COMMUNICATIONS IN STATISTICS - SIMULATION AND COMPUTATION®
2020, VOL. 49, NO. 7, 1746–1767
https://doi.org/10.1080/03610918.2018.1506028

Taylor & Francis
Taylor & Francis Group



# Comparison of maximum likelihood approach, Diggle–Kenward selection model, pattern mixture model with MAR and MNAR dropout data

Nan Chen[a*], Meijuan Li[b,c], and Hongyun Liu[a,d]

[a]Faculty of Psychology, Beijing Normal University, Beijing, China; [b]Collaborative Innovation Center of Assessment toward Basic Education Quality, Beijing Normal University, Beijing, China; [c]Beijing Research Center for Education Supervision and Quality Assessment, Beijing Academy of Educational Sciences, Beijing, China; [d]Beijing Key Laboratory of Applied Experimental Psychology, National Demonstration Center for Experimental Psychology Education (Beijing Normal University), Faculty of Psychology, Beijing Normal University, Beijing, China

**ABSTRACT**

In longitudinal studies, missing data are ubiquitous. This article made a comparison of three model-based techniques for handling different types of missing data (i.e., missing at random (MAR)-based maximum likelihood (ML) approach, missing not at random (MNAR)-based Diggle–Kenward (DK) selection model and MNAR-based pattern mixture (PM) model) in longitudinal studies through a Monte Carlo simulation study. Two influential factors were considered: the dropout rates (5%, 10%, 20%, and 40%) and the sample sizes (100, 300, 500, and 1000) under MAR and MNAR missingness mechanisms respectively. The results indicated that the model selection was a crucial issue when researchers were dealing with missing data in longitudinal studies because under MNAR mechanism, DK method outperformed MAR-based ML approach, but PM method performed worse than MAR-based method did. The differences of the parameter estimation among three methods became more significant as the sample size and the dropout rate increased.

**ARTICLE HISTORY**
Received 28 September 2017
Accepted 23 July 2018

**KEYWORDS**
Diggle–Kenward selection model; Latent growth curve model; Maximum likelihood approach; Missing not at random (MNAR); Pattern mixture model

## 1. Introduction

Longitudinal data have been widely used in different domains, such as psychology, education, and medicine. Compared with cross-sectional studies, a longitudinal study is advantageous in that researchers are able to detect development or changes in the characteristics of the target population at both the group level and the individual level, which suggests a more likely cause-and-effect relationship (Singer and Willett 2003). However, in longitudinal studies where the individuals have repeated observations, missing data remains a common problematic issue. Individuals don't always finish the whole set of experiments or assessments because of their own traits or other external factors such as feeling bored or confused. Missing data might negatively affect the quality of statistical data in many different ways, such as reducing sample size, biasing

**CONTACT** Hongyun Liu hyliu@bnu.edu.cn Beijing Normal University, No. 19 Xin Jie Kou Wai Street, Hai Dian District, Beijing 100875, China.
*IQVIA, Beijing, China.
Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/lssp.

parameter estimates (Collins, Schafer, and Kam 2001), and reducing statistical power (Acock 2005). It can even distort research results, especially when systematic differences exist between individuals who are completely observed and those with missing observations (Langkamp, Lehman, and Lemeshow 2010).

When it comes to the mechanisms of missingness, the following three terms are mentioned: missing completely at random (MCAR), missing at random (MAR), and missing not at random (MNAR). A MAR mechanism holds when the probability of missingness only depends on observed data, but not on unobserved data. MCAR is a special case of MAR, in which missingness does not depend on the observed data either. An MNAR occurs, however, when the propensity for missingness does depend on unobserved data (Schafer and Graham 2002). During the past decades, various attempts to deal with missing data have received considerable attention, especially the relatively "modern" analytic approaches that assume MAR, that is, Maximum Likelihood (ML) and Multiple Imputation (MI) procedures (e.g., Carpenter, Kenward, and Vansteelandt 2006; Enders 2010; Little and Rubin, 2002; Schafer and Graham 2002). For example, in a longitudinal study on changes of life quality throughout a clinical trial for a new cancer medication, patients with rapidly decreasing scores of life quality are more likely to leave the study because they die or become too ill to participate. This is an example of the MNAR mechanism, because the missing probability of the outcome variable depends on the unobserved value of the outcome variable (i.e. quality of life). The missingness mechanism in this example is characterized as MNAR, as long as the probability of a missing value depends on the unobserved outcome variable (i.e. quality of life). However, if patients refuse the study due to moving house or less education, the missingness mechanism is MAR because there is no relationship between the propensity of missing data and the incomplete outcome variable. Although the MAR mechanism is widely assumed and applied and MAR-based approaches represent the current state of art (Schafer and Graham, 2002), certain situations exist where the assumptions of the MAR mechanism are incorrect. An MNAR mechanism occurs when the probability of missing data on a variable $Y$ is related to the values of $Y$ itself, even after controlling for other variables. Under the MNAR mechanism, the analyses based on a sampling distribution might be invalid. Therefore, there will be a sizable bias in estimates of parameters and standard errors if the model fails to account for this kind of mechanism, which will lead to worse confidence interval estimates (Schafer and Graham 2002; Little and Rubin 2002).

The model-based approaches could be used to deal with MNAR missing data. As Little and Rubin (2002) and Schafer and Graham (2002) recommended, the selection model and the pattern mixture model could be employed. The two models both integrate a part describing the missing mechanism into the substantive model but are different in the decomposition of the joint distribution combining the response variable ($Y$) and the missing data indicator ($d$). In addition, both models could be applied with different types of latent growth models (McArdle and Epstein 1987; Meredith and Tisak 1990). In view of the possible differences in response patterns or development trends causing missingness, some extended models containing latent class variables were proposed (Beunckens et al. 2008; Muthén et al. 2011; Roy 2003). These MNAR-based approaches for dealing with missing data have been developed for years and are widely used at present (Enders 2011; Muthén et al. 2011; Power et al. 2012). Moreover, programs for handling missing data have

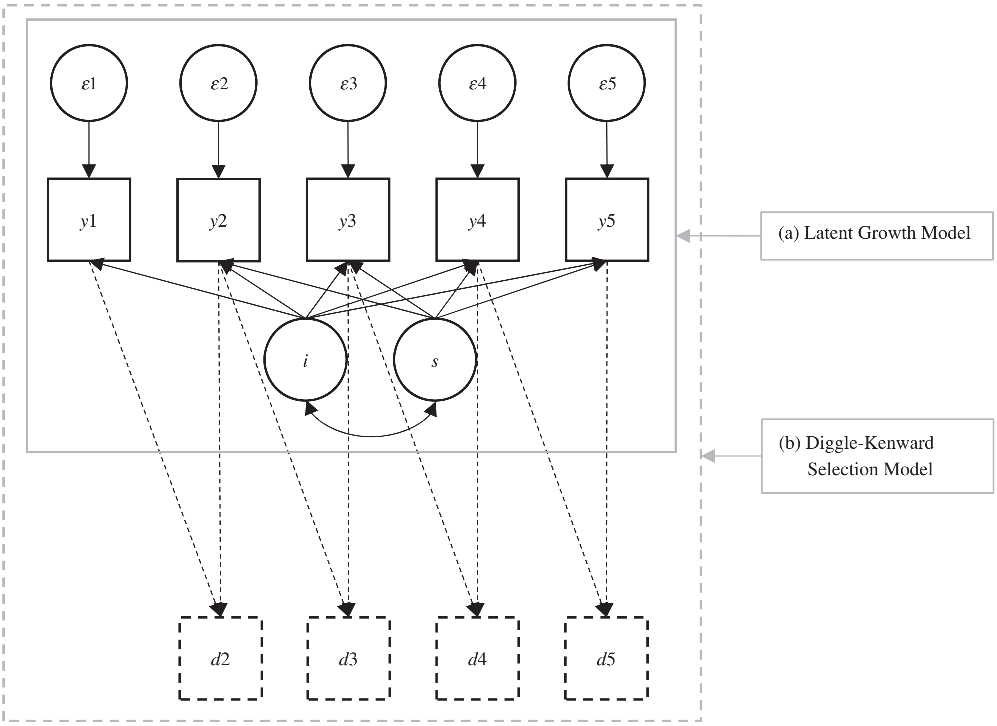



**Figure 1.** (a) Latent growth model (shown with solid line); (b) Diggle–Kenward Selection Model (shown with dotted line based on LGM).

already been embedded into some statistical softwares, such as Mplus (Muthén and Muthén 1998–2012), AMOS, and LISREL (Jöreskog and Sörbom 1996). For instance, Mplus provides different choices of models dealing with missing data when latent growth model (LGM) is employed, so as to generate dummy variables $d_t$ automatically.

Although different techniques have been explored, researchers proposed different viewpoints in regard to how to select the method in practice, especially for an MNAR mechanism. Some researchers hold the view that ignoring the MNAR mechanism in data analyses will produce greater deviation in estimation; however, there is also a viewpoint holding that a better MAR model still outperforms a worse MNAR model, even if the MAR assumption is violated (Schafer, 2003). However, there are few studies to compare the performance of different MNAR-based approaches in the current technical literatures.

