EXAMINATION OF A BAYESIAN JOINT MODELING APPROACH FOR HANDLING MISSING MODERATORS IN META-REGRESSION

A Dissertation

by

BRANDIE SEMMA

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Christopher Thompson
Co-Chair of Committee,	Wen Luo
Committee Members,	Oi-Man Kwok
	Lei-Shih Chen
	Myeongsun Yoon
Head of Department,	Fuhui Tong

December 2020

Major Subject: Educational Psychology

ABSTRACT

Meta-regression is used to understand the role of moderators in a meta-analytic model. However, during data extraction it is common for the data to not be clearly presented, incomplete, or missing. Consequently, missing study and participant characteristics arise, which can make it difficult to estimate meta-regression models. This dissertation examines a Bayesian conditional joint modeling (CJM) method for handling missing moderators in meta-regression using a series of conditional distributions. The use of CJM has been proposed in the meta-analysis literature to predict missing moderators (Hemming, Hutton, Maguire, & Marson, 2010). However, its performance has yet to be empirically studied. This dissertation investigated the CJM approach through a simulation study. Results suggest that the CJM approach performed similarly to listwise deletion when estimating the missing moderator, but performed better when estimating the overall true effect-size.

DEDICATION

To my mother, father, brothers, and my beautiful seester.

ACKNOWLEDGEMENTS

I would like to sincerely thank my committee chair, Christopher Thompson for his amazing mentorship, guidance, and constant support. His enthusiasm and willingness to go above and beyond in providing me with feedback and advice brought out the best in me. I would not have achieved as much as I did without his encouragement and guidance. My appreciation also goes to my co-chair, Wen Luo, who challenged my thinking and helped to make this dissertation complete.

I would also like to thank and acknowledge the members of my committee, Dr. Oi-Man Kwok, Dr. Myeongsun Yoon, and Dr. Lei-Shih Chen, for their guidance and support throughout my entire academic career.

I will also never forget the good times I had during my PhD studies and amazing friends who became like family. I would not have been able to get through the toughest times without them: Maria McCameron, Lucy Chen, Mirim Kim, and Ashley Mazjun, Martin Mendoza, Sandy Smith, Mary Whirely, Alyce Odasso, and Fatma Altinsoy. I want to give a special thanks to April Haas, who was there through all the good and the bad. She kept me motivated and pushed me to finish, even when I felt like I couldn’t finish.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a dissertation committee consisting of Christopher Thompson, Wen Luo, Oi-Man Known and Myeongsun Yoon of the Department of Educational Psychology and Lei-Shih Chen of the Department of Health and Kinesiology.

All other work conducted for the dissertation was completed by the student independently.

Funding Sources

Graduate study was supported by the College of Education and Human Development Strategic Research Award from Texas A&M University

NOMENCLATURE

$i = 1, \ldots, K$	Number of studies
K	Number of effect sizes
T_i	Observed effect size in study i
θ_i	True effect size in study i
θ	Overall true effect size
ε_i	Sampling error associated with study i
δ_i	Random error associated with study i
v_i	Within-study variance parameter for study i
τ^2	Between-studies variance parameter
η_i	Total variability of the observed study effect size in study i
w_i	Unconditional random-effects weight for study i
Q	Q statistic
j_i	Fixed-effect weight for study i
a_i	Unconditional random-effects weight for study i
β	Regression coefficient
$p(\cdot)$	Probability function
$p(A)$	Prior distribution
$p(B\|A)$	Likelihood function
$p(A\|B)$	Posterior distribution
$p(B)$	Normalizing constant
β_0	Model intercept
$\beta_1, \ldots, \beta_p$	Regression coefficients for p predictors
$x_{i1}, \ldots, x_{ip}$	Study characteristics for p predictors
φ_i	Random error in the conditional random-effects model for study i
τ_x^2	Conditional between-studies variance parameter
w_i^*	Conditional random-effects weight for study i
Y	Generic explanatory variable
Y_{obs}	Observed data
Y_{mis}	Missing data
R	Binary variable or a matrix of values indicating score missingness
$r = 1$	Observed score
$r = 0$	Missing score
MCAR	Missing completely at random
MNAR	Missing not at random
MAR	Missing at random
$\emptyset$	Unknown parameters
$\forall$	For all
$\prod$	Multiplication operator
L_i	Likelihood function for study i
r	Number of explanatory variables in the missing data model
$\boldsymbol{V_o}$	Score vector for case o
$\boldsymbol{\mu}$	Vector of population means

$\boldsymbol{\Sigma}$	Covariance matrix
m	Number of imputations
Y_t^*	Imputed values at I-step t
ϑ	Parameter of interest/model parameter
Y_t^*	Imputed values at I-step t
$\boldsymbol{\vartheta}_{t-1}^*$	Mean vector and the covariance matrix
θ_t^*	Simulated parameter values from P-step t
$\boldsymbol{X}$	Matrix of moderators
$Y_{\_j}$	Missing explanatory variable
$\boldsymbol{\alpha}_j$	Vetor of parameters for the p*th* distribution
y	Response variable
τ_{xp}	Variance for γ_p
d_i	Standardized mean difference in study *i*
$\bar{Y}_i^T$	Treatment group mean in study *i*
$\bar{Y}_i^C$	Control group mean in study *i*
S_i^P	Pooled standard deviation in study *i*
sd_T^2	Treatment group standard deviation in study *i*
sd_C^2	Control group standard deviation in study *i*
n_C	Sample size of the control group
n_T	Sample size of the treatment group
$s_{d_i}^2$	Variance of the standardized mean difference in study *i*
g	Hedges' g
J	Correction factor for Hedges' g
df	Degrees of freedom
γ_p	Joint moderator distribution for predictor p
η^2	Eta-squared from ANOVA model
E	Coverage
B	Number of replications

TABLE OF CONTENTS

LIST OF FIGURES

Page

LIST OF TABLES

Page

CHAPTER I

INTRODUCTION

The occurrence of missing data is a common issue when reviewing articles for a meta-analysis (Pigott, 2019). Missing data are especially problematic when including variables to be analyzed in a moderator analysis. Moderators are typically included in a meta-analysis to explain the presence of effect-size heterogeneity (Thompson & Higgins, 2002; Thompson & Sharp, 1999; Tipton, Pustejovsky, & Ahmadi, 2019). One reason missing moderators can be an issue is that it becomes difficult to include several moderators in a single statistical model, which can help control for confounding factors. Current practices tend to exclude the primary article for analysis if there is a missing moderator (Pigott, 2019). Unfortunately, excluding the article results in loss of information, which may cause biased estimates. Commonly used methods for missing moderators, such as listwise and pairwise deletion, tend to yield biased estimates of the observed effect sizes and regression coefficients (Pigott, 2001). There still remains a need for missing data methods in the presence of missing moderators (Tipton et al., 2019).

Bayesian methods are becoming increasingly popular in meta-analytical research (Van de Schoot, Winter, Ryan, Zondervan-Zwijnenburg, & Depaoli, 2017). One reason for its rise in popularity is due to its flexibility in answering a range of questions. Other reasons include the use of prior distributions and the ability to make direct probability statements. Bayesian methods also allow researchers to account for the uncertainty due to missing data.

The purpose of this dissertation is to present a Bayesian joint modeling approach for handling missing moderators in meta-analysis. I examine a method that uses Bayesian estimation to predict missing moderators, conditional on at least one fully observed moderator. The next immediate sections of this dissertation detail the relevant literature, introduces the Bayesian conditional joint model, and explains the methodology to complete the simulation.

Chapter II begins by introducing the role of meta-analyses in research and reviews two statistical models used to estimate parameters, namely the unconditional random-effects model and the conditional random-effects model. Chapter II also addresses several methods of heterogeneity estimation in meta-analysis and introduces Bayesian estimation. The chapter details general missing data theory, as well as in meta-analysis, explaining different mechanisms and presenting two classical procedures of handling missing data (listwise and pairwise deletion). The chapter goes on to explain two common methods for handling missing data (maximum likelihood and multiple imputation). Last, an overview of the current methods available to handle missing moderators in meta-analysis is provided.

Chapter III examiness a Bayesian conditional joint modeling (CJM) approach for handling missing moderators and outlines the methodology of the dissertation. The CJM approach was created by Lipsitz and Ibrahim (1996) and applied to missing moderators, as proposed by Hemming et al. (2010). The chapter details the way the CJM approach is evaluated using computer simulations. Various simulation conditions, as well as imputation and evaluation procedures, are described.

Chapter IV summarizes the results of the computer simulation. Results compare the CJM to listwise deletion. Additionally, results are also compared across conditions, focusing on the statistical bias, coverage, and mean squared error. Chapter V discusses findings, implications, and presents areas of further research.

CHAPTER II

LITERATURE REVIEW

Meta-Analysis

Meta-analysis is a set of statistical methods used to combine results from studies of the same topic (Glass, 1976). The power to detect an effect and the generalizability of the results increases by combining the results (Cooper, Hedges, & Valentine, 2009). As an example, in the medical sciences literature meta-analysis can be used to assess the clinical effectiveness of healthcare interventions (Haidich, 2010). In the social sciences literature, for example, meta-analysis can be used as evidence to inform educational initiatives and policies affecting mental health care (Davis, Mengersen, Bennett, & Mazerolle, 2014).

One component of meta-analysis is to synthesize research results to obtain an overall effect-size estimate for a population of studies. Study results are combined using effect sizes from primary studies to estimate an overall effect size. There are several types of effect sizes, including standardized mean difference, bivariate correlation, and odds ratio (Borenstein, Hedges, Higgins, & Rothstein, 2011). The focus of this dissertation is on the standardized-mean-difference effect-size metric.

Three popular statistical models for meta-analysis are the fixed-effect (FE) model, unconditional random-effects (URE) model, and conditional random effects (CRE; sometimes referred to as a mixed-effects) model. Model choice is important because it impacts computations, helps define the goals of analyses, and influences interpretations of results (Borenstein, Hedges, Higgins, & Rothstein, 2010). A FE model assumes only one source of variation in the observed effect sizes, that of variation within a primary study. Put another way, there is only one level of sampling in a FE model, therefore there is only one source of variance (Borenstein et al., 2011). In a notational form, the FE model can be stated as

$$T_i = \theta_i + \varepsilon_i \text{ where } \varepsilon_i \sim N(0, v_i), \tag{1}$$

where T_i is the observed effect size for study i, and θ_i is the true effect size for study i. The term ε_i is the difference between the true effect size and the observed overall effect size and is the only source of variation in the FE model. This source of variation is equal to the within-study error variance. The term v_i is the known variance for study i.

In an URE model, effect sizes are assumed to be realizations from a common distribution of effect sizes. In contrast to the FE model, the URE model has two sources of variance (within-study variance and between-studies variance). This involves the examination of variability of effect sizes across studies (i.e., heterogeneity). Effect sizes in a meta-analysis may be considered homogenous if they share a common underlying true effect size. The error assumption in a FE model is often unrealistic to assume in social science research (Higgins, Thompson, & Spiegelhalter, 2009); effect sizes may vary depending on 1) study characteristics, and/or 2) sampling error, and/or 3) at random. It is important to assess heterogeneity in meta-analysis because identifying and describing the variation among the effect sizes provides a deeper understanding of the topic of interest (Davis et al., 2014; Haidich, 2010). Different true effects among the studies can be due to the presence of study characteristics. For example, studies may differ on the average age of participants or intervention type. This dissertation focuses on the conditional random-effects (CRE) model, in which I will go into more detail later.

Unconditional Random-effects Model

Like the FE, in a URE model true effect sizes also differ from each other because of sampling error. However, unlike the FE model, in the URE model true effect sizes also differ because they are assumed to be a sample from a population of effect sizes. Put another way, true effect sizes also differ due to random error. As I will go into detail when explaining the CRE model, studies are expected to have some variability in terms of their participant, treatment, and/or other study-specific characteristics.

In both the FE and URE models, each study provides information about a different effect size, presumably from a common population. The URE model attempts to represent these effect sizes in the overall estimate (Borenstein et al., 2010). The URE model assumes that each true effect-size variance is partitioned into two components. In a multilevel framework, there is a within-study level (level 1) and a between-studies level (level 2): error due to sampling error (level 1) and error due to heterogeneity between studies (level 2). Provided $i = 1, \ldots, K$ statistically independent effect-size estimates, the URE model can be defined as

$$\begin{aligned} T_i &= \theta_i + \varepsilon_i \text{ where } \varepsilon_i \sim N(0, v_i) \\ \theta_i &= \theta + \delta_i \text{ where } \delta_i \sim N(0, \tau^2), \end{aligned} \tag{2}$$

where T_i is the observed effect size for study i, θ_i is the true effect size in study i, and θ is the overall true effect. The error term ε_i represents random deviations from the true effect size and are assumed independent with mean zero and known variance, v_i. The error term δ_i is the random error in study i. The variance of the random error δ_i is the random-effects variance parameter, τ^2. The term τ^2 measures the amount of heterogeneity in the distribution of true effects (i.e., between-studies variance). Both error components are assumed to be normally distributed.

The total variability of an effect-size estimate under the URE model, η_i, can be expressed as

$$\eta_i = v_i + \tau^2. \tag{3}$$

When τ^2 is equal to zero, the model simplifies to the FE model with $\eta_i = v_i$ (Konstantopoulos & Hedges, 2019). Weighted least squares regression can be used to estimate the URE model with weights

$$w_i^* = \frac{1}{v_i + \tau^2}. \tag{4}$$

Conditional Random-effects Model

Exploring sources of effect-size heterogeneity is a crucial step in meta-analysis (Hedges & Pigott, 2004; Konstantopoulos & Hedges, 2019; Viechtbauer, 2007). In the URE model, the between-studies variance represents the excess variation in observed effects after accounting for sampling error. A second type of random-effects model considered in this dissertation is a CRE model. The CRE model evaluates studies together in attempt to understand similarities and differences in the reported outcomes. The CRE model includes moderators, which may explain the possible causes of heterogeneity, as well as sampling and a random/residual error component.

By including moderators, true effect sizes partially depend on a set of study characteristics. Equation (2) can be rewritten to include moderators,

$$\theta_i = \beta_0 + \beta_1 X_{i1} + \cdots + \beta_p X_{ip} + \varphi_i, \tag{5}$$

where β_0 is the intercept, $\beta_1, \ldots, \beta_p$ are regression parameters for p moderators, $X_{i1}, \ldots, X_{ip}$ are respective study characteristics, and the variance of $\varphi_i = v_i + \tau_X^2$. As opposed to the URE model, the between-studies variance in the CRE model accounts for p moderators and is denoted as τ_X^2. As a single-level model, the CRE model can be written as

$$T_i = \beta_0 + \beta_1 X_{i1} + \cdots + \beta_p X_{ip} + \varphi_i + \varepsilon_i, \tag{6}$$

$$\varepsilon_i \sim N(0, v_i)$$

$$\varphi_i \sim N(0, \tau_X^2).$$

One approach for examining moderators is meta-regression, which accommodates continuous and categorical predictors (Borenstein et al., 2011; Konstantopoulos & Hedges, 2019). Here, the weights, w_i^*, are equal to

$$w_i^* = \frac{1}{v_i + \tau_X^2}. \tag{7}$$

The goal of meta-regression is to examine how a set of moderators relate to the overall effect. Meta-regressions are also used to assess the influence of study characteristics on the effect-size variability (Borenstein et al., 2011; Terri D. Pigott, 2012; Tipton et al., 2019).

Between-Studies Variance Estimation

In this section, I present several common methods for estimating between-studies variability from a URE model, including the DerSimonian and Laird method (1986) and restricted maximum likelihood estimation.

