bias	-0.539	2.650	11.228	232.379	30.921	-6.940	1.531	3.519	6.104	5.666
MSE	0.149	0.019	0.023	0.122	0.029	0.026	0.138	0.018	0.156	0.019
coverage	0.620	0.930	0.894	0.829	0.882	0.990	0.560	0.925	0.651	0.938
means s										
% bias	-167.157	-60.980	-96.875	-162.376	-72.079	-100.783	-172.079	-64.356	-172.178	-63.267
SE % bias	2.063	0.923	4.429	54.472	2.694	2.906	7.317	3.467	4.697	0.743
MSE	0.033	0.007	0.013	0.031	0.009	0.013	0.035	0.007	0.035	0.007
coverage	0.235	0.786	0.512	0.294	0.662	0.654	0.295	0.775	0.267	0.765
var i										
% bias	27.925	3.713	0.424	-2.128	-9.377	-13.458	-9.660	-9.117	55.605	20.543
SE % bias	-9.871	-3.198	39.847	-6.686	-19.147	-65.928	6.540	1.185	-2.597	-3.224
MSE	1.415	0.406	0.492	0.998	0.541	0.277	0.523	0.295	2.347	0.580
coverage	0.962	0.936	0.905	0.825	0.752	0.926	0.935	0.925	0.855	0.905
var s										
% bias	28.406	18.841	-1.455	-25.606	9.545	26.690	130.909	121.364	98.333	76.970
SE % bias	10.886	8.453	4.120	-22.809	-36.230	2.474	53.941	49.725	6.427	-1.490
MSE	0.005	0.003	0.004	0.007	0.007	0.003	0.009	0.008	0.010	0.007
coverage	0.999	0.999	0.904	0.556	0.634	0.984	0.350	0.358	0.859	0.852
1000										
cov i with s										
% bias	28.550	45.191	47.857	40.160	45.040	33.713	32.320	22.880	20.240	18.240
SE % bias	-6.429	-4.348	5.696	-47.432	-60.058	4.762	-2.667	-4.792	-12.864	-11.047
MSE	0.009	0.008	0.009	0.039	0.022	0.006	0.007	0.005	0.009	0.005
coverage	0.864	0.810	0.794	0.524	0.438	0.988	0.893	0.890	0.884	0.875
means i										
% bias	164.828	28.916	45.042	157.143	44.384	54.335	164.581	28.670	164.778	29.310
SE % bias	-0.942	-1.568	5.339	43.890	-35.653	11.271	-1.912	-1.549	-1.641	-1.368
MSE	0.119	0.007	0.012	0.128	0.019	0.018	0.119	0.007	0.119	0.007
coverage	0.022	0.807	0.633	0.320	0.584	0.989	0.021	0.822	0.024	0.820
means s										
% bias	-165.098	-63.627	-96.367	-162.772	-74.554	-100.555	-168.812	-72.079	-167.129	-74.158
SE % bias	-1.887	-0.429	2.023	3.504	-33.815	2.466	0.366	0.431	-3.371	-2.193
MSE	0.029	0.005	0.010	0.028	0.007	0.011	0.030	0.006	0.029	0.006
coverage	0.000	0.208	0.017	0.039	0.172	0.230	0.000	0.128	0.000	0.100
var i										
% bias	1.680	-8.046	-8.421	-6.015	-21.184	-19.900	-2.586	-4.043	4.994	-0.901
SE % bias	-8.003	-3.796	21.875	-45.077	-48.628	32.132	-7.273	-6.613	-12.558	-9.741
MSE	0.158	0.084	0.092	0.553	0.270	0.155	0.129	0.066	0.171	0.071
coverage	0.939	0.870	0.868	0.609	0.457	1.000	0.923	0.908	0.906	0.907
var s										
% bias	-23.623	-15.942	-24.491	-31.212	-3.939	12.971	29.242	36.515	-13.788	2.424
SE % bias	-4.247	-6.478	1.995	-65.929	-74.324	2.006	31.210	19.048	-10.811	-15.625
MSE	0.001	0.001	0.001	0.005	0.004	0.001	0.001	0.001	0.001	0.001
coverage	0.834	0.877	0.844	0.296	0.345	1.000	0.896	0.791	0.892	0.869

Note. LD = listwise deletion; D = direct missing data estimation; Bayes $IW(l,p+1)$ = proper prior setting; Bayes $IW(0,-p-1)$ = improper prior setting.

rates were generally unacceptable when MAR could not be fully met. This was the case regardless of the estimation approach.

