PSYCHOMETRIKA—VOL. 78, NO. 1, 154–184
JANUARY 2013
DOI: 10.1007/S11336-012-9301-5

METHODS FOR MEDIATION ANALYSIS WITH MISSING DATA

ZHIYONG ZHANG AND LIJUAN WANG

UNIVERSITY OF NOTRE DAME

Despite wide applications of both mediation models and missing data techniques, formal discussion of mediation analysis with missing data is still rare. We introduce and compare four approaches to dealing with missing data in mediation analysis including listwise deletion, pairwise deletion, multiple imputation (MI), and a two-stage maximum likelihood (TS-ML) method. An R package `bmem` is developed to implement the four methods for mediation analysis with missing data in the structural equation modeling framework, and two real examples are used to illustrate the application of the four methods. The four methods are evaluated and compared under MCAR, MAR, and MNAR missing data mechanisms through simulation studies. Both MI and TS-ML perform well for MCAR and MAR data regardless of the inclusion of auxiliary variables and for AV-MNAR data with auxiliary variables. Although listwise deletion and pairwise deletion have low power and large parameter estimation bias in many studied conditions, they may provide useful information for exploring missing mechanisms.

Key words: mediation analysis, missing data, MI, TS-ML, bootstrap, auxiliary variables.

1. Introduction

In behavioral and social sciences, mediation analysis is a widely used technique in exploring the underlying mechanism of observed relationships, and the issue of missing data is hardly avoidable even in a well-designed study (e.g., Enders, 2003; MacKinnon, 2008). However, formal discussion of mediation analysis with missing data is still rare. Figure 1 depicts the path diagram of a mediation model with one mediator. In the figure, X, M, and Y represent the independent or input variable, the mediation variable (mediator), and the dependent or outcome variable, respectively. In this model, the total effect of X on Y, $c' + ab$, consists of the direct effect c' and the indirect effect ab. The indirect effect is also called the mediation effect because it is the effect of X on Y through the mediation of M. The residual variances of M and Y are denoted by σ^2_{eM} and σ^2_{eY}.

The missing data problem is a challenge for any statistical analysis. Mediation analysis is not an exception. Missing data either reduce the efficiency of statistical inference and/or render it incorrect (e.g., Little & Rubin, 2002). Different general strategies have been developed to deal with missing data. For example, in listwise deletion an entire case is excluded from analysis if any single value is missing. In pairwise deletion, when data are missing for either (or both) variable(s) for a subject, the case is excluded from the computation of the covariance between these two variables. Multiple imputation (MI) first replaces missing data with plausible values and then analyzes the imputed data as complete data (e.g., Schafer, 1997). The maximum likelihood (ML) method obtains model parameters by maximizing the likelihood function based on all available data (e.g., Little & Rubin, 2002).

Although there are approaches to dealing with missing data for structural equation modeling (SEM) in general, such as aforementioned MI and ML, thorough discussion and treatment of missing data are still lacking in mediation analysis. A common practice is to analyze complete data through listwise deletion or pairwise deletion (e.g., Chen, Aryee, & Lee, 2005;

Requests for reprints should be sent to Zhiyong Zhang, University of Notre Dame, Notre Dame, IN, USA. E-mail: zzhang4@nd.edu

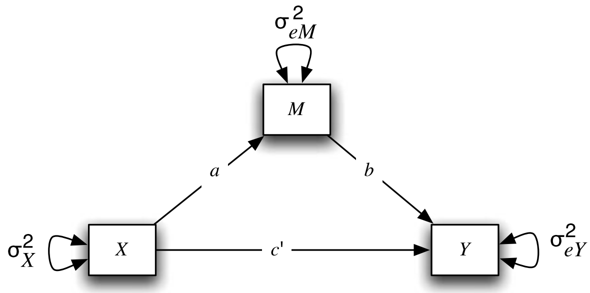

FIGURE 1.
Path diagram demonstration of a mediation model.

Preacher & Hayes, 2004). A limited number of studies have taken advantage of missing data handling routines in certain software when conducting mediation analysis but have not explicitly discussed possible influences of missing data in estimating and testing mediation effects (e.g., Bauer, Preacher, & Gil, 2006). Therefore, the aim of this study is to investigate methods that can be applied in mediation analysis when there are missing data and to introduce free software for implementing these methods.

In the rest of the paper, we will discuss and compare four methods for dealing with different types of missing data for mediation analysis including listwise deletion, pairwise deletion, multiple imputation (MI), and a two-stage maximum likelihood (TS-ML) method. We will first discuss how to estimate mediation effects (obtain point estimates of mediation effects) through the four methods. Then, we will show how to obtain confidence intervals of the mediation effects using a bootstrap method for inference. After that, two empirical examples will be provided to demonstrate the application of the introduced methods. Finally, we will conduct several simulation studies with finite samples to evaluate and compare the performance of those methods under different missing data mechanisms. An R package `bmem` that implements the four methods in this study will also be provided and illustrated.

2. A General Mediation Model with the Bentler–Weeks Representation

In this study, we specify a mediation model in the structural equation modeling framework to allow for the inclusion of latent variables and greater flexibility in modeling. A general mediation model with latent variables can be written as a Bentler–Weeks model (Bentler & Weeks, 1980)

$$\boldsymbol{\eta} = \boldsymbol{\beta}\boldsymbol{\eta} + \boldsymbol{\gamma}\boldsymbol{\xi}, \tag{1}$$

where $\boldsymbol{\eta}$ is a vector of observed or latent endogenous (dependent) variables, $\boldsymbol{\xi}$ is a vector of observed or latent exogenous (independent) variables, $\boldsymbol{\beta}$ is a matrix of coefficients determining the relationship among endogenous variables, and $\boldsymbol{\gamma}$ is a coefficient matrix governing the relationship between endogenous variables and exogenous variables. From a mediation analysis point of view, the coefficients $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ represent direct relations. The indirect effects are functions of $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$.

For example, for the mediation model shown in Figure 1, the Bentler–Weeks representation is

$$\begin{array}{ccccccccc} \boldsymbol{\eta} & = & \boldsymbol{\beta} & \boldsymbol{\eta} & + & \boldsymbol{\gamma} & \boldsymbol{\xi} \\ \begin{bmatrix} M \\ Y \end{bmatrix} & = & \begin{bmatrix} 0 & 0 \\ b & 0 \end{bmatrix} & \begin{bmatrix} M \\ Y \end{bmatrix} & + & \begin{bmatrix} a & 1 & 0 \\ c' & 0 & 1 \end{bmatrix} & \begin{bmatrix} X \\ e_M \\ e_Y \end{bmatrix} \end{array}. \tag{2}$$

In this example, a, b, and c' are direct relations. The mediation effect ab is the product of a and b. Furthermore, we can express the total effect of X on Y as $ab + c'$.

To estimate the model, we first specify a matrix $\mathbf{G}$ that differentiates the observed variables from the latent variables. The covariance matrix implied by the model can then be written as (Bentler & Weeks, 1980)

$$\boldsymbol{\Sigma}(\boldsymbol{\beta}, \boldsymbol{\gamma}) = \mathbf{G}(\mathbf{I}_1 - \mathbf{B})^{-1}\boldsymbol{\Gamma}\boldsymbol{\Phi}\boldsymbol{\Gamma}'\big[(\mathbf{I}_1 - \mathbf{B})'\big]^{-1}\mathbf{G}', \tag{3}$$

where, $\boldsymbol{\Phi}$ is the covariance matrix for the exogenous variables, $\boldsymbol{\Gamma}' = [\boldsymbol{\gamma}', \mathbf{I}_2]$, and

$$\mathbf{B} = \begin{bmatrix} \boldsymbol{\beta} & \mathbf{0}_1 \\ \mathbf{0}_2 & \mathbf{0}_3 \end{bmatrix}, \tag{4}$$

with the dimensions of the identity matrix $\mathbf{I}$ and the zero matrices $\mathbf{0}$ determined by a given model.

Assume that there are o_1 and l_1 manifest and latent endogenous variables, respectively, in the general model of Equation (1). Furthermore, assume that there are o_2 and l_2 manifest and latent exogenous variables, respectively. Then, $\boldsymbol{\eta}$ is an $(o_1 + l_1) \times 1$ vector, and $\boldsymbol{\xi}$ is an $(o_2 + l_2) \times 1$ vector. $\boldsymbol{\beta}$ is an $(o_1 + l_1) \times (o_1 + l_1)$ matrix, and $\boldsymbol{\gamma}$ is an $(o_1 + l_1) \times (o_2 + l_2)$ matrix. $\mathbf{G}$ is an $(o_1 + o_2) \times (o_1 + o_2 + l_1 + l_2)$ matrix. $\mathbf{I}_1$ and $\mathbf{B}$ are $(o_1 + o_2 + l_1 + l_2) \times (o_1 + o_2 + l_1 + l_2)$ matrices. $\boldsymbol{\Gamma}$ is an $(o_1 + o_2 + l_1 + l_2) \times (o_2 + l_2)$ matrix, and $\boldsymbol{\Phi}$ is an $(o_2 + l_2) \times (o_2 + l_2)$ matrix. $\mathbf{I}_2$ is an $(o_2 + l_2) \times (o_2 + l_2)$ identify matrix, $\mathbf{0}_1$ is an $(o_1 + l_1) \times (o_2 + l_2)$ matrix, $\mathbf{0}_2$ is an $(o_2 + l_2) \times (o_1 + l_1)$ matrix, and $\mathbf{0}_3$ is an $(o_2 + l_2) \times (o_2 + l_2)$ matrix. For example, for the specific mediation model in Equation (2), $o_1 = 2$, $o_2 = 1$, $l_1 = 0$, and $l_2 = 2$. Therefore, the matrices involved are

$$\mathbf{G} = [\mathbf{I}_{3\times 3} \quad \mathbf{0}_{3\times 2}], \quad \mathbf{I}_1 = \mathbf{I}_{5\times 5}, \quad \mathbf{I}_2 = \mathbf{I}_{3\times 3}, \quad \mathbf{0}_1 = \mathbf{0}_{2\times 3}, \quad \mathbf{0}_2 = \mathbf{0}_{3\times 2}, \quad \mathbf{0}_3 = \mathbf{0}_{3\times 3}. \tag{5}$$

The covariance matrix $\boldsymbol{\Sigma}$ for the Bentler–Weeks model can also be obtained using the RAM notation in which the $\mathbf{G}$ matrix serves as the filter matrix, $\boldsymbol{\Phi}$ is a part of the symmetric matrix, and $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are elements of the asymmetric matrix (McArdle & Boker, 1990). Let $\mathbf{S}$ denote the sample covariance matrix for the observed variables. The parameter estimates for $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$, $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\gamma}}$, can be obtained by minimizing the discrepancy function

$$F = \operatorname{tr}\big(\mathbf{S}\boldsymbol{\Sigma}^{-1}\big) - \log\big|\mathbf{S}\boldsymbol{\Sigma}^{-1}\big| - p, \tag{6}$$

where p is the total number of observed variables. For a sought-after mediation effect, it can be constructed from $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\gamma}}$ as illustrated in our applications. For convenience, we denote a mediation effect as $I(\boldsymbol{\beta}, \boldsymbol{\gamma})$ to allow the description of a more complicated mediation effect (e.g., the total mediation effect with multiple mediators). For example, for the mediation effect in Equation (2), it is $I(\boldsymbol{\beta}, \boldsymbol{\gamma}) = ab$.

Statistical approaches to test mediation effects for complete data have been discussed in the literature (e.g., MacKinnon, 2008). One approach is to testing the null hypothesis $H_0: I(\beta, \gamma) = 0$. If a large sample is available, the normal approximation method can be used, which constructs the standard error of $I(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\gamma}})$ through the delta method. For example, for $I(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\gamma}}) = \hat{a}\hat{b}$ in the simple mediation model, $\widehat{s.e.}(\hat{a}\hat{b}) = \sqrt{\hat{b}^2\hat{\sigma}_a^2 + 2\hat{a}\hat{b}\hat{\sigma}_{ab} + \hat{a}^2\hat{\sigma}_b^2}$ with parameter estimates $\hat{a}$ and $\hat{b}$, their estimated variances $\hat{\sigma}_a^2$ and $\hat{\sigma}_b^2$, and covariance $\hat{\sigma}_{ab}$ (Sobel, 1982, p. 298). Researchers have shown that the distribution of a mediation effect may not be normal especially when the sample size is small (Bollen & Stine, 1990; MacKinnon, Lockwood, Hoffman, West, & Sheets, 2002). Therefore, bootstrap methods have been recommended to obtain the empirical distributions and confidence intervals of mediation effects (e.g., MacKinnon, Lockwood,

& Williams, 2004; Preacher & Hayes, 2008; Shrout & Bolger, 2002). MacKinnon et al. (2004) further showed that the bias-corrected (BC) confidence intervals have good Type I error rate and the largest power among many different confidence intervals evaluated.