The purpose of this article is to compare three model-based approaches of missing data handlings, namely MAR-based maximum likelihood approach, the MNAR-based Diggle–Kenward selection model and the pattern mixture model, under different conditions (i.e., varying sample size, different dropout rate, and the missing mechanisms) through Monte Carlo simulations. Finally, an illustration on a real dataset is used to demonstrate the general procedures for dealing with missing data, and some suggestions are provided.

## 2. Model-based missing data techniques in the framework of LGM

The article is based on the longitudinal data analysis, and a brief overview of the latent growth model is warranted before proceeding. Figure 1 shows a path diagram with solid lines of a linear growth model with five equally spaced repeated assessments. The unit factor loadings for the intercept latent variable reflect the fact that the intercept is a constant component of each individual's idealized growth trajectory, and the loadings for the slope capture the changing of latent variables. A number of resources are available to readers who want additional details on latent growth models (Bollen and Curran 2006; Singer and Willett 2003).

### 2.1. MAR-based maximum likelihood approach

The MAR-based maximum likelihood (ML) approach defines a model of observed data and makes inferences using likelihood functions. When the maximum likelihood approach is applied for handling missing data, the factorization of the joint distribution $f(Y_i, d_i)$ of $Y$ and $d$ is simplified as the product of two independent distributions and it is unnecessary to estimate the parameters that indicate missingness (Enders 2011).

Methodologists currently regard maximum likelihood as a state-of-art missing data technique (Schafer and Graham 2002) because ML estimates are unbiased and more efficient than the estimates provided by other methods (e.g., listwise, pairwise and single imputation) under an MAR mechanism. Even under an MCAR mechanism, ML is still superior to traditional techniques because it maximizes statistical power by borrowing information from the observed data (Enders and Bandalos 2001). Despite these desirable properties, the ML method assumes that the missing data mechanism is MAR or MCAR. When the missing data mechanism is MNAR, whether the MAR-based ML approach can still get reliable results needs further research.

### 2.2. Diggle–Kenward selection model

The selection model (Glynn, Laird, and Rubin 1986; Little 1993, 1995) factorize the joint distribution of $Y$ and $d$ as follows:

$$f(Y_i, d_i) = f_{D|Y}(d_i|Y_i) f_Y(Y_i), \tag{1}$$

where $Y$ is the response variable and $d$ is the corresponding binary missing data indicator. The marginal distribution of $Y$, $f_Y(Y_i)$, corresponds to the substantive analysis, and the conditional distribution of $d$ given $Y$, $f_{D|Y}(d_i|Y_i)$, describes the probability of missing data through regression. In the selection model, response variables directly or indirectly predict the probability of missingness.

The Diggle–Kenward selection model (Diggle and Kenward 1994) combines a growth curve model (as the substantive model) with a set of logistic or probit regression models (as the conditional distribution) that directly relate the missing data probability to the outcome variable. More specifically, the missingness probability at time $t$ is predicted by the outcome variables at $t-1$ and $t$. The probit regression model used to describe the missingness probability is as follows:

$$\begin{aligned} p_i(y_1, y_2, \ldots y_{i-1}, y_i) &= p\left(d_i = 1|y_1, y_2, \ldots y_{i-1}, y_i\right) \\ &= \Phi(\beta_1 y_{i-1} + \beta_2 y_i - c), \end{aligned} \tag{2}$$

where $y_i$ is the outcome variable at time $i$, $d_i$ is the missing data indicator at time $i$,$p_i(y_1, y_2, \cdots, y_{i-1}, y_i)$ represents the conditional probability of missingness at time $i$, and $c$ is the threshold of indicator $d$. The indicator $d$ can be encoded in several ways, the most common of which is coding it as a discrete-time survival indicator. The value of the indicator is 0 in the periods before dropout occurs, 1 at the time when dropout occurs, and missing in the periods after the occurrence (Muthén and Masyn 2005). When neither of the regression coefficients, $\beta_1$ and $\beta_2$, is equal to 0, Equation (2) generates an MNAR mechanism. In addition, an MAR mechanism holds if $\beta_1 \neq 0$ and $\beta_2 = 0$, when the probability of dropout at time $t$ is a function of the observed $y$ value at the previous time point $t-1$, not the current one. However, it should be noted that $\beta_2 \neq 0$ cannot be seen as a hypothesis rejecting the MAR assumption due to different assumptions between Diggle–Kenward selection model and common MAR models besides the mechanism assumption (Muthén et al. 2011). To illustrate in details, Figure 1 shows a path diagram of a linear Diggle–Kenward selection model in the dotted line box.

### 2.3. Pattern mixture model

The pattern mixture model factorize the joint distribution of $Y$ and $d$ as follows:

$$f(Y_i, d_i) = f_{Y|D}(Y_i|d_i)f_D(d_i). \tag{3}$$

Different from the Diggle–Kenward selection model, here the marginal distribution of $d$, $f_D(d_i)$, describes the probability of missing data, and the conditional distribution of $Y$ given $d$, $f_{Y|D}(Y_i|d_i)$, describes the substantive model under a given missing data pattern, where the pattern is determined by the time point of missingness or the missing proportion. The parameter estimates are obtained under each pattern and then integrated to a final estimation by averaging. To obtain the estimates of some inestimable pattern-specific parameters, it is necessary to make some restrictions on these parameters (Enders 2010). However, incorrect restrictions might produce a substantive bias in estimation.

### 2.4. Model assumptions

The Diggle–Kenward selection model requires assumptions of multivariate normality for the repeated measured outcome variables to achieve identification. Without the distributional assumptions, the logistic regression part of the model is not estimable because $y$ is always missing whenever $d$ equals one. The pattern mixture model also assumes normal distributions for the outcomes, but the normality is conditional on missing patterns. Specifically, a pattern mixture analysis stratifies the sample into subgroups that share a same missing data pattern and a growth model is estimated separately within each pattern. It can be seen that the factorization of the joint distribution is different from the MAR-based model since the two MNAR-based models require a submodel for the missing data indicators.

## 3. Simulation study

To explore the factors that affect the precision of parameter estimation, a Monte Carlo simulation was applied to generate longitudinal datasets with different types of missingness.

### 3.1. Design and data generation

*First, generate a simulated complete longitudinal dataset,* which had five waves for $n$ individuals. The observed data on each wave was simulated depending on a latent growth curve model with parameters set as follows:

the intercept,i $\sim \mathrm{N}(-1,\ 0.50)$

the linear slope, s $\sim \mathrm{N}(0.5,\ 0.02)$ and

$$\text{the residuals,} \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \end{pmatrix} \sim \mathrm{N}\left( 0,\ \begin{matrix} 0.50 \\ 0.48 \\ 0.42 \\ 0.32 \\ 0.18 \end{matrix} \right).$$

*Then, generate dropout data,* considering different missingness mechanisms. The probability of dropout was a function of the observed $y$ values in the current and the previous time points under the Diggle–Kenward selection model assumption (MNAR) but only of the observed $y$ values in the last time point under the ML approach assumption (MAR). While models are built for generating MNAR missingness, selection models are generally adopted instead of mixture models mainly because researchers are more interested in considering the variables in the whole target population, rather than that in sub-populations defined by missing data patterns. (Fitzmaurice et al. 2008; Lu and Zhang 2014).

The model for generating missing data was defined as follows: Define a dummy variable $d_t$ to describe whether the target variable $y$ is observed at time point $t$: $d_t = 0$ means that $y$ is observed at time $t$, $d_t = 1$ represents $y$ is missing at time $t$ and all missing after time $t$. The missing mechanism defined by the Probit regression model can be expressed as a formula:

$$\begin{aligned} p_t(y_1, y_2, \ldots, y_{t-1}, y_t) &= p\left(d_k = 1 | y_1, y_2, \ldots, y_{t-1}, y_t\right) \\ &= \Phi(\beta_1 y_{t-1} + \beta_2 y_t - c), \end{aligned} \tag{4}$$

where $y_t$ represents the target variable at time $t$, $p_t(y_1, y_2, \ldots\infty, y_{t-1}, y_t)$ is the conditional probability of missingness at time $t$, $c$ is the threshold of the categorical variable $d$ and its specific value is calculated from the dropout rate. For the nonrandom missingness mechanism, we set $\beta_1 = -0.5$ and $\beta_2 = 1$. For the random missingness mechanism, we set $\beta_1 = -0.5$ and $\beta_2 = 0$ (Gad and Ahmed 2007; Mazumdar et al. 2007; Soullier, de La Rochebrochard, and Bouyer 2010).

In alignment with the previous studies (Gad and Ahmed 2007; Kristman, Manno, and Cote 2005; Langkamp, Lehman, and Lemeshow 2010; Marshall, Altman, and Holder 2010; Mazumdar et al. 2007; Newman 2003; Soullier, de La Rochebrochard, and Bouyer 2010; Yuan, Yang-Wallentin, and Bentler 2012), three most important factors were considered: (1) the missingness mechanism (MNAR and MAR); (2) sample size (100, 300, 500 and 1000); (3) the dropout rate (5%, 10%, 20% and 40%). Proportions of complete cases can range from 100% (no missing value) to less than 50% in practice (Daniels and Hogan 2008). According to Collins, Schafer, and Kam (2001), when the proportion of the cases with missing values is less than 10% and the correlation between two variables is greater than 0.4, the bias in parameter estimates is negligible even under an MNAR mechanism. And Kristman, Manno, and Cote (2005) found that the MNAR mechanism produced problematic estimates when the dropout rate was 25% or 40%. Therefore, we regard 5% as a small or trivial missingness proportion and 40% as a large proportion. In summary, a total of $2 \times 4 \times 4 = 32$ conditions were investigated, and for each condition, 500 replicated datasets were generated using R language.