3.3.1. WLSMV and MLR

Although MI tended to reduce bias for WLSMV estimation relative to WLSMV-D, use of MI increased bias somewhat with MLR relative to MLR-D. Sample size also improved estimates when WLSMV was used with MI, no such improvement was observed with a larger sample size when

MLR was used with MI. Standard errors, however, were often accurately estimated for both WLSMV and MLR estimators with direct missing data estimation for $N=1,000$. WLSMV-MI had slightly higher bias in standard errors than WLSMV-D, and MLR-MI had slightly lower bias in standard errors than MLR-D. Although estimation with both Bayesian priors had unacceptable bias for all parameters with $N=200$, both Bayes estimates had lower bias than WLSMV or MLR with direct estimation or with MI estimation for $N=1,000$. Bias was still unacceptably high for most parameters, however.

3.3.2. Bayes

Bayes with direct missing data had somewhat lower bias than the other estimation approaches. Results for the two types of priors were similar, but Bayes $IW(I, p-1)$ had lower bias for the slope variance and covariance estimates than Bayes $IW(I, p+1)$.

3.4. Pure MNAR

3.4.1. WLSMV and MLR

Bias in nearly all parameters was unacceptably high in this condition, regardless of the estimation approach. As with partial MNAR, MI had a tendency to increase bias for WLSMV and MLR estimation for most parameters in both sample sizes. Bias did not decrease dramatically for most parameters in any of the estimation approaches for $N = 1,000$.

3.4.2. Bayes

Bias for WLSMV-D, Bayes $IW(I, p+1)$, and Bayes $IW(I, p-1)$ were better than the other estimation approaches and were roughly comparable. In a few instances, parameter bias was within acceptable ranges (intercept variance for Bayes $IW(I, p+1)$ and Bayes $IW(I, p-1)$, slope variance for Bayes $IW(I, p-1)$). Standard errors had low bias and were mostly within acceptable limits for the direct missing data handling with each of the estimation approaches for $N = 1,000$. Standard errors for Bayesian direct missing data estimation for slope variances were unacceptably biased, though modestly so (19.05% for Bayes $IW(I, p+1)$ and -15.63% for Bayes $IW(I, p-1)$), as were standard errors for the intercept-slope covariance with Bayes $IW(I, p-1)$ (-11.05%). Coverage was generally too low, especially for average slope estimates, a pattern similar across estimators. Coverage was generally better for $N = 200$, which was poor for listwise deletion and MI for each of the estimators but often approached the nominal level for direct missing data estimation for each of the estimators.

4. Discussion

Although the results from this study were complex, several general conclusions can be drawn. As was expected from prior theory and findings (e.g., Enders & Bandalos, 2001), listwise deletion led to the highest levels of bias, was the least efficient, and had the poorest coverage and significance rates across all missing data mechanism conditions, particularly when data were MNAR. Although, when data did not fully meet the MAR assumption (the conditional MNAR and pure MNAR conditions), none of the approaches to estimating missing data resulted in a complete absence of parameter or standard error bias. A sample size of 200 was not adequate for any of the estimation approaches, even when the MAR assumption was met. This is markedly different from the small sample size ($N = 100$) success of ML and Bayesian data estimation for MAR attrition patterns using continuous growth curve indicators, except when priors were fully incorrect (Kim et al., 2020). In the larger

sample size condition ($N = 1,000$) of the present study with categorical estimators, WLSMV performed best and MLR performed the least well under MAR, although MLR estimates were improved when analyzing multiply imputed data. This is contrary to expectations because WLSMV, in part, relies on pairwise deletion (Asparouhov & Muthén, 2021) and MLR is considered a full information estimation (DeMars, 2012). Both Bayes and MI, which can be considered to have a Bayesian underpinning (Enders, 2022), may improve estimation when missing data do not fully meet the MAR restriction, as the unrestricted model for imputation makes minimal assumptions about the full data distribution (Linero & Daniels, 2018).