DerSimonian and Laird Estimator

DerSimonian and Laird (1986) proposed a method-of-moments estimator to estimate the heterogeneity variance in a random-effects meta-analysis, which is a partial function of Cochran's Q statistic:

$$Q = \sum w_i (T_i - \theta)^2, \tag{8}$$

where w_i is equal to $1/\hat{v}_i$. The DerSimonian and Laird estimator is given by

$$\hat{\tau}_{MM}^2 = \frac{Q - (k-1)}{\sum_{i=1}^k \widehat{w}_i - \sum_{i=1}^k w_i^2 / \sum_{i=1}^k \widehat{w}_i}. \tag{9}$$

This method is unbiased if the within-study variances are known and does not make assumptions regarding the distribution of τ^2. Due to its simplicity and non-iterative estimation approach, $\hat{\tau}^2_{MM}$ method is one of the most commonly used estimators in URE models (Brockwell & Gordon, 2001; Sidik & Jonkman, 2007; Thompson & Sharp, 1999). However, it lacks efficiency and tends to underestimate the standard error of the overall effect-size estimate (Bodnar, Link, Arendacká, Possolo, & Elster, 2017; Brockwell & Gordon, 2001).

Restricted Maximum Likelihood Estimator

Restricted maximum likelihood (REML) assumes a normal distribution for the random effects. The log-likelihood function of REML stems from the maximum likelihood method (MLE; Hardy & Thompson, 1996). The difference between REML and MLE is that REML does not directly estimate the parameter θ. Consequently, no assumptions are made about θ, which can be a benefit of REML over MLE (Viechtbauer, 2005).

The REML estimator is given by

$$\hat{\tau}^2_{REML} = \max\left\{0, \frac{\sum_{i=1}^{k} a_i^2 \left(\left(\hat{\theta}_i - \hat{\theta}_{RE}\right)^2 - \hat{\theta}_i^2\right)}{\sum_{i=1}^{k} a_i^2} + \frac{1}{\sum_{i=1}^{k} a_i}\right\}, \quad (10)$$

where $a_i = 1/(\hat{v}_i + \hat{\tau}^2_{REML})$. The parameter estimate of $\hat{\tau}^2_{REML}$ is computed through an iterative procedure (Hardy & Thompson, 1996; Konstantopoulos & Hedges, 2019; Sidik & Jonkman, 2007; Raudenbush, 2009). REML obtains estimates by maximizing the likelihood as a function of τ^2 alone.

Bayesian Estimation

Bayesian inference is a statistical approach that is based on Bayes' rule for probabilities. In a Bayesian approach to statistical analysis, probability is interpreted as a subjective knowledge of uncertainty and is regarded as a degree of belief (de Finetti, 1974; Jackman, 2009). Bayes rule can be expressed as

$$p(A|B) = \frac{p(B|A)\, p(A)}{p(B)}, \quad (11)$$

where $p(A)$ is the prior distribution, $p(B|A)$ is the likelihood function, $p(A|B)$ is the posterior distribution, and $p(B)$ is a normalizing constant to ensure the distribution integrates to unity (Turner & Higgins, 2019).

In frequentist methods, parameters are unknown quantities that are to be estimated by a statistic (e.g., mean, standard deviation). These parameters, though unknown, are assumed to be fixed. However, in a Bayesian framework, unknown parameters contain uncertainty and are described by probability distributions (Gelman, 2003; Jackman, 2009; Schmid, 2001; Van de Schoot et al., 2014). The probability distributions of the unknown parameters are also referred to as prior distributions and contain our state of knowledge or beliefs about the parameters. The unknown parameters in a URE meta-analysis are the within-study variances v_i, between-studies variance, τ^2, and the population treatment effect θ; though, only prior information is available about τ^2 and θ (Abrams & Sansó, 1998; R. M. Turner & Higgins, 2019). The within-study variances v_i are usually not considered to be unknown parameters because they are assumed to be known and are replaced by the observed within-study variances from each study (Abrams & Sansó, 1998). When conducting meta-regression, the regression coefficients of the moderators are also unknown parameters.

Observed data are used in the likelihood function, which determines the probability of the data, conditional on the parameters. Conceptually, the prior distribution is combined with the likelihood function and the unconditional distribution of the data to create new information about the parameters. This combination produces a posterior distribution, which describes the parameters after observing the data (Schmid, 2001; Van de Schoot et al., 2014). Probability statements about the parameters are taken from the posterior distribution.

There are several advantages to using a Bayesian approach in meta-analysis. First, Bayesian inference produces a distribution for the parameters of interest, which allows for direct probability statements (Sutton & Abrams, 2001; Turner & Higgins, 2019). For example, if we are interested in testing the difference in means between sex (male and female), we can compute the probability that the average effect size for males is larger or smaller than the average for females. A Bayesian approach uses credible intervals to draw these inferences and make these statements.

Second, because the Bayesian approach can incorporate external evidence on the between-studies variance, Higgins & Whitehead (1996) suggest incorporating real data from previous studies rather than using subjective opinion to formulate the prior distribution of the between-studies variance. This allows researchers to "borrow strength" from other studies, which in turn reduces the imprecision and enables predictions of effects in future studies (Sutton & Abrams, 2001; Turner, Jackson, Wei, Thompson, & Higgins, 2015).

Third, the Bayesian approach allows for all parameter uncertainty to be investigated in the analysis (Sutton & Abrams, 2001), not just the between-studies variance. Obtaining information about the parameters through the posterior distribution is advantageous because researchers can analyze the posterior estimates to determine if the results are meaningful, or to investigate possible explanations for certain results.

A fourth advantage is that the Bayesian approach allows for a more flexible modeling framework. Using prior distributions naturally allows researchers to account for uncertainty due to missing data (Turner, Dias, Ades, & Welton, 2015)

A potential drawback to this approach stems from the use of prior distributions. Different prior distributions can be used to generate varying results, especially if the observed data are sparse (Kruschke & Liddell, 2018). Additionally, the overall computational method can be complex to implement and time consuming to write and to conduct (Sutton & Abrams, 2001).

Missing Data Theory

A common problem in quantitative research is that of missing data; many statistical methods assume complete information for all the variables in an analysis (Enders, 2010; Peugh & Enders, 2004). For example, missing data may occur when research participants refuse or forget to answer a survey question. In a meta-analysis, missing data can occur when primary studies fail to provide the necessary quantitative information (e.g., means or standard deviations) to calculate effect sizes.

There are several reasons why missing data is an issue. First, the presence of missing data may reduce statistical power. Second, missing data may introduce statistical bias in the estimation of parameters. Third, it can reduce the representativeness of the samples. Overall, missing data can threaten statistical validity (Kang, 2013).

A “complete dataset” can be thought of as having two components: observed data (Y_{obs}) and missing data (Y_{mis}). The complete dataset can also be represented as a data matrix that contains values of a collection of explanatory variables. Each row represents individual units, cases, observations, subjects, etc. and columns represent characteristics or explanatory variables that are measured for each unit.

In missing data theory, it is assumed that there is a score on every explanatory variable for each case (Enders, 2010). For example, if a participant responds to a 10-item instrument (e.g., 10 explanatory variables), it is assumed there will be a score on each of the 10 items. However, in practice suppose some portion of data are missing. In the above example, it is possible the participant may respond to only a portion of the 10 explanatory variables.

Rubin (1976) conceptualizes missingness as a binary variable, R, with a probability distribution. This variable indicates whether a score on an explanatory variable is observed or missing (r = 1 if a score is observed or r = 0 if the score is missing). However, provided multivariate data, R becomes a matrix of missing data indicators, denoted as **R**. What is most important for missing data methods is the relationship (or lack thereof) of the missing data to other explanatory variables in the dataset (Enders, 2017). Below is an example of an **R** matrix with three variables (or items) and four participants:

$$\begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{12}$$

Columns indicate items 1 through 3 on some questionnaire and rows indicate participants 1 to 4 who answered the questionnaire. Specifically, participant 1 responded to items 1 and 3, however did not respond to item 2. Participants 2 and 3 responded to the first two items but did not answer the third item, and participant 4 did not respond to any items.

Missing Data Mechanisms

Missing data mechanisms represent the relationship between observed explanatory variables and the probability of missing data (Enders, 2010; Little & Rubin, 2019; Rubin, 1976). This explanation is not causal. Missing data mechanisms specify what conditions must be present to accurately estimate the missing parameters, for example, missing effect sizes in a meta-analytic context (Little & Rubin, 2019; Rubin, 1976a).

Rubin (1976) defines three types of missing data mechanisms: missing completely at random (MCAR), missing at random (MAR), and missing not at random (MNAR). The nature of the relationship between **R** (missing data matrix) and the observed data is what determines the missing data mechanisms. The type of missing data mechanism influences the appropriate statistical analysis of the dataset (Enders, 2010; Little & Rubin, 2019; Rubin, 1976).

Missing Completely at Random (MCAR) Mechanism

The MCAR mechanism is the probability of missing data for an explanatory variable, Y, that does not depend on either observed explanatory variables or Y itself (Rubin, 1976). That is, the missing data are completely independent of the combined observed data (i.e., **R** is independent of Y_{obs} and Y_{mis}). This can be expressed as the probability equivalence

$$p(R|Y, \emptyset) = p(R|\emptyset) \; \forall \; \{Y, \emptyset\}, \tag{13}$$

where $\emptyset$ represents the set of unknown parameter(s) (Little & Rubin, 2019). There is no assumed systematic difference between the missing data for an explanatory variable and the observed data for the explanatory variable. The missing observations are a random sample of all observations, had the data been complete (Bhaskaran & Smeeth, 2014).

In a meta-analytic context, the MCAR mechanism stipulates that studies with completely observed data can be considered to be a random sample of the studies originally identified for the synthesis (Pigott, 2019). For example, primary authors make different decisions when reporting information in a research article; the result of these differences can be values of a specific explanatory variable being MCAR (Pigott, 2019).

Missing at Random (MAR) Mechanism

If the probability of missing a value on an explanatory variable, Y, is related to observed or unobserved explanatory variables, but not the actual value of Y, then the missing mechanism is said to be MAR (Rubin, 1976). This can be expressed as a probability equivalence:

$$p(R|Y, \emptyset) = p(R|Y_{obs}, \emptyset)\ \forall\ \{Y_{mis,} \emptyset\}. \quad (14)$$

The probability of missingness depends on observed data by some unknown parameter(s) $\emptyset$, that relates Y_{obs} to **R** (Enders, 2010; Little & Rubin, 2019). This relationship implies that **R** is dependent on Y_{obs}, but not Y_{mis}. That is, the probability of a missing observation is independent of the missing value for the explanatory variable *Y*, but is dependent on the values of completely observed explanatory variables (Bhaskaran & Smeeth, 2014; Pigott, 2012).

The MAR assumption is relevant to meta-analysis as well:

> "Some studies may report the income level of subjects as a function of the percent of students who qualify for free lunch, while others report income level as the average income reported by parents. The differences between these studies could be due to the discipline of the primary author – studies in education tend to use the percent of studies with low income in a school while large-scale studies may have the resources to conduct a survey of parents to obtain a more direct measure of income. A missing value for a particular measure of income in a particular study is not necessarily related to the value of income itself but to the choices of the primary author and constraints on the published version of the study" (Pigott, 2012, p. 91-92).

A systematic relationship between one or more measured explanatory variables and the probability of missing data exists because the missing values for *Y* are not random. There is no way that the probability of the missing data on *Y* is solely a function of other measured explanatory variables. Thus, there is no way to test or verify the MAR mechanism. MAR is described as ignorable missingness in the missing data literature because likelihood-based analyses of missing data, such as maximum likelihood estimation and multiple imputation, do not require an estimate of the missing data distribution (Enders, 2010). Later I discuss likelihood-based missing data methods.

Missing Not at Random (MNAR) Mechanism

The third missingness mechanism, missing not at random (MNAR), occurs when the probability distribution of **R** depends on the missing values (Enders, 2010; Little & Rubin, 2019). This mechanism assumes the probability of the missing data on *Y* relates to the observed values of *Y* itself, even after considering other variables. Like the MAR mechanism, there is no way to verify that scores are MNAR without knowing the values of the missing variables. The probability distribution of the MNAR mechanism is

$$p(R|Y_{obs}, Y_{mis}, \emptyset). \tag{15}$$

As one example in meta-analysis, this can occur when effect sizes are not reported in a study because they are not statistically significant (Pigott, 2019). Specific values are more likely to be missing than other values because they are censored based on statistical significance.

Classical Missing Data Techniques

This section provides an overview of two classical approaches to handling missing data: listwise deletion and pairwise deletion. Advantages and disadvantages to these methods are highlighted.

Listwise deletion

A popular missing data handling method is listwise deletion (i.e., complete case analysis). This method involves discarding cases where one or more explanatory variables are missing (Enders, 2010; Little & Rubin, 2019; Schafer & Graham, 2002). Listwise deletion is advantageous because a set of fully observed cases is used for all analyses. Listwise deletion also allows for the comparability of univariate statistics because all are computed from a common sample (Little & Rubin, 2019).

However there are also several disadvantages to this approach (Little & Rubin, 2019). Listwise deletion is inefficient because it excludes data that may be informative analyses. Estimates may be biased if the missing information is MAR (i.e., complete cases systematically differ from the full sample). Furthermore, a decrease in the number of cases used for the analysis can lead to loss of statistical power. Overall, listwise deletion is a limiting technique and is generally not recommended (Graham, 2009; Little & Rubin, 2019).

Pairwise Deletion

Pairwise deletion (i.e., available case analysis) eliminates cases on an analysis-by-analysis basis, which allows for the use of as much observed data as possible. Consequently, the sample of data changes from variable to variable and analysis to analysis with respect to the pattern of missingness. Although this method uses more information than listwise deletion, it is still an inefficient method. Like listwise deletion, pairwise deletion requires the data to be MCAR and can produce biased estimates if the assumption is not met (Enders, 2010; Little & Rubin, 2019).

Additionally, using different subsets of cases may cause issues when estimating correlations or covariances among the variables. This is particularly important in case of multivariate analyses that use a covariance matrix as the input data (e.g., regression). If each correlation/covariance is estimated on the basis of having data for all variables, then there is no guarantee that the correlation matrix will be positive definite (Graham, 2009). Non-positive definite matrices can lead to estimation problems for some multivariate statistical analyses (Graham, 2009). A non-positive definite variance-covariance matrix occurs when a correlation or covariance matrix contains combinations of estimates that would have been mathematically impractical. Due to these issues, pairwise deletion is also generally not recommended (Graham, 2009; Little & Rubin, 2019).

Modern Missing Data Techniques

Likelihood-based analyses, such as MLE and multiple imputation (MI), have been shown to be more effective than the two classical procedures mentioned above (Allison, 2000; Graham, 2009; Schafer, 1997). These methods are supported by a growing number of empirical research studies that demonstrate their effectiveness (e.g., Anderson, 1957; Buuren & Groothuis-Oudshoorn, 2011; Carpenter, Kenward, & White, 2007; Collins, Schafer, & Kam, 2001; Dempster, Laird, & Rubin, 1977; Furukawa, Barbui, Cipriani, Brambilla, & Watanabe, 2006; Gold & Bentler, 2000; Grund, Lüdtke, & Robitzsch, 2018). MLE and MI require less assumptions about the cause of missing data and produce parameter estimates with less statistical bias and greater power (Enders, 2001; 2010). Furthermore, these methods also do not require information about the unknown parameter $\emptyset$ if the data are assumed to be MAR or MCAR (Little & Rubin, 2019; Rubin, 1976).

Maximum Likelihood Estimation (MLE)

MLE begins by specifying a distribution for the population of data, which is commonly assumed to be normally distributed for areas in the social sciences (Enders, 2010). In most cases, MLE is applied to multivariate data because datasets typically contain more than one explanatory variable. Thus, MLE examines the probability of a score from a multivariate normal distribution because it generalizes the curve of the distribution of data to multiple explanatory variables (Enders, 2010).

The probability density function describes the relative likelihood that an explanatory variable takes on a given value. The probability density function is directly estimated using only the observed data when the dataset contains missing values (Grund, Lüdtke, & Robitzsch, 2019). However, the natural logarithm of the individual likelihood values is used during estimation due to rounding errors and computational simplicity (Enders, 2001; 2010).