3. Methods for Estimating and Testing Mediation Effects with Missing Data

In this section, we discuss how to conduct mediation analysis when there are missing data. Little and Rubin (2002) distinguished three kinds of missing data mechanisms: missing completely at random (MCAR), missing at random (MAR), and missing not at random (MNAR; see also, e.g., Rubin, 1976; Schafer, 1997). Let D denote all the data that can be potentially observed on variables in a model. For example, for a simple mediation model, $D = (X, M, Y)$. D_{obs} and D_{miss} denote data that are actually observed and data that are not observed, respectively. Let R denote an indicator matrix with the same dimension as D. If a datum in D is missing, the corresponding element in R is equal to 1, otherwise 0. Finally, let A denote data for auxiliary variables that may be related to the missingness of D but are not a part of the model. Suppose that we are interested in the change of mathematical ability in a test–retest experiment. D would include the mathematical test scores from the initial test and the retest. In addition to the data on mathematical ability, we may also collect data on reading ability of the participants. Then, A would include the data on reading ability.

The missing mechanism is MCAR if $\Pr(R|D_{obs}, D_{miss}, A, \boldsymbol{\theta}) = \Pr(R|\boldsymbol{\theta})$, where the vector $\boldsymbol{\theta}$ represents unknown model parameters. This suggests that missing data D_{miss} are a simple random sample of D and the missingness is not related to D_{obs}, D_{miss}, or A. For example, in the test–retest experiment, some participants may miss their retests simply because of traffic jam or random illness. This kind of missingness can be viewed as MCAR. The missing mechanism is MAR if $\Pr(R|D_{obs}, D_{miss}, A, \boldsymbol{\theta}) = \Pr(R|D_{obs}, \boldsymbol{\theta})$, meaning that the probability that a datum is missing is related to the data actually observed D_{obs} but not to the missing data D_{miss} or A. In the test–retest example, data may be complete for the initial test. However, some participants may not attend the retest because they did not perform well in the initial test. Such missingness could be explained by the observed mathematical ability data at the initial test and therefore the missing mechanism is MAR. Popular missing data methods and techniques in general assume that missing data are MCAR or MAR (e.g., Little & Rubin, 2002; Schafer, 1997).

Finally, the missing mechanism is MNAR if the missing probability of a datum is related to the missing data D_{miss} and/or A while A are not included in the data analysis. If missingness is only related to A and A are observed and included in the data analysis, then the overall missing mechanism becomes MAR. For the test–retest experiment, some participants may not complete the tests because of reading deficits. Therefore, if we include the test scores on reading ability into data analysis, the missingness can be well explained. However, if missingness is related to D_{miss}, the inclusion of A may not explain or change the missing mechanism. For the test–retest example, the data for the initial test are complete. However, in the retest, participants may refuse to complete it believing that they were not doing well during the retest. Therefore, such missingness is directly related to their potential retest scores and cannot be addressed using the data from the initial test or the reading ability test. For convenience, if the MNAR mechanism can be explained by the inclusion of auxiliary variables such as A, we call it auxiliary variables dependent MNAR (AV-MNAR), and otherwise, we call it model variables dependent MNAR (MV-MNAR).

In the following, we will discuss four methods for dealing with missing data in mediation analysis. These methods include listwise deletion, pairwise deletion, MI, and TS-ML. Listwise deletion and pairwise deletion are traditional methods, which have been found to perform poorly in dealing with missing data for many models. In practice, however, they are still widely used in mediation analysis with missing data. In addition, we may be able to obtain some useful sensitivity information regarding missing mechanisms by comparing results from multiple missing data analysis methods. Therefore, we decide to include the evaluation of listwise and pairwise deletion in this study. Both MI and TS-ML are flexible and well-performing methods for dealing with missing data and allow the inclusion of auxiliary variables. We set our discussion in a general setting with p observed variables in the mediation model denoted by $D_1, D_2, \dots, D_p$ and a set of q ($q \geq 0$) auxiliary variables $A_1, A_2, \dots, A_q$. By augmenting the auxiliary variables with the mediation model variables, we have a total of $p+q$ variables denoted by $Z = (D_1, D_2, \dots, D_p, A_1, \dots, A_q)'$. Note that q could be 0 indicating that no auxiliary variable is included. Although auxiliary variables could have missing data, we recommend to include the auxiliary variables with complete data. The general theme of the four methods for mediation effect estimation is to first estimate the saturated covariance matrix for the mediation model variables and then obtain model parameter estimates by minimizing the discrepancy function in Equation (6) using the estimated saturated covariance matrix.

3.1. Listwise Deletion and Pairwise Deletion

For listwise deletion, a whole case is deleted if any single value of the model variables from the case is missing. The saturated covariance matrix for the mediation model variables is then calculated based on the complete cases. In calculating the covariance matrix, the sample size of the complete data is used. For pairwise deletion, the covariance between any two variables is estimated using complete data of the two variables and is then used to form a covariance matrix for all model variables. As a result, different covariances in the matrix may be based on different sample sizes, and the estimated covariance matrix may not be positive-definite (Little & Rubin, 2002). With the estimated covariance matrix, the parameter estimates and the mediation effects can be obtained as $\hat{\boldsymbol{\beta}}$, $\hat{\boldsymbol{\gamma}}$, and $I(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\gamma}})$. Neither listwise deletion nor pairwise deletion can take advantage of the use of auxiliary variables.

3.2. Multiple Imputation (MI)

Multiple imputation (e.g., Little & Rubin, 2002; Schafer, 1997) is a procedure to fill each missing value with a set of plausible values that represent the uncertainty about the right value to impute. The multiple imputed data sets are then analyzed by using standard procedures for complete data, and the results from these analyses are combined to obtain point estimates of model parameters. Multiple imputation has been implemented in software such as SAS and LISREL.

Specifically for mediation analysis, the following steps can be implemented.

1. Assuming that $Z = (D_1, D_2, \dots, D_p, A_1, \dots, A_q)'$ is from a multivariate normal distribution, generate K sets of values for each missing value. Because the auxiliary variables are augmented with the model variables in Z, the information in auxiliary variables is automatically utilized in imputing the missing data. Combine the generated values with the observed data to get K sets of complete data for $(D_1, D_2, \dots, D_p)$. K is the number of imputations.
2. For each of the K sets of complete data, apply the complete data mediation analysis to obtain an estimate of the mediation effect $I(\hat{\boldsymbol{\beta}}_k, \hat{\boldsymbol{\gamma}}_k)$, $k = 1, \dots, K$.
3. The combined estimation (the final point estimate from MI) for the mediation effect is the average of the K complete data mediation effect estimates:

$$I(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\gamma}}) = \frac{1}{K} \sum_{k=1}^{K} I(\hat{\boldsymbol{\beta}}_k, \hat{\boldsymbol{\gamma}}_k). \tag{7}$$

3.3. *Two-Stage Maximum Likelihood (TS-ML) Method Using the EM Algorithm*

The expectation-maximization (EM) algorithm is a popular technique for handling missing data (Little & Rubin, 2002; Schafer, 1997). The EM algorithm is an iterative method consisting of the expectation step (E-step) and the maximization step (M-step). It starts with a guess as to the values of unknown parameters. For example, parameter estimates from listwise deletion can be used as starting values. In the E-step, expectations of missing data conditional on the unknown parameters are calculated. In the M-step, new parameter estimates are obtained through maximization routines by plugging in the expectations of missing data from the E-step. These two steps are repeated iteratively until changes in the parameter estimates are small enough.

For mediation analysis with missing data, a two-stage ML approach using the EM algorithm is adopted in this study. In the first stage, the EM algorithm is applied to estimate the saturated mean and covariance matrix of Z without assuming a particular mediation model. In the second stage, model parameters are estimated using the saturated covariance matrix obtained from the first stage. The two-stage strategy has been applied in the literature to deal with missing data under different situations. For example, Yuan and Bentler (2000) applied the two-stage method for mean and covariance structure analysis with nonnormal missing data. Enders (2003) employed the two-stage approach to obtain internal consistency reliability estimates with item-level missing data. A main reason for us to adopt the two-stage method instead of the full information likelihood (FIML) method is that Savalei and Bentler (2009) and Savalei and Falk (in press) showed that the two-stage method can perform as well as or even outperform FIML. Furthermore, it is straightforward to include auxiliary variables in the two-stage method. To include auxiliary variables, one only needs to augment them with the model variables directly as shown below.

Let $z_{i,obs}$ and $z_{i,miss}$ represent observed data and missing data of individual i, respectively. Furthermore, let U denote the mean vector, and S denote the covariance matrix of the augmented data Z $(D_1, D_2, \ldots, D_p, \ldots, A_q)$. Then the following procedure utilizing the EM algorithm can be implemented to obtain parameter estimates.

1. Let d denote a small number, such as 10^{-6}, for convergence criterion. Start with $U^{(0)}$ and $S^{(0)}$ that are obtained from listwise deletion. Therefore, $U^{(0)}$ and $S^{(0)}$ are the sample mean and covariance matrix of the complete observed data.
2. At iteration t, one has $U^{(t)}$ and $S^{(t)}$.
3. At iteration $t+1$,
   1. In E-step, calculate the expectation of $z_{i,miss}^{(t)}$ based on $U^{(t)}$ and $S^{(t)}$ (given in Appendix A).
   2. In M-step, replace missing data using the expectation of $z_{i,miss}^{(t)}$ and then obtain $U^{(t+1)}$ and $S^{(t+1)}$.
4. Calculate the maximum absolute relative difference between $U^{(t)}$, $S^{(t)}$ and $U^{(t+1)}$, $S^{(t+1)}$ as $e = \max_{j,k}[|\frac{u_j^{(t+1)} - u_j^{(t)}}{u_j^{(t)}}|, |\frac{s_{jk}^{(t+1)} - s_{jk}^{(t)}}{s_{jk}^{(t)}}|]$, where u_j is the jth element of U, and s_{jk} is the element of S on the jth row and kth column. If $e \leq d$, stop the iteration. Otherwise, go to the next iteration.
5. Let $\hat{U}$ and $\hat{S}$ denote the estimated saturated mean vector and covariance matrix of Z from the EM algorithm after convergence. Denote $\hat{U}_M$ and $\hat{S}_M$ as the saturated mean vector and covariance matrix of $D_1, D_2, \ldots, D_p$ in the mediation model. Then the mediation model parameter estimates can be obtained by minimizing Equation (6) with $\hat{S}_M$ replacing $\mathbf{S}$.
6. The mediation effects are estimated by $I(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\gamma}})$.

The method discussed here will be referred to as TS-ML in this study.

3.4. *Testing Mediation Effects Through the Bootstrap Method*

After obtaining point estimates of the mediation effects using one of the above discussed methods, we can use bootstrap methods to obtain empirical distributions and confidence intervals of mediation effects. Bootstrap methods (Efron, 1979, 1994; Yung, 1996) were first employed in mediation analysis by Bollen and Stine (1990) and have been studied in a variety of research contexts (e.g., MacKinnon et al., 2004; Preacher & Hayes, 2008; Shrout & Bolger, 2002). This method has no distribution assumption on the indirect effects $I(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\gamma}})$. Instead, it approximates the distributions of $I(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\gamma}})$ using their bootstrap empirical distributions.

The original discussion on applying bootstrap methods for missing data analysis can be found in Efron (1994). Specifically for mediation analysis, the following procedure can be used.

1. Using the *original data set* (sample size $= N$) as a population, draw a bootstrap sample of N persons randomly with replacement from the original data set. The bootstrap sample usually contains missing data as well.
2. With the bootstrap sample, estimate model parameters and mediation effects using one of the above methods: listwise deletion, pairwise deletion, MI, or TS-ML.
3. Repeat Steps 1 and 2 for a total of B times. B is the number of bootstrap samples.
4. Empirical distributions of model parameters and mediation effects are then obtained using the B sets of bootstrap estimates. Thus, confidence intervals of model parameters and mediation effects can be constructed.

Using the bootstrap sample estimates, one can obtain the bootstrap standard errors and confidence intervals of model parameters and mediation effects conveniently. Let $\boldsymbol{\theta}$ denote a vector of model parameters and mediation effects. For example, for the simple mediation model $\boldsymbol{\theta} = (a, b, c', \sigma_{eM}^2, \sigma_{eY}^2, ab)^t$, with data from each bootstrap, we can obtain $\hat{\boldsymbol{\theta}}^b$, $b = 1, \dots, B$. The standard error of the kth component $\hat{\theta}_k$ of $\boldsymbol{\theta}$ can be calculated as

$$\widehat{s.e.}(\hat{\theta}_k) = \sqrt{\sum_{b=1}^{B} \left(\hat{\theta}_k^b - \bar{\hat{\theta}}_k\right)^2 / (B-1)} \tag{8}$$

with

$$\bar{\hat{\theta}}_k = \sum_{b=1}^{B} \hat{\theta}_k^b / B. \tag{9}$$

Many methods for constructing confidence intervals from $\hat{\boldsymbol{\theta}}^b$ have been proposed such as the percentile interval, the bias-corrected (BC) interval, and the bias-corrected and accelerated (BCa) interval (Efron, 1987; MacKinnon et al., 2004). In the present study, we focus on the BC interval because MacKinnon et al. (2004) showed that the BC confidence intervals perform better than the percentile interval and the BCa interval in complete data mediation analysis. The BC confidence interval corrected the bias between the estimated mediation effect from the original sample and the average mediation effect estimate from the bootstrap samples (Efron & Tibshirani, 1993).