### 3.2. Estimation precision

The performance of each method was evaluated according to three criteria, namely, the convergence rate, root mean square error and the 95% coverage probability.

*The convergence rate.* The convergence rate can be calculated by:

$$Convergence\ Rate = \frac{R}{R_{total}} \times 100\%, \tag{5}$$

where $R$ is the number of replications that successfully converged, and$R_{total}$ is the total number of replications (here, $R_{total}$=500).

*Root mean square error (RMSE).* RMSE describes the difference between a parameter estimate and its true value. A lower RMSE means a smaller difference between the estimated value and the true value. RMSE is calculated as follows:

$$RMSE(\hat{\theta}) = \sqrt{\frac{1}{R}\sum_{r=1}^{R}\left(\hat{\theta}_r - \theta\right)^2}, \tag{6}$$

where $\hat{\theta}_r$ is the parameter estimate in the $r$th replication, $\theta$ is the corresponding true value, and R is the number of replications that successfully converged.

*The coverage probability (CP) of 95% confidence interval.* This indicator reflects the parameter estimation precision and the corresponding estimated standard error to a certain extent. The 95% CP is calculated as follows:

$$CP = \frac{1}{R}\sum_{r=1}^{R} CI_{95}\left(\hat{\theta}_r\right), \tag{7}$$

where $CI_{95}(\hat{\theta}_r) = 1$ if the value $\hat{\theta}_r$ falls within the 95% confidence interval; otherwise, $CI_{95}(\hat{\theta}_r) = 0$.

### 3.3. Results

In longitudinal studies, the primary parameters of interest are the latent variable means ($\mu_i$ and $\mu_s$) and variances ($\sigma_i^2$ and $\sigma_s^2$). The $\mu_i$ and $\mu_s$ describe the development trend of a sample, and the $\sigma_i^2$ and $\sigma_s^2$ indicate the individual variations. The estimates of these four parameters under different conditions were obtained using Mplus 7 (Muthén and Muthén 2012). In addition, the detailed syntax is presented in Appendix.

#### 3.3.1. The convergence rate

Among all conditions, the average convergence rates of the MAR-based ML method and the MNAR-based DK method were 99.8% and 99.4% respectively, which indicated that the difference between two methods was really trivial and the convergence was not a problem for both methods. But it should be noted that the MNAR-based PM model had the problem of convergence as its convergence rates did not reach above 90% in most conditions and were even as low as 46.0% (when the sample size was 100 and the MAR dropout rate was 5%). The replications that failed to converge were removed prior to further analyses.

#### 3.3.2. Root mean square error (RMSE)

Figure 2 shows the comparison of RMSE values for $\mu_i$ estimation with three different methods at different dropout rates under MNAR and MAR missingness mechanisms. The precision of the parameter estimates increased as the sample size became larger, whichever method was applied. As can be seen from Figure 2A, $RMSE_{DK} < RMSE_{PM} < RMSE_{ML}$. The differences among three models became more significant as the dropout rate increased. In addition, there was an interaction between the MNAR dropout rate and the sample size, that is, increased sample size magnified the differences among the three approaches. Only when the sample size was small and the dropout rate was low, the PM method and the ML method were acceptable alternatives. However, Figure 2B indicates that the three methods performed very similarly and the MAR dropout rate made little difference in the RMSE results.

For the slope mean $\mu_s$, Figure 3A shows that $RMSE_{DK} < RMSE_{ML} < RMSE_{PM}$. When the MNAR dropout rate was less than 10%, the differences of $\mu_s$ estimation among the three methods were small. However, these differences became larger and the estimation precision became worse as the dropout rate increased. There was also an interaction between the dropout rate and the sample size and the DK method displayed more advantages with larger sample size. Figure 3B indicates that $RMSE_{ML} < RMSE_{DK} < RMSE_{PM}$, and the differences in RMSE values between ML method and DK method were small, especially for a large sample size or a low dropout rate.

Figures 4 and–5 exhibit the RMSE values for the estimates of $\sigma_i^2$ and $\sigma_s^2$ using these three methods under both MNAR and MAR missingness mechanisms. It could be observed that in the same condition, the DK method and the ML method performed similarly. For $\sigma_i^2$, the RMSE values of the PM method were smaller than those of the DK method and the ML method in the conditions with a small sample size. For $\sigma_s^2$, RMSE values of the PM model were larger than those of the other two methods with the increased dropout rate. A larger dropout rate and a lower sample size (e.g., $n = 100$) would lead to a worse estimation of $\sigma_i^2$ and $\sigma_s^2$.

#### 3.3.3. The 95% coverage probability (95% CP)

Table 1 contains the results of 95% CP of the four parameters using the three methods under different conditions.

Under the MNAR missingness mechanism, the 95% CP for $\mu_i$ using the DK method, the PM method and the ML method were 92.67%, 86.01% and 79.49% respectively. The results obtained by the DK method were better than those of the PM method, while the results of the ML method were the worst. The 95% CP for $\mu_i$ of the DK method changed little when the dropout rate increased, but the 95% CP of the PM method and

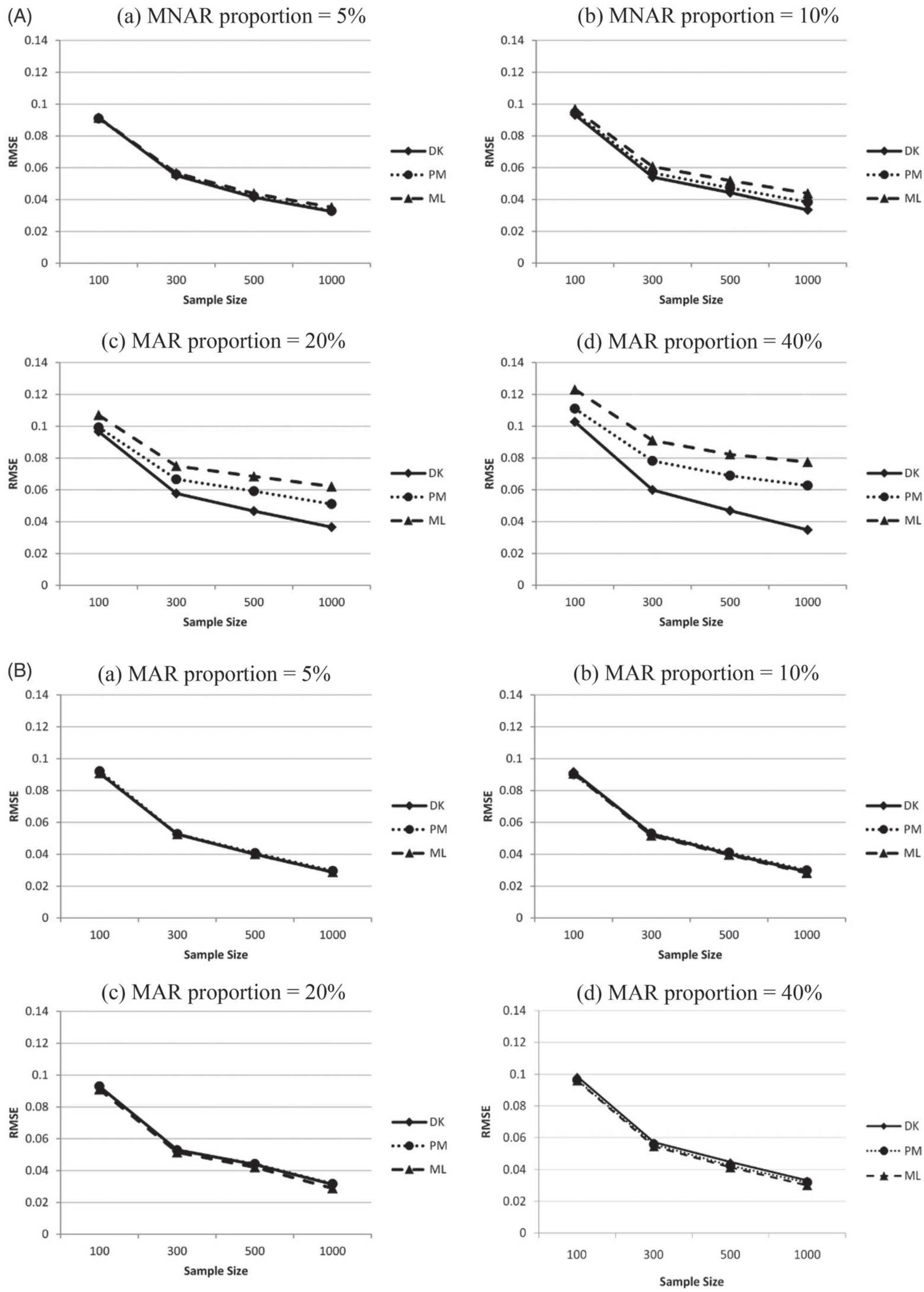



**Figure 2.** RMSEs for the estimates of $\mu_i$ by different methods under different missingness mechanisms.

the ML method decreased substantially. However, under the MAR missingness mechanism, the 95% CP for $\mu_i$ using the three methods were 94.39%, 94.33% and 94.83% respectively, implying that the results had little difference across different methods.