The goal of MLE is to identify population parameter values that have the highest probability of reproducing a particular sample of data (Enders, 2001, 2010; Gold & Bentler, 2000; Graham, Hofer, & MacKinnon, 1996; Kenward & Molenberghs, 1998). Different log-likelihood values are estimated based on the unique combination of parameter estimates. MLE attempts to identify the set of estimates that produces the largest log-likelihood. This involves an iterative process which tests different values for the unknown parameters until the model converges to a set of parameter values that maximize the likelihood (Dempster et al., 1977). The log-likelihood for each case is

$$Log(L_0) = -\frac{r}{2}\log(2\pi) - \frac{1}{2}\log|\boldsymbol{\Sigma}| - \frac{1}{2}(\boldsymbol{V}_o - \boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{V}_o - \boldsymbol{\mu}), \qquad (16)$$

where r is the number of explanatory variables, $\boldsymbol{V}_o$ is the score vector for case o, and $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are the population mean vector and covariance matrix for case o (Enders, 2010). The mean vector ($\boldsymbol{\mu}$) contains the means of each variable and the covariance matrix ($\boldsymbol{\Sigma}$) consists of the variances of the variables along the main diagonal and the covariances between each pair of variables on the off-diagonals.

Values of the log-likelihood that are closest to zero reflect a higher relative probability of drawing a sample of scores and closer proximity to the population mean than values those are further away from zero. The sample log-likelihood quantifies the fit between the data and the parameter estimates. It also provides a basis for choosing among a set of potential parameter values (Enders, 2010).

Multiple Imputation (MI)

The key idea behind MI (Rubin, 1986) is to use the distribution of the observed data to estimate a set of potential values for the missing data (Graham, Olchowski, & Gilreath, 2007; White, Royston, & Wood, 2011). One advantage of this technique is that it allows for unbiased parameter estimates (Graham et al., 2007). MI can be thought of as having three phases (Enders, 2017; Rubin, 1987): imputation phase, analysis phase, and pooling phase.

Imputation phase

The imputation phase first creates several versions of a sample dataset. The researcher selects the number of imputations *m* (i.e., copies of the data). However, it is suggested that at least 20 datasets ($m = 20$) are imputed (Graham et al., 2007). Each version contains different estimates of the missing values (Enders, 2017). The estimated parameters are then used to generate a new set of parameters for the next iteration of imputation.

Conceptually, we can think of the imputation phase as having two steps (Enders, 2010): *imputation* step (I-step) and *posterior* step (P-step). The I-step builds a set of regression equations using an estimate of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$. The development of MI is rooted in a Bayesian framework (Rubin, 1987), where each I-step is a random draw from the full conditional distribution. Also known as the posterior predictive distribution, the full conditional distribution is the distribution of possible unobserved values, conditional on the observed values. This can be summarized as

$$Y_t^* \sim p(Y_{mis}|Y_{obs}, \boldsymbol{\vartheta}_{t-1}^*), \tag{17}$$

where Y_t^* represents the imputed value at the t^{th} I-step and $\boldsymbol{\vartheta}_{t-1}^*$ contains the parameter values that generate the imputation regression equations (i.e., $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$ from the P-step). That is, $\boldsymbol{\vartheta}_{t-1}^*$ contains the parameter values from the imputed value before I-step *t*. Essentially, imputed values at an I-step are taken from a distribution that makes the missing values conditional on the observed data and the parameter estimates of $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$ from the previous I-step.

Missing values are then replaced by imputed values that are sampled from their posterior predictive distribution, given the observed data. The P-step generates different parameter values using data from the previous I-step via adding a random residual term to each element of the mean vector, $\hat{\boldsymbol{\mu}}$, and covariance matrix, $\hat{\boldsymbol{\Sigma}}$:

$$\boldsymbol{\vartheta}_{\boldsymbol{t}}^* \sim p(\vartheta | Y_{obs}, Y_i^*), \qquad (18)$$

where $\boldsymbol{\theta}_{\boldsymbol{t}}^*$ are the new estimated values from the P-step *t* (Enders, 2010). The values of $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$ are taken as a random draw from their respective posterior predictive distributions.

Analysis and pooling phases

After the imputation phase, the set of *m* datasets contains different estimates of the missing values. Each imputed dataset is analyzed to provide *m* sets of parameter estimates and their standard errors. The pooling phase combines the estimated missing values into a single set of results (Enders, 2010).

According to Rubin (1987), the average can be computed to pool the estimates from the analysis phase. Standard errors are pooled by computing the within-imputation variance and the between-imputation variance (Enders, 2010). The within-imputation variance is the average of the *m* sample variances and estimates the variability in the sample as if there were no missing data. The between-imputation variance estimates the variability across the *m* imputations and examines the sampling error due to missing data. The total sampling variance is the sum of the within and between-imputation variances. This three-step process is used for all MI procedures. However, the imputation phase can incorporate several different algorithms, depending on the specific missing data problem (Enders, 2010). For example, a longitudinal dataset requires a different imputation algorithm when compared to a cross-sectional dataset. Common algorithms for MI are expectation maximization (Dempster et al., 1977) and data augmentation (Tanner & Wong, 1987). Additionally, there are several methods of MI, including fully conditional specification (FCS; Van Buuren, Brand, Groothuis-Oudshoorn, & Rubin, 2006) and joint modeling (JM; Liu, Taylor, & Belin, 2000).

Joint Modeling (JM)

Joint modeling (Schafer, 1997) involves specifying a single imputation for all explanatory variables with missing data, where missing data are specified as having a multivariate distribution. Imputations are created for all explanatory variables simultaneously, which are drawn from conditional distributions by Markov Chain Monte Carlo (MCMC) techniques (Grund, Lüdtke, & Robitzsch, 2018a; Kline, Andridge, & Kaizar, 2017; Van Buuren & Groothuis-Oudshoorn, 2011).

MCMC techniques combine two properties: *Monte-Carlo* and *Markov Chain*. Monte Carlo examines random samples from a distribution to be able to describe the distribution (e.g., estimate the mean of the distribution). A Markov Chain is when samples from a probability distribution are obtained randomly through a sequential process. Each random sample in the Markov chain is used to generate the next random sample. Samples taken from a Markov chain only depend on the one before it; they do not directly depend on samples before the previous one. MCMC is constructed to establish a stationary distribution (Lin, 2010). The properties of the stationary distribution are estimated by examining random samples from the distribution (Van Ravenzwaaij, Cassey, & Brown, 2018).

As stated in Enders et al. (2016), “the term ‘joint model’ comes from the fact that the incomplete variables are assumed to follow a common distribution” (p. 225). The main idea of the JM approach is to define a joint multivariate model for all the explanatory variables in the dataset (Quartagno, Grund, & Carpenter, 2019). The JM approach first specifies a parametric multivariate density, $p(Y|\vartheta)$ for the data Y, given model parameters ϑ (Schafer, 1997). Next, imputations are drawn from the posterior predictive distribution, $p(Y_{mis}|Y_{obs})$. Bayesian methods, specifically Gibbs sampling, are typically used to fit and impute the missing data via a data augmentation algorithm (Tanner & Wong, 1987). New values for the parameters in the model are repeatedly drawn from a conditional distribution (Quartagno et al., 2019).

One advantage of the JM approach is that it works well with multilevel data structures (Andridge, 2011; Liu et al., 2000; Mistler & Enders, 2017). However, it can be difficult to build a joint model when there are a large number of explanatory variables or when variables are not normally distributed (Kline et al., 2017).

Fully Conditional Specification (FCS)

Fully conditional specification (FCS) imputes multivariate missing data on a variable-by-variable basis. This is completed by specifying a multivariate distribution through a set of conditional densities, one for each incomplete explanatory variable (Van Buuren, 2007). The conditional density is used to impute the missing explanatory variable provided the matrix of moderators, missing explanatory variable, Y_{miss}, and the Y missingness indicator, **R**. FCS begins with simple random draws from the marginal distribution, which is a distribution containing data from a single parameter in the dataset. Values are imputed by iterating over the conditional densities (Liu & De, 2015).

There are several other advantages to the FCS approach over the JM approach. One advantage is that the FCS approach may be easier to generalize models under nonignorable missing data mechanisms (Van Buuren et al., 2006). Additionally, Van Buuren et al. (2006) states that it may be easier to communicate this model to others because each variable has a separate imputation model. However, FCS is not without its disadvantages. One disadvantage of FCS is that each conditional density must be specified separately, which may be an issue if a dataset has many of variables (Van Buuren et al., 2006).

FCS has been proposed using several names, including regression switching (Van Buuren, Boshuizen, & Knook, 1999), sequential regressions (Raghunathan, Lepkowski, Hoewyk, & Solenberger, 2001), ordered pseudo-Gibbs sampler (Heckerman, Chickering, Meek, Rounthwaite, & Kadie, 2000), partially incompatible MCMC (Rubin, 2003), iterated univariate imputation (Gelman, 2004), multiple imputation by chained equations (Van Buuren & Groothuis-Oudshoorn, 2011), and fully conditional specification (Van Buuren, 2007).

Missing Data in Meta-Analysis

There are three main ways missing data can occur in meta-analysis: 1) studies are missing from the review due to publication bias or reporting bias, 2) effect sizes are missing due to inadequate quantitative reporting, including those needed to compute sample variances, and 3) moderators for regression models of effect-size variation are missing (Pigott, 2012; 2019). I will focus on the third scenario, missing moderator(s).

Missing Moderators

Meta-regression can be used to explore the extent to which study characteristics (i.e., moderators) explain the variation between studies. Reviewers generally code what information is most relevant pertaining to these characteristics. However, a lot of variation in terms of reporting in primary studies exists, thus missing study characteristics can occur.

Classical procedures for handling missing data tend to have several issues (Pigott, 2012; 2001). In a meta-analysis, using only complete case analysis (i.e., listwise deletion) limits the number of studies available for analysis. If studies that do not provide adequate information are deleted, then reviewers assume the remaining studies for the analysis are representative of those originally gathered for review (i.e., studies are MCAR; Pigott, 2001). However, this is a strong and unlikely realistic assumption. Instead, information contained in the complete studies is lost and statistical power decreases (Kim & Curry, 1977; Pigott, 2012; Rubin, 1987). Lastly, Pigott (2001) showed that a listwise deletion inflates standard errors and underestimates parameter estimates when utilized in a meta-analysis.

Estimation problems can also occur if a reviewer implements an available case analysis (i.e., pairwise deletion; Pigott, 2001). One statistical problem that may arise is a non-positive definite variance-covariance matrix. Again, a non-positive definite variance-covariance matrix occurs when the correlation or covariance matrix has combinations of estimates that are not plausible. For example, if a correlation exceeds one.

Another issue concerns the within-study sample size when computing standard errors for the parameter of interest (Allison, 2000). This is because each parameter can be estimated with a different subset of studies. Pigott (2001) argues that the performance of an available case analysis is not reliable because each estimated value is based on a different subset of the original data set. If data are MCAR, then the subset is representative of the original data. However, if the data are MAR, then the subset is not representative of the original data and will produce biased estimates (Pigott, 2012)

Pigott (2001; 2012; 2019) recommends the use of multiple imputation for handling missing moderators in meta-analysis, arguing that multiple imputation may be more flexible when compared to MLE. Other researchers have also assessed the issue of missing moderators in meta-analysis by utilizing a joint modeling approach (Hemming et al., 2010).

Consistent with the JM approach, Hemming et al. (2010) imputed one explanatory variable at a time, conditional on the other included variables. The method by Hemming et al. (2010) was illustrated to contain a mixture of continuous and discrete variables. Moderators are specified to have a joint density, which is modeled as a factorization of a meta-regression model and a conditional factorization of the density for the moderators.

Hemming et al. (2010) followed the conditional joint model (CJM) specification by Lipsitz and Ibrahim (1996), which suggested a conditional approach where the joint distribution of *p*-dimensional moderator vector $\boldsymbol{X}_i = (X_{i1}, \ldots, X_{ip})$ is modeled as

$$\begin{gathered} p\left(X_{i1}, \ldots, X_{ip} \middle| \boldsymbol{\alpha}\right) \\ = p(X_{ip}|X_{i1}, \ldots, X_{ip-1}, \boldsymbol{\alpha}_p) \times p\left(X_{ip-1} \middle| X_{i1}, \ldots, X_{ip-2}, \boldsymbol{\alpha}_{p-1}\right) \times \ldots \times p(X_{i1}|\boldsymbol{\alpha}_i), \end{gathered} \tag{19}$$

where $\boldsymbol{\alpha}_p$ is the vector of parameters for the *p*th conditional distribution, $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_p)$. The purpose of this model is to estimate the parameter vector of the regression coefficients of the conditional distribution of the response variable given the moderator vector X.

Equation (19) outlines this model, in which there is a conditional distribution of the response variable given the moderator vector X, and then a sequence of one-dimensional conditional distributions are specified for *X (*Lipsitz and Ibrahim 1996). This model was proposed exclusively for missing categorical data models.

This work was extended to include a Monte Carlo version using the EM algorithm to obtain the maximum likelihood estimate and was implemented through the use of the Gibbs sampler (Ibrahim, Chen, & Lipsitz, 1999). Only moderators that are missing are modeled. The moderators that are completely observed are used in the conditional distribution when constructing the distribution of the missing moderators (Ibrahim, Chen, Lipsitz, & Herring, 2005).

The CJM approach has been shown to perform better than listwise deletion (Ibrahim et al., 1999). Hemming et al. (2010) compared the CJM approach to listwise and pairwise deletion. Deviance Information Criterion (DIC) was used to compare the models, where models with smaller DIC values among models with the same data are considered to have better fit. Although Hemming et al. (2010) stated their proposed model performed the best, DIC values were unclear and hard to interpret. However, comparable models were not presented with the DIC values (Hemming et al. 2010) (2010). Thus, it may not be possible to directly compare the models because the sample sizes are different.

Lastly, a simulation was not performed which may have been helpful when comparing the CJM model to listwise deletion. Simulation studies produce empirical results about statistical methods in certain situations or scenarios (Morris, White, & Crowther, 2019). This is beneficial when evaluating new methods, such as the proposed missing moderator method, and comparing alternative methods.

The focus of this dissertation was to examine the conditional joint modeling approach illustrated by Hemming et al. (2010) and proposed by Ibrahim et al., (1999) under known conditions using computer simulations. Specifically, I aim to answer the following two questions:

1. How does the CJM method perform under various conditions?
2. How does the CJM method compare to listwise deletion?

CHAPTER III

METHOD

As mentioned in the literature review, few methods exist for handling missing moderators in meta-analysis other than classical approaches. One approach, proposed by (Ibrahim et al., 1999) and later illustrated by Hemming et al. (2010) used a CJM approach. The purpose of this chapter is to present the key components of the research including the simulation conditions, data generation, and the evaluation procedures used to assess the effectiveness of the CJM imputation.

Data Generation

A simulation study is a valuable tool when examining new statistical methods. In simulation studies, data are created by pseudo-random sampling from known probability distributions (Morris et al., 2019). Data were generated to simulate multiple meta-regression models with a continuous outcome variable and two covariates for this dissertation.