The $100(1-2\alpha)$ % BC interval for the kth element of $\boldsymbol{\theta}$ can be constructed using the percentiles $\hat{\theta}_k^b(\tilde{\alpha}_l)$ and $\hat{\theta}_k^b(\tilde{\alpha}_u)$ of $\hat{\theta}_k^b$. Here

$$\tilde{\alpha}_l = \Phi\left(2z_0 + z^{(\alpha)}\right) \tag{10}$$

and

$$\tilde{\alpha}_u = \Phi\left(2z_0 + z^{(1-\alpha)}\right), \tag{11}$$

where Φ is the standard normal cumulative distribution function, $z^{(\alpha)}$ is the α percentile of the standard normal distribution, and

$$z_0 = \Phi^{-1}\left[\frac{\text{number of times that } \hat{\theta}_k^b < \hat{\theta}_k}{B}\right] \tag{12}$$

with $\hat{\theta}_k$ denoting the kth parameter estimate or mediation effect from the original sample.

3.5. *Implementation*

To ease the application of missing data techniques in mediation analysis, we implement the four estimation methods for estimating mediation effects and the bootstrap method for confidence interval construction in an R package called `bmem`. The package `bmem` is built on the popular free R package `sem` (Fox, 2006) for structural equation modeling, which uses the flexible RAM notation to specify a model. The package `bmem` is freely available on the R website.[1] It is also incorporated in a newly developed free online SEM program, WebSEM, that provides a graphical interface for conducting mediation analysis through drawing path diagrams (Zhang & Yuan, 2012).

4. Two Empirical Examples

In this section, we illustrate the application of the four missing data methods discussed in the preceding section through two examples.

4.1. *Example 1*

Research has found that parents' education levels influence adolescent mathematics achievement directly and indirectly. For example, Davis-Kean (2005) showed that parents' education levels are related to children's academic achievement through parents' beliefs and behaviors. To test a similar hypothesis, we investigate whether home environment mediates the relationship between mothers' education and children's mathematical achievement.

Data used in this example are randomly sampled from the National Longitudinal Survey of Youth, the 1979 cohort (NLSY79, Center for Human Resource Research, 2006). Data were collected in 1986 from 76 families on mothers' education level (ME), home environment (HE), children's mathematical achievement (Math) and reading recognition ability. The HE variable is an observation measure of how mothers interact with and support children at home (Cladwell & Bradley, 1979). We hypothesize that mothers with more eduction create a better home environment (e.g., more positive interactions with children at home) that helps the development of children's mathematical ability. Therefore, for the mediation analysis, mothers' education is the input variable, home environment is the mediator, and children's mathematical achievement is the outcome variable. The reading recognition variable is used as an auxiliary variable. The path diagram for the mediation model is given in Figure 2. Note that the mediation effect is $a * b$.

In the data set, there are 10 % missing data in HE and 12 % missing data in Math. The results from listwise deletion, pairwise deletion, MI, and TS-ML are provided in Table 1 with R code for the analysis provided in Appendix B. The results based on 1000 bootstraps show the following. First, model parameter and mediation effect estimates from different methods are different. For example, for the direct effect cp, it is positive and significant from listwise deletion but is

[1] To install the package for the first time use, one can issue the command `install.packages("bmem")` in R console.

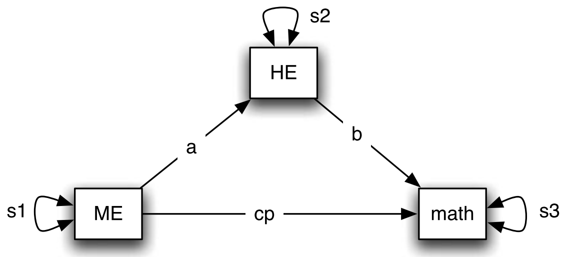

FIGURE 2.
HE as a mediator between ME and math.

negative and insignificant from pairwise deletion, MI, and TS-ML. Second, after including the auxiliary variable, the results for MI and TS-ML are different from those excluding the auxiliary variable. For example, without the auxiliary variable, the mediation effect is significant. But with the auxiliary variable, the mediation effect becomes insignificant. Therefore, the auxiliary variable, reading recognition, plays an important role when evaluating the mediation effect of home environment on the relationship between mothers' education and children's mathematical achievement with missing data. Third, MI and TS-ML are more time consuming than listwise and pairwise deletion. For example, TS-ML took about three times the computation time of listwise and pairwise deletion, whereas MI took about 70 times the computation time of listwise and pairwise deletion.

4.2. *Example 2*

Data used in this example are from the Advanced Cognitive Training for Independent and Vital Elderly study (ACTIVE; Jobe, Smith, Ball, Tennstedt, Marsiske, Willis, & Kleinman, 2001). A subset of $N = 115$ participants were measured on seven variables including age, education level, the Hopkins Verbal Learning Test (HVLT; Brandt, 1991), the word series (WS; Gonda & Schaie, 1985) test, the letter series (LS; Thurstone & Thurstone, 1949) test, the letter sets (LT; Ekstrom, French, Harman, & Derman, 1976) test, and the Everyday Problems Test (EPT; Willis & Marsiske, 1993). By design, HVLT is a measure of memory ability, and WS, LS, and LT are measures of reasoning ability. EPT measures the participants' ability to solve problems associated with daily living.

Using this set of data, we illustrate how to investigate the possible mechanism by which age and education are related to EPT through the mediation model given in Figure 3. In this model, there are two mediators, the memory ability measured by HVLT and the reasoning ability (R) indicated by WS, LS, and LT. Thus, the current mediation model involves two mediators, and one of them is a latent variable. For this data analysis, no auxiliary variables are used.

In the data, there are 13 %, 1 %, and 2 % missing data in HVLT, LS, and LT, respectively. The results based on the four missing data analysis methods with 1000 bootstraps are presented in Table 2 with R code for the analysis provided in Appendix C.[2] Overall, the four methods obtained very similar parameter and mediation effect estimates. The conclusions on mediation effects that can be drawn from the four methods are identical. For example, the negative relationship between age and EPT was completely mediated by memory ability and reasoning ability because both indirect effects, from age to EPT via HVLT ($a * b$) and from age to EPT through R ($d * h$), and the total indirect effect ($a * b + d * h$) were significant, whereas the direct effect cp

[2]The model fitted the data well according to the chi-square test using the Bollen-Stine bootstrap method implemented in `bmem`.

TABLE 1.
Results from testing mediation effect of HE in Example 1.

Parameter	Listwise deletion			Pairwise deletion			MI						TS-ML					
							Without AV			With AV			Without AV			With AV		
	Est	L	U	Est	L	U	Est	L	U	Est	L	U	Est	L	U	Est	L	U
a	0.31	0.13	0.56	0.26	0.07	0.51	0.18	−0.07	0.41	0.15	−0.12	0.42	0.19	−0.05	0.43	0.15	−0.09	0.42
b	0.63	0.06	1.22	0.99	0.35	1.73	2.62	0.87	4.31	2.72	1.05	3.93	2.68	1.11	4.15	2.87	1.33	4.26
cp	0.89	0.08	1.47	−0.12	−2.16	0.97	−0.42	−1.95	0.71	−0.31	−1.76	0.67	−0.47	−1.96	0.60	−0.39	−1.66	0.70
$s1$	3.24	2.03	5.18	3.02	2.05	4.78	3.02	2.04	4.71	3.02	2.15	4.87	2.98	2.12	4.63	2.98	2.09	4.59
$s2$	2.47	1.69	3.54	2.57	1.81	3.69	3.15	2.14	5.29	3.77	2.34	7.50	2.98	2.00	4.78	3.55	2.04	6.94
$s3$	19.9	13.7	30.1	92.2	46.6	144.5	71.2	43.6	116.9	62.7	37.2	110.6	71.7	43.8	127.4	63.1	35.9	108.6
$a * b$	0.20	0.03	0.55	0.26	0.06	0.76	0.47	0.00	1.33	0.41	−0.15	1.27	0.51	0.05	1.34	0.43	−0.13	1.38
	72 s			72 s			4636 s			4999 s			175 s			227 s		

Note. The parameters correspond to the paths in Figure 2. MI: Multiple imputation. TS-ML: two-stage maximum likelihood. AV: auxiliary variable. Est: parameter estimate. L and U: lower and upper limits for the BC confidence interval. The BC confidence intervals are based on 1000 replications of bootstrap. The last row in the table provides the computation time in seconds.

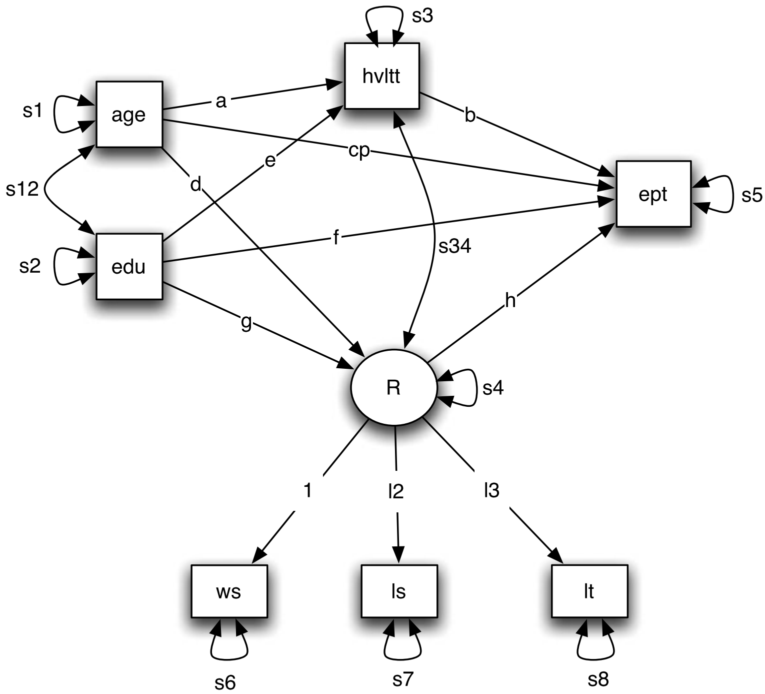

FIGURE 3.
A multiple mediator model with a latent variable.

was not significant. Furthermore, the positive relationship between education and EPT was also completely mediated by memory ability and reasoning ability.

The computation time for listwise deletion, pairwise deletion and TS-ML are very close, about 620 seconds, approximately 0.6 second for each bootstrap replication. However, the computation time required for MI is about 38 times that of the other three methods. This is about the size of the multiple imputations used in MI ($K = 40$ in this example).

5. Simulation Comparison of Listwise Deletion, Pairwise Deletion, MI, and TS-ML

In the previous section, we have demonstrated the use of listwise deletion, pairwise deletion, MI, and TS-ML through two examples. In the first example, the four methods led to different conclusions, while in the second example they obtained very similar results. Furthermore, in Example 1, the inclusion of an auxiliary variable was shown to affect the mediation analysis. Therefore, it is important to investigate when the four methods will obtain the same or different results and what methods are recommended for use under certain conditions. In this section, we conduct several simulation studies to investigate the influence of the use of different methods and auxiliary variables on mediation analysis with missing data.

5.1. Simulation Design

To better control for confounding factors and be comparable with the existing literature, we focus our simulation on the simple mediation model depicted in Figure 1. The population parameters for this model are set at $\sigma_X^2 = \sigma_{eM}^2 = \sigma_{eY}^2 = 1$ and $c' = 0$. By changing the values of a and b we control the effect size of mediation effects. In the simulation, we investigate

TABLE 2.
Results for the multiple mediator model with a latent variable.

Parameter	Listwise deletion			Pairwise deletion			MI			TS-ML		
	Est	L	U	Est	L	U	Est	L	U	Est	L	U
a	−0.37	−0.56	−0.19	−0.36	−0.53	−0.13	−0.35	−0.51	−0.15	−0.35	−0.53	−0.17
cp	0.00	−0.19	0.17	0.03	−0.15	0.20	0.03	−0.13	0.19	0.03	−0.15	0.19
d	−0.06	−0.11	−0.02	−0.06	−0.10	−0.03	−0.06	−0.10	−0.02	−0.06	−0.10	−0.02
e	0.46	0.14	0.74	0.49	0.16	0.83	0.49	0.17	0.80	0.49	0.16	0.77
f	0.32	−0.09	0.69	0.30	−0.07	0.66	0.29	−0.08	0.64	0.29	−0.08	0.64
g	0.21	0.11	0.30	0.22	0.14	0.30	0.22	0.13	0.30	0.22	0.14	0.31
b	0.35	0.10	0.55	0.37	0.11	0.60	0.39	0.13	0.59	0.40	0.16	0.60
h	2.36	1.32	3.55	2.36	1.34	3.49	2.38	1.40	3.45	2.36	1.26	3.40
$a*b$	−0.13	−0.25	−0.05	−0.13	−0.28	−0.04	−0.14	−0.26	−0.06	−0.14	−0.27	−0.06
$d*h$	−0.15	−0.29	−0.05	−0.15	−0.28	−0.05	−0.15	−0.28	−0.05	−0.15	−0.28	−0.05
$e*b$	0.16	0.04	0.37	0.18	0.04	0.45	0.19	0.04	0.37	0.19	0.06	0.45
$g*h$	0.48	0.23	0.83	0.52	0.29	0.88	0.52	0.29	0.86	0.52	0.27	0.81
$a*b+d*h$	−0.28	−0.43	−0.13	−0.28	−0.44	−0.13	−0.28	−0.42	−0.14	−0.28	−0.43	−0.14
$e*b+g*h$	0.64	0.32	1.00	0.70	0.40	1.06	0.71	0.41	1.08	0.71	0.41	1.06
$s1$	32.49	25.72	42.09	31.99	24.95	39.82	31.99	25.17	40.54	31.71	24.79	39.66
$s2$	7.86	5.80	10.40	7.77	5.98	10.05	7.77	5.88	9.81	7.70	5.91	9.90
$s12$	−2.81	−5.69	−0.12	−2.80	−5.45	−0.34	−2.80	−5.63	−0.38	−2.78	−5.25	−0.02
$s3$	22.27	18.19	28.67	23.27	17.85	29.64	23.46	18.96	29.67	23.16	18.27	29.63
$s4$	14.71	11.38	19.95	16.44	13.43	22.06	15.67	12.52	20.56	15.53	12.14	20.66
$s5$	3.23	1.72	5.45	3.47	1.85	5.46	3.36	1.84	5.14	3.32	1.83	5.20
$s6$	3.86	1.56	6.23	4.00	2.09	6.40	3.84	1.91	6.03	3.81	1.84	6.19
$s7$	3.60	2.72	4.64	3.82	2.99	4.95	3.75	2.89	4.90	3.72	2.84	4.87
	624 s			625 s			23692 s			608 s		