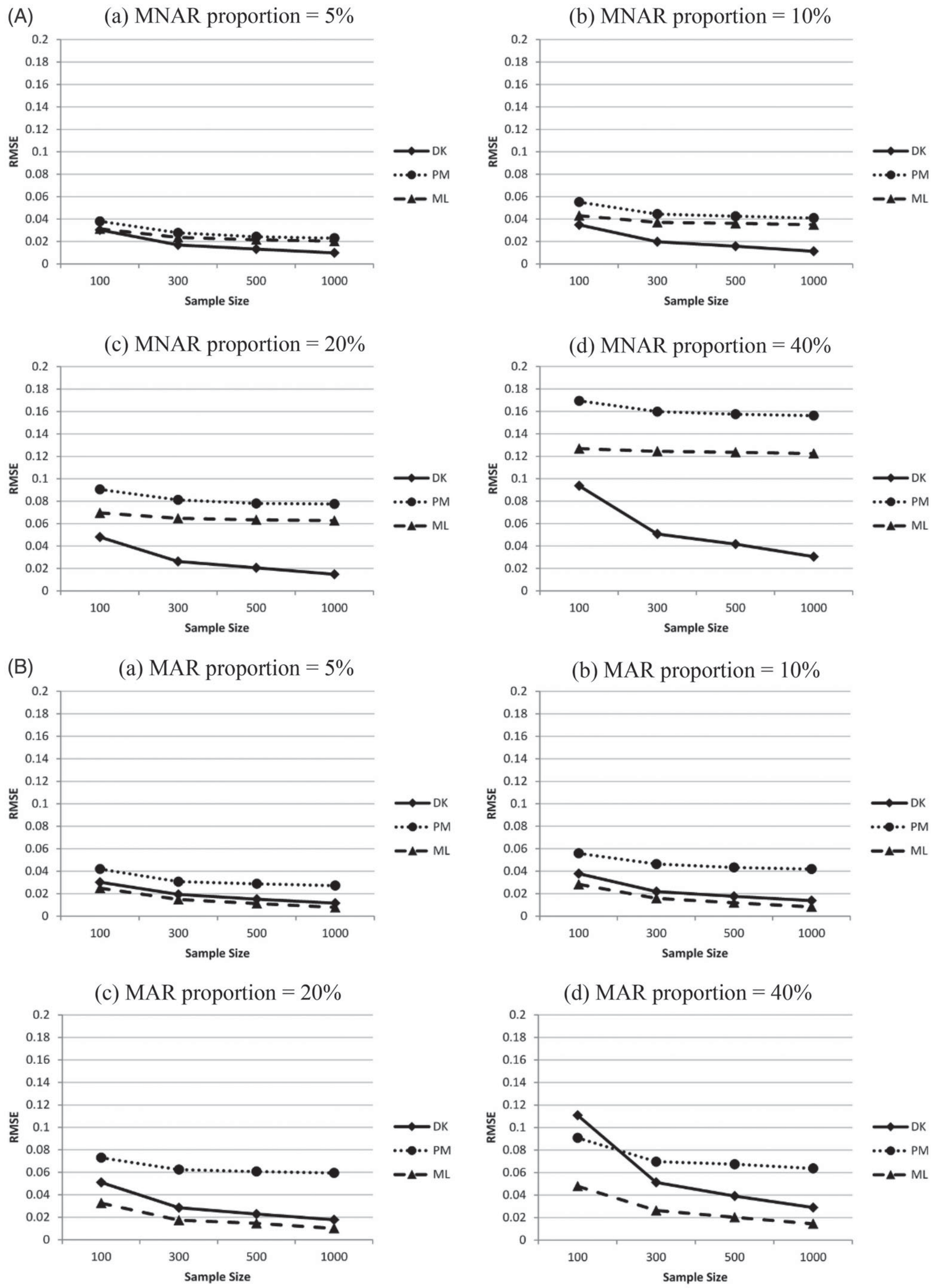



**Figure 3.** RMSEs for the estimates of $\mu_s$ by different methods under different missingness mechanisms.

Under the MNAR missingness mechanism, the 95% CP for $\mu_s$ of the DK method, the PM method and the ML method were 89.47%, 35.27% and 30.15% respectively. The DK method outperformed the PM method and the ML method substantially.

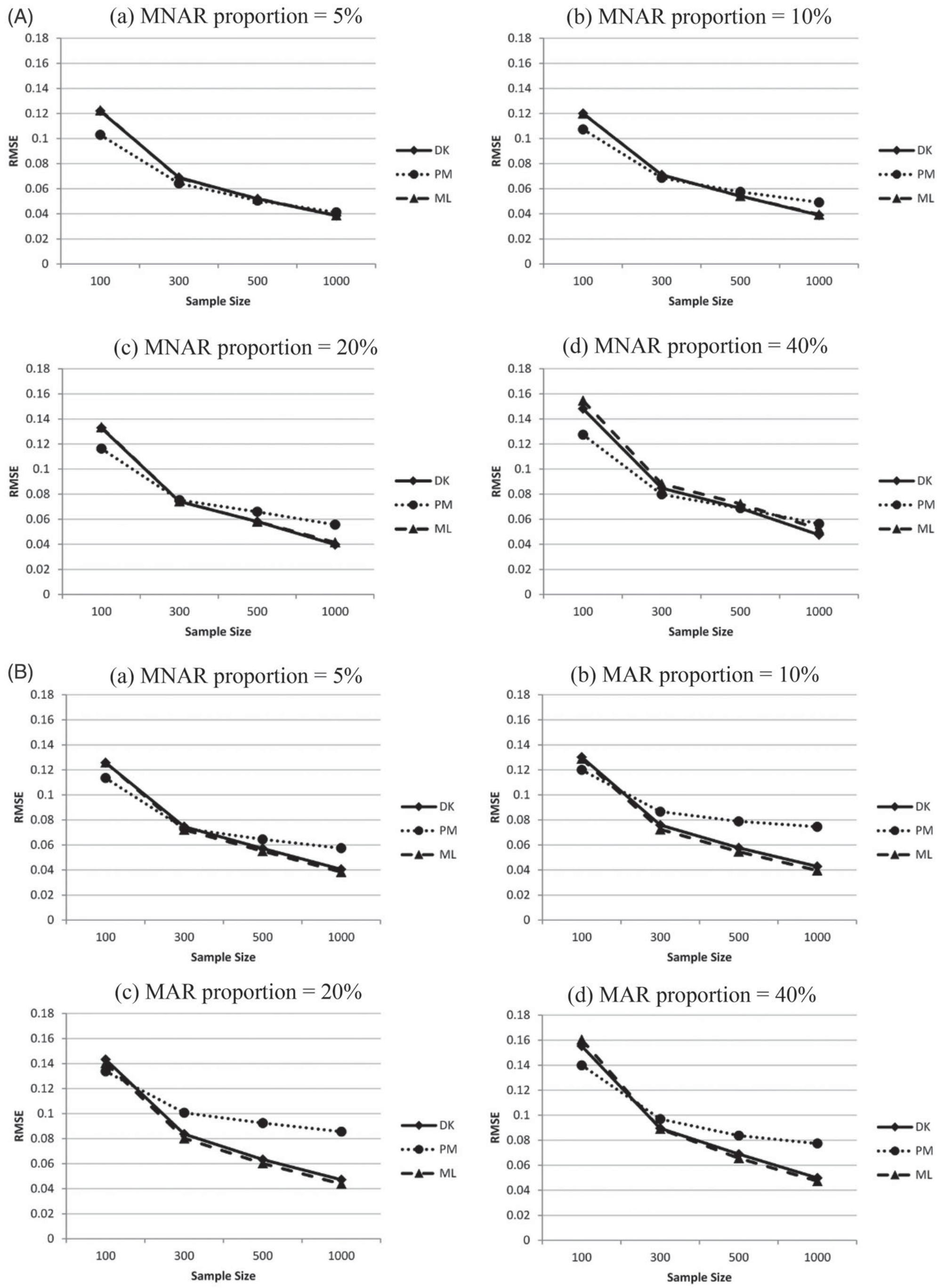



**Figure 4.** RMSEs for the estimates of $\sigma_i^2$ by different methods under different missingness mechanisms.

While under the MAR missingness mechanism, the 95% CP for $\mu_s$ of the DK method, the PM method and the ML method were 87.59%, 44.15% and 94.68% respectively.