When data did not meet the MAR assumption conditionally—in either the partial MNAR or full MNAR condition—WLSMV estimates with MI were no longer superior to Bayesian estimation. In fact, for both of these missing data mechanism conditions, Bayesian estimation with missing data estimation produced the least biased estimates and standard errors, with most values within acceptable limits for the larger sample size condition. One exception was the average slope, which was unacceptably biased high. Differences between the proper and improper Bayesian priors settings were generally not large, although improper priors were more likely to reduce bias values to within acceptable limits for both the partial MNAR and full MNAR conditions. Although the results may appear inconsistent with the smaller simulation conducted by Asparouhov and Muthén (2021), who found an advantage for proper priors, only the MAR mechanism was examined in that study. It also should be noted that missingness in that study was related to each indicator, rather than to the rate of change in the indicators over time, and may lead to parameter bias for slope estimates that could differ from the bias that occurs with attrition patterns that are associated with rates of change in the dependent variable. In our study, for conditions in which an auxiliary variable could fully or partially account for the relationship between the probability of attrition and change in the dependent variable, bias was lower than when values were purely MNAR, and this reinforces the recommendation to incorporate auxiliary variables even if MNAR is not likely to be met (Collins et al., 2001).

Based on our study's findings and prior work with complete data, LGC models with binary variables should use a sample size greater than 200 and WLSMV estimation provides the best option (Bandalos, 2014; Boulet, 1996; Finch, 2017; Newsom & Smith, 2020). Use of WLSMV in conjunction with MI was expected to reduce or eliminate bias when MCAR was not met (Asparouhov & Muthén, 2021), given that WLSMV is considered a limited information estimator (DeMars, 2012) and partially relies on pairwise deletion (Enders & Bandalos, 2001; Jia & Wu, 2019). Indeed, this was the case—the combination of MI with WLSMV led to unbiased estimates if the MAR assumption was met with the inclusion of an auxiliary variable. MI with WLSMV, however, did not improve estimates when data were MNAR. It also was surprising that, under MAR, WLSMV outperformed MLR when MI was used, given that marginal



maximum likelihood is considered a full likelihood estimator that should perform better when the MAR assumption is met. This is perhaps reflective of the ability for MI to better account for the auxiliary variable than direct missing data estimation approach with the extra DV model specification and an overcorrection when used with MLR (Graham et al., 1997).

Perhaps the most remarkable of the findings was the performance of Bayesian estimation when values were MNAR. Bayesian estimation, with some advantages of the use of improper priors, often produced unbiased and otherwise the least biased estimates compared to other estimators or missing data handling approaches. It was not unexpected that parameter bias would occur for MNAR, either in part or in full, but Bayesian estimation produced surprisingly low bias for most parameter estimates and standard errors, outperforming the other two estimation approaches. The average slope was still overestimated by more than one-third, and the slope variance was considerably overestimated when the proper prior setting was used. Yet, the use of Bayesian estimation with improper priors produced acceptable estimates and standard errors, except for the aforementioned overestimate of the slope and a modest overestimate of the average intercept (by approximately 15%), covariance standard error (approximately 12%), and the variance standard error (approximately 15%). This is a relatively impressive result given that there are few optimal solutions for dealing with MNAR (Enders, 2011). The advantage of Bayesian estimation, and, in particular, Bayesian estimation with improper priors, may occur because there are fewer restrictions on the complete data distribution (Asparouhov & Muthén, 2021). Even improper Wishart priors are well-known to still provide information (Gelman, 2006). Caution is warranted in interpreting these results, because biases will depend on the direction and magnitude of the selection characteristics of the MNAR mechanism. Future work might compare the relative performance of the approaches examined here, or the combination of these estimation methods with various SEM modeling approaches for MNAR that have been proposed (see Newsom, 2024, for a review).