Note. The parameters correspond to the paths in Figure 3. MI: Multiple imputation. TS-ML: two-stage maximum likelihood. Est: parameter estimate. L and U: lower and upper limits for the BC confidence interval. The BC confidence intervals are based on 1000 replications of bootstrap. The last row in the table provides the computation time in seconds.

no mediation ($a = b = 0$, $ab = 0$), medium ($a = b = 0.39$, $ab = 0.1521$), and large ($a = b = 0.59$, $ab = 0.3481$) effect sizes (MacKinnon et al., 2004).[3] Two auxiliary variables (A_1 and A_2) are considered in the simulation, where the correlation between A_1 and M and the correlation between A_2 and Y are both 0.5. Three other factors that could influence mediation analysis are also considered, including the sample size $N = 50$, 100, and 200, the percentage of missing data at 10 %, 20 %, and 40 %, and missing data mechanisms: MCAR, MAR, and MNAR.

Specifically, missing data are generated in the following way. First, $R = 1000$ sets of complete data are generated with a given sample size and a given effect size. Missing data are then generated from the 1000 sets of complete data. To facilitate the comparisons among different missing mechanisms, missing data are only allowed in M and Y. Thus, data on X, A_1, and A_2 are complete. For MCAR, s % ($s = 10, 20, 40$) of data are randomly set as missing for M and Y independently. For MAR data, the missing probability of Y and M depends only on X. Specifically, if X is smaller than its sth percentile, M is missing and if X is larger than its $(100 - s)$th percentile, Y is missing. Two types of MNAR data are generated. For AV-MNAR, it is assumed

[3] We have also conducted simulations where $c' = 0.5$ with no ($a = b = 0$), medium ($a = b = 0.39$), and large ($a = b = 0.59$) effect sizes. For the condition of no mediation effects, we also considered $a = 0.39$ & $b = 0$ and $a = 0$ & $b = 0.39$ and no different patterns were observed. The results from these simulations can be seen on the authors' website at http://rpackages.psychstat.org/examples/bmem/Supplenment%20results.pdf.

TABLE 3.
Parameter estimation bias (in %) of listwise deletion, pairwise deletion, and TS-ML.

c	$a=b$	Rate	$\sigma_X^2=\sigma_{eM}^2=\sigma_{eY}^2=1$						$\sigma_X^2=\sigma_{eM}^2=\sigma_{eY}^2=2$					
			MAR			MNAR			MAR			MNAR		
			LW	PW	TSML	LW	PW	TSML	LW	PW	TSML	LW	PW	TSML
0	0	10 %	0	0	0	0	0	0	0	0	0	0	0	0
0	0	20 %	0	0	0	0	0	0	0	0	0	0	0	0
0	0	30 %	0	0	0	0	0	0	0	0	0	0	0	0
0	0.14	10 %	0	−29	0	−47	−48	−88	0	−40	0	−47	−62	−88
0	0.14	20 %	0	−42	0	−63	−64	−89	0	−49	0	−63	−72	−88
0	0.14	30 %	0	−51	0	−73	−74	−88	0	−55	0	−73	−77	−88
0	0.39	10 %	0	−32	0	−42	−49	−92	0	−43	0	−42	−61	−92
0	0.39	20 %	0	−45	0	−57	−63	−92	0	−52	0	−57	−69	−93
0	0.39	30 %	0	−54	0	−67	−71	−92	0	−58	0	−67	−74	−93
0	0.59	10 %	0	−35	0	−37	−51	−98	0	−47	0	−37	−61	−100
0	0.59	20 %	0	−49	0	−51	−62	−100	0	−55	0	−51	−68	−101
0	0.59	30 %	0	−57	0	−60	−69	−101	0	−61	0	−60	−72	−102
0.5	0	10 %	0	0	0	0	0	0	0	0	0	0	0	0
0.5	0	20 %	0	0	0	0	0	0	0	0	0	0	0	0
0.5	0	30 %	0	0	0	0	0	0	0	0	0	0	0	0
0.5	0.14	10 %	0	−32	0	−49	−49	−85	0	−44	0	−49	−61	−85
0.5	0.14	20 %	0	−46	0	−66	−63	−85	0	−53	0	−66	−70	−86
0.5	0.14	30 %	0	−55	0	−75	−72	−86	0	−59	0	−75	−76	−85
0.5	0.39	10 %	0	−34	0	−42	−47	−87	0	−46	0	−42	−57	−88
0.5	0.39	20 %	0	−48	0	−57	−58	−88	0	−55	0	−57	−64	−88
0.5	0.39	30 %	0	−57	0	−67	−66	−88	0	−61	0	−67	−69	−89
0.5	0.59	10 %	0	−37	0	−35	−46	−89	0	−50	0	−35	−55	−91
0.5	0.59	20 %	0	−52	0	−49	−56	−91	0	−59	0	−49	−61	−93
0.5	0.59	30 %	0	−61	0	−59	−63	−93	0	−64	0	−59	−66	−94

Note. Rate: missing data rate; LW: listwise deletion; PW: pairwise deletion; TS-ML: two-stage maximum likelihood; MNAR represents the condition where M and Y missing depends on M and Y, respectively, and auxiliary variables are not considered.

that missingness of M depends on A_1 and missingness of Y depends on A_2. To be precise, if A_1 is smaller than its sth percentile, M is missing, and if A_2 is smaller than its sth percentile, Y is missing. Note that for AV-MNAR, if auxiliary variables A_1 and A_2 are included in an analysis, overall data become MAR. For MV-MNAR, missingness in M and Y depends on M and Y themselves, respectively. If M is smaller than its sth percentile, M is missing, and if Y is smaller than its sth percentile, Y is missing. For both MAR and MNAR, missing rates are also set at 10 %, 20 %, and 40 %. For our simulation design, the missing data rates range from 10 % to 80 % casewisely, and there will be 6.7 % to 26.7 % missing data points. One rationale of using the described methods for generating missing data is that the generated missing data are from truncated multivariate normal distributions; thus, we can use the properties of truncated multivariate normal distributions to analytically or numerically compute the asymptotic parameter estimation bias for some conditions (see the results in Table 3).

All data are generated and analyzed in R using our `bmem` package. For each of the four methods, the BC confidence intervals are obtained using 1000 bootstrap samples. For MI, 100 sets of data are imputed. For each missing mechanism, data are analyzed using listwise deletion, pairwise deletion, MI, and TS-ML. For MI and TS-ML, the generated data are analyzed both without and with the two auxiliary variables.

Comparisons of different methods will be conducted based on mediation effects because they are usually the focus of mediation analysis. First, we will study whether the mediation effects can be estimated accurately. The accuracy will be measured using the parameter estimation bias when the mediation effect is 0 and the relative parameter estimation bias when the mediation effect is not 0. For presentation purposes, we refer to both measures as bias. Let $\delta = ab$ denote the true mediation effect in the population, and $\hat{\delta}_r = \hat{a}_r \hat{b}_r$, $r = 1, \ldots, 1000$, denote the mediation effect estimate of the rth replication. The bias is calculated as

$$\text{Bias} = \begin{cases} 100 \times (\frac{1}{1000} \sum_{r=1}^{1000} \hat{\delta}_r), & \delta = 0, \\ 100 \times (\frac{\sum_{r=1}^{1000} \hat{\delta}_r}{1000\delta} - 1), & \delta \neq 0. \end{cases} \tag{13}$$

Note that the bias is rescaled through multiplying by 100. When comparing multiple methods, smaller bias indicates that the point estimate is more accurate. Throughout this study, we also use 5 % as a subjective cutoff for evaluating the accuracy of parameter estimates of a single method alone. If the bias is smaller than 5 %, it indicates that the bias is small. Second, we will evaluate which methods can obtain correct confidence intervals. Let $\hat{l_r}$ and $\hat{u_r}$ denote the lower and upper limits of the 95 % confidence interval for the mediation effect in the rth replication. The coverage probability is

$$\text{coverage} = \frac{\#(\hat{l}_r < \delta < \hat{u}_r)}{1000}, \tag{14}$$

where $\#(\hat{l}_r < \delta < \hat{u}_r)$ is the total number of replications with confidence intervals covering the population mediation effect. A good 95 % confidence interval should give coverage probability equal or close to 0.95. Third, statistical power in detecting mediation effects will be investigated and compared. Power is calculated by

$$\text{power} = \frac{\#(\hat{l}_r > 0) + \#(\hat{u}_r < 0)}{1000}, \tag{15}$$

where $\#(\hat{l}_i > 0)$ is the total number of replications with the lower limits of confidence intervals larger than 0, and $\#(\hat{u}_r < 0)$ is the total number of replications with the upper limits smaller than 0. Note that when the population mediation effect is 0, this becomes the Type I error rate. Given accurate parameter estimates and correct coverage probabilities, a better method should produce greater statistical power or Type I error rates closer to 0.05.

5.2. *Expected (Asymptotic) Results*

Our simulation setup enables us to study parameter estimation bias in ab numerically using properties of truncated multivariate normal distributions (Leppard & Tallis, 1989). The covariance matrices of mediation model variables under different missing mechanisms can be obtained through numerical integration according to different methods for dealing with missing data (Tallis, 1961; Tang & Bentler, 1997; Wilhelm & Manjunath, 2010). In Table 3, we present the parameter estimation bias for listwise deletion, pairwise deletion, and TS-ML under a variety of conditions.[4] Several observations can be made from Table 3. First, where there were no mediation effects, no bias in the estimates of ab was observed from listwise deletion, pairwise deletion,

[4]We only considered MAR and MNAR data because for MCAR the covariance matrices are expected to be the same with and without missing data and, thus, biased parameter estimates are not expected. Numerical results from MI cannot be easily obtained but it is expected that results from MI would be similar to those from TS-ML.

and TS-ML regardless of the missing mechanisms. Second, for MAR data, neither listwise deletion[5] nor TS-ML has biased parameter estimates, although pairwise deletion demonstrates quite large bias when mediation effects exist. Third, for MNAR data, all methods underestimate the mediation effect when mediation effect exists.

The results in Table 3 hold for the population or a sample with an infinite sample size. However, the performance of each method for finite samples may vary. Therefore, in the following, we will present results from four finite sample simulation studies to compare the performance of listwise deletion, pairwise deletion, MI, and TS-ML under MCAR, MAR, and MNAR missing mechanisms.

5.3. Finite Sample Simulation Study 1: Analysis of MCAR Data

In this simulation, we compared the performance of listwise deletion, pairwise deletion, MI and TS-ML for analyzing MCAR data in mediation analysis. To investigate the influences of the use of auxiliary variables on MI and TS-ML, we first analyzed the data without auxiliary variables and then repeated the data analysis with auxiliary variables. The results from listwise and pairwise deletion are based on analysis without auxiliary variables because they do not allow the use of auxiliary variables. The estimated relative bias, coverage probabilities, and power/Type I error with different missing data percentages and sample sizes under the investigated conditions are summarized in Tables 4, 5, and 6.

First, when there was no mediation effect ($a = b = 0$), the mediation effect estimates from the four methods were very close to the population value 0 with the absolute bias smaller than 1.5 % even when the sample size was as small as 50 and the missing data rate was as large as 40 %. Overall, the Type I error rate was smaller than the alpha level 5 % for listwise deletion and pairwise deletion while it was larger than 5 % for MI and TS-ML. Thus, it appeared that listwise deletion and pairwise deletion were more conservative while MI and TS-ML were more liberal in rejecting the null hypothesis of no mediation effect.

Second, with medium and large mediation effects, the relative bias of the parameter estimates from all methods under the studied MCAR conditions was always smaller than or close to the typical 5 % rule of thumb indicating that all methods estimated mediation effects very well. Coverage probabilities for listwise deletion and pairwise deletion seemed to be consistently smaller than the nominal level 95 % ranging from 85.8 % to 93.9 %, although with a larger sample size they became closer to 95 %. On the contrary, coverage probabilities for both MI and TS-ML were closer to the nominal level 95 %, ranging from 93.1 % to 96.5 %. Furthermore, it is evident that statistical power of MI and TS-ML was much larger than that of listwise and pairwise deletion. It is also clear that power increased with sample size and mediation effect size but decreased with the amount of missing data. Overall, it seems that the inclusion of auxiliary variables can booster power in detecting mediation effects for MI and TS-ML.