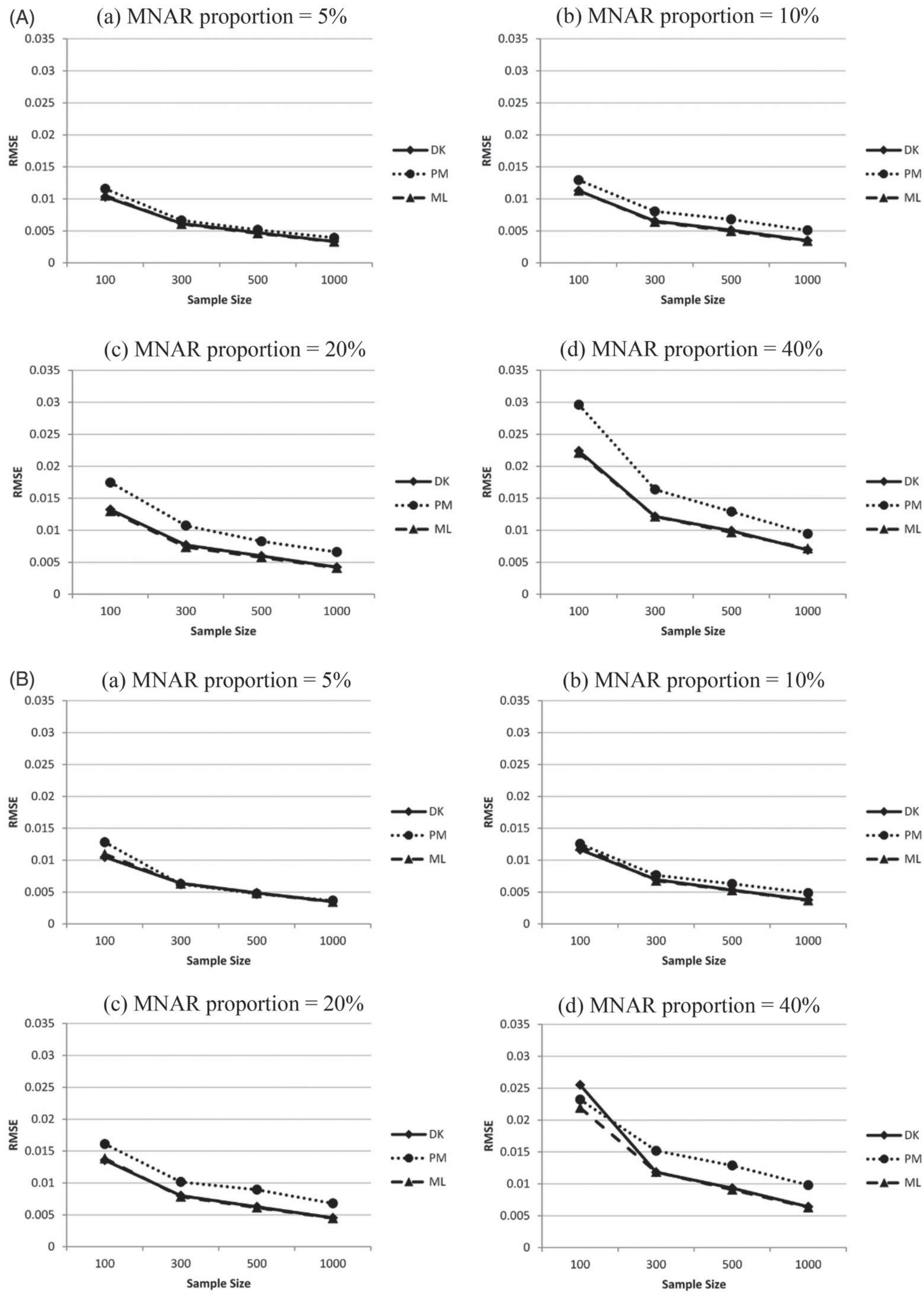



**Figure 5.** RMSEs for the estimates of $\sigma_s^2$ by different methods under different missingness mechanisms.

The 95% CP for $\sigma_i^2$ using the three methods were 94.37%, 80.41% and 94.60% under the MNAR missingness mechanism, while they were 93.17%, 63.28% and 94.47% under the MAR missingness mechanism. In most cases, the probabilities of the DK method

**Table 1.** 95% Coverage probabilities for parameter estimates under MNAR and MAR missingness mechanisms.

| | | Dropout percentage of MNAR (%) | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 5 | | | 10 | | | 20 | | | 40 | | |
| | | DK | PM | ML | DK | PM | ML | DK | PM | ML | DK | PM | ML |
| $\mu_i$ | 100 | 94.33 | 94.30 | 93.67 | 94.80 | 94.10 | 92.87 | 93.27 | 93.00 | 90.67 | 93.67 | 90.80 | 86.53 |
| | 300 | 93.07 | 93.33 | 92.47 | 94.00 | 93.27 | 90.73 | 93.73 | 89.00 | 82.40 | 93.13 | 83.17 | 73.53 |
| | 500 | 93.87 | 93.83 | 92.73 | 92.27 | 90.73 | 87.33 | 92.47 | 82.93 | 72.80 | 93.80 | 75.17 | 61.00 |
| | 1000 | 91.20 | 90.80 | 88.47 | 90.07 | 85.57 | 79.00 | 88.27 | 69.73 | 52.73 | 90.73 | 56.37 | 34.93 |
| $\mu_s$ | 100 | 93.07 | 90.20 | 87.07 | 92.40 | 81.30 | 74.27 | 89.73 | 59.53 | 46.60 | 83.50 | 33.17 | 25.33 |
| | 300 | 93.87 | 76.30 | 73.20 | 92.73 | 51.87 | 41.33 | 90.73 | 15.90 | 8.47 | 85.27 | 3.10 | 1.67 |
| | 500 | 93.27 | 67.63 | 61.13 | 91.80 | 29.33 | 19.67 | 90.87 | 6.37 | 2.53 | 83.40 | 0.33 | 0.20 |
| | 1000 | 91.67 | 42.17 | 37.33 | 90.73 | 6.97 | 3.47 | 90.20 | 0.13 | 0.07 | 78.33 | 0.00 | 0.00 |
| $\sigma_i^2$ | 100 | 92.87 | 86.43 | 93.40 | 94.73 | 85.00 | 95.33 | 92.87 | 83.67 | 93.87 | 92.67 | 85.00 | 94.07 |
| | 300 | 94.47 | 86.63 | 94.33 | 94.27 | 82.97 | 94.80 | 95.27 | 81.13 | 95.60 | 93.87 | 83.50 | 93.80 |
| | 500 | 95.47 | 86.50 | 95.40 | 95.53 | 79.43 | 95.87 | 94.73 | 74.27 | 95.47 | 93.13 | 81.63 | 93.60 |
| | 1000 | 94.87 | 82.37 | 95.07 | 95.13 | 70.77 | 94.93 | 95.60 | 64.73 | 95.67 | 94.47 | 72.50 | 92.40 |
| $\sigma_s^2$ | 100 | 91.40 | 93.13 | 93.60 | 91.93 | 92.67 | 94.47 | 89.00 | 91.60 | 94.07 | 82.03 | 94.50 | 93.13 |
| | 300 | 95.00 | 90.17 | 94.80 | 94.33 | 89.03 | 94.47 | 93.67 | 87.73 | 95.20 | 92.13 | 90.80 | 93.80 |
| | 500 | 94.87 | 89.03 | 95.27 | 94.20 | 85.33 | 94.87 | 95.13 | 88.67 | 95.67 | 93.33 | 90.13 | 94.87 |
| | 1000 | 95.27 | 88.60 | 94.87 | 95.33 | 84.73 | 95.53 | 95.00 | 84.33 | 94.93 | 92.93 | 90.00 | 93.60 |
| | | Dropout percentage of MAR (%) | | | | | | | | | | | |
| | | 5 | | | 10 | | | 20 | | | 40 | | |
| | | DK | PM | ML | DK | PM | ML | DK | PM | ML | DK | PM | ML |
| $\mu_i$ | 100 | 94.47 | 94.10 | 94.67 | 94.73 | 95.50 | 94.80 | 94.07 | 93.90 | 94.13 | 94.07 | 94.70 | 94.47 |
| | 300 | 94.40 | 94.00 | 94.80 | 95.53 | 95.10 | 95.73 | 94.93 | 94.90 | 95.40 | 93.87 | 94.23 | 94.47 |
| | 500 | 95.13 | 94.17 | 94.67 | 95.47 | 94.83 | 95.93 | 93.67 | 93.83 | 93.73 | 94.20 | 95.00 | 95.67 |
| | 1000 | 94.20 | 93.90 | 94.27 | 94.87 | 94.33 | 95.07 | 93.67 | 93.53 | 94.60 | 93.00 | 93.27 | 94.93 |
| $\mu_s$ | 100 | 93.27 | 88.80 | 94.80 | 91.33 | 78.13 | 94.20 | 90.67 | 75.40 | 93.27 | 82.07 | 79.10 | 93.60 |
| | 300 | 90.47 | 71.17 | 94.47 | 91.00 | 46.73 | 94.67 | 90.27 | 38.00 | 95.53 | 84.47 | 56.53 | 93.67 |
| | 500 | 90.27 | 52.63 | 95.13 | 89.87 | 28.20 | 95.87 | 87.80 | 15.67 | 94.47 | 83.73 | 35.17 | 95.00 |
| | 1000 | 87.00 | 24.43 | 95.33 | 85.27 | 5.30 | 95.53 | 84.07 | 1.73 | 94.67 | 79.87 | 9.40 | 94.67 |
| $\sigma_i^2$ | 100 | 92.60 | 81.20 | 93.73 | 92.40 | 78.27 | 93.67 | 90.80 | 74.90 | 93.27 | 93.33 | 80.90 | 93.33 |
| | 300 | 92.73 | 78.33 | 93.93 | 93.60 | 68.90 | 95.27 | 92.67 | 61.13 | 93.40 | 93.27 | 74.77 | 94.47 |
| | 500 | 93.13 | 72.67 | 94.67 | 95.47 | 56.90 | 96.00 | 93.13 | 49.53 | 94.93 | 94.73 | 68.77 | 95.60 |
| | 1000 | 93.47 | 55.37 | 94.67 | 93.53 | 35.47 | 94.87 | 91.93 | 26.60 | 94.07 | 93.87 | 48.73 | 95.67 |
| $\sigma_s^2$ | 100 | 91.33 | 95.10 | 94.07 | 88.13 | 91.60 | 92.93 | 87.27 | 92.03 | 94.13 | 80.17 | 94.23 | 91.67 |
| | 300 | 94.13 | 93.07 | 93.93 | 94.07 | 87.87 | 95.07 | 95.07 | 86.47 | 95.67 | 90.53 | 88.13 | 93.67 |
| | 500 | 95.20 | 90.87 | 95.40 | 94.07 | 86.83 | 94.07 | 94.33 | 83.27 | 94.67 | 92.60 | 85.40 | 93.87 |
| | 1000 | 94.27 | 89.57 | 94.60 | 94.60 | 84.17 | 94.93 | 93.53 | 82.27 | 94.40 | 93.60 | 83.93 | 94.40 |