Mean differences were used as the effect size of interest. First, the standardized mean difference (SMD), or d, was estimate for study i, where $i = 1,\ldots,K$:

$$d_i = \frac{\bar{Y}_i^T - \bar{Y}_i^C}{S_i^P}, \qquad (20)$$

where $\bar{Y}_i^T$ and $\bar{Y}_i^C$ are the treatment and control means, respectively, and S_i^P is the pooled standard deviation. Let n^C and n^T be the sample sizes in the two groups, and $(S^T)^2$ and $(S^C)^2$ be the sample standard deviations of the two groups, respectively. Then,

$$S_i^P = \sqrt{\frac{(n_C - 1)(S^C)^2 + (n_T - 1)(S^T)^2}{n_C + n_T - 2}}. \qquad (21)$$

The variance of the i^{th} value of d is estimated by

$$S_{d_i}^2 = \frac{n_i^T + n_i^C}{n_i^T n_i^C} + \frac{d_i^2}{2(n_i^T + n_i^C)}. \tag{22}$$

However, d is slightly biased (Borenstein et al., 2011) and tends to overestimate the absolute bias of the mean differences, particularly provided small within-study sample sizes. A correction factor, J, was applied to convert d to Hedges' g:

$$J = 1 - \frac{3}{4(df - 1)}, \tag{23}$$

where

$$df = (n_C + n_T - 2). \tag{24}$$

The computation for Hedges' g is

$$g = J * d. \tag{25}$$

Hedges' g was used as the overall effect size, denoted as β_0 in Equation (10), and was set as a medium effect size of 0.5 (i.e., β_0 = 0.5; Lopez-Lopez, Botella, Sanchez-Meca, & Marin-Martinez, 2013). The use of a fixed effect size is common in mixed-effects meta-regressions in simulation studies (see Lopez-Lopez et al., 2013; Viechtbauer, López-López, Sánchez-Meca, & Marín-Martínez, 2015). The within-study sample sizes were also fixed to $n_C = n_T = 250$ per group, or N_{Total} = 500 for ease of simulation computing.

Similar to Viechtbauer et al. (2015), the standardized coefficient values for the covariates X_{i1} and X_{i2} were set to reflect low ($\beta_p = 0.05$), medium ($\beta_p = 0.20$), and high ($\beta_p = 0.50$) values. The study characteristic values ($\boldsymbol{X}_{ip}$) were generated from a multivariate normal distribution with a sample size equal to the number of effect sizes (K) and a mean equal to a vector of zeros. The standard deviation component of this distribution was equal to a 2x2 correlation matrix,

$$\begin{bmatrix} 1 & 0.30 \\ 0.30 & 1 \end{bmatrix}, \quad (26)$$

where 0.30 is the correlation between X_{i1} and X_{i2}. The generated values of x_{i1} and x_{i2} were used to produce values for the individual effect sizes, θ_i,

$$\theta_i = (\beta_1 \times X_{i1}) + (\beta_2 \times X_{i2}) + \omega, \quad (27)$$

with

$$\omega \sim N(0,1). \quad (28)$$

The effect size θ_i in Equation (27) was used in the generation for the control and treatment means. The treatment mean was generated from a normal distribution with a sample size equal to 250 and the mean equal to 0.50 plus the generated effect size θ_i and the generated random effect. The variance of the treatment mean generation was equal to one. The control group mean was also generated from a normal distribution with a sample size equal to 250, mean equal to zero plus the generated effect size θ_i, and a variance of one.

The random effect was generated from a normal distribution with a sample size equal to *K*, a mean of 0, and a variance equal to the between-studies variance, τ^2, that was set to reflect low (0.01), medium (0.40) and high (1.00) values. The sample size of the meta-analysis, *K*, was manipulated to be 20 or 80 to simulate small and large meta-analyses.

Lastly, the conditions were reflected to contain 15%, 30% or 50% missing values for X_2. All data were generated and analyzed using R version 1.2.5033 (R Studio Team, 2020). This procedure was used to create *k* rows of data that contained the mean and SD of the control and treatment groups, effect sizes and variances, and the mean of the study characteristics X_1 and X_2. This can be thought of as a single individual meta-analytic study. Each individual meta-analysis was a "complete dataset" as it contained all the simulated values.

Missing Data and Meta-Analysis Procedure

A total of 3 ($\boldsymbol{\tau}^2$) × 2 (*K*) × 3 ($\boldsymbol{\beta}_1$) × 3 ($\boldsymbol{\beta}_2$) × 3 (% Missing) × 2 (Missing Procedures) provided 324 simulation conditions. For each of these conditions, 500 datasets were simulated. The simulation was conducted in R (R Core Team, 2020) and JAGS (Plummer, 2003), using several packages including *metafor* (Viechtbauer, 2010), *rjags* (Plummer, 2019), *coda* (Plummer, Best, Cowles, & Vines, 2006), and *jagsUI* (Kellner, 2019). **Error! Reference source not found.** Table 1 outlines all the conditions in the simulation.

Missing Data Procedure

Missing data for X_2 were generated with a missing data mechanism of MAR. This implies that the missing data mechanism can be ignored in estimating the missing parameters. Each generated dataset was specified to have 15%, 30%, and 50% missing. The same data set was used for all three percent missing conditions. This was completed utilizing the *mice* (Van Buuren & Groothuis-Oudshoorn, 2011) package in R. As stated earlier, these datasets were analyzed using two procedures: 1) listwise deletion, and 2) CJM procedure.

Overall Meta-Analysis Procedure

A meta-analysis was conducted on each dataset utilizing the *metafor* package (Viechtbauer, 2010) in R. Study characteristics X_1 and X_2 were included as moderators in the analysis. The between-studies variance was estimated using restricted maximum likelihood in *metafor* and Bayesian estimation in *rjags*.

Missing Data Methods

Listwise Deletion

The listwise method was implemented in the *metafor* package in R, which automatically eliminates all studies with missing data. The remaining full cases were analyzed using restricting maximum likelihood.

Conditional Joint Modeling

The CJM method was implemented using Bayesian techniques. Several R packages were used, including *rjags* (Plummer, 2019), *coda* (Plummer, Best, Cowles, & Vines, 2006), and *jagsUI* (Kellner, 2019). Estimates were based on iterations of length 100,000 after an initial burn-in length of 50,000. MCMC convergence was established to be enough after examination of traceplots, autocorrelations, and the Geweke statistic.

Choices of prior distributions were based on Hemming et al. (2010). Priors for β_0, β_1, and β_2 were specified to have a normal distribution with a mean of 0 and a variance of 100. The conditional densities for the moderators X_1 and X_2 were specified to have normal distributions with a mean γ_{10} for X_1 and $\gamma_{20} + \gamma_{21}X_1$ for X_2 and a variance

$$\frac{1}{\sigma_{xp} * \sigma_{xp}}, \tag{29}$$

where

$$\sigma_{xp} \sim U(0, 100) \tag{30}$$

The terms γ_{10} and γ_{20} were specified to have normal priors with a mean of 0 and a variance of 100. A conditional joint distribution was only considered for X_2 because it was the only moderator that had missing values. The moderator X_2 was modeled by $p(X_2|X_1)$, where the missing values of X_2 were conditional on X_1.

Evaluation Procedure

After obtaining the estimated missing values, data sets were evaluated using several assessment and diagnostic methods: relative bias, mean squared error (MSE), and 95% coverage. Parameter estimates accuracy was evaluated through relative bias and MSE. Relative bias measures the degree to which the average of the simulated value of interest exceeds the true value of the parameter. This was calculated as

$$\frac{\mathrm{E}[\hat{\upsilon}] - \upsilon,}{\mathrm{E}[\hat{\upsilon}]} \times 100, \quad (31)$$

where $\hat{\upsilon}$ reflects the average of the simulated parameters of interest for each missing condition (i.e., β_0 or Hedges' g, β_1, and β_2), and υ represents the true value of the parameter, which is the fixed value used in the simulation design. A negative relative bias indicates that the estimated value is smaller than the true value of the parameter (i.e., underestimated). An overestimated parameter has a positive relative bias, where the estimated value is larger than the true parameter value.

MSE represents efficiency when the parameter estimate is unbiased. MSE values can be compared to determine the relative efficiency. These values must be from the same parameter but can be from different simulation conditions. MSE is calculated as

$$\mathrm{E}[(\hat{\upsilon} - \upsilon)^2]. \quad (32)$$

MSE is a measure of how close an estimated value is to the true value (Walther & Moore, 2005). A large MSE indicates less accuracy in the estimate.

Lastly, coverage is determined by the proportion of 95% confidence or credible intervals that contained the true population value for each parameter. This is defined as

$$Pr(\hat{\upsilon}_{Low} \leq \upsilon \leq \hat{\upsilon}_{Upper}), \quad (33)$$

where $\hat{\upsilon}_{Low}$ and $\hat{\upsilon}_{Upper}$ are the lower and upper limits of the interval. Coverage measures how well the parameter and standard errors are estimated. According to Burton, Altman, Royston, & Holder (2006), coverage should not fall outside two standard errors (SE) of the nominal coverage probability, which is calculated as

$$\sqrt{\frac{E(1-E)}{B}}, \tag{34}$$

where E is equal to 0.95 and B is the number of replications in the simulation (i.e., 500). In this dissertation, adequate coverage should be between 0.94 and 0.96.

Each of these evaluation procedures are discussed for both the effect size, g, and the regression coefficient for the missing moderator, β_2. Furthermore, a mixed-effects analysis of variance (ANOVA), with two within-subject factors and four between-subject factors, was used to determine the effects of the design factors on bias and MSE. Missing data method and missing percentage were treated as within-subjects' factors because each simulated meta-analysis dataset was analyzed using both methods (i.e. listwise deletion and the CJM approach) and was simulated to have 15%, 30%, and 50% missing data.

Furthermore, five two-way interactions were calculated:

1. $\tau^2 \times missing\ data\ method,$
2. $K \times missing\ data\ method,$
3. $\beta_1 \times missing\ data\ method,$
4. $\beta_2 \times missing\ data\ method,$
5. $Missing\ percentage \times missing\ data\ method$

The mixed-effects ANOVA also included 10 three-way interactions:

1. $K \times \tau^2 \times missing\ data\ method,$

2. $K \times \beta_1 \times missing\ data\ method,$

3. $K \times \beta_2 \times missing\ data\ method,$

4. $K \times missing\ percentage \times missing\ data\ method$

5. $\tau^2 \times \beta_1 \times missing\ data\ method$

6. $\tau^2 \times \beta_2 \times missing\ data\ method$

7. $\tau^2 \times missing\ percentage \times missing\ data\ method$

8. $\beta_1 \times \beta_2 \times missing\ data\ method$

9. $\beta_1 \times missing\ percentage \times missing\ data\ method$

10. $\beta_2 \times missing\ percentage \times missing\ data\ method$

Following Luo and Kwok (2012), partial eta-squared (η_p^2) greater than 0.01 effect size will be reported for the ANOVA rather than the *p*-value of the *F* test.

CHAPTER IV

RESULTS

Performance of Bayesian Estimation

Several diagnostic measures were used to assess the convergence performance of the Bayesian CJM approach. The Geweke diagnostic was used to assess the adequacy of the burn-in length. Values between the first chain and the second chain suggested adequate burn-in length. Additionally, autocorrelation was used to examine if values were statistically independent during the MCMC iterations. Assessment of autocorrelation revealed a rapid decline in values as the number of lags increased, suggesting minimal-to-no issues via MCMC convergence.

Relative Bias

Relative bias for β_2

Simulation results assessing estimated relative bias for the true parameter of β_2 are provided in Table 2. Specifically, Table 2 shows differences among relative bias for both missing data methods across values of τ^2, *K*, β_1, β_2, and percentage missing. Overall, the CJM approach and listwise deletion showed considerable negative bias. Further examination of Table 2 showed that relative bias ranged from -1.357 and -0.023. Unexpectedly, these values both occurred when listwise deletion was implemented as the missing data method with 50% missing data at *K* = 20. The lowest value (i.e., -1.357) occurred when τ^2= 1.00 and β_1 was 0.5 β_2 was 0.05 and the maximum value (i.e., -0.023) occurred when τ^2= 0.40, β_1 = 0.05, and β_2= 0.05.

A two-way interaction between *K* and missing data method accounted for approximately 11% of the error variance, η_p^2 = 0.107. As shown in Figure 1, relative bias was better for listwise deletion than it was for the CJM approach when there were 20 studies in the meta-analysis. Relative bias at 20 studies was around -0.42 for the CJM approach and -0.38 for listwise deletion. However, when there was an increase in *K*, the CJM approach performed better than listwise deletion, with relative bias improving to approximately -0.3 for the CJM approach and to approximately -0.35 for listwise deletion.

The interaction between the true value of τ^2 and the missing data method interaction accounted for about 11% of the error variance, η_p^2 = 0.111. Figure 2 shows that relative bias decreased for both missing data methods when there was an increase in τ^2 to 0.04. However, relative bias increased when τ^2 increased from 0.4 to 1.

Additionally, a two-way interaction between the missing data method and the true value of β_1 accounted for approximately 4% of the error variance, η_p^2 = 0.038; while a two-way interaction between the missing data method and the true value of β_2 explained approximately 3% of the error variance, η_p^2 = 0.029.

Figure 3 illustrates the interaction between the missing data method and the missing percentage, η_p^2 = 0.045. Relative bias for both missing data methods was approximately -0.3 at 15% missing data and -0.35 at 30% missing data. When missing percentage increased to 50%, relative bias increased to -0.6 for the CJM approach and approximately -0.4 for listwise deletion.

A three-way interaction between the true value of τ^2, the number of studies *K*, and the missing data method explained approximately 7% of the error variance, η_p^2 = 0.071. Figure 4 illustrates this three-way interaction and shows that relative bias slightly decreased for both missing data methods when the true value of τ^2 increased from 0.01 to 0.40, regardless of the number of studies *K*. Though, there was an increase in relative bias when the true value of τ^2 increased from 0.40 to 1.0, except when the CJM approach was implemented with 80 studies. Instead, the CJM approach with 80 studies slightly decreased in relative bias.

A three-way interaction between the number of studies *K*, true value of β_1, and the missing data method explained 3% of the error variance, η_p^2 = 0.030. Figure 5 shows that when *K* = 20 for both missing data methods, relative bias increased in conjunction with the increase of the true value of β_1. When K = 80 for both missing data methods, average relative bias did not change when there was an increase in the true value of β_1. However, when examining Table 2, this trend did not appear as prevalent. This may be because when β_1= 0.5, relative bias reaches as high as -0.730.

Examination of a three-way interaction between *K*, the true value of β_2, and the missing data methods suggested that relative bias decreased as there was an increase in the true value of β_2. This interaction explained 2% of the error variance, η_p^2 = 0.022.

Additionally, the interaction between *K*, the missing percentage, and the missing data method also explained 2% the error variance, η_p^2 = 0.023. Overall, relative bias increased as there was an increase in the missing percentage, except when there were 80 studies, when implementing listwise deletion. When the CJM method had 80 studies included in the meta-analysis, relative bias was lower than when the CJM method included only 20 studies. However, this trend was not observed when examining changes in missing percentage as both *K* and missing percentage increased for listwise deletion. When there was 50% missing data, listwise deletion with 80 studies decreased in relative bias. This trend can be seen in Figure 6.

Results suggest that the three-way interaction between τ^2, β_1, and missing data method explained approximately 14% of the error variance (η_p^2 = 0.136) and that three-way interaction between τ^2, β_2, and missing data method explained approximately 17% of the error variance, (η_p^2= 0.171). However, examination of Table 2 indicated that this may be due to the extreme values found when β_2= 0.5. ANOVA results also suggested that the three-way interaction between β_1, β_2, and missing data method explained 10% (η_p^2= 0.10) of the error variance, and that the interaction between β_1, missing percentage, and missing data method explained approximately 3% (η_p^2= 0.030) of the error variance.

Relative bias for Hedges' g

Like the performance of β_2, the CJM approach and listwise deletion underestimated the true parameter of Hedges' g, on average. Table 3 showcases the differences among relative bias for both missing data methods across values of τ^2, *K*, β_1, β_2, and percentage missing. Examination of Table 3 showed that relative bias was larger for listwise deletion. Additionally, the largest relative bias values occured when τ^2 = 1.00 and listwise deletion was implemented.