5.4. Finite Sample Simulation Study 2: Analysis of MAR Data

In this simulation, we compare the performance of listwise deletion, pairwise deletion, MI and TS-ML for analyzing MAR data in mediation analysis. The estimated relative bias, coverage probabilities, and power/Type I error rates are summarized in Tables 7, 8, and 9.

First, when there is no mediation effect, the parameters were still estimated well with bias less than 1.5 % regardless of the use of different methods. In addition, similar to MCAR results,

[5]This result is due to the MAR missing data manipulation method of this study. It does not imply that listwise deletion works the same way for all MAR data. However, for some specific MAR data such as the one we generate here, listwise deletion does not produce biased results.

TABLE 4.
Bias/relative bias under MCAR situations. All values are in percent scale.

Mediation			Listwise	Pairwise	Without AVs		With AVs	
					MI	TS-ML	MI	TS-ML
None	10 %	50	−0.205	−0.264	−0.408	−0.267	−0.297	−0.169
		100	−0.252	−0.249	−0.23	−0.27	−0.163	−0.186
		200	−0.141	−0.154	−0.214	−0.149	−0.221	−0.17
	20 %	50	−0.159	−0.071	−0.48	−0.172	0.019	−0.14
		100	−0.015	−0.041	−0.072	0.014	−0.381	−0.042
		200	−0.14	−0.148	−0.124	−0.128	−0.234	−0.116
	40 %	50	−0.588	−0.876	−0.421	−0.379	−1.464	−0.301
		100	−0.264	−0.273	−0.319	−0.176	−0.397	−0.151
		200	−0.159	−0.108	−0.437	−0.127	−0.065	−0.119
Medium	10 %	50	−1.253	0.766	−6.005	−0.408	0.095	−0.521
		100	−1.445	−1.636	−2.707	−1.824	−2.152	−1.583
		200	−0.644	−0.298	−1.647	−0.586	−1.544	−0.484
	20 %	50	0.422	1.875	−2.216	0.624	2.322	0.898
		100	−2.949	−1.04	−3.126	−1.621	−6.8	−1.815
		200	−1.005	−0.789	−0.639	−0.458	−0.847	−0.484
	40 %	50	−4.272	0.302	−2.026	−0.513	−4.51	−2.192
		100	−3.896	−2.071	−3.766	−2.142	−3.171	−1.295
		200	−1.19	0.415	−4.561	−0.833	−1.974	−1.066
Large	10 %	50	0.246	1.5	−1.52	0.393	0.935	0.868
		100	−0.358	−0.102	−0.299	−0.297	−0.419	−0.354
		200	−0.024	0.047	−4.734	−0.088	−4.991	−0.064
	20 %	50	−1.525	0.268	−2.111	−0.281	−2.99	−0.276
		100	0.263	1.745	0.383	0.174	−0.31	−0.108
		200	−0.083	0.384	0.469	0.087	−0.198	0.056
	40 %	50	−3.407	5.173	−4.508	−0.335	−4.805	−0.871
		100	−0.94	1.765	1.255	0.56	−0.506	0.969
		200	−0.294	0.347	−0.357	−0.256	−0.394	−0.312

Note. When there is no mediation effect, the bias is calculated by the difference between the mean parameter estimates and the population value. For medium and large mediation effects, the bias is the relative bias. AVs: auxiliary variables.

overall, listwise deletion and pairwise deletion seemed to be more conservative while MI and TS-ML appeared to be more liberal in rejecting the null hypothesis of no mediation effects.

Second, several results can be observed for medium and large mediation effects. (1) For listwise deletion, parameter estimation bias was generally larger than for MI and TS-ML but much smaller than for pairwise deletion. Especially when the sample size and/or effect size were larger, the estimation bias became smaller or close to 5 %, which is consistent with our expected results in Table 3. Coverage probabilities were also close to the nominal level of 95 %. However, the statistical power of listwise deletion was much smaller than that of the other three methods. For example, with $N = 200$, large mediation effect, and 40 % missing data, the power for listwise deletion was only 5.9 % while the power was 77 %, 91.3 %, and 93.8 % for pairwise deletion, MI and TS-ML without auxiliary variables, respectively.

(2) There was a large amount of bias in the parameter estimates for pairwise deletion. For example, with a sample size of 100 and 10 % missing data, the pairwise deletion method already

TABLE 5.
Coverage probabilities under MCAR situations. All values are in percent scale.

Mediation			Listwise	Pairwise	Without AVs		With AVs	
					MI	TS-ML	MI	TS-ML
None	10 %	50	97.6	97.6	92.9	92.8	93.6	93.5
		100	96.2	96.1	91.9	91.7	91.9	91.4
		200	96.1	96.1	92.7	92	93.8	93.1
	20 %	50	97.9	98.6	92.6	93.5	93.5	93.6
		100	96.9	97	91.6	91.4	91.3	91.2
		200	96.2	96.1	91.7	92.7	92	92.5
	40 %	50	99.1	98.9	94.9	95.1	94.8	95.4
		100	98.6	97.6	90.3	92.1	93.6	93.6
		200	97.4	96.8	90.5	91.9	91.4	91.1
Medium	10 %	50	89	89	96.2	93.6	94.6	93.7
		100	90.6	91.1	96.5	94.9	95	95.1
		200	93.2	93.8	94.7	95.4	95.1	95.1
	20 %	50	89.2	88.5	94.4	94.9	95	94.9
		100	91	90.6	94.2	94.3	93.7	95.1
		200	92.6	92	94.2	95.3	95.6	95.6
	40 %	50	92.2	88.1	93.8	93.6	93.9	94.9
		100	90.4	88.9	95	94.8	94.2	94.7
		200	91.3	92.3	96	96	95.5	96.2
Large	10 %	50	90.6	90.2	93.1	93.1	94.7	93.9
		100	92.8	92.2	94	94.2	94.1	94.2
		200	93.9	93.4	96.6	94.4	95.6	95
	20 %	50	91.7	90	95.3	94.9	95.1	95.1
		100	93	92.8	94.5	95	94	94
		200	93.9	93.8	94.5	95.6	94.5	95
	40 %	50	90	85.8	95.8	95.6	95	96.3
		100	90	89.7	94.2	94.1	94.9	94.9
		200	92.3	90.1	94.4	94.4	94.1	93.5

Note. AVs: auxiliary variables.

underestimated the mediation effect about 30 %. The estimation bias further increased to about 60 % with 40 % missing data. Furthermore, the bias did not seem to decrease with the increase in sample size. Coverage probabilities for pairwise deletion were also underestimated, and the underestimation became even worse with larger sample size and effect size. Finally, statistical power of pairwise deletion was much smaller than that of MI and TS-ML.

(3) MI and TS-ML still performed well for MAR data. Bias in the mediation effect estimates from MI and TS-ML, regardless of the inclusion of auxiliary variables under different missing data proportion conditions, was smaller than or close to 5 % for MAR data. The coverage probabilities were close to the nominal level 0.95 ranging from 92.5 % to 98.1 % and were closer to the nominal value when sample size increased. Statistical power for the two methods was comparable. After including the auxiliary variables, power was boosted without influencing mediation effect estimates and coverage probabilities. Furthermore, with more missing data, the increase in power was larger. By comparing the results from MCAR and MAR, it is evident that power for MAR was smaller than that for the corresponding MCAR with the same amount of missing data.

TABLE 6.
Power/Type I error under MCAR situations. All values are in percent scale.

Mediation			Listwise	Pairwise	Without AVs		With AVs	
					MI	TS-ML	MI	TS-ML
None	10 %	50	2.4	2.4	7.1	7	6.4	6.3
		100	3.8	3.9	8.1	8.2	8.1	8.5
		200	3.9	3.9	7.3	7.9	6.2	6.8
	20 %	50	2.1	1.4	7.4	6.5	6.5	6.3
		100	3.1	3	8.4	8.5	8.7	8.7
		200	3.8	3.9	8.3	7.2	8	7.3
	40 %	50	0.9	1.1	5.1	4.9	5.2	4.5
		100	1.4	2.4	9.7	7.8	6.4	6
		200	2.6	3.2	9.5	8.1	8.6	8.7
Medium	10 %	50	29.7	29.4	46.5	53.6	54.8	55.3
		100	74.5	75.2	87.6	88.8	89.3	89.3
		200	99.1	99.1	99.6	99.9	99.8	99.9
	20 %	50	19.4	20.8	40.3	44.7	46.8	45.8
		100	56.1	61.8	78.7	80.5	76.8	83.8
		200	95.7	97	99.5	99	99.5	99.4
	40 %	50	5.3	5.8	20.7	20.9	14.4	21.4
		100	21.8	27.8	50.7	54.7	58.4	60.3
		200	67.3	77.1	85.1	89.8	93.1	94.2
Large	10 %	50	77.6	77	90.5	91.8	92.7	92.9
		100	98.7	98.5	99.6	99.6	99.9	99.7
		200	100	100	100	100	100	100
	20 %	50	57.1	58.3	78.8	82.1	84	83.9
		100	96.6	95.6	98.1	98.8	98.5	99
		200	100	100	100	100	100	100
	40 %	50	19.6	16.5	46.5	54.9	44.8	58.4
		100	67.8	67.4	92.1	90	93.2	94.1
		200	99.1	97.3	100	99.5	100	100

Note. The results are Type I error rates for the no mediation condition and power for the medium and large mediation conditions. AVs: auxiliary variables.

Overall, for MAR data, listwise deletion had smaller power and pairwise deletion had larger parameter estimation bias. MI and TS-ML still performed well regardless of auxiliary variables. Thus, in analyzing MAR data, MI and TS-ML outperformed listwise deletion and pairwise deletion.

5.5. *Finite Sample Simulation Study 3: Analysis of Auxiliary Variables Dependent MNAR (AV-MNAR) Data*

In this simulation, we compare the performance of listwise deletion, pairwise deletion, MI and TS-ML in analyzing AV-MNAR data. Note that by including the auxiliary variables in the analysis, the overall missing mechanism became MAR. The results from these methods are summarized in Tables 10, 11, and 12.

First, when there was no mediation effect ($a = b = 0$), the parameters were still estimated very well under the current AV-MNAR simulation design. Overall, listwise deletion and pairwise

TABLE 7.
Bias/relative bias under MAR situations. All values are in percent scale.

Mediation			Listwise	Pairwise	Without AVs		With AVs	
					MI	TS-ML	MI	TS-ML
None	10 %	50	−0.301	−0.664	0.079	0.155	0.16	0.095
		100	−0.232	−0.553	−0.053	−0.062	−0.058	−0.072
		200	−0.119	−0.369	−0.177	−0.136	−0.171	−0.147
	20 %	50	−0.415	−0.781	−0.193	0.061	0.352	0.045
		100	−0.296	−0.868	−0.206	−0.045	−0.301	−0.075
		200	−0.181	−0.51	−0.138	−0.132	−0.314	−0.161
	40 %	50	−0.788	−0.994	−0.397	−0.619	−0.889	0.285
		100	−0.355	−0.764	−1.338	−0.153	−0.339	−0.079
		200	−0.386	−0.546	−0.834	−0.281	0.074	−0.34
Medium	10 %	50	−5.876	−28.364	3.252	1.996	3.396	2.189
		100	−4.574	−31.909	−1.801	−0.916	1.203	−0.736
		200	−0.542	−31.782	−1.242	−0.792	−3.664	−0.968
	20 %	50	−15.663	−41.884	3.829	2.163	4.7	1.361
		100	−13.33	−45.039	−1.448	−0.916	−1.698	−0.972
		200	−1.771	−45.189	1.683	−0.815	0.864	−1.268
	40 %	50	−51.794	−57.941	−4.956	−4.245	1.809	0.769
		100	−21.738	−59.97	−2.878	−0.248	12.313	0.924
		200	−6.608	−61.194	−1.269	−2.499	−0.497	−2.498
Large	10 %	50	4.072	−33.402	0.42	0.337	0.436	−0.066
		100	−3.573	−34.516	−0.707	−0.225	−2.919	−0.339
		200	−0.26	−34.928	−2.132	−0.05	−3.365	−0.121
	20 %	50	6.158	−46.806	2.328	1.49	3.225	0.968
		100	−5.247	−47.766	−0.243	−0.12	−0.701	−0.041
		200	−1.378	−48.542	−0.408	−0.414	−0.495	−0.506
	40 %	50	10.195	−62.553	−5.288	−1.754	−2.159	−3.998
		100	−8.108	−62.348	−1.611	−0.587	−0.208	−0.728
		200	−5.1	−63.226	−1.221	−0.051	1.784	0.399

Note. The same as Table 4.

deletion again seemed to be more conservative whereas MI and TS-ML appeared to be more liberal in rejecting the null hypothesis of no mediation effects.

Second, with medium and large mediation effects, mediation effect estimates from all methods had large bias without the inclusion of the auxiliary variables, although the bias for MI (ranging from 9.8 % to 32 %) and TS-ML (ranging from 9.4 % to 23.1 %) was generally slightly smaller than that of listwise (ranging from 12.5 % to 27.4 %) and pairwise deletion (ranging from 15.9 % to 32.2 %). Moreover, coverage probabilities were underestimated, especially for listwise (ranging from 72.9 % to 89.5 %) and pairwise deletion (ranging from 62.3 % to 85.2 %). These results were not surprising because all methods employed here assume that the missing mechanisms were ignorable.