*Note.* ML = MAR-based ML approach; DK = Diggle–Kenward Selection Model; PM = Pattern Mixture Model.

and the ML method reached 90%. The results of the DK method were affected little by the dropout rate, while the estimation of the pattern mixture model were largely undermined by the high level of dropout rate.

The differences of three methods under different conditions were relatively small for the parameter estimation for $\sigma_i^2$. Also, for all four parameters, a larger sample size would lead to lower 95% CP.

## 4. An illustrative example

In this section, we used the publicly available data in the Early Childhood Longitudinal Study, Kindergarten Class of 1998–1999 (ECLS-K) (Tourangeau et al. 2009) to illustrate how to address MNAR missingness in practice. The sampled children were followed from kindergarten to 8th grade, which was sponsored by National Center for Education

Statistics (NCES). To be specific, information was collected in the fall and the spring of kindergarten (1998–1999), the fall and spring of 1st grade (1999–2000), the spring of 3rd grade (2002), the spring of 5th grade (2004), and the spring of 8th grade (2007). Because only 30% of the subjects participated in the investigation in the fall of 1999, the data at this wave were discarded and only the remaining data at the other six waves (i.e., wave 1, wave 2, wave 4, wave 5, wave 6 and wave 7) were analyzed. The development trend of reading IRT scores was focused in the current analyses.

First, examine the missing data magnitude and the sample size of the dataset. The ECLS-K had a sample of 21,409 children, of which 545 had no records in the reading IRT scores at all waves. Thus, the remaining 20,864 subjects were analyzed in the current study. Of the 20,864 subjects, 11,639 (55.8%) had a dropout missing data pattern, 2,124 (10.2%) had non-dropout intermittent missingness, and only 7,101 (34.0%) had complete data for all six waves. Figure 6a shows the reading score means of the entire sample, non-dropouts and dropouts at each wave respectively.

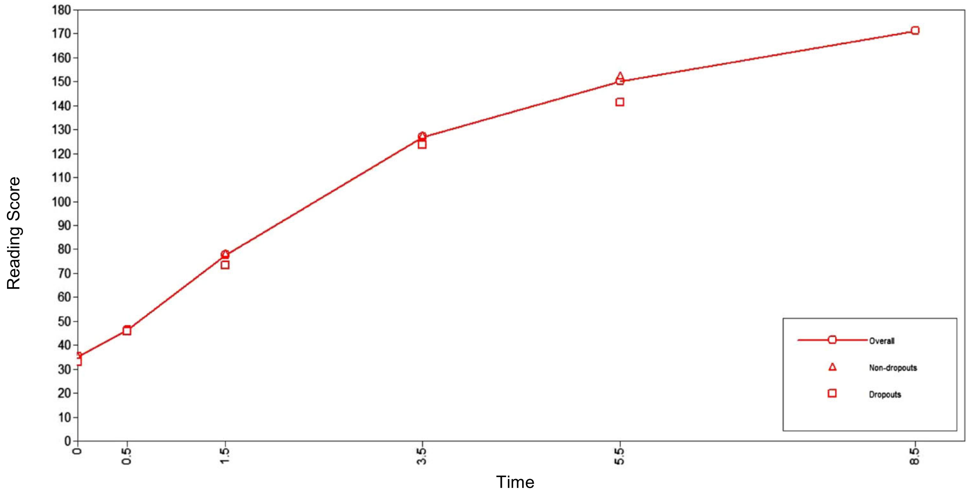
Second, explore the missing data mechanism. Little's MCAR test inferred that the missing data of this dataset did not satisfy the MCAR mechanism, as $\chi^2 = 1780.554$, $df = 182$, $p < 0.001$. Considering that (i) the MNAR missingness is common in longitudinal studies, (ii) there is a high dropout rate in this dataset, and (iii) the reading scores of the subjects who dropped out are different from that of the non-dropout ones as shown in Figure 6a, we could regard the dropout missing data mechanism as MNAR and the intermittent missing data mechanism as MAR.

Finally, a sensitivity analysis with several models should be conducted. For this purpose, the model-based ML approach, the DK method and the PM method were implemented using the Mplus software with a latent growth model. In the DK method, we distinguished the dropouts in MNAR mechanism and the intermittent missingness in MAR mechanism through assigning different values to the discrete-time survival indicators mentioned before. In the PM method, we just defined the missing pattern as dropout and intermittent missingness and there was no need to specify restrictions.

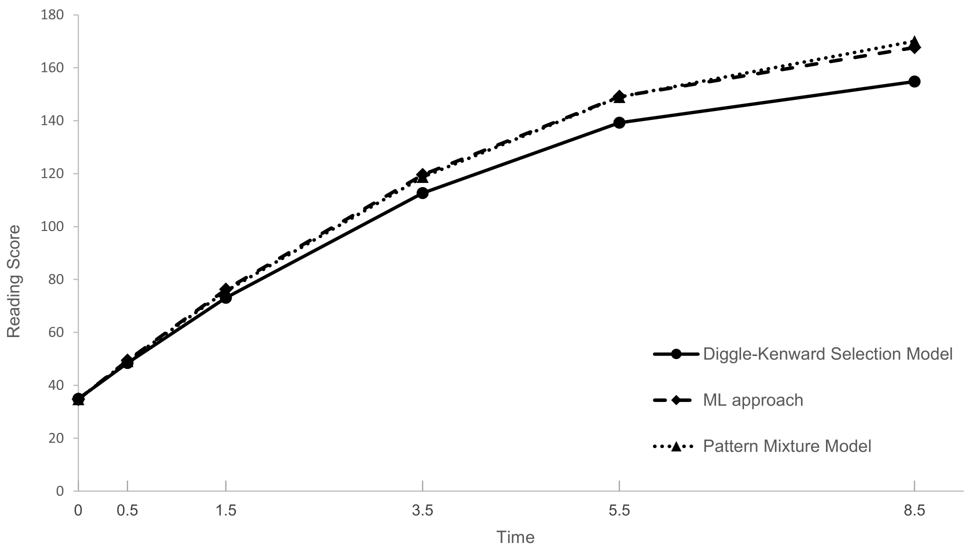
The estimated means and variances of latent variables are shown in Table 2.

For the estimates of $\mu_i$ and its *SE*, the three methods differed little from each other. For the estimates of $\mu_s$ and $\mu_q$, the PM method and the ML method had similar results that were both larger than those of the DK method. For the estimates of ${\sigma_i}^2$, ${\sigma_s}^2$, and ${\sigma_q}^2$ and their *SE*s, the results of the PM method and the ML approach were almost the same, which were lower than those of the DK method.

Figure 6b illustrates that the estimated means of DK method at each wave were lower than those of the PM method and the ML method. This discrepancy became more obvious when the MNAR missingness magnitude increased over time. As stated before, the subjects who had lower reading scores or weaker reading abilities were more likely to drop out from the experiment, so the dropout pattern satisfied the MNAR mechanism. Furthermore, considering that the reading scores were most likely to be normally distributed, the DK method had the most reasonable assumptions for this dataset, and the analysis results coincided with this. Above all, we concluded that the results of the DK method under MNAR mechanism were the most acceptable.