As shown by Table 3, relative bias slightly decreased when *K* increased to 80 studies for the CJM approach. However, when *K* was 80 studies for listwise deletion, there was barely any change in relative bias. Additionally, relative bias for the CJM approach was -0.04, on average, but was about -0.16, on average, for listwise deletion. This interaction between *K* and the missing data method explained almost 99% of the error variance, η_p^2 = 0.988.

Estimated relative bias values for varying degrees of τ^2 for the two missing data methods are given in Figure 7, which illustrated that as the value of τ^2 increased for each missing data method, the relative bias became more negative. This finding was supported with the ANOVA results, which indicated the interaction between the true value of τ^2 and the missing data method explained 97% of the error variance, η_p^2 = 0.967. Specifically, as the value of τ^2 increased, the relative bias increased for both missing data methods. In comparison, when τ^2 = 0.4, values of relative bias were approximately 0.00 for the CJM approach and -0.1 for listwise deletion. Greater differences were observed when τ^2 increased to 1.0. Specifically, the CJM approach had relative bias around -0.1, while listwise deletion had relative bias values around -0.4, on average.

ANOVA results also indicated that the percentage of missing data paired with the missing data method explained about 29% of the error variance, η_p^2 = 0.291. As Figure 8 illustrates, relative bias values for the CJM approach were less than -0.1 for all conditions of missing percentage. However, relative bias values became more negative when missing percentage increased for listwise deletion. At 30% missing, relative bias values for listwise deletion were approximately -0.15, and at 50% missing, relative bias values were approximately -0.25.

The two-way interaction between missing data method and the true value of β_1 explained approximately 11% of the error variance, η_p^2 = 0.109. Examination of Table 3 showed that there was no considerable change for the CJM approach when there was an increase in β_1. Relative bias for the CJM approach was around -0.05 for all conditions of β_1. Instead, differences in relative bias emerged for the listwise deletion approach. As illustrated by Figure 9, relative bias slowly decreased as the true value of β_1 increased. Relative bias was approximately -0.2 when β_1 was 0.05 for listwise deletion, then around -0.18 when β_1 = 0.2, and then around -0.15 when β_1 = 0.5.

Examination of Table 3 suggested an interaction between *K*, missing data method, and τ^2, which was supported by the ANOVA results, η_p^2 = 0.233. Relative bias slightly increased for both missing data methods, when τ^2 increased from 0.01 to 0.4, regardless of number of studies *K*, and when τ^2 increased from 0.4 to 1.0. Figure 10 illustrates the three-way interaction, which also shows a considerable difference in relative bias between listwise deletion and the CJM approach. Specifically, relative bias for the CJM approach at *K* = 20 and at *K* = 80 increased from approximately 0.00 to approximately -.05 when τ^2 increased to 0.4. In comparison, relative bias for listwise deletion increased from approximately 0 to -0.1 for listwise deletion at *K* = 20 and at K = 80. When τ^2 increased to 1.0, relative bias for both conditions of K increased to -0.1, on average, for the CJM approach, which is an increase of about .05. Relative bias for listwise deletion increased to -0.4, which is an increase of about 0.3.

Table 3 shows that the CJM approach had lower relative bias values than listwise deletion when examining by K and the true value of β_2. This three-way interaction explained 7% of the error variance, η_p^2 = 0.072. In Table 3, when β_1= 0.05 and β_2 = 0.05 at K = 20, relative bias values for the CJM approach are - 0.001, - 0.022, and - 0.079, and for listwise deletion are higher at -0.005, -0.073, and -0.229. Figure 11 illustrates this interaction, showing that the CJM approach had lower relative bias, overall, than does listwise deletion. Additionally, Figure 11 also illustrates that relative bias decreased as the true value of β_2 increased, for both missing data methods and conditions of K.

Examination of Table 3 suggested a three-way interaction between K, missing data method, and missing percentage, which was supported by the ANOVA results, η_p^2 = 0.926. Table 3 shows that, overall, listwise deletion had higher relative bias than did the CJM approach at all conditions of K and missing percentage. Figure 12 also illustrates this interaction, which shows that the CJM approach at both conditions of K did not have relative bias values that were greater than -0.1. However, when K = 80 for the CJM approach, relative bias values were slightly larger than were relative bias values when K = 20 for the CJM approach. Generally, as both missing percentage and K increased, relative bias values increased for the CJM approach.

Similar results were found when examining listwise deletion, however with greater relative bias values. As both missing percentage and K increased for listwise deletion, relative bias values got considerably more negative. As illustrated by Figure 12, when missing percentage is 15, relative bias values for listwise deletion at both K conditions were -0.1. As missing percentage increased to 30, relative bias was about -0.15 for both *K* conditions, and then about -0.25 when missing percentage was 50.

The three-way interaction between missing data method, the true value of β_1 and τ^2 explained approximately 20% of the error variance, η_p^2 = 0.195. Table 3 shows that as τ^2 and β_1 increased, relative bias improved for both missing data methods, and that listwise deletion had more negative relative bias values than did the CJM approach. For example, when τ^2 when 0.1 and β_1 = 0.05, relative bias ranged from -0.002 to approximately 0.000 for the CJM approach and ranged from -0.004 to -0.007 for listwise deletion. This interaction is also illustrated by Figure 13, which shows that, on average, relative bias values for the CJM approach at any value of τ^2 did not reach greater than -0.1, but was as large as -0.4 for listwise deletion when τ^2 = 1.0. Furthermore, Figure 13 shows that as the true value of β_1 increased, relative bias slightly decreased for all conditions of τ^2 and missing data methods.

A similar relationship occurred when examining the three-way interaction between missing data method, τ^2, and β_2 which was supported by the ANOVA results, η_p^2 = 0.128. Figure 14 shows this three-way interaction, illustrating that relative bias was largest when τ^2 = 1.0 and the missing data method was listwise deletion. Specifically, Figure 14 shows that relative bias decreased as β_2 increased, but that relative bias increased as τ^2 increased.

Examination of Table 3 suggested a three-way interaction between missing data, τ^2, and missing percentage, which was supported by the ANOVA results, η_p^2 = 0.800. Specifically, as missing percentage increased, relative bias increased when τ^2 was 0.40 for the CJM approach, and when τ^2 was 0.40 and 1.0 for listwise deletion. This three-way interaction is also shown in Figure 15.

ANOVA results suggested that a three-way interaction between missing data method, β_1, and β_2 explained approximately 8% of the error variance, η_p^2 = 0.084. Specifically, as β_2 increased for the CJM approach, there was a slight decrease in relative bias regardless of the true value of β_1. That is, relative bias was relatively the same for all values of β_1 and β_2 for the CJM approach. However, relative bias was slightly smaller for larger values of β_1 for listwise deletion. Additionally, relative bias also decreased as β_2 increased. That is, as β_1 and β_2 both increased, relative bias for listwise deletion decreased.

Lastly, ANOVA results suggested a three-way interaction between missing data method, as β_1, and missing percentage, η_p^2 = 0.043. Table 3 shows that when there was 15% missing data and as β_1 = 0.05, the CJM approach had lower relative bias values than did listwise deletion. Specifically, when was β_1 = 0.05, relative bias values ranged from – 0.099 to 0.015 for the CJM approach but ranged from -0.635 to 0.000 for listwise deletion. As missing percentage increased, no noticeable change was observed in relative bias values for the CJM approach, regardless of as β_1 values. Not only did relative bias increase when missing percentage increased, but that relative bias was larger for smaller values of as β_1.

Mean Squared Error

MSE for β_2

The simulation results for assessing the estimated MSE for the true parameter of β_2 for the various conditions are provided in Table 4. ANOVA results indicated that several conditions explained the variation in observed MSE values for β_2. First, the interaction between true value of τ^2 and missing data method explained approximately 46% of the error variance, η_p^2 = 0.462. As illustrated by Table 4, when the true value of τ^2 equaled 0.001, MSE values were the highest for both the CJM approach (*M* = 0.021, *SD* = 0.040) and listwise deletion (*M* = 0.021, *SD* =

0.040). The average MSE value decreased when τ^2 increased to 0.40 with a mean of 0.010 (*SD* = 0.011) for the CJM approach and a mean of 0.010 (*SD* = 0.12) for listwise deletion. However, MSE values increased for both missing data methods when τ^2 increased to 1.0 with a mean of 0.012 for both the CJM approach and listwise deletion (CJM, *SD* = 0.014; listwise deletion, *SD* = 0.13).

Second, the true value of β_1 explained approximately 57% of the error variance in MSE values for β_1, η_p^2 = 0.570. On average, MSE values were smaller when the true value of β_1 was low. The average MSE when β_1 = 0.050 was 0.010 (*SD* = 0.011) for the CJM approach and was 0.011 (*SD* = 0.12) for listwise deletion. MSE stayed relatively the same when β_1 was 0.200 with an average MSE of 0.010 (*SD* = 0.012) for the CJM approach and 0.011 (*SD* = 0.012) for listwise deletion, but increased when β_1 increased to 1 for both the CJM (*M* = 0.023, *SD* = 0.040) and listwise deletion (*M* = 0.022, *SD* = 0.040). Overall analysis indicated that MSE values for this two-way interaction were relatively the same for both missing data methods.

Furthermore, the true value of β_2 explained approximately 25% of the error variance in MSE values for β_2, η_p^2 = 0.254. Like β_1, MSE values were smaller when the true value of β_2 was low. The average MSE when β_2 = 0.05 was approximately 0.000 (*SD* ≈ 0.000) for the CJM approach and was approximately 0.000 (*SD* = 0.001) for listwise deletion. MSE slightly increased when β_2 increased to 0.200 with an average MSE of 0.005 (*SD* = 0.002) for the CJM approach and 0.005 (*SD* = 0.001) for listwise deletion. MSE increased again when β_2 increased to 1 for both the CJM approach (*M* = 0.038, *SD* = 0.034) and listwise deletion (*M* = 0.038, *SD* = 0.033). Overall analysis indicated that MSE values for this two-way interaction were relatively the same for both missing data methods.

ANOVA results suggested that the three-way interaction between missing data method, τ^2, and β_1 explained approximately 33% of the error variance, η_p^2 = 0.331. As shown by Table 4 and Figure 16, when τ^2 was 0.01 for both missing data methods, MSE values were approximately 0.01 when β_1 = 0.05 and 0.2. However, when β_1 increased to 0.5 and τ^2 was 0.01, MSE values for the CJM approach and listwise deletion increased to about 0.045. Moreover, when τ^2 increased to 0.04 for both missing data methods, MSE values were also approximately 0.01 for all conditions of β_1. When β_1 increased to 0.5 and τ^2 was 0.1.0, MSE values for the CJM approach and listwise deletion slightly increased from 0.01 to about 0.0125.

Similarly, ANOVA results suggested that the three-way interaction between missing data method, τ^2, and β_2 explained approximately 28% of the error variance, η_p^2 = 0.279. Table 4 shows that regardless of the value of τ^2, MSE for the CJM approach increased as the true value of β_2 increased. Though, Figure 17 shows that MSE values varied for different conditions of τ^2 and β_1 for listwise deletion. Specifically, MSE values were similar at β_2= 0.05 and 0.2 for all conditions of τ^2. However, when β_2 increased to 0.5 and τ^2 was 0.01, MSE values increased substantially increased from below 0.01 to about 0.06 when β_2 increased to 0.5.

Lastly, examination of Table 4 suggested a three-way interaction between missing data method, β_1 and β_2, which was supported by the ANOVA results, η_p^2 = 0.329. Particularly, MSE values increased when β_1 and β_2 increased for both missing data methods. The increase in MSE values was relatively similar for both missing data methods, until β_1 and β_2 were 0.50, then MSE was approximately 0.01 greater for listwise deletion than it was for the CJM approach.

MSE for Hedges' g

Table 5 provides the performance of MSE values for Hedges' g for the various conditions. ANOVA results indicated differences among MSE values for several conditions for Hedges' g.

First, the interaction between *K* and missing data method explained approximately 97% of the error variance, η_p^2 = 0.967.

From Table 5, when comparing MSE values by *K* and the missing data method, listwise deletion had considerably higher MSE values than did the CJM approach. When *K = 20*, listwise deletion had an average MSE value of 0.016 (*SD* = 0.028), while the CJM approach had an average MSE of 0.001 (*SD* = 0.002). When *K* increased to 80, listwise deletion still had higher MSE values than the CJM approach with a mean of 0.017 (*SD* = 0.028) for listwise deletion and a mean of 0.001 (*SD* = 0.001) for the CJM approach.

Examination of Table 5 also suggested a two-way interaction between the missing data method and the true value of τ^2, which was supported by the ANOVA, η_p^2 = 0.950. As can be seen by Figure 18, MSE values for both the CJM approach and listwise deletion were approximately 0 when τ^2 was 0.01. However, MSE values began to increase for the listwise deletion missing data method when the value of τ^2 increased to 0.4 and 1.0. When τ^2 = 0.40, the average MSE was 0.000 (*SD* ≈ 0.000) for the CJM approach and was 0.004 (*SD* = 0.003) for listwise deletion. When τ^2 = 1.00, the average was 0.002 (*SD* = 0.002) and 0.045 (*SD* = 0.033) for the CJM approach and listwise deletion, respectively.

A two-way interaction between the missing data method and the true value of β_1 accounted for approximately 7% the error variance, η_p^2 = 0.073. Figure 19 shows the differences in MSE for both missing data methods across the varying values of β. MSE values for all three β_1 conditions for the CJM approach were approximately 0.00, however MSE values for listwise deletion were approximately 0.002. Additionally, Figure 19 shows a slight decrease in MSE values for listwise deletion when there was an increase in β_1.

ANOVA results also suggested a two-way interaction between the missing data method and the percentage missing, η_p^2 = 0.032. Examination of Table 5 across missing percentage between the missing data methods suggested that MSE values increased for listwise deletion as missing percentage increased. MSE values were lowest for both missing data methods when missing percentage was 15%. The average MSE for the CJM approach at 15% missing data was approximately 0.000 (*SD* = 0.001) and for listwise deletion was 0.004 (S*D* = 0.006). When missing percentage increased to 30%, the CJM approach slightly increased to approximately 0.001 (*SD* = 0.001), while listwise deletion increased to 0.013 (*SD* = 0.016). When missing percentage increased to 50%, the average MSE for the CJM approach stayed the same at 0.001 (*SD* = 0.003) but increased again for listwise deletion to 0.032 (*SD* = 0.041).

The three-way interaction between *K*, the missing data method, and τ^2 accounted for approximately 4% of the error variance, η_p^2 = 0.037. Table 5 shows that overall MSE values were smaller for the CJM approach than they were for listwise deletion regardless of *K* or τ^2. Figure 20 shows differences among MSE values for the CJM approach when *K* was 20 versus when *K* was 80. Specifically, MSE values were slightly larger the CJM approach when *K* was 20 than when *K* was 80 for the CJM approach. Additionally, MSE increased for the CJM approach when τ^2 increased to 1. In terms of listwise deletion, Figure 20 shows minimal differences between the various conditions of *K*, however MSE values increased whenever there was an increase in τ^2 values. There was a substantial jump of about 0.045 in MSE for listwise deletion when there was an increase in τ^2 from 0.4 to 1.0.

The three-way interaction between *K*, the missing data method, and β_2, accounted for approximately 2% of the error variance, η_p^2 = 0.015. Illustrated by Figure 21, we can see that average MSE values for the CJM approach got slightly smaller as there was an increase in the true value of β_2, regardless of the value of *K*. MSE values were also slightly larger when *K* was 20 for the CJM approach. When examining listwise deletion, MSE values were larger than they were for the CJM approach. Additionally, MSE values were larger when *K* was 80 for listwise deletion than when *K* was 20. Like the CJM approach, MSE values slightly decreased as β_2 increased regardless of *K*. However, most notable in Figure 21 is that MSE values were approximately 0 for listwise deletion when β_2 = 0.05 and there were 20 studies. When β_2 increased to 0.2, MSE values jumped to around 0.015.