With the inclusion of the auxiliary variables, MI and TS-ML estimated mediation effect parameters well with bias smaller than or close to 5 % except when the missing rate was as high as 40 % and the sample size was as small as 50. The coverage probabilities were also close to the nominal level of 0.95. This was expected because by including the auxiliary variables that could

TABLE 8.
Coverage probabilities under MAR situations. All values are in percent scale.

Mediation			Listwise	Pairwise	Without AVs		With AVs	
					MI	TS-ML	MI	TS-ML
None	10 %	50	100	98.1	92.4	93.1	93.1	93
		100	98.1	96.7	91.5	91.1	92.7	91.8
		200	96.7	95	90.9	91.6	91.5	91.8
	20 %	50	100	98	94.2	94.9	94	95
		100	98.5	98.2	91.1	92	92.2	93
		200	98.9	95.9	90.4	91.1	91.6	92.3
	40 %	50	100	99	98.2	98	96.5	98.4
		100	100	97.3	90.9	94	92.4	94.4
		200	99.8	96.8	86.2	92.5	90.7	92
Medium	10 %	50	91.4	79.4	93.9	93.9	93.9	94.4
		100	91.7	72.4	94.9	95	95	95.6
		200	93.5	65.4	96.3	96.4	95.9	95.7
	20 %	50	94.5	71.6	93.1	92.5	93.7	93.7
		100	93.9	57.3	95.6	95	95	95.8
		200	94.5	40.4	96.2	96.2	96.3	96.1
	40 %	50	94	65.4	93.2	93.2	93	98.1
		100	98.6	43.2	92.5	93.8	93.5	95
		200	96.4	27.5	94.9	95.9	94.9	96.2
Large	10 %	50	91.9	74.2	95.1	93.8	94.7	93.7
		100	93.1	59.7	94.3	94.7	96	94.6
		200	94.2	36.4	93.7	93.9	94.5	94.3
	20 %	50	92.7	54.6	94.4	95.2	95.5	95.6
		100	94.4	33.6	95.4	95.8	94.1	95.8
		200	95.5	9.8	94.1	94.4	94	94.3
	40 %	50	94	39.7	94.6	94.4	94.7	94.8
		100	97.1	21.4	95.2	95.6	96.1	95.7
		200	96.7	7.1	93.2	95.7	94.9	95.6

Note. The same as Table 5.

explain the missingness of mediation model data, the overall data became MAR. Furthermore, power for MI and TS-ML was comparable under the studied conditions.

5.6. *Finite Sample Simulation Study 4: Analysis of Model Variables Dependent MNAR (MV-MNAR) Data with Auxiliary Variables*

In this simulation, we compare the performance of listwise deletion, pairwise deletion, MI and TS-ML in analyzing MV-MNAR data for mediation effects. The results from these methods are provided in Tables 13, 14, and 15.

Overall, the parameter estimation bias from the four methods was still small when there was no mediation effects. However, when mediation effects existed, none of the four methods performed well. Notably, both the parameters and coverage probabilities were underestimated under all studied conditions. For both MI and TS-ML, the inclusion of auxiliary variables improved both parameter estimates and coverage probabilities. However, large bias in parameter estimation and coverage probabilities still prevented meaningful inference from MI and TS-ML even after the inclusion of auxiliary variables for MV-MNAR data.

TABLE 9.
Power/Type I error under MAR situations. All values are in percent scale.

Mediation			Listwise	Pairwise	Without AVs		With AVs	
					MI	TS-ML	MI	TS-ML
None	10 %	50	0	1.9	7.6	6.6	6.9	6.7
		100	1.9	3.3	8.5	8.8	7.3	8.1
		200	3.3	5	9.1	8.4	8.5	8.2
	20 %	50	0	2	5.8	5.1	6	4.8
		100	1.5	1.8	8.9	7.9	7.8	7
		200	1.1	4.1	9.6	8.7	8.4	7.4
	40 %	50	0	1	1.8	2	3.5	1.6
		100	0	2.7	9.1	6	7.6	5.5
		200	0.2	3.2	13.8	7.4	9.3	8
Medium	10 %	50	14.6	25.8	50.1	48.2	51.7	51.4
		100	42.6	64.5	82.5	83.1	85.5	85.4
		200	86.1	98.1	99.8	99.6	99.3	99.8
	20 %	50	5.3	14.1	36.2	33.6	41.4	36.3
		100	14.2	43.3	68.2	70.3	74.3	74.9
		200	41.8	89	97.4	95.9	98.9	98.8
	40 %	50	0.1	3.1	5.7	7.5	7.5	13
		100	0.9	6.1	24.5	24	32.3	31.5
		200	3.1	28.3	62.3	57.2	74.5	70.1
Large	10 %	50	50.2	67.2	86.3	87.2	88.7	87.9
		100	88.5	96.8	99.2	99.4	100	99.5
		200	99.5	100	100	100	100	100
	20 %	50	18.9	44.3	75.7	75	80.1	78.1
		100	47	88.1	97	97.6	97.2	98.5
		200	80.5	99.9	100	99.9	100	100
	40 %	50	1	17.2	16.1	19.1	22.3	21.2
		100	2.5	24.2	63.1	63.8	71.4	70.8
		200	5.9	77	91.3	93.8	98.2	97.7

Note. The same as Table 6.

5.7. *Summary of Results*

To summarize, when there was no mediation effect, all methods, including listwise deletion, pairwise deletion, MI and TS-ML, had little bias in their mediation parameter estimates, regardless of the missing mechanisms under the studied conditions. Overall, listwise deletion and pairwise deletion seemed to be more conservative with Type I error rates smaller than 5 %, whereas MI and TS-ML appeared to be more liberal with Type I error rates larger than 5 % in rejecting the null hypothesis of no mediation effects. The remaining summary is based on the conditions where mediation effects existed.

For MCAR, all methods, including listwise deletion, pairwise deletion, MI and TS-ML, could recover mediation effect parameters very well under all studied conditions, even with a small sample size and a large proportion of missing data. Although listwise and pairwise deletion had slightly underestimated coverage probabilities, both MI and TS-ML obtained good coverage probabilities without and with auxiliary variables. Furthermore, listwise deletion had smaller power than pairwise deletion, which in turn had smaller power than MI and TS-ML. For MAR data, neither listwise deletion nor pairwise deletion should be applied because of either low power

TABLE 10.
Bias/relative bias under AV-MNAR situations. All values are in percent scale.

Mediation			Listwise	Pairwise	Without AVs MI	Without AVs TS-ML	With AVs MI	With AVs TS-ML
None	10 %	50	0.097	0.057	−0.045	0.086	0.051	0.001
		100	−0.066	−0.091	−0.052	−0.083	−0.087	−0.12
		200	−0.183	−0.197	−0.124	−0.191	−0.225	−0.21
	20 %	50	0.127	0.039	−0.25	0.05	0.095	−0.088
		100	−0.032	−0.038	−0.202	−0.068	−0.485	−0.068
		200	−0.201	−0.212	−0.236	−0.226	−0.389	−0.223
	40 %	50	0.307	−0.188	0.07	0.274	−1.204	0.06
		100	0.086	−0.12	−0.188	0.029	−0.035	−0.044
		200	−0.094	−0.143	−0.411	−0.103	0.059	−0.141
Medium	10 %	50	−14.346	−16.001	−12.745	−11.756	−1.597	−1.824
		100	−14.378	−16.151	−12.393	−12.157	−2.492	−2.095
		200	−14.001	−15.945	−12.57	−11.824	−2.063	−1.141
	20 %	50	−20.252	−21.767	−19.593	−17.446	−2.889	−3.489
		100	−19.829	−22.274	−18.536	−17.169	−5.941	−2.745
		200	−19.697	−21.923	−15.857	−16.898	−2.39	−1.785
	40 %	50	−26.311	−29.148	−24.482	−21.958	−13.973	−4.012
		100	−27.376	−30.318	−23.673	−23.089	−4.694	−3.654
		200	−26.397	−28.928	−32.026	−22.587	−3.455	−2.013
Large	10 %	50	−12.47	−16.122	−10.262	−9.434	−0.094	−0.209
		100	−12.564	−16.495	−9.837	−9.57	−0.84	−0.398
		200	−12.787	−16.956	−12.164	−9.74	−0.92	−0.349
	20 %	50	−17.243	−21.82	−14.882	−13.542	0.114	−0.167
		100	−18.567	−23.343	−15.339	−14.523	−3.403	−0.725
		200	−18.817	−24.162	−15.516	−14.673	−1.201	−0.405
	40 %	50	−25.457	−29.972	−21.067	−18.883	−9.428	−1.813
		100	−24.669	−30.481	−19.76	−19.665	−1.715	−0.902
		200	−25.613	−32.204	−21.007	−19.996	−1.493	−0.416

Note. The same as Table 4.

or large parameter estimation bias. MI and TS-ML still recovered mediation effect parameters well and their coverage probabilities were close to the nominal level of 0.95. Statistical power was comparable for MI and TS-ML. Furthermore, the inclusion of auxiliary variables did not affect estimation bias and coverage probabilities but boosted the power for detecting mediation effects. For analyzing AV-MNAR data, MI and TS-ML performed equally well through the inclusion of auxiliary variables although neither of them could recover mediation effect parameters without the use of appropriate auxiliary variables. For MV-MNAR, neither MI nor TS-ML could fully correct the bias in the mediation effect estimates or obtain accurate coverage probabilities through the inclusion of auxiliary variables. However, it seemed that the inclusion of auxiliary variables could improve the mediation effect and coverage probability estimates when the auxiliary variables were correlated with model variables.

TABLE 11.
Coverage probabilities under AV-MNAR situations. All values are in percent scale.

Mediation			Listwise	Pairwise	Without AVs		With AVs	
					MI	TS-ML	MI	TS-ML
None	10 %	50	98.4	98.2	93.1	92.5	93.3	93.2
		100	96.4	96	91.9	91.9	91.6	91.4
		200	96	95.9	92.1	92.8	93	92.1
	20 %	50	98.7	98.4	94.5	93.7	93.6	93.4
		100	96.2	96.1	91.4	91.6	92.5	90.5
		200	96.5	96.2	91.4	92.2	91.4	91.6
	40 %	50	98.5	98.8	93.4	94.5	92.5	93.8
		100	98.3	97.8	91.7	93.3	91.7	92.2
		200	98	97.1	92.1	93.1	91	91.5
Medium	10 %	50	85.3	84	93.5	93	94.4	93.6
		100	86.8	85.2	93.2	93.1	94.7	94.7
		200	87.1	84.9	93.1	92.2	94.8	95.1
	20 %	50	83.8	81.9	91.6	92.4	94.6	94.1
		100	84.5	82.5	92	93.4	94.4	94.9
		200	81.7	81.1	93.4	92.6	94.9	95.6
	40 %	50	89.5	82.2	91.5	91.9	92	93.2
		100	83.3	77.4	93	92.9	94.9	95.4
		200	81	75.9	85.7	90.8	95.5	95.7
Large	10 %	50	87	85	93.4	93.2	93.8	93.8
		100	87.3	83.1	93.6	94	94	94.1
		200	84	77.4	84.8	90.5	94.5	94.9
	20 %	50	84.6	79.3	92.3	92.3	93.5	93.9
		100	83.7	75.6	90.5	91.5	95.2	94.3
		200	77.9	67.7	89.1	88.3	94.2	94.4
	40 %	50	80.8	72.3	91.7	93.1	91.7	95.5
		100	78.8	69.7	90.7	90.2	94.2	95.3
		200	72.9	62.3	83.4	85.5	95.4	94.9

Note. The same as Table 5.

6. Discussion

In this study, we introduced and compared four methods for analyzing missing data in mediation analysis, including listwise deletion, pairwise deletion, MI and TS-ML. After outlining the procedure of each method, we first demonstrated their applications through the analysis of two real data sets and then compared the performance of these methods through simulation studies under MCAR, MAR, and MNAR missing mechanisms without and with auxiliary variables. Factors considered in the simulation studies include sample sizes, effect sizes, and amount of missing data.

Based on both expected (asymptotic) results from population or infinite samples and simulation results from finite samples, it is shown that the four methods, including listwise deletion and pairwise deletion, produced *ab* estimates without bias and with reasonable coverage probabilities across all studied missing mechanism conditions when there was no mediation effect. However, when mediation existed, the four methods produced different results. All four methods can obtain similar parameter estimates under MCAR, but listwise and pairwise deletion had less power. MI

TABLE 12.
Power/Type I error under AV-MNAR situations. All values are in percent scale.