*(a) Sample Means of Overall, Non-Dropouts and Dropouts*




*(b) Estimated means of Different Models*

**Figure 6.** Sample means and estimated means at each wave.

## 5. Discussion

The purpose of this study was to evaluate different model-based methods for dealing with missing data under different conditions within the framework of latent growth curve models. The findings would provide insights into the strategies in handling MNAR missingness within longitudinal datasets.

Under the MAR mechanism, there is little difference between the results of the MAR-based ML method and the MNAR-based DK method for the parameters of interest in the latent growth model, especially in the case of a large sample size. The results for the 95% CP also suggest that there is no evident difference between the two methods. This indicates that even when the data are MAR, the use of the MNAR-based DK method will not

**Table 2.** Results for parameter estimates and standard errors of LGM.

| | Diggle–Kenward Selection Model | | ML approach | | Pattern Mixture Model | |
|---|---|---|---|---|---|---|
| Parameter | Estimate | *SE* | Estimate | *SE* | Estimate | *SE* |
| $\mu_i$ | 34.871 | 0.076 | 34.720 | 0.078 | 34.737 | 0.082 |
| $\mu_s$ | 27.909 | 0.089 | 30.266 | 0.077 | 29.651 | 0.100 |
| $\mu_q$ | −1.624 | 0.009 | −1.721 | 0.008 | −1.615 | 0.017 |
| $\sigma_i^2$ | 108.077 | 1.219 | 106.782 | 1.218 | 106.183 | 3.357 |
| $\sigma_s^2$ | 80.504 | 1.368 | 66.166 | 1.097 | 63.702 | 1.128 |
| $\sigma_q^2$ | 0.637 | 0.018 | 0.557 | 0.015 | 0.464 | 0.022 |

produce extremely biased estimates. But the results of the PM method were worse than those of the DK method, because the way of generating missingness data is beneficial to the Diggle–Kenward selection model.

Under the MNAR mechanism that the DK model is based on, there is little difference between the results of ML method and DK method for the variances of the intercept and slope latent variables. The results for the 95% CP also show that there is no significant difference between the two methods. However, the ML method would produce highly biased estimates for the means of the intercept and slope even under a large sample size. It should be noted that when the dropout rate is small (such as 5%), there is little difference between the ML method and the DK method. The results state that ignoring the MNAR mechanism in data analyses will lead to large biases in estimation, especially for a high dropout rate. The PM method still gets the worst results because of the violation of assumptions of the missingness mechanism. The differences between the ML method and the PM method indicate that a better MAR model outperforms a worse MNAR model even if the MAR assumption is violated. This supports the viewpoint of Schafer (2003).

It is a complex issue to choose an appropriate model for handling missing data in longitudinal datasets. If MNAR missing data may exist in the longitudinal analysis, applying a correct approach is necessary. Only when there is a sufficiently low percentage of MNAR missing data can other methods with inappropriate assumptions (such as the ML approach and the PM model) be acceptable alternatives. If the dropout rate of MNAR is greater than 20%, the model should be chosen with some caution. However, in practice, since it is often difficult to accurately verify the missing mechanism, one should maximize the chance for the MAR mechanism. Firstly, sensitivity analyses should be used to assess the robustness of inferences to departures from MAR assumptions. Because each specific scientific context may require different considerations when researchers carry out such analyses, no standard method exists, nor should it be prescribed as this is still an active area of research. Researchers perform sensitivity analyses through a methodology relying on the MNAR pattern-mixture model (Moreno-Betancur and Chavance 2013), post-modeling based on bootstrap type method (Jamshidian and Mata 2008), post-modeling based on an asymptotic approximation (Jamshidian and Yuan 2013) and so on. Secondly, Graham (2003) proposed that adding variables that were correlated with variables containing missingness, regardless of whether they were related to missingness, could substantially improve estimation (precision and efficiency). And Yuan and Lu (2008) further found that when missing data mechanisms were unknown, including auxiliary variables related to the target variables in the analysis would make the missing data mechanism more likely to be MAR. Lastly, based on a series of model fit indices, such as AIC, BIC, CAIC, SSBIC and DIC (Lu, Zhang, and

Cohen 2013; Lu and Zhang 2014), the best-fit model can be selected. Moreover, regarding estimation methods, more methods should be considered, such as 2-stage ML and Bayesian approach. When the population distribution is non-normal or unknown, the 2-stage ML would be a valid statistical procedure (Yuan and Lu 2008). The Bayesian method is also flexible enough to estimate a variety of models with different missing data mechanisms, contaminated data, and mixture structures (Dunson 2000; Lu and Zhang 2014).

Although the findings of this study supplement the body of literature in the field of MNAR dropout missingness with longitudinal data, there are some limitations in the simulation study design. The first limitation is the latent growth model utilized, in which latent class variables, quadratic latent variables and covariates were not included. Second, only three types of methods under different assumptions were compared in this study while there are more approaches within these three families of models. Finally, the results are beneficial to DK model as DK model is used to generate dropout data, instead of PM model. Today, although most of researches generated MNAR missingness mechanism based on the selection model, we recommend continuing to explore the data generation mechanism of MNAR.

## Funding

Supported by National Natural Science Foundation of China (31571152); Special Found for Beijing Common Construction Project (019-105812); National Education Examination Research Program (GJK2017015).

## Appendix

**Monte Carlo simulation syntax of ML**

```
TITLE:
      ML-mnar-mar-n-40201000
DATA:
      FILE = mnarmar40201000replist.dat;
      TYPE = MONTECARLO;
VARIABLE:
      NAMES = m1-m5 y1-y5;
      USEVARIABLES = y1-y5;
      MISSING = ALL(999);

ANALYSIS:
      ESTIMATOR = ML;

MODEL:
      i s|y1@0 y2@1 y3@2 y4@3 y5@4;
      i WITH s*0;
      i*0.5 s*.02;
      [i*-1 s*0.5];

      y1*.5 y2*.48 y3*.42 y4*.32 y5*0.18;
```

```
SAVEDATA:
    RESULTS are ml40201000results.dat;
```

**Monte Carlo simulation syntax of DK selection model**

```
TITLE:
    DK-mnar-mar-n-40201000
DATA:
    FILE = mnarmar40201000replist.dat;
    TYPE = MONTECARLO;
VARIABLE:
    NAMES = m1-m5 y1-y5;
    USEVARIABLES = y1-y5 d2-d5;
    MISSING = ALL(999);
    CATEGORICAL = d2-d5;
DATA MISSING:
    NAMES = y1-y5;
    TYPE = SDROPOUT;
    BINARY = d2-d5;
ANALYSIS:
    ESTIMATOR = ML;
    LINK = PROBIT;
    ALGORITHM = INTEGRATION;
    INTEGRATION = MONTECARLO;
    PROCESSORS =2;

MODEL:
    i s|y1@0 y2@1 y3@2 y4@3 y5@4;
    i WITH s*0;
    i*0.5 s*.02;
    [i*-1 s*0.5];

    y1*.5 y2*.48 y3*.42 y4*.32 y5*0.18;

    [d2$1*1.1134];
    [d3$1*1.3506];
    [d4$1*1.5745];
    [d5$1*1.7842];

    d2 on y1*-0.5 (11)
    y2*1 (12);
    d3 on y2*-0.5 (11)
    y3*1 (12);
    d4 on y3*-0.5 (11)
    y4*1 (12);
```

```
      d5 on y4*-0.5 (11)
      y5*1 (12);

SAVEDATA:
      RESULTS are dk40201000results.dat;
```

**Monte Carlo simulation syntax of PM model**

```
TITLE:
      PM-mnar-mar-n-40201000
DATA:
      FILE = mnarmar40201000replist.dat;
      TYPE = MONTECARLO;
VARIABLE:
      NAMES = m1-m5 y1-y5;
      USEVARIABLES = y1-y5 group;
      MISSING = ALL(999);

      classes = patn(2);
      knownclass = patn(group =1 group =2);

DEFINE:
      group =1;
      if(m5 eq 2)then group =2;
      if(m5 eq 2 and m4 eq 2)then group =2;
      if(m5 eq 2 and m4 eq 2 and m3 eq 2)then group =2;
      if(m5 eq 2 and m4 eq 2 and m3 eq 2 and m2 eq 2)then group =2;

ANALYSIS:
      type = mixture;

MODEL:
      %overall%

      i s|y1@0 y2@1 y3@2 y4@3 y5@4;
      i WITH s @ 0;

      i*0.5 s*.02;
      [i*-1 s*0.5];

      y1*.5 y2*.48 y3*.42 y4*.32 y5*0.18;

   [patn#1] (p1logit);
%patn#1%
[i](mi1);
[s](ms1);
```

```
i(vi1);
s(vs1);

%patn#2%
[i](mi2);
[s](ms2);
i(vi2);
s(vs2);

MODEL CONSTRAINT:
    new(pi1 pi2 mi*-1 ms*0.5 vi*0.5 vs*0.02);

    pi1 = exp(p1logit)/(exp(p1logit) +
      exp(0));

    pi2 = exp(0)/(exp(p1logit) +
      exp(0));

    mi = mi1*pi1 + mi2*pi2;
    ms = ms1*pi1 + ms2*pi2;
    vi = vi1*pi1 + vi2*pi2;
    vs = vs1*pi1 + vs2*pi2;

OUTPUT:
    tech4;

SAVEDATA:
    RESULTS are pm40201000results.dat;
```

## References


Acock, A. C. 2005. Working with missing values. *Journal of Marriage and Family* 67 (4):1012–1028. doi:10.1111/j.1741-3737.2005.00191.x

Beunckens, C., G. Molenberghs, G. Verbeke, and C. Mallinckrodt. 2008. A latent-class mixture model for incomplete longitudinal Gaussian data. *Biometrics* 64 (1):96–105. doi:10.1111/j.1541-0420.2007.00837.x

Bollen, K. A., and P. J. Curran. 2006. *Latent curve models: a structural equation perspective. Wiley series in probability and mathematical statistics*. New York: Wiley.