A three-way interaction between *K*, the missing data method, and missing percentage accounted for approximately 93% of the error variance η_p^2 = 0.934. Figure 22 shows that both the largest MSE values for both missing data methods occurred when *K* = 80 and there was 50% missing data. The CJM approach had lower MSE values than did listwise deletion.

ANOVA results also suggested that the three-way interaction between missing data method, τ^2, and β_1 explained approximately 21% of the error variance, η_p^2 = 0.206. As shown by Table 5 and Figure 23, MSE values were largest when listwise deletion was implemented and τ^2 was 1.0. Figure 23 also illustrates that as τ^2 increased for listwise deletion, there was also an increase in MSE values. However, as β_1 increased, MSE values decreased for all conditions of τ^2 for listwise deletion. In terms of the CJM approach, MSE values were lower than 0.01 regardless of τ^2 value or β_1.

Similarly, ANOVA results suggested that the three-way interaction between missing data method, τ^2, and β_2 explained approximately 18% of the error variance, η_p^2 = 0.176. As shown by Table 5 and Figure 24, MSE values were largest when listwise deletion was implemented and τ^2 was 1.0. Figure 24 also illustrates that as τ^2 increased for listwise deletion, there was also an increase in MSE values. However, as β_1 increased, MSE values decreased for all conditions of

τ^2 for listwise deletion. In terms of the CJM approach, MSE values were lower than 0.01 regardless of τ^2 value or β_1.

A three-way interaction between missing data method, τ^2, and missing percentage accounted for approximately 90% of the error variance, η_p^2 = 0.901. As illustrated by Figure 25, MSE values were largest when listwise deletion was implemented with a τ^2 of 1.0 and 50% missing data. When missing percentage was 15%, MSE for listwise deletion at all values of τ^2 ranged from approximately 0.00 to 0.01. When missing percentage was 15% for the CJM approach, MSE values were approximately 0 for all values of τ^2. When missing percentage increased to 30%, MSE values slightly increased for the CJM approach but were still below 0.01. MSE values for listwise deletion at 30% missing data were below 0.01 for listwise deletion when τ^2 was 0.01 and 0.40 but were approximately 0.035 when τ^2 was 1.0. Values of MSE were larger for listwise deletion when τ^2 was 0.40 than they were for listwise deletion when τ^2 was 0.01. As the value of missing percentage increased to 50, MSE values for all conditions of missing data method and τ^2 increased but were not below 0.01 for any condition other than listwise deletion when τ^2 was 1.

Examination of Table 5 suggested a three-way interaction between missing data method, β_1 and β_2, which was supported by the ANOVA results, η_p^2 = 0.069. In general, Figure 26 shows that MSE values were larger for listwise deletion than they were for the CJM approach. Regardless of β_1 or β_2 values, MSE was approximately 0 for the CJM approach. However, when β_1 and β_2 were both small for listwise deletion, MSE values were the largest. For example, when β_1 was 0.05 and β_2 was 0.05, Figure 26 shows that the average MSE for listwise deletion was around 0.019. However, when β_1 was 0.50 and β_2 was 0.05, MSE values for listwise deletion were around 0.015. As β_2 increased, MSE values decreased for all values of β_1 for listwise deletion.

Examination of Table 5 suggested a three-way interaction between missing data method, β_1 and missing percentage, which was supported by the ANOVA results, η_p^2 = 0.062. Particularly, MSE values increased when β_1 and missing increased for listwise deletion but not for the CJM approach. MSE values were approximately 0 for all values of β_1 and missing for the CJM approach. In comparison, MSE values for listwise deletion across the varying values of β_1 and missing percentage ranged from approximately 0.005 to 0.035.

Coverage

Coverage rates for β_2

Table 5 provides coverage performance for β_2 for the various conditions. ANOVA results indicated differences among coverage for several conditions for β_2. First, the interaction between *K* and missing data method explained approximately 24% of the error variance, η_p^2 = 0.236.

From Table 5, when comparing coverage by *K* and the missing data method, listwise deletion had considerably lower coverage than did the CJM approach. When $K = 20$, listwise deletion had an average coverage of 0.849 (*SD* = 0.092), while the CJM approach had an average coverage of 0.875 (*SD* = 0.099). When *K* increased to 80, listwise deletion still had lower coverage than the CJM approach with a mean of 0.830 (*SD* = 0.110) for listwise deletion and a mean of 0.0841 (*SD* = 0.108) for the CJM approach.

Examination of Table 5 also suggested a two-way interaction between the missing data method and the true value of τ^2, which is supported by the ANOVA, η_p^2 = 0.794. As can be seen by Figure 27, coverage for both the CJM approach and listwise deletion were around 0.848 (*SD* = 0.115). and 0.836 (*SD* = 0.112) when τ^2 was 0.01. Coverage slightly decreased when τ^2

increased to 0.4 for listwise deletion (*M* = 0.830, *SD* = 0.106) but stayed the same for the CJM approach (*M* = 0.848, *SD* = 0.109). When τ^2 increased to 1, coverage increased to approximately 0.877 (*SD* = 0.088) and 0.852 (*SD* = 0.087) for the CJM approach and listwise deletion, respectively.

A two-way interaction between the missing data method and the true value of β_2 accounted for approximately 97% of the error variance, η_p^2 = 0.968. Table 5 shows that listwise deletion had relatively lower coverage than did the CJM approach throughout all values of β_1. Furthermore, a two-way interaction between the missing data method and the true value of β_2 accounted for approximately 64% of the error variance, η_p^2 = 0.641. As can be seen by Figure 28, the relationship between β_2 regarding coverage was similar for both missing data methods. Specifically, coverage rates decreased as the true value of β_2 increased.

ANOVA results also suggested a two-way interaction between the missing data method and the percentage missing, η_p^2 = 0.057. Examination of Table 5 across missing percentage between the missing data methods suggested a slight increase in coverage as missing percentage increased, and that coverage was slightly higher for the CJM across all conditions of missing percentage than listwise deletion. Average coverage for the CJM approach at 15% missing data was 0.85 (*SD* = .109) and for listwise deletion was 0.832 (*SD* = 0.108) for listwise deletion. When missing percentage increased to 30%, the CJM approach had a slight increase of coverage to approximately 0.855 (*SD* = 0.107), while listwise deletion increased to 0.835 (*SD* = 0.108). When missing percentage increased to 50%, the average coverage for the CJM approach increased to 0.868 (*SD* = 0.100) and for listwise deletion to 0.851 (*SD* = 0.090).

The three-way interaction between *K*, the missing data method, and τ^2 accounted for approximately 16% of the error variance, η_p^2 = 0.155. Table 5 shows that overall coverage was smaller for listwise deletion than for the CJM approach by *K* and τ^2. Figure 29 shows differences among coverage for the CJM approach when *K* was 20 versus when *K* was 80. Specifically, MSE values were slightly larger the CJM approach when K was 20 than when *K* as 80 for the CJM approach. Additionally, coverage increased for the CJM approach when τ^2 increased to 1. In terms of listwise deletion, Figure 29 shows minimal differences between the various conditions of *K*, however coverage increased whenever there was an increase in τ^2, and listwise deletion with 20 studies had slightly larger coverage than did listwise deletion with 80 studies.

The three-way interaction between *K*, the missing data method, and β_1 accounted for approximately 3% of the error variance, η_p^2 = 0.027. Illustrated by Figure 30, the average coverage for the CJM approach increased as β_1 increased when *K* = 20. When β_1 = 0.05 at *K* = 20, the average coverage for the CJM approach was 0.855 (*SD* = 0.120), at β_1 = 0.20, the average coverage for the CJM approached increased to 0.863 (*SD* = 0.107) and then increased again to 0.906 (*SD* = 0.056) when β_1 increased to 0.50 and *K* = 20. In comparison at *K* = 20 and β_1 = 0.05, the average coverage for listwise deletion was 0.859 (*SD* = 0.114), which was slightly larger than the CJM approach given the same conditions. However, coverage decreased to 0.83 (*SD* = 0.098) for listwise deletion when β_1 increased to 0.20, and then increased again to 0.866 (*SD* = 0.055) when β_1 increased to 0.50 at *K* = 20.

When β_1 = 0.05 and *K* = 80 the average coverage was higher for the CJM approach than when β_1 = 0.05 and *K* = 20. Specifically, the average coverage at these conditions was 0.862 (*SD* = 0.112), and then decreased to 0.833 (*SD* = 0.111) when β_1 = 0.20. Average coverage increased for the CJM approach at *K* = 80 when β_1 = 0.50 to 0.847 (*SD* = 0.104). In comparison, average coverage was smaller for listwise deletion when *K* = 20 and β_1 = 0.05, with a mean of 0.831 (*SD*

= 0.112). As β_1 increased to 0.20, coverage decreased to 0.822 (*SD* = 01.114), and then increased to 0.837 (*SD* = 0.109) when β_1 increased to 0.50.

Similarly, the three-way interaction between *K*, the missing data method, and β_2 accounted for approximately 29% of the error variance, η_p^2 = 0.285. Illustrated by Figure 31, coverage for the CJM approach and listwise deletion decreased as there was an increase in the true value of β_2, regardless of the value of *K*. However, coverage was higher for the CJM approach and listwise deletion when *K* = 20. Coverage ranged from approximately 0.95 to about 0.70.

ANOVA results suggested that the three-way interaction between missing data method, *K*, and missing percentage explained approximately 2% the error variance, η_p^2 = 0.016. As shown by Table 6 coverage values increase as missing percentage also increased, but as *K* increased, coverage values decreased. For example, when *K* = 20 and there was 15% missing data, the average coverage for the CJM approach was 0.865 (*SD* = 0.105) and the average coverage for listwise deletion was 0.840 (*SD* = 0.103). When *K* increased to 80 and there was15% missing data, the average coverage for the CJM approach was 0.835 (*SD* = 0.112) and the average coverage for listwise deletion was 0.824 (*SD* = 0.115). When missing percentage increased and *K* = 20, the average for the CJM approach was 0.874 (*SD* = 0.102) and the average coverage for listwise deletion was 0.848 (*SD* = 0.097). When *K* was increased to 80, average coverage decreased at 30% missing for both missing data methods to 0.837 (*SD* = 0.111) and 0.823 (*SD* = 0.118) for the CJM approach and listwise deletion, respectively. Lastly, when missing was increased to 50%, the average coverage at *K* = 20 for the CJM approach is 0.886 (*SD* = 0.93) and is 0.858 (*SD* = 0.078) for listwise deletion. When *K* is increased to 80, average coverage was

0.850 (*SD* = 0.105) and 0844 (*SD* = 0.213) for the CJM approach and listwise deletion, respectively.

ANOVA results also indicated that the three-way interaction between missing data method, τ^2, and β_1 explained approximately 4% of the error variance, η_p^2 = 0.036. As shown by Table 5, coverage increased as τ^2 increased, on average. Additionally, coverage was largest for the CJM approach for all conditions of β_1, specifically when τ^2 was equal to 1. Overall analysis shows that coverage was between 0.80 and 0.90 for all values of β_1 and τ^2.

Figure 32 illustrates the three-way interaction between missing data method, τ^2, and β_2, which explained approximately 46% of the error variance, η_p^2 = 0.459. Specifically, the lower values of β_2 and τ^2 had the highest coverage for both missing data methods. When β_2 was 0.05 and τ^2 was 0.01, the average coverage for the CJM approach was 0.971 (*SD* = 0.013) and for listwise deletion was 0.95 (*SD* = 0.010). When τ^2 increases to 0.40, the average coverage when β_2 = 0.20 was 0.886 (SD = 0.029) for the CJM approach and 0.863 (*SD* =0.025) for listwise deletion. In contrast, when τ^2 = 0.40 and β_2 = 0.05, the average coverage decreased to 0.955 (*SD* = 0.010) for the CJM approach and decreased to 0.935 (*SD* = 0.016) for listwise deletion. Overall, Figure 32 shows that coverage was lowest when β_2 was largest and τ^2 is lowest.

The three-way interaction between the missing data method, τ^2 and missing percentage explained approximately 5% of the error variance, η_p^2 = 0.051. The largest coverage occurred in the CJM approach when τ^2 was 1.0 and missing percentage was 50%. However, all coverage for this three-way interaction was between 0.80 and 0.90.

Examination of Table 5 also suggested a three-way interaction between missing data method, β_1 and missing percentage, which was supported by the ANOVA results, η_p^2 = 0.165. Particularly, coverage increased as missing percentage and β_1 increased for both missing data methods. Figure 33 illustrates that the CJM approach had larger coverage for all values of β_1 and missing percentage. When missing percentage was 15% for the CJM approach, coverage was approximately 0.85 for all values of β_1. As missing percentage increased to 30, coverage increased for the CJM approach when β_1 was 0.50 but was relatively the same for the CJM approach when β_1 was 0.05 or 0.20. When missing percentage increased to 50%, coverage for the CJM approach at β_1 = 0.05 and 0.20 became larger than when β_1 was 0.50 for the CJM approach.

In terms of listwise deletion, coverage was lowest when missing percentage was 15. Coverage also increased for all values of β_1 as there was an increase in missing percentage. When missing percentage increased to 50%, coverage was slightly above 0.85 when β_1 was 0.05 and 0.5, but not when β_1 was 0.20.

ANOVA results also suggested a three-way interaction between missing data method, β_2 and missing percentage, which was supported by the ANOVA results, η_p^2 = 0.017. Particularly, coverage increased as missing percentage and β_2 increased for both missing data methods. Figure 34 illustrates that the CJM approach and listwise deletion had larger coverage rates for values of β_1 and missing percentage. When missing percentage was 15% for the CJM approach and listwise deletion, coverage was approximately 0.95 when β_2 was 0.05. As missing percentage increased to 30 and 50, coverage was relatively the same for both missing data methods when β_2 was 0.05. When β_2 was 0.20 for both missing data methods, coverage was between 0.80 and 0.90 across all values of missing percentage. Lastly, when β_2 was 0.50 for both missing data methods, coverage was around 0.70 at 15% missing and 30% missing, and then slightly increased to about 0.75 when missing percentage is 50.

Coverage for Hedges' g

Table 7 provides coverage performance for g for the various conditions. ANOVA results indicated differences among coverage for several conditions for Hedges' g. First, the interaction between *K* and missing data method explained approximately 24% of the error variance, η_p^2 = 0.236.

From Table 7, when comparing coverage by *K* and the missing data method, listwise deletion had considerably lower coverage than did the CJM approach. When K = 20, listwise deletion had an average coverage of 0.863 (*SD* = 0.091), while the CJM approach had an average coverage of 0.922 (*SD* = 0.059). When *K* increased to 80, listwise deletion still had lower coverage than the CJM approach with a mean of 0.724 (*SD* = 0.254) for listwise deletion and a mean of 0.028 (*SD* = 0.122) for the CJM approach.

Examination of Table 7 also suggested a two-way interaction between the missing data method and the true value of τ^2, which was supported by the ANOVA, η_p^2 = 0.714. As can be seen by Figure 35, coverage for both the CJM approach and listwise deletion was over 0.95, with an average of 0.971 (*SD* = 0.011) for the CJM approach and 0.955 (*SD* = 0.014) for listwise deletion when τ^2 was 0.01. Coverage values decreased for both missing data methods when τ^2 increased 0.4 with an average of 0.897 (*SD* = 0.049) for the CJM approach and an average of 0.812 (*SD* = 0.106) for listwise deletion. When τ^2 increased to 1, coverage continued to decrease to approximately 0.757 (*SD* = 0.088) and 0.614 (*SD* = 0.230) for the CJM approach and listwise deletion, respectively.

ANOVA results also suggested a two-way interaction between the missing data method and the percentage missing, η_p^2 = 0.240. Examination of Table 7 across missing percentage between the missing data methods suggested that overall coverage for the CJM approach was approximately the same for all values of missing percentage. However, as can be seen by Figure 36, coverage decreased as there was an increase in missing percentage. When missing percentage was 15 for listwise deletion, coverage was approximately 0.875, but was around 0.755 when missing was 30% and was around 0.725 when missing was around 50%.