Mediation			Listwise	Pairwise	Without AVs		With AVs	
					MI	TS-ML	MI	TS-ML
None	10 %	50	1.6	1.8	6.9	7.5	6.7	6.7
		100	3.6	4	8.1	8.1	8.4	8.4
		200	4	4.1	7.9	7	7	7.8
	20 %	50	1.3	1.6	5.5	6.2	6.4	6.4
		100	3.8	3.9	8.6	8.4	7.5	9.4
		200	3.5	3.8	8.6	7.8	8.6	8.3
	40 %	50	1.5	1.2	6.6	5.4	7.5	6.3
		100	1.7	2.2	8.3	6.7	8.3	7.7
		200	2	2.9	7.9	6.9	9	8.5
Medium	10 %	50	23.4	24.2	47.1	48.8	55.5	55.4
		100	66.8	68	85.9	85.5	89.3	89.9
		200	98.2	98.9	99.7	99.5	99.8	99.9
	20 %	50	13.8	15.3	34.1	35.2	46.3	43.3
		100	47	49.8	72.6	76.6	78.9	83.2
		200	92	92.6	97.6	97.1	98.7	99.4
	40 %	50	3.9	3.8	18.2	18.5	18.4	25.1
		100	16.2	18.5	48.5	46.4	62.4	59.5
		200	54	61.5	74.6	82.9	93.6	92.9
Large	10 %	50	69.5	68	87.1	89	92.9	92.6
		100	98.7	98	99.7	99.7	99.8	99.7
		200	100	100	100	100	100	100
	20 %	50	49.4	44.5	75.1	78.7	86.5	84.8
		100	93.2	90.7	98.3	98.6	99.6	99.2
		200	100	100	100	100	100	100
	40 %	50	16.2	12.2	46.5	49.9	52.8	61.1
		100	56.5	50.1	86.8	88	94.4	93.8
		200	96.2	92.3	99.1	99.5	99.9	99.8

Note. The same as Table 6.

and TS-ML could recover mediation model parameters almost equally well and had comparable power in detecting mediation effects for MCAR, MAR, and AV-MNAR with the inclusion of auxiliary variables. The inclusion of auxiliary variables did not affect parameter estimation bias and coverage probabilities but boosted the power of detecting mediation effects for MCAR and MAR data. This is because adding auxiliary variables that are correlated with the mediation variables can reduce noise in the mediation variables with missing data. However, neither method can deal with AV-MNAR data without the inclusion of auxiliary variables. Furthermore, for MV-MNAR data analysis, the inclusion of auxiliary variables can improve parameter estimation, although the parameter estimation bias did not seem to be fully corrected.

Overall, when mediation effects existed, the four methods may obtain comparable results for MCAR and MV-MNAR, although the results might not be correct for MV-MNAR; and the results were different across some methods for MAR and AV-MNAR. In Example 2, the results from all four methods were similar to each other. Therefore, it can be viewed as an indication that the missing data in Example 2 were more likely to be either MCAR or MV-MNAR than MAR and AV-MNAR. In Example 1, the parameter estimates from different methods were quite

TABLE 13.
Bias/relative bias under MV-MNAR situations. All values are in percent scale.

Mediation			Listwise	Pairwise	Without AVs		With AVs	
					MI	TS-ML	MI	TS-ML
None	10 %	50	0.084	0.215	−0.061	0.072	0.05	0.042
		100	0.123	0.195	0.104	0.105	0.066	0.077
		200	−0.018	0.065	−0.09	−0.024	−0.09	−0.043
	20 %	50	0.338	0.466	0.106	0.383	0.345	0.359
		100	0.241	0.306	0.082	0.265	−0.063	0.237
		200	0.035	0.123	−0.001	0.035	−0.086	0.024
	40 %	50	0.61	0.58	0.35	1.328	−0.257	1.325
		100	0.373	0.412	0.223	0.693	0.442	0.72
		200	0.183	0.297	0.034	0.363	0.264	0.372
Medium	10 %	50	−41.645	−44.547	−41.762	−40.166	−30.343	−31.768
		100	−42.667	−46.162	−41.252	−41.322	−33.54	−33.164
		200	−42.9	−46.232	−41.683	−41.592	−38.689	−33.348
	20 %	50	−55.962	−59.896	−57.281	−56.001	−43.726	−46.189
		100	−57.873	−61.73	−58.471	−57.863	−50.396	−48.575
		200	−58.022	−61.666	−57.254	−58.385	−49.072	−49.235
	40 %	50	−72.653	−76.659	−73.057	−70.11	−69.016	−62.179
		100	−73.475	−76.185	−73.871	−72.611	−65.7	−64.787
		200	−73.26	−75.905	−74.743	−74.07	−66.746	−66.257
Large	10 %	50	−36.181	−43.34	−34.192	−33.341	−26.032	−26.662
		100	−36.389	−44.179	−34.244	−34.201	−27.651	−27.348
		200	−36.494	−44.437	−34.405	−34.57	−27.891	−27.572
	20 %	50	−50.882	−58.137	−50.323	−49.276	−39.931	−41.1
		100	−50.399	−58.498	−50.063	−49.792	−42.367	−41.204
		200	−51.067	−59.141	−49.633	−50.756	−41.789	−42.198
	40 %	50	−68.439	−73.577	−69.266	−66.677	−64.484	−59.067
		100	−67.495	−73.885	−68.446	−68.224	−60.059	−60.149
		200	−67.799	−74.085	−69.561	−69.16	−61.215	−60.924

Note. The same as Table 4.

different, which suggests that the missing data were not likely to be MCAR. Furthermore, in simulation Studies 3 and 4, for AV-MNAR and MV-MNAR, analyses showed that when auxiliary variables were related to missingness, the results from data analysis with and without the inclusion of auxiliary variables would be different. Given the fact that the conclusions in Example 1 changed after the inclusion of the auxiliary variable, the missing mechanism in Example 1 was more likely to be AV-MNAR or MV-MNAR than MCAR or MAR. However, the conclusion drawn here should be used with caution and needs to be verified with extra information. To be more confident in understanding missingness mechanisms, one can (1) borrow information from previous research findings to assist the decision making. For example, some previous research may have found that some variables were related to missingness of the outcome variable and thus those variables can be included as auxiliary variables; (2) conduct some rigorous statistical tests (e.g., Yuan, 2009) to test whether the auxiliary variables influence the modeling; (3) conduct more sensitivity analyses with simulations. For our Example 1, reading cognition has been found to be related to mathematical ability in previous research (e.g., Grimm, 2008). Therefore, we are

TABLE 14.
Coverage probabilities under MV-MNAR situations. All values are in percent scale.

Mediation			Listwise	Pairwise	Without AVs		With AVs	
					MI	TS-ML	MI	TS-ML
None	10 %	50	98.5	98.5	94.8	95.3	96	94.7
		100	97.6	97.1	93.1	92.8	92.8	93.1
		200	96.4	96.7	92.7	93.8	93.5	92.8
	20 %	50	99.4	99.1	95.2	96	94.7	95.5
		100	97.8	97.5	93.5	93.6	92.2	92.8
		200	97.2	96.9	91.2	91.8	91.4	92
	40 %	50	99.3	99.5	97.2	96.7	95.6	96.9
		100	99.6	99.3	95.5	96.4	95	95.1
		200	98	98.3	91.9	93.1	92.3	91.8
Medium	10 %	50	70.8	67.2	85.7	85.2	89.9	88.1
		100	61.4	54.5	76.2	75.9	82.8	82
		200	45.9	36.3	57.9	58.9	65.2	71.4
	20 %	50	61.5	56.1	75.3	77.1	83.3	81.5
		100	46.1	36.1	58.3	59.5	64.9	69.8
		200	25.6	15.9	34.9	28.7	47.5	45.4
	40 %	50	63.7	47.5	68.7	68.4	68.2	74.8
		100	35.9	24.9	45.4	45	55.7	54.8
		200	16.9	19.6	47.5	50.3	67	66.1
Large	10 %	50	68.1	57.2	79.4	82.3	87.7	86
		100	54	38.9	68	66.1	76.6	76.9
		200	32.3	14.2	43.8	41.9	57.5	59.3
	20 %	50	54	37.9	67.6	65.9	80.4	76.2
		100	34.3	18.9	42.2	41.3	54.5	57
		200	12.5	3.1	17.6	12.9	30.5	27.2
	40 %	50	43.1	27.7	52.3	52.6	60.5	63.2
		100	21.6	8.8	26.6	21.7	40.5	34
		200	5.9	1.6	3.2	3	7.2	6.9

Note. The same as Table 5.

more confident that reading cognition is a valid auxiliary variable to explain data missingness in mathematical ability.

Listwise deletion and pairwise deletion are well known not to be efficient and/or effective in dealing with missing data. However, these methods are still very widely used in empirical research because of their availability in popular software (e.g., Jelicic, Phelps, & Lerner, 2009). Our simulation studies suggested that listwise deletion and pairwise deletion should not be used in handling missing data for mediation analysis, especially in the cases of MAR and MNAR data. These two methods may lead to either biased parameter estimates and/or lower statistical power in mediation analysis. Therefore, we discourage the use of listwise and pairwise deletion. However, as illustrated in the preceding paragraph, comparing results from listwise and pairwise deletion to results from MI and TS-ML may help us better understand the missing mechanism and thus can be used as interim results for the sensitivity study purpose. Furthermore, our simulation results indicated that listwise deletion performed better than pairwise deletion. This is partially because of the data generation method in our simulation design. Generally speaking, pairwise deletion may not always outperform listwise deletion because the results from pairwise

TABLE 15.
Power/Type I error under MV-MNAR situations. All values are in percent scale.

Mediation			Listwise	Pairwise	Without AVs		With AVs	
					MI	TS-ML	MI	TS-ML
None	10 %	50	1.5	1.5	5.2	4.6	4	5.2
		100	2.4	2.9	6.9	7.1	7.2	6.9
		200	3.6	3.3	7.3	6.1	6.5	7.1
	20 %	50	0.6	0.9	4.8	4	5.3	4.4
		100	2.2	2.5	6.5	6.4	7.8	7.1
		200	2.8	3.1	8.8	8	8.6	8
	40 %	50	0.7	0.5	2.8	3.1	4.4	2.9
		100	0.4	0.7	4.5	3.6	5	4.9
		200	2	1.7	8.1	6.6	7.7	7.8
Medium	10 %	50	14.6	14.7	30.7	34	39.4	41.3
		100	46.5	46.6	71.4	71.9	79.7	80.7
		200	90	90.6	96.1	96.9	95.7	98.5
	20 %	50	7.1	7.5	19.2	21.6	33	28.9
		100	20.7	19.9	47.4	48.5	56.6	60.7
		200	65	63.7	86.6	84.7	94.1	91.6
	40 %	50	1.3	0.7	4.8	7	7	10.5
		100	4.1	4.4	20	19.7	31.1	30.3
		200	16.9	19.6	47.5	50.3	67	66.1
Large	10 %	50	55.2	51.3	80	80.1	84.6	84.4
		100	94	92	97.9	98.2	99.1	99.3
		200	100	100	100	100	100	100
	20 %	50	28	27.5	55.5	60.1	71.4	70.2
		100	76.2	73.2	91.9	93.2	95.5	96.4
		200	99.4	98.4	99.8	99.5	100	100
	40 %	50	6.6	5.3	21.8	26.1	23.7	34.4
		100	24.4	23.5	62.2	61	67.9	74.4
		200	73.1	68.4	89.9	90.7	95.6	95.8

Note. The same as Table 6.

deletion may not be stable as a result of the way it constructs the covariance matrix (Azen & Van Guilder, 1981; Little & Rubin, 2002).

The performances of MI and TS-ML were comparable under all studied situations. The basic idea of MI is intuitive and straightforward—first filling in missing data using plausible values and then analyzing the imputed data as complete data. One practical issue related to MI is to determine how many imputations are needed. Based on our experience, we found that different numbers of imputations are needed for different proportions of missing data. For example, 40 imputations seemed to be sufficient for a mediation analysis with 10 % missing data whereas 80 imputations were needed with 40 % missing data. Both MI and TS-ML can be applied effortlessly to include auxiliary variables. Rubin (1996) suggested that one should include as many variables as one can when conducting multiple imputation. In addition, TS-ML performs as well as (sometimes better than) MI but consumes significantly less computation time as shown in real data analysis examples.

We have implemented all four methods (with or without auxiliary variables) discussed in this study in a free R package `bmem`. Intensive simulations have been conducted to ensure its robustness under a variety of situations. In addition to the binaries of the package for Windows, Mac, and Linux, the source codes are also available to researchers who are willing to modify or improve the package. As shown earlier for the empirical examples, by comparing results from different methods, more (although not conclusive) information regarding the missing mechanism can be obtained through sensitivity analysis. The package `bmem` makes it possible and easy to analyze the same set of data using all four methods discussed in this study to gain more insight on the missing data mechanism.

Several different perspectives related to mediation analysis with missing data can be evaluated in the future. First, in discussing MI and TS-ML, we have assumed that both mediation model variables and auxiliary variables are normally distributed. However, it is not rare that a study may involve non-normal data such as categorical data. Thus, the performance of the two methods for non-normal data can be investigated in the future. Second, in simulation Study 4, we found that the inclusion of auxiliary variables did not work very well for MV-MNAR data. In the literature, selection models have been proposed in dealing with MV-MNAR data for certain models (e.g., Best, Spiegelhalter, Thomas, & Brayne, 1996; Lu, Zhang, & Lubke, 2011). How to apply selection models to mediation analysis with missing data and how they perform should be investigated in the future. Third, the current mediation model only focuses on the cross-sectional data analysis. Some researchers have suggested that the time variable should be considered in mediation analysis and have developed longitudinal mediation models (e.g., Cole & Maxwell, 2003; MacKinnon, 2008). For longitudinal research, missing data could be a more serious problem than in cross-sectional research. Future research can investigate missing data techniques in longitudinal mediation analysis.