Carpenter, J. R., M. G. Kenward, and S. Vansteelandt. 2006. A comparison of multiple imputation and doubly robust estimation for analyses with missing data. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 169 (3):571–584. doi:10.1111/j.1467-985X.2006.00407.x

Collins, L. M., J. L. Schafer, and C. M. Kam. 2001. A comparison of inclusive and restrictive strategies in modern missing data procedures. *Psychological Methods* 6 (4):330–351. doi:10.1037/1082-989X.6.4.330

Daniels, M. J., and J. W. Hogan. 2008. *Missing data in longitudinal studies*. Boca Raton, FL: Chapman & Hall/CRC.

Diggle, P., and M. G. Kenward. 1994. Informative drop-out in longitudinal data-analysis. *Applied Statistics-Journal of the Royal Statistical Society Series C* 43(1):49–93. doi:10.1037/1082-989X.6.4.330

Dunson, D. B. 2000. Bayesian latent variable models for clustered mixed outcomes. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 62 (2):355–366.

Enders, C. K. 2010. *Applied missing data analysis*. New York: The Guilford Press.

Enders, C. K. 2011. Missing not at random models for latent growth curve analyses. *Psychological Methods* 16 (1):1–16. doi:10.1037/a0022640

Enders, C. K., and D. L. Bandalos. 2001. The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal* 8 (3):430–457. doi:10.1207/S15328007SEM0803_5

Fitzmaurice, G., M. Davidian, G. Verbeke, and G. Molenberghs. 2008. *Longitudinal data analysis*. Boca Raton, FL: Chapman & Hall/CRC Press.

Gad, A. M., and A. S. Ahmed. 2007. Sensitivity analysis of longitudinal data with intermittent missing values. *Statistical Methodology* 4 (2):217–226. doi:10.1016/j.stamet.2006.08.001

Glynn, R. J., N. M. Laird, and D. B. Rubin. 1986. Selection modeling versus mixture modeling with nonignorable nonresponse. In *Drawing inferences from self-selected samples*, H. Wainer, 115–142. New York: Springer.

Graham, J. W. 2003. Adding missing-data-relevant variables to FIML-based structural equation models. *Structural Equation Modeling* 10 (1):80–100.

Jamshidian, M., and M. Mata. 2008. Postmodeling sensitivity analysis to detect the effect of missing data mechanisms. *Multivariate Behavioral Research* 43 (3):432–452.

Jamshidian, M., and K. H. Yuan. 2013. Data-driven sensitivity analysis to detect missing data mechanism with applications to structural equation modelling. *Journal of Statistical Computation and Simulation* 83 (7):1344–1362. doi:10.1080/00949655.2012.660486

Jöreskog, K. G., and D. Sörbom. 1996. *LISREL 8: user's reference guide*. Chicago, IL: Scientific Software International.

Kristman, V. L., M. Manno, and P. Cote. 2005. Methods to account for attrition in longitudinal data: do they work? a simulation study. *European Journal of Epidemiology* 20 (8):657–662. doi:10.1007/s10654-005-7919-7

Langkamp, D. L., A. Lehman, and S. Lemeshow. 2010. Techniques for handling missing data in secondary analyses of large surveys. *Academic Pediatrics* 10 (3):205–210. doi:10.1016/j.acap.2010.01.005

Little, R. J. A. 1993. Pattern-mixture models for multivariate incomplete data. *Journal of the American Statistical Association* 88 (421):125–134. doi:10.2307/2290705

Little, R. J. A. 1995. Modeling the drop-out mechanism in repeated-measures studies. *Journal of the American Statistical Association* 90 (431):1112–1121. doi:10.2307/2291350

Little, R. J. A., and D. B. Rubin. 2002. *Statistical analysis with missing data*. 2nd ed. Hoboken, NJ: Wiley.

Lu, Z., and Z. Zhang. 2014. Robust growth mixture models with non-ignorable missingness: Models, estimation, selection, and application. *Computational Statistics and Data Analysis* 71:220–240. doi:10.1016/j.csda.2013.07.036

Lu, Z., Z. Zhang, and A. Cohen. 2013. Bayesian methods and model selection for latent growth curve models with missing data. In *New developments in quantitative psychology*, eds. Millsap., R. E., L. A. van der Ark, D. M. Bolt, & C. M. Woods, 275–304. New York: Springer.

Marshall, A., D. G. Altman, and R. L. Holder. 2010. Comparison of imputation methods for handling missing covariate data when fitting a cox proportional hazards model: a resampling study. *BMC Medical Research Methodology* 10 (1):1–10. doi:10.1186/1471-2288-10-112

Mazumdar, S., G. Tang, P. R. Houck, M. A. Dew, A. E. Begley, J. Scott, B. H. Mulsant, and C. F. Reynolds. 2007. Statistical analysis of longitudinal psychiatric data with dropouts. *Journal of Psychiatric Research* 41 (12):1032–1041. doi:10.1016/j.jpsychires.2006.09.007

McArdle, J. J., and D. Epstein. 1987. Latent growth curves within developmental structural equation models. *Child Development* 58 (1):110–133. doi:10.2307/1130295

Meredith, W., and J. Tisak. 1990. Latent curve analysis. *Psychometrika* 55 (1):107–122. doi:10.1007/BF02294746

Moreno-Betancur, M., and M. Chavance. 2016. Sensitivity analysis of incomplete longitudinal data departing from the missing at random assumption: Methodology and application in a clinical trial with drop-outs. *Statistical Methods in Medical Research* 25 (4):1471–1489. doi:10.1177/0962280213490014

Muthén, B., T. Asparouhov, A. M. Hunter, and A. F. Leuchter. 2011. Growth modeling with non-ignorable dropout: alternative analyses of the STAR*D antidepressant trial. *Psychological Methods* 16 (1):17–33. doi:10.1037/a0022634

Muthén, B., and K. Masyn. 2005. Discrete-time survival mixture analysis. *Journal of Educational and Behavioral Statistics* 30 (1):27–58.

Muthén, L. K., and B. O. Muthén. 1998. *Mplus user's guide*. 7th ed. Los Angeles, CA: Muthén & Muthén.

Newman, D. A. 2003. Longitudinal modeling with randomly and systematically missing data: a simulation of ad hoc, maximum likelihood, and multiple imputation techniques. *Organizational Research Methods* 6 (3):328–362.

Power, R. A., B. Muthén, N. Henigsberg, O. Mors, A. Placentino, J. Mendlewicz, W. Maier, P. McGuffin, C. M. Lewis, and R. Uher. 2012. Non-random dropout and the relative efficacy of escitalopram and nortriptyline in treating major depressive disorder. *Journal of Psychiatric Research* 46 (10):1333–13338. doi:10.1016/j.jpsychires.2012.06.014

Roy, J. 2003. Modeling longitudinal data with nonignorable dropouts using a latent dropout class model. *Biometrics* 59 (4):829–836. doi:10.1111/j.0006-341X.2003.00097.x

Schafer, J. L. 2003. Multiple imputation in multivariate problems when the imputation and analysis models differ. *Statistica Neerlandica* 57 (1):19–35. doi:10.1111/1467-9574.00218

Schafer, J. L., and J. W. Graham. 2002. Missing data: our view of the state of the art. *Psychological Methods* 7 (2):147–177. doi:10.1037//1082-989X.7.2.147

Singer, J. D., and J. B. Willett. 2003. *Applied longitudinal data analysis: Modeling change and event occurrence*. New York: Oxford University Press.

Soullier, N., E. de La Rochebrochard, and J. Bouyer. 2010. Multiple imputation for estimation of an occurrence rate in cohorts with attrition and discrete follow-up time points: a simulation study. *BMC Medical Research Methodology* 10 (1):1–7. doi:10.1186/1471-2288-10-79

Tourangeau, K., C. Nord, T. Le, A. G. Sorongon, and M. Najarian. 2009. *Early childhood longitudinal study, kindergarten class of 1998-99 (ecls-k): combined user's manual for the ecls-k eighth-grade and k-8 full sample data files and electronic codebooks (NCES 2009-004)*. Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Yuan, K.-H., and L. Lu. 2008. SEM with missing data and unknown population distributions using two-stage ML: Theory and its application. *Multivariate Behavioral Research* 43 (4):621–652. doi:10.1080/00273170802490699

Yuan, K.-H., F. Yang-Wallentin, and P. M. Bentler. 2012. ML versus MI for missing data with violation of distribution conditions. *Sociological Methods & Research* 41 (4):598–629. doi:10.1177/0049124112460373