5. Limitations

Although this study constitutes the most extensive simulation conducted to date investigating estimation for growth curve models with binary indicators, several issues were not addressed. Our study only included one rate of attrition, but we fully expect that biases would increase with higher levels of attrition. Moreover, the models tested were limited to linear changes only and only included binary variables (Jia & Wu, 2019) that were relatively symmetric (at baseline, indicators had a 45%/55% ratio). Without a comparison to non-robust estimates, it is unclear if these adjustments averted biases or over-adjusted standard errors, particularly in the smallest sample size condition (Maydeu-Olivares, 2017). A linear relationship between growth parameters and change was assumed, and we cannot generalize to more complex relationships between auxiliary variables and attrition or

growth parameters (Thoemmes & Mohan, 2015). Moreover, because we did not expect attrition to differentially affect bias for different sample sizes, we did not include an extensive array of sample size conditions.

6. Recommendations for Researchers

There are several recommendations that can be offered for researchers. As others have demonstrated (e.g., Collins et al., 2001), our findings support the recommendation that inclusion of auxiliary variables can improve the chances of meeting MAR assumptions and reduce the bias that occurs when values are MNAR. Use of MI in conjunction with WLSMV produces the most efficient and least biased results when data meet the MAR assumptions, but if there is doubt about meeting this assumption, then Bayesian estimates, particularly with improper priors will provide better estimation under MNAR.

Acknowledgements

We thank Craig Enders for helpful suggestions at the initial stages of this project and the participants at the 2023 9th Modern Modeling Methods Conference, Storrs, CT, for their questions and feedback.

References

- Albert, J. H., & Chib, S. (1993). Bayes analysis of binary and polychotomous response data. *Journal of the American Statistical Association*, 88, 669–679. <https://doi.org/10.1080/01621459.1993.10476321>
- Asparouhov, T., & Muthén, B. (2021). *Bayesian analysis of latent variable models using Mplus, version 5* (Technical report). <https://www.statmodel.com/download/BayesAdvantages18.pdf>
- Bandalos, D. L. (2014). Relative performance of categorical diagonally weighted least squares and robust maximum likelihood estimation. *Structural Equation Modeling: A Multidisciplinary Journal*, 21, 102–116. <https://doi.org/10.1080/10705511.2014.859510>
- Bock, R. D., & Aitkin, M. (1981). Marginal maximum likelihood estimation of item parameters. An application of an EM algorithm. *Psychometrika*, 46, 443–459. <https://doi.org/10.1007/BF02293801>
- Boulet, J. R. (1996). *The effect of nonnormal ability distributions on IRT parameter estimation using full-information and limited-information methods* [Unpublished master's thesis]. University of Ottawa.
- Collins, L. M., Schafer, J. L., & Kam, C. M. (2001). A comparison of inclusive and restrictive strategies in modern missing data procedures. *Psychological Methods*, 6, 330–351. <https://doi.org/10.1037/1082-989X.6.4.330>
- DeMars, C. (2012). A comparison of limited-information and full-information methods in Mplus for estimating item response theory parameters for nonnormal populations. *Structural Equation Modeling: A Multidisciplinary Journal*, 19, 610–632. <https://doi.org/10.1080/10705511.2012.713272>
- Enders, C. K. (2011). Missing not at random models for latent growth curve analyses. *Psychological Methods*, 16, 1–16. <https://doi.org/10.1037/a0022640.supp>
- Enders, C. K. (2022). *Applied missing data analysis* (2nd ed.). Guilford.
- Enders, C. K., & Bandalos, D. L. (2001). The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal*, 8, 430–457. https://doi.org/10.1207/S15328007SEM0803_5
- Finch, W. H. (2017). Investigation of parameter estimation accuracy for growth curve modeling with categorical indicators: Impact of