The three-way interaction between *K*, the missing data method, and τ^2 accounted for 68% of the error variance, η_p^2 = 0.680. Figure 37 shows differences among coverage between *K*, missing data method, and τ^2. When *K* = 20, coverage was between 0.75 and 0.99 for both listwise deletion and CJM approach, regardless of τ^2 value. In general, coverage decreased as τ^2 increased. However, coverage decreased substantially when τ^2 increased when *K* was 80. Specifically, coverage was as low as 0.45 for listwise deletion when *K* = 80 and τ^2 = 1.0. In comparison, when *K* = 80 and τ^2 = 1.0, coverage was around 0.78 for the CJM approach.

A three-way interaction between *K*, missing data method, and β_1 accounted for approximately 4% of the error variance, η_p^2 = 0.036. Figure 38 shows when *K* = 20, coverage was above 0.80 for both listwise deletion and the CJM approach, regardless of β_1 value. When *K* = 20, there was barely any change in coverage for listwise deletion and the CJM approach as there was an increase in β_1. This is also true for the CJM approach when *K* = 80. However, when *K* = 80 for listwise deletion, coverage increased as the true value of β_1 increased. Coverage rates were as low as 0.70 for the CJM approach at β_1= 0.05, and as high as 0.75 when β_1 = 0.50.

Figure 39 illustrates the three-way interaction between *K*, missing data method, and missing percentage. When *K* = 20 and missing percentage increased, coverage did not change for the CJM approach. However, coverage decreased for listwise deletion. Coverage decreased from around 0.90 at 15% missing to about 0.85 at 50% missing for listwise deletion. When *K* = 80, coverage was a little bit above 0.80 for the CJM approach. However, like when *K* = 20 for listwise deletion, coverage decreased as missing percentage increased for listwise deletion when *K* = 80. Specifically, coverage was as high as 0.98 at 15% missing for listwise deletion and as low as 0.60 at 50% missing when *K* = 80. This three-way interaction accounted for approximately 24% of the error variance, η_p^2 = 0.241.

ANOVA results indicated that a three-way interaction between missing data method, τ^2, and β_1 accounted for about 7% of the error variance, η_p^2 = 0.069. Figure 40 illustrates that β_1 only influenced the coverage within listwise deletion and not the CJM approach. Similarly, a three-way interaction between missing data, τ^2, and β_2 suggested that only τ^2 influenced the coverage within listwise deletion and not the CJM approach. This interaction accounted for 3%, η_p^2 = 0.028.

A three-way interaction between missing data method, τ^2, and missing percentage accounted for approximately 23% of the error variance, η_p^2 = 0.230. As missing percentage increased for listwise deletion, coverage decreased substantially. This occurred specifically when τ^2 was 0.40 and 1.0.

CHAPTER V

CONCLUSIONS

Summary of Results

One of the most widely employed methodology to handle missing moderators when performing meta-analysis is listwise deletion (Pigott, 2001). However, listwise deletion tends to have several issues including bias estimation and a loss of statistical power (Pigott, 2019). Being able to accurately predict missing moderators in meta-analysis is beneficial. With the growth of evidence based practice, meta-analyses are typically considered at the top of the hierarchy of reliable evidence (Berlin & Golub, 2014). Consequently, it is important to be able to account for missing data.

This dissertation examined a Bayesian joint model for handling missing moderators, which is based on a sequence of conditional distributions, namely conditional joint modeling (CJM). The use of CJM has been used in the meta-analysis literature to predict missing moderators (Hemming et al., 2010). However, its performance has yet to be empirically studied in the context of meta-analysis.

This study primarily aimed to answer 1) how the CJM method performs under various conditions, and 2) how the CJM method performed compared to listwise deletion. CJM was examined under various simulation conditions and then analyzed using three forms of assessment: relative bias, mean squared error, and coverage. Relative bias was calculated to measure how consistently the estimated parameter under or overestimated the true value of the parameters of interest. Results from the simulation show that the CJM method severely underestimated the true value of β_2. Relative bias values ranged from as low as -0.956 to as high as -0.03815.

In terms of the comparison between listwise and CJM, the CJM approach performed similarly to listwise deletion when estimating β_2. Though, there were a few instances where one method outperformed the other method. For example, CJM and listwise deletion performed similarly when there was a small or medium amount of missing data (relative bias was between -0.35 and -0.30). However, listwise deletion produced more precise estimates of β_2 than did CJM when missing percentage increased to 50%. Listwise deletion also performed better in estimating β_2 when the number of studies in the meta-analysis was small (i.e., K = 20); but, CJM performed better than listwise deletion when K increased to 80. As indicated by the mixed-effects ANOVA, CJM tended to perform better than listwise deletion across all other conditions when the number of studies in the meta-analysis was large.

Most of the relative biases for Hedges' g were also underestimated when implementing CJM, with values ranging from as low as -0.221 to as high as 0.020. Overall, in terms of Hedges' g, CJM outperformed listwise deletion across all simulation conditions.

MSE was calculated to quantify the differences between the estimated and true values of the parameters of interest, indicating how close the estimates were to the true value(s). Examination of MSE when estimating β_2 indicated that 1) CJM did not accurately predict the missing regression coefficient, and 2) CJM and listwise deletion performed similarly. Overall, estimated β_2 was the least accurate when the true values of the regression coefficients were large and there was a small amount of between-studies variance.

Examination of MSE when estimating Hedges' g suggested that CJM was more accurate than listwise deletion across all conditions. In all the scenarios of simulation, MSE was never larger than 0.0 1when CJM was implemented as the missing data method. However, accuracy declined as there was an increase in the between-studies variance for both missing data methods.

This implies that when there is a large amount of between-studies variance in a meta-analytic study and either CJM or listwise deletion is implemented, there is less accuracy in the prediction of the overall true effect size.

The pattern of the results also suggest that *K* is an important factor for accuracy of Hedges' g when measured by MSE. First, there was considerably larger MSE values for listwise deletion than there was for the CJM approach when comparing by values of *K*. Second, the CJM approach also performed better when examining the three-way interaction between *K*, τ^2, and the missing data method. Listwise deletion did not perform well as there was an increase in between-studies variance and the number of studies. Additionally, the CJM approach also performed better when examining the three-way interaction between *K*, β_2, and missing data method.

The CJM approach generally performed better than listwise deletion, with respect to coverage for the coefficient of the missing moderator, β_2. However, examination of Table 6 illustrates a severe amount of under-coverage for both missing data methods. For example, when the regression coefficient of β_2 was high (i.e., greater than 0.2), the average coverage for both missing data methods was at, or, below 0.85. When β_2 increased, there was a decrease in coverage, which reached as low as 70% when β_2 was at its largest of 0.50.

The CJM approach also generally performed better than listwise deletion, with respect to coverage for Hedges' g. Results also suggested that 1) greater between-studies variance resulted in lower coverage for both missing data methods, 2) a larger number of studies resulted in lower coverage for both missing data methods, and 3) the interaction between between-studies variance and number of studies resulted in lower coverage for both missing data methods

In general, results suggested that both methods for both parameters had inflated type 1 error rates. This is because there was a large amount of conditions in which coverage was lower than the acceptable value of 0.94 (i.e., 94%). Inflated type I errors occur because there is greater confidence in the estimates due to more simulations incorrectly finding a significant result (Burton et al., 2006; Luo & Kwok, 2012).

Implications and Limitations

A limitation is noted when drawing conclusions from this research. An error message occurred when trying to compute the meta-analysis in the *metafor* package for 50% missing data. It is worth nothing that this estimation issue occurred randomly. That is, sometimes the *metafor* package computed the meta-analysis with 50% missing and sometimes it did not. No pattern was found as to why this occurred.

However, we can see from these results how statistical analyses can be affected if studies with missing moderators are omitted from the meta-analysis. The major implication of this dissertation is that simply dropping a moderator with missing values from the analysis can lead to biased regression coefficients of the remaining moderators, which can change the conclusions of meta-analytic studies. Furthermore, dropping a study with missing values may result in losing important information that may impact the overall effect size.

Although the CJM approach was not perfect, it still performed better than listwise deletion when estimating Hedges' g. As modeled, the CJM approach did not contain any missing values for Hedges' g. The point of this dissertation was to be able to accurately estimate β_2 when there were missing values in the moderator. The Bayesian model (i.e., CJM approach) was able to compute g even if the study had missing data because all the cases were available to compute g.

However, in the listwise deletion approach, any study that had missing data was deleted. This includes all the full information that could have been used to calculate g. This difference in calculation and method may explain the differences in results when comparing Hedges' g.

Future Research and Conclusion

According to Li, Yu, and Rubin (2012) the CJM approach is "theoretically valid" when the missing data and data analysis are considered at the same time. However, if the method is only used to impute missing data, then the order of the moderator variables can lead to completely different joint distributions and resulting imputations. Other research (e.g., Murray, 2018) has also stated that different orders of the moderating variables tend to lead to different joint distributions.

This dissertation focused on the utilizing the CJM approach to impute missing data. The different orders of the moderating variables were not examined given that only two moderators, one with missing data, were utilized. Future research should examine more than two moderators, with different orderings of the joint distribution.

Other research suggests that estimates can be biased if the covariate distribution is misspecified (Ibrahim et al., 2012; Zhang & Rockette, 2005). Ibrahim et al. (1999) mention the CJM approach has the potential to be misspecified, suggesting that "great care be taken in specifying the [moderator] distributions" (p. 596). Ibrahim et al. (1999) also suggest conducting a sensitivity analysis. The sensitivity analysis should be used to examine the robustness of the regression coefficient estimates under different prior distributions specified for the moderators.

Though misspecification was not examined in this dissertation, a sensitivity analysis of the different choice of prior distributions can help future research. As mentioned in this dissertation, one advantage of using Bayesian methods is the ability to "borrow strength" from other studies. A sensitivity analysis of priors may help reduce the imprecision when used to impute missing data. Sensitivity analyses are recommended by 1) changing the order of the condition in the covariate distributions, and 2) examining how the different choice in priors may influence the outcome.

Although this dissertation focused on two missing data methods, listwise deletion method and the CJM approach, many other methods are available to analyze missing data. For example, multiple imputation is generally recommended for use in meta-analysis (Pigott, 2019); however, there exists a range of algorithms for multiple imputation procedures. Other algorithms should also be examined for handling missing moderators in meta-analysis.

In summary, analysis of moderators in meta-analysis should be interpreted with caution when missing data exists, especially if the research implemented listwise deletion. Overall, current recommendations for handling missing moderators still suggest the use of maximum likelihood or multiple imputation (Pigott, 2019). However, the specifics of what algorithm or approach to take is still an area that needs to be examined.

REFERENCES

Abrams, K., & Sansó, B. (1998). Approximate Bayesian inference for random effects meta-analysis. *Statistics in Medicine, 17*(2), 201–218. DOI: 10.1002/(SICI)1097-0258(19980130)17:2<201::AID-SIM736>3.0.CO;2-9

Allison, P. D. (2000). Multiple imputation for missing data: A cautionary tale. *Sociological Methods & Research, 28*(3), 301–309. DOI: https://doi.org/10.1177/0049124100028003003

Anderson, T. W. (1957). Maximum likelihood estimates for a multivariate normal distribution when some observations are missing. *Journal of the American Statistical Association*, *52*(278), 200–203. DOI: 10.2307/2280845

Berlin, J. A., & Golub, R. M. (2014). Meta-analysis as evidence: Building a better pyramid. *JAMA*, *312*(6), 603–606. DOI: 10.1001/jama.2014.8167

Bhaskaran, K., & Smeeth, L. (2014). What is the difference between missing completely at random and missing at random? *International Journal of Epidemiology*, *43*(4), 1336–1339. DOI: 10.1093/ije/dyu080

Bodnar, O., Link, A., Arendacká, B., Possolo, A., & Elster, C. (2017). Bayesian estimation in random effects meta-analysis using a non-informative prior. *Statistics in Medicine*, *36*(2), 378–399. DOI: 10.1002/sim.7156

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2010). A basic introduction to fixed-effect and random-effects models for meta-analysis. *Research Synthesis Methods*, *1*(2), 97–111. DOI: 10.1002/jrsm.12

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2011). *Introduction to Meta-Analysis*. John Wiley & Sons.

Brockwell, S. E., & Gordon, I. R. (2001). A comparison of statistical methods for meta-analysis. *Statistics in Medicine*, *20*(6), 825–840. DOI: 10.1002/sim.650

Burton, A., Altman, D. G., Royston, P., & Holder, R. L. (2006). The design of simulation studies in medical statistics. *Statistics in Medicine*, *25*(24), 4279–4292. DOI: 10.1002/sim.2673

Carpenter, J. R., Kenward, M. G., & White, I. R. (2007). Sensitivity analysis after multiple imputation under missing at random: A weighting approach. *Statistical Methods in Medical Research*, *16*(3), 259–275. DOI: 10.1177/0962280206075303

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

Cooper, H., Hedges, L. V., & Valentine, J. C. (2009). *The Handbook of Research Synthesis and Meta-Analysis* (2nd ed.). Russell Sage Foundation.

Davis, J., Mengersen, K., Bennett, S., & Mazerolle, L. (2014). Viewing systematic reviews and meta-analysis in social research through different lenses. *SpringerPlus*, *3*, 511. DOI: 10.1186/2193-1801-3-511

De Finetti, B. (1974). Bayesianism: Its unifying role for both the foundations and applications of statistics. *International Statistical Review / Revue Internationale de Statistique*, *42*(2), 117–130. DOI: 10.2307/1403075

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, *39*(1), 1–22. DOI: 10.1111/j.2517-6161.1977.tb01600.x

DerSimonian, R., & Laird, N. (1986). Meta-analysis in clinical trials. *Controlled Clinical Trials*, *7*(3), 177–188.

Enders, C. K. (2001). A primer on maximum likelihood algorithms available for use with missing data. *Structural Equation Modeling: A Multidisciplinary Journal*, *8*(1), 128–141. DOI: 10.1207/S15328007SEM0801_7

Enders, C. K. (2010). *Applied Missing Data Analysis*. Guilford Press.

Enders, C. K. (2017). Multiple imputation as a flexible tool for missing data handling in clinical research. *Behaviour Research and Therapy*, *98*, 4–18.

Enders, C. K., Mistler, S. A., & Keller, B. T. (2016). Multilevel multiple imputation: A review and evaluation of joint modeling and chained equations imputation. *Psychological Methods*, *21*(2), 222–240. DOI: 10.1037/met0000063

French, J. L. (2004). Generalized additive models for cancer mapping with incomplete covariates. *Biostatistics*, *5*(2), 177–191. DOI: 10.1093/biostatistics/5.2.177

Furukawa, T. A., Barbui, C., Cipriani, A., Brambilla, P., & Watanabe, N. (2006). Imputing missing standard deviations in meta-analyses can provide accurate results. *Journal of Clinical Epidemiology*, *59*(1), 7–10. DOI: 10.1016/j.jclinepi.2005.06.006

Gelman, A. (2003). A Bayesian formulation of exploratory data analysis and goodness-of-fit testing*. *International Statistical Review*, *71*(2), 369–382. DOI: 10.1111/j.1751-5823.2003.tb00203.x

Gelman, A. (2004). Parameterization and Bayesian modeling. Journal of the American Statistical Association, 99(466), 537-545.

Glass, G. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher*, *5*(10), 3–8.