Appendix A. Expectation for the EM Algorithm

The E-step of the EM algorithm is to fill in the missing data using their expectations

$$E\left(z_{ij}|z_{obs}, U^{(t)}, S^{(t)}\right) = z_{ij}^{(t)}; \quad i = 1, \dots, N, \ j = 1, 2, \dots, p+3, \tag{A.1}$$

and

$$E\left(z_{ij}z_{ik}|z_{obs}, U^{(t)}, S^{(t)}\right) = z_{ij}^{(t)} z_{ik}^{(t)} + c_{ijk}^{(t)} \tag{A.2}$$

where

$$z_{ij}^{(t)} = \begin{cases} z_{ij} & \text{if } z_{ij} \text{ is observed} \\ E(z_{ij}|z_{obs}, U^{(t)}, S^{(t)}) & \text{if } z_{ij} \text{ is missing} \end{cases} \tag{A.3}$$

and

$$c_{ijk}^{(t)} = \begin{cases} \mathrm{Cov}(z_{ij}, z_{ik}|z_{obs}, U^{(t)}, S^{(t)}) & \text{if both } z_{ij} \text{ and } z_{ik} \text{ are missing} \\ 0 & \text{otherwise} \end{cases} \tag{A.4}$$

with $j, k = 1, 2, \dots, p+3$ and z_{obs} denoting the observed data. The expectation $E(z_{ij}|z_{obs}, U^{(t)}, S^{(t)})$ and covariance $\mathrm{Cov}(z_{ij}, z_{ik}|z_{obs}, U^{(t)}, S^{(t)})$ are readily available from the conditional distribution of the multivariate normal distribution with mean $U^{(t)}$ and covariance $S^{(t)}$.

Appendix B. R Code for Example 1

The R code lines below were used to obtain the results for Example 1. The statements following "#" are annotations for the R code lines. Line 2 loads our R library, and Line 7 reads data for the first example into R. Lines 9 to 15 specify the path model in Figure 2. The three-term phrase `start -> end, parname, st` denotes a single-headed path from the variable `start` to the variable `end`. The parameter on this path is represented by `parname`, and its starting value for estimation purpose is `st`. The starting value `st` can be set as `NA` to ask the program to choose a starting value. Similarly, `<->` represents a double-headed arrow denoting variance or covariance in the path diagram. For more information on how to specify a path model, see Fox (2006). Line 18 specifies the mediation effect or indirect effect to be estimated. More than one indirect effect can be given as shown in Appendix C for Example 2. On Line 21, the model parameters are estimated using the `bmem` function through the listwise deletion method by setting `method='list'`. By changing the `method` argument to `'pair'`, `'tsml'`, and `'mi'`, respectively, other missing data handling methods can be used to get parameter estimates. In the `bmem` function, the first argument is the data set to be used, `ex1`, in this example. The second argument is the model to be estimated. The third argument supplies the indirect effects. The fourth argument selects the variables to be used in the mediation model. This argument distinguishes the variables used in the mediation model from the auxiliary variables.

```
1  ## Load the library 'bmem'
2  library(bmem)
3  ## Read Example 1 data into R
4  ## In the data set ex1.txt, the first three variables
5  ## are mediation model variables and the fourth variable
6  ## is an auxiliary variable
7  ex1<-read.table('ex1.txt', head=T)
8  ## Specify the model for Example 1
9  ex1m<-specifyModel()
10  ME -> HE, a, NA
11  HE -> math, b, NA
12  ME -> math, cp, NA
13  ME <-> ME, s1, NA
14  HE <-> HE, s2, NA
15  math <-> math, s3, NA
16
17 ## Specify the indirect effect to be estimated
18 indirect<-c('a*b')
19
20 ## listwise deletion
21 bmem(ex1,ex1m,indirect,1:3,method='list')
22 ## Pairwise deletion
23 bmem(ex1,ex1m,indirect,1:3,method='pair')
24 ## TS-ML
25 bmem(ex1,ex1m,indirect,1:3,method='tsml')
26 ## MI
27 bmem(ex1,ex1m,indirect,1:3,method='mi',m=40)
```

Appendix C. R Code for Example 2

In this example, multiple mediation effects are specified on Line 27. For example, `a*b` is the indirect effect from age to EPT through HVLT and `d*h` is the indirect effect from age to EPT through R. Note that `a*b+d*h` is the total indirect effect from age to EPT.

```
ex2<-read.table('ex2.txt', header=T)
## In the data set ex2.txt,  the seven variables are all
## mediation model variables
ex2m<-specifyModel()
  age -> hvltt, a, NA
  age -> ept,   cp, NA
  age -> R,     d, NA
  edu -> hvltt, e, NA
  edu -> ept,   f, NA
  edu -> R,     g, NA
  hvltt -> ept, b, NA
  R   -> ept,   h, NA
  R   -> ws,   NA, 1
  R   -> ls,   l2, NA
  R   -> lt,   l3, NA
  age <-> age, s1, NA
  edu <-> edu, s2, NA
  age <-> edu, s12,NA
  hvltt <-> hvltt, s3, NA
  R   <-> R,   s4, NA
  hvltt <-> R, s34, NA
  ept <-> ept, s5, NA
  ws  <-> ws,  s6, NA
  ls  <-> ls,  s7, NA
  lt  <-> lt,  s8, NA

indirect<-c('a*b','d*h','e*b','g*h','a*b+d*h', 'e*b+g*h')

bmem(ex2,ex2m,indirect,1:7,'list','bc')
bmem(ex2,ex2m,indirect,1:7,'pair','bc')
bmem(ex2,ex2m,indirect,1:7,'mi','bc', m=40)
bmem(ex2,ex2m,indirect,1:7,'tsml','bc')
```

References

Azen, S., & Van Guilder, M. (1981). Conclusions regarding algorithms for handling incomplete data. In *Proceedings of the survey research methods section* (pp. 53–56).

Bauer, D.J., Preacher, K.J., & Gil, K.M. (2006). Conceptualizing and testing random indirect effects and moderated mediation in multilevel models: new procedures and recommendations. *Psychological Methods*, *11*(2), 142–163.

Bentler, P.M., & Weeks, D.G. (1980). Linear structural equations with latent variables. *Psychometrika*, *45*, 289–308.

Best, N.G., Spiegelhalter, D.J., Thomas, A., & Brayne, C.E. (1996). Bayesian analysis of realistically complex models. *Journal of the Royal Statistical Society. Series A*, *159*, 323–342.

Bollen, K.A., & Stine, R.A. (1990). Direct and indirect effects: classical and bootstrap estimates of variability. *Sociological Methodology*, *20*, 115–140.

Brandt, J. (1991). The Hopkins verbal learning test: development of a new memory test with six equivalent forms. *Clinical Neuropsychology*, *5*, 125–142.

Chen, Z.X., Aryee, S., & Lee, C. (2005). Test of a mediation model of perceived organizational support. *Journal of Vocational Behavior*, *66*(3), 457–470.

Center for Human Resource Research (2006). *NLSY79 child & young adult data users guide: a guide to the 1986–2004 child data* (Computer software manual). Columbus.

Cladwell, B.M., & Bradley, R.H. (1979). *Home observation for measurement of the environment*. Little Rock: University of Arkansas.

Cole, D.A., & Maxwell, S.E. (2003). Testing mediational models with longitudinal data: questions and tips in the use of structural equation modeling. *Journal of Abnormal Psychology*, *112*, 558–577.

Davis-Kean, P.E. (2005). The influence of parent education and family income on child achievement: the indirect role of parental expectations and the home environment. *Journal of Family Psychology*, *19*, 294–304.

Efron, B. (1979). Bootstrap methods: another look at the jackknife. *The Annals of Statistics*, *7*(1), 1–26.
Efron, B. (1987). Better bootstrap confidence intervals. *Journal of the American Statistical Association*, *82*(397), 171–185.
Efron, B. (1994). Missing data, imputation, and the bootstrap. *Journal of the American Statistical Association*, *89*(426), 463–478.
Efron, B., & Tibshirani, R. (1993). *An introduction to the bootstrap*. New York: CRC Press.
Ekstrom, R.B., French, J.W., Harman, H.H., & Derman, D. (1976). *Kit of factor-referenced cognitive tests*. Princeton: Educational Testing Service.
Enders, C.K. (2003). Using the expectation maximization algorithm to estimate coefficient alpha for scales with item-level missing data. *Psychological Methods*, *8*, 322–337.
Fox, J. (2006). Structural equation modeling with the sem package in *r*. *Structural Equation Modeling*, *13*, 465–486.
Gonda, J., & Schaie, K.W. (1985). *Schaie-Thurstone mental abilities test: word series test*. Palo Alto: Consulting Psychologists Press.
Grimm, K.J. (2008). Longitudinal associations between reading and mathematics. *Developmental Neuropsychology*, *33*, 410–426.
Jelicic, H., Phelps, E., & Lerner, R.M. (2009). Use of missing data methods in longitudinal studies: the persistence of bad practices in developmental psychology. *Developmental Neuropsychology*, *45*, 1195–1199.
Jobe, J.B., Smith, D.M., Ball, K., Tennstedt, S.L., Marsiske, M., Willis, S.L., & Kleinman, K. (2001). Active: a cognitive intervention trial to promote independence in older adults. *Controlled Clinical Trials*, *22*(4), 453–479.
Leppard, P., & Tallis, G.M. (1989). Evaluation of the mean and covariance of the truncated multinormal. *Applied Statistics*, *38*, 543–553.
Little, R.J.A., & Rubin, D.B. (2002). *Statistical analysis with missing data* (2nd ed.). New York: Wiley-Interscience.
Lu, Z., Zhang, Z., & Lubke, G. (2011). Bayesian inference for growth mixture models with non-ignorable missing data. *Multivariate Behavioral Research*, *46*, 567–597.
MacKinnon, D.P. (2008). *Introduction to statistical mediation analysis*. London: Taylor & Francis.
MacKinnon, D.P., Lockwood, C.M., Hoffman, J.M., West, S.G., & Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods*, *7*, 83–104.
MacKinnon, D.P., Lockwood, C.M., & Williams, J. (2004). Confidence limits for the indirect effect: distribution of the product and resampling methods. *Multivariate Behavioral Research*, *39*(1), 99–128.
McArdle, J.J., & Boker, S.M. (1990). *Rampath*. Hillsdale: Lawrence Erlbaum.
Preacher, K.J., & Hayes, A.F. (2004). SPSS and sas procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, & Computers*, *36*, 717–731.
Preacher, K.J., & Hayes, A.F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, *40*, 879–891.
Rubin, D.B. (1976). Inference and missing data. *Biometrika*, *63*(3), 581–592.
Rubin, D.B. (1996). Multiple imputation after 18+ years. *Journal of the American Statistical Association*, *91*, 473–489.
Savalei, V., & Bentler, P.M. (2009). A two-stage approach to missing data: theory and application to auxiliary variables. *Structural Equation Modeling*, *16*, 477–497.
Savalei, V., & Falk, C. (in press). Robust two-stage approach outperforms robust FIML with incomplete nonnormal data. *Structural Equation Modeling*.
Schafer, J.L. (1997). *Analysis of incomplete multivariate data*. London: Chapman & Hall/CRC.
Shrout, P.E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: new procedures and recommendations. *Psychological Methods*, *7*, 422–445.
Sobel, M.E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. In S. Leinhardt (Ed.), *Sociological methodology* (pp. 290–312). San Francisco: Jossey-Bass.
Tallis, G.M. (1961). The moment generating function of the truncated multinormal distribution. *Journal of the Royal Statistical Society. Series B*, *23*, 223–229.
Tang, M.-L., & Bentler, P.M. (1997). Maximum likelihood estimation in covariance structure analysis with truncated data. *British Journal of Mathematical & Statistical Psychology*, *50*(2), 339–349. doi:10.1111/j.2044-8317.1997.tb01149.x. Available from http://dx.doi.org/10.1111/j.2044-8317.1997.tb01149.x.
Thurstone, L.L., & Thurstone, T.G. (1949). *Examiner manual for the SRA primary mental abilities test (form 10–14)*. Chicago: Science Research Associates.
Wilhelm, S., & Manjunath, B.G. (2010). tmvtnorm: truncated multivariate normal and Student t distribution [Computer software manual]. Available from http://CRAN.R-project.org/package=tmvtnorm (R package version 1.2-3).
Willis, S.L., & Marsiske, M. (1993). *Manual for the everyday problems test*. University Park: Pennsylvania State University.
Yuan, K.-H. (2009). Identifying variables responsible for data not missing at random. *Psychometrika*, *74*, 233–256.
Yuan, K.-H., & Bentler, P.M. (2000). Three likelihood-based methods for mean and covariance structure analysis with nonnormal missing data. *Sociological Methodology*, *30*, 165–200.
Yung, Y.-F. (1996). Bootstrapping techniques in analysis of mean and covariance structures. In G.A. Marcoulides & R.E. Schumacker (Eds.), *Advanced structural equation modeling: issues and techniques* (pp. 195–226). Mahwah: Erlbaum.
Zhang, Z., & Yuan, K.-H. (2012). WebSEM manual [Computer software manual]. Available from https://websem.psychstat.org.

Manuscript Received: 16 DEC 2011
Final Version Received: 2 MAY 2012
Published Online Date: 7 DEC 2012