- number of measurement occasions and number of categories. *Methodology*, 13, 98–112. <https://doi.org/10.1027/1614-2241/a000134>
- Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, 7, 457–472. <https://doi.org/10.1214/ss/1177011136>
- Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). *Bayesian Analysis*, 1, 515–534. <https://doi.org/10.1214/06-BA117A>
- Graham, J. W., Hofer, S. M., Donaldson, S. I., MacKinnon, D. P., & Schafer, J. L. (1997). Analysis with missing data in prevention research. In K. Bryant, M. Windle, & S. West (Eds.), *The science of prevention: Methodological advances from alcohol and substance abuse research* (pp. 325–366). American Psychological Association.
- Gomer, B., & Yuan, K. H. (2021). Subtypes of the missing not at random missing data mechanism. *Psychological Methods*, 26, 559–598. <https://doi.org/10.1037/met0000377>
- Hallquist, M. N., & Wiley, J. F. (2018). MplusAutomation: An R package for facilitating large-scale latent variable analyses in Mplus. *Structural Equation Modeling: A Multidisciplinary Journal*, 25, 621–638. <https://doi.org/10.1080/10705511.2017.1402334>
- Heeringa, S. G., & Connor, J. H. (1995). Sample design and methods for the health and retirement survey [Technical Report]. Statistical Design Group, Survey Research Center, University of Michigan. <https://hrs.isr.umich.edu/publications/biblio/5310>
- Hoogland, J. J., & Boomsma, A. (1998). Robustness studies in covariance structure modeling: An overview and meta-analysis. *Sociological Methods & Research*, 26, 329–367. <https://doi.org/10.1177/0049124198026003003>
- Jia, F., & Wu, W. (2019). Evaluating methods for handling missing ordinal data in structural equation modeling. *Behavior Research Methods*, 51, 2337–2355. <https://doi.org/10.3758/s13428-018-1187-4>
- Kim, J., & Whittaker, T. A. (2025). Latent growth model with categorical indicators: Comparison of the different estimators. *Structural Equation Modeling: A Multidisciplinary Journal*, 32, 400–414. <https://doi.org/10.1080/10705511.2024.2416064>
- Kim, S. Y., Huh, D., Zhou, Z., & Mun, E. Y. (2020). A comparison of Bayesian to maximum likelihood estimation for latent growth models in the presence of a binary outcome. *International Journal of Behavioral Development*, 44, 447–457. <https://journals.sagepub.com/doi/full/10.1177/0165025419894730>
- Linero, A. R., & Daniels, M. J. (2018). Bayesian approaches for missing not at random outcome data: The role of identifying restrictions. *Statistical Science: a Review Journal of the Institute of Mathematical Statistics*, 33, 198–213. <https://doi.org/10.1214/17-STS630>
- Masyn, K. E., Petras, H., & Liu, W. (2014). Growth curve models with categorical outcomes. In G. Bruinsma & D. Weisburd (Eds.), *Encyclopedia of criminology and criminal justice* (pp. 2013–2025). Springer.
- Maydeu-Olivares, A. (2017). Maximum likelihood estimation of structural equation models for continuous data: Standard errors and goodness of fit. *Structural Equation Modeling: A Multidisciplinary Journal*, 24, 383–394. <https://doi.org/10.1080/10705511.2016.1269606>
- Muthén, B. (1996). Growth modeling with binary responses. In A. V. Eye & C. Clogg (Eds.), *Categorical variables in developmental research: Methods of analysis* (pp. 37–54). Academic Press.
- Muthén, L. K., & Muthén, B. O. (1998–2017). *Mplus user's guide* (8th ed.). Muthén & Muthén.
- Newsom, J. T. (2024). *Longitudinal structural equation modeling: A comprehensive introduction* (2nd ed.). Routledge.
- Newsom, J. T., & Smith, N. A. (2020). Performance of latent growth curve models with binary variables. *Structural Equation Modeling: A Multidisciplinary Journal*, 27, 888–907. <https://doi.org/10.1080/10705511.2019.1705825>
- Rubin, D. B. (1976). Inference and missing data. *Biometrika*, 63, 581–592. <https://doi.org/10.1093/biomet/63.3.581>
- Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. John Wiley & Sons.
- Satorra, A., & Bentler, P. M. (1994). Corrections to test statistics and standard errors in covariance structure analysis. In A. von Eye & C. C. Clogg (Eds.), *Latent variables analysis: Applications for developmental research* (pp. 399–419). Sage.
- Thoemmes, F., & Mohan, K. (2015). Graphical representation of missing data problems. *Structural Equation Modeling: A Multidisciplinary Journal*, 22, 631–642. <https://doi.org/10.1080/10705511.2014.937378>
- Wicklin, R. (2013). *Simulating data with SAS*. SAS Institute, Inc.
- Ye, F. (2016). Latent growth curve analysis with dichotomous items: Comparing four approaches. *The British Journal of Mathematical and Statistical Psychology*, 69, 43–61. <https://doi.org/10.1111/bmsp.12058>