Gold, M. S., & Bentler, P. M. (2000). Treatments of missing data: A Monte Carlo comparison of RBHDI, iterative stochastic regression imputation, and expectation-maximization. *Structural Equation Modeling*, *7*(3), 319–355. DOI: 10.1207/S15328007SEM0703_1

Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual Review of Psychology*, *60*(1), 549–576. DOI: 10.1146/annurev.psych.58.110405.085530

Graham, J. W., Hofer, S. M., & MacKinnon, D. P. (1996). Maximizing the usefulness of data obtained with planned missing value patterns: An application of maximum likelihood procedures. *Multivariate Behavioral Research*, *31*(2), 197–218. DOI: 10.1207/s15327906mbr3102_3

Graham, J. W., Olchowski, A. E., & Gilreath, T. D. (2007). How many imputations are really needed? Some practical clarifications of multiple imputation theory. *Prevention Science*, *8*(3), 206–213. DOI: 10.1007/s11121-007-0070-9

Grund, S., Lüdtke, O., & Robitzsch, A. (2018a). Multiple imputation of missing data at level 2: A comparison of fully conditional and joint modeling in multilevel designs. *Journal of Educational and Behavioral Statistics*, *43*(3), 316–353. DOI: 10.3102/1076998617738087

Grund, S., Lüdtke, O., & Robitzsch, A. (2018b). Multiple imputation of missing data for multilevel models: Simulations and Recommendations. *Organizational Research Methods*, *21*(1), 111–149. DOI: 10.1177/1094428117703686

Grund, S., Lüdtke, O., & Robitzsch, A. (2019). Missing data in multilevel research. In S. E. Humphrey & J. M. LeBreton, *The Handbook of Multilevel Theory, Measurement, and Analysis* (pp. 365–386). Washington, DC: American Psychological Association. DOI: 10.1037/0000115-017

Haidich, A. B. (2010). Meta-analysis in medical research. *Hippokratia*, *14*(Suppl 1), 29–37.

Hardy, R. J., & Thompson, S. G. (1996). A likelihood approach to meta-analysis with random effects. *Statistics in Medicine*, *15*(6), 619–629.

Heckerman, D., Chickering, D. M., Meek, C., Rounthwaite, R., & Kadie, C. (2000). Dependency networks for inference, collaborative filtering, and data visualization. *Journal of Machine Learning Research*, *1*, 49–75.

Hedges, L. V., & Pigott, T. D. (2004). The power of statistical tests for moderators in meta-analysis. *Psychological Methods*, *9*(4), 426–445. DOI: 10.1037/1082-989x.9.4.426

Hemming, K., Hutton, J. L., Maguire, M. G., & Marson, A. G. (2010). Meta-regression with partial information on summary trial or patient characteristics. *Statistics in Medicine*, *29*(12), 1312–1324. DOI: 10.1002/sim.3848

Hemming, K., & Hutton, J. L. (2012). Bayesian sensitivity models for missing covariates in the analysis of survival data. *Journal of Evaluation in Clinical Practice*, *18*(2), 238–246. DOI: 10.1111/j.1365-2753.2010.01569.x

Higgins, J. P. T., Thompson, S. G., & Spiegelhalter, D. J. (2009). A re-evaluation of random-effects meta-analysis. *Journal of the Royal Statistical Society. Series A, (Statistics in Society)*, *172*(1), 137–159. DOI: 10.1111/j.1467-985X.2008.00552.x

Higgins, J. P. T., & Whitehead, A. (1996). Borrowing strength from external trials in a meta-analysis. *Statistics in Medicine*, *15*, 2733–2749.

Ibrahim, J. G., Chen, M.H., & Lipsitz, S. R. (1999). Monte Carlo EM for missing covariates in parametric regression models. *Biometrics*, *55*(2), 591–596. DOI: 10.1111/j.0006-341X.1999.00591.x

Ibrahim, J. G., Chen, M. H., Lipsitz, S. R., & Herring, A. H. (2005). Missing-data methods for generalized linear models: A comparative review. *Journal of the American Statistical Association*, *100*(469), 332–346. DOI: 10.1198/016214504000001844

Ibrahim, J. G., Chu, H., & Chen, M.H. (2012). Missing data in clinical studies: Issues and methods. *Journal of Clinical Oncology*, *30*(26), 3297–3303. DOI: 10.1200/JCO.2011.38.7589

Jackman, S. (2009). *Bayesian Analysis for the Social Sciences* (1st ed.). West Sussex, United Kingdom: John Wiley & Sons, Ltd.

Kang, H. (2013). The prevention and handling of the missing data. *Korean Journal of Anesthesiology*, *64*(5), 402. DOI: 10.4097/kjae.2013.64.5.402

Kellner, K. (2019). *JagsUI: A warapper around "rjags" to streamline "JAGS" analyses*. Retrieved from https://CRAN.R-project.org/package=jagsUI

Kenward, M. G., & Molenberghs, G. (1998). Likelihood based frequentist inference when data are missing at random. *Statistical Science*, *13*(3), 236–247. DOI: 10.1214/ss/1028905886

Kim, J. O., & Curry, J. (1977). The treatment of missing data in multivariate analysis. *Sociological Methods & Research*, *6*(2), 215–240. DOI: 10.1177/004912417700600206

Kline, D., Andridge, R., & Kaizar, E. (2017). Comparing multiple imputation methods for systematically missing subject-level data. *Research Synthesis Methods*, *8*(2), 136–148. (2015-57304-001). DOI: 10.1002/jrsm.1192

Konstantopoulos, S., & Hedges, L. V. (2019). Statistically analyzing effect sizes: Fixed-and random-effects models. In Cooper, H., Hedges, L.V., & Valentine. J.C. (Eds.), *The Handbook of Research Synthesis and Meta-Analysis* (3rd ed.). New York: Russell Sage Foundation.

Kruschke, J. K., & Liddell, T. M. (2018). Bayesian data analysis for newcomers. *Psychonomic Bulletin & Review*, *25*(1), 155–177. DOI: 10.3758/s13423-017-1272-1

Li, F., Yu, Y., & Rubin, D. B. (2012). Imputing missing data by fully conditional models: Some cautionary examples and guidelines. *Duke University Department of Statistical Science Discussion Paper*, *11*(14), 1–35.

Lin, T. H. (2010). A comparison of multiple imputation with EM algorithm and MCMC method for quality of life missing data. *Quality & Quantity*, *44*(2), 277–287. DOI: 10.1007/s11135-008-9196-5

Lipsitz, S., & Ibrahim, J. G. (1996). A conditional model for incomplete covariates in parametric regression models. *Biometrika, 83*(4), 916–922. DOI: 10.1093/biomet/83.4.916

Little, R. J. A., & Rubin, D. B. (2002). *Statistical Analysis with Missing Data* (2nd ed). John Wiley & Sons, Inc.

Little, R. J. A., & Rubin, D. B. (2019). *Statistical Analysis with Missing Data* (3rd ed., Vol. 793). John Wiley & Sons, Incorporated.

Liu, M., Taylor, J. M. G., & Belin, T. R. (2000). Multiple imputation and posterior simulation for multivariate missing data in longitudinal studies. *Biometrics*, *56*(4), 1157–1163. DOI: 10.1111/j.0006-341X.2000.01157.x

Liu, Y., & De, A. (2015). Multiple imputation by fully conditional specification for dealing with missing data in a large epidemiologic study. *International Journal of Statistics in Medical Research*, *4*(3), 287–295. DOI: 10.6000/1929-6029.2015.04.03.7

Lopez-Lopez, J. A., Botella, J., Sanchez-Meca, J., & Marin-Martinez, F. (2013). Alternatives for mixed-effects meta-regression models in the reliability generalization approach: A simulation study. *Journal of Educational and Behavioral Statistics*, *38*(5), 443–469. DOI: 10.3102/1076998612466142

Luo, W., & Kwok, O. (2012). The consequences of ignoring individuals' mobility in multilevel growth models: A Monte Carlo study. *Journal of Educational and Behavioral Statistics, 37*(1), 31–56. DOI: 10.3102/1076998610394366

Morris, T. P., White, I. R., & Crowther, M. J. (2019). Using simulation studies to evaluate statistical methods. *Statistics in Medicine*, *38*(11), 2074–2102. DOI: 10.1002/sim.8086

Murray, J. S. (2018). Multiple imputation: A review of practical and theoretical findings. *Statistical Science, 33*(2), 142–159. DOI: 10.1214/18-STS644

Peugh, J. L., & Enders, C. K. (2004). Missing data in educational research: A review of reporting practices and suggestions for improvement. *Review of Educational Research*, *74*(4), 525–556. DOI: 10.3102/00346543074004525

Pigott, T. D. (2012). *Advances in Meta-Analysis*. Springer Science & Business Media. DOI: 10.1007/978-1-4614-2278-5

Pigott, T. D. (2019). Missing data in meta-analysis. In *The Handbook of Research Synthesis and Meta-Analysis* (3rd ed.). New York, NY: Russell Sage Foundation.

Pigott, T. D. (2001). Missing predictors in models of effect size. *Evaluation & the Health Professions*, *24*(3), 277–307. DOI: 10.1177/01632780122034920

Plummer, M. (2003). JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. *Working Papers*, 8.

Plummer, M. (2019). *rjags: Bayesian graphical models using MCMC*. Retrieved from https://CRAN.R-project.org/package=rjag

Plummer, M., Best, N., Cowles, K., & Vines, K. (2006). CODA: Convergence diagnosis and output analysis for MCMC. *R News, 6*(1), 7–11.

Quartagno, M., Grund, S., & Carpenter, J. (2019). jomo: A flexible package for two-level joint modelling multiple imputation. *The R Journal*, *11*(2), 1–24.

Raghunathan, T. E., Lepkowski, J. M., Hoewyk, J. V., & Solenberger, P. (2001). A multivariate technique for multiply imputing missing values using a sequence of regression models. *Survey Methodology*, (12), 11.

Raudenbush, S. (2009). Analyzing effect sizes: Random effects models. In Cooper, H., Hedges, L.V., & Valentine J.C (Eds.), *The Handbook of Research Synthesis and Meta-analysis* (2nd ed., pp. 295–315). New York: Russell Sage Foundation.

Rubin, D. B. (1976). Inference and Missing Data. *Biometrika*, *63*(3), 581–592. JSTOR. DOI: 10.2307/2335739

Rubin, D. B. (1986). Basic ideas of multiple imputation for nonresponse. *Survey Methodology*, *12*(1), 37–47.

Rubin, D. B. (1987). *Multiple Imputation for Nonresponse in Surveys*. Wiley.

Rubin, D. B. (2003). Nested multiple imputation of NMES via partially incompatible MCMC. *Statistica Neerlandica*, *57*(1), 3–18. DOI: 10.1111/1467-9574.00217

Schafer, J. L. (1997). *Analysis of Incomplete Multivariate Data*. Chapman and Hall/CRC. DOI: 10.1201/9781439821862

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, *7*(2), 147–177. (2002-13431-001). DOI: 10.1037/1082-989X.7.2.147

Schmid, C. H. (2001). Using Bayesian inference to perform meta-analysis. *Evaluation & the Health Professions*, *24*(2), 165–189. DOI: 10.1177/01632780122034867

Van de Schoot, R., Kaplan, D., Denissen, J., Asendorpf, J. B., Neyer, F. J., & Van Aken, M. A. (2014). A gentle introduction to Bayesian analysis: Applications to developmental research. Child development, 85(3), 842-860. DOI: 10.1111/cdev.12169

Sidik, K., & Jonkman, J. N. (2007). A comparison of heterogeneity variance estimators in combining results of studies. *Statistics in Medicine*, *26*(9), 1964–1981. DOI: 10.1002/sim.2688

Stubbendick, A. L., & Ibrahim, J. G. (2003). Maximum likelihood methods for nonignorable missing responses and covariates in random effects models. *Biometrics*, *59*(4), 1140–1150. DOI: 10.1111/j.0006-341X.2003.00131.x

Sutton, A. J., & Abrams, K. R. (2001). Bayesian methods in meta-analysis and evidence synthesis. *Statistical Methods in Medical Research*, *10*(4), 277–303. DOI: 10.1177/096228020101000404

Tanner, M. A., & Wong, W. H. (1987). The calculation of posterior distributions by data augmentation. *Journal of the American Statistical Association*, *82*(398), 528–540. DOI: 10.2307/2289457

Thompson, S. G., & Higgins, J. P. T. (2002). How should meta-regression analyses be undertaken and interpreted? *Statistics in Medicine*, *21*(11), 1559–1573. DOI: 10.1002/sim.1187

Thompson, S. G., & Sharp, S. J. (1999). Explaining heterogeneity in meta-analysis: A comparison of methods. *Statistics in Medicine*, *18*(20), 2693–2708.

Tipton, E., Pustejovsky, J. E., & Ahmadi, H. (2019). A history of meta-regression: Technical, conceptual, and practical developments between 1974 and 2018. *Research Synthesis Methods*, *10*(2), 161–179. DOI: 10.1002/jrsm.1338

Turner, N. L., Dias, S., Ades, A. E., & Welton, N. J. (2015). A Bayesian framework to account for uncertainty due to missing binary outcome data in pairwise meta-analysis. *Statistics in Medicine*, *34*(12), 2062–2080. DOI: 10.1002/sim.6475

Turner, R. M., & Higgins, J. P. T. (2019). Bayesian meta-analysis. In Cooper, H., Hedges, L.V., & Valnetine, J.C. (Eds.), *The Handbook of Research Synthesis and Meta-Analysis* (3rd ed.). New York, NY: Russell Sage Foundation.

Turner, R. M., Jackson, D., Wei, Y., Thompson, S. G., & Higgins, J. P. T. (2015). Predictive distributions for between-study heterogeneity and simple methods for their application in Bayesian meta-analysis. *Statistics in Medicine*, *34*(6), 984–998. DOI: 10.1002/sim.6381

Van Buuren, S., Boshuizen, H. C., & Knook, D. L. (1999). Multiple imputation of missing blood pressure covariates in survival analysis. *Statistics in Medicine*, *18*(6), 681–694. DOI: 10.1002/(SICI)1097-0258(19990330)18:6<681::AID-SIM71>3.0.CO;2-R

Van Buuren, S., Brand, J. P. L., Groothuis-Oudshoorn, C. G. M., & Rubin, D. B. (2006). Fully conditional specification in multivariate imputation. *Journal of Statistical Computation and Simulation*, *76*(12), 1049–1064. DOI: 10.1080/10629360600810434

Van Buuren, S. (2007). Multiple imputation of discrete and continuous data by fully conditional specification. *Statistical Methods in Medical Research*, *16*(3), 219–242. DOI: 10.1177/0962280206074463

Van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software*, *45*(3). DOI: 10.18637/jss.v045.i03

Van de Schoot, R., Winter, S. D., Ryan, O., Zondervan-Zwijnenburg, M., & Depaoli, S. (20170608). A systematic review of Bayesian articles in psychology: The last 25 years. *Psychological Methods*, *22*(2), 217. DOI: 10.1037/met0000100

Van Ravenzwaaij, D., Cassey, P., & Brown, S. D. (2018). A simple introduction to Markov Chain Monte–Carlo sampling. *Psychonomic Bulletin & Review*, *25*(1), 143–154. (2016-13114-001). DOI: 10.3758/s13423-016-1015-8

Viechtbauer, W. (2005). Bias and efficiency of meta-analytic variance estimators in the random-effects model. *Journal of Educational and Behavioral Statistics*, *30*(3), 261–293. DOI: 10.3102/10769986030003261

Viechtbauer, W. (2007). Accounting for heterogeneity via random-effects models and moderator analyses in meta-analysis. *Zeitschrift Für Psychologie / Journal of Psychology*, *215*(2), 104–121. DOI: 10.1027/0044-3409.215.2.104

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*(3), 1–48.

Viechtbauer, W., López-López, J. A., Sánchez-Meca, J., & Marín-Martínez, F. (2015). A comparison of procedures to test for moderators in mixed-effects meta-regression models. *Psychological Methods*, *20*(3), 360–374. DOI: 10.1037/met0000023

Walther, B. A., & Moore, J. L. (2005). The concepts of bias, precision and accuracy, and their use in testing the performance of species richness estimators, with a literature review of estimator performance. *Ecography*, *28*(6), 815–829. DOI: 10.1111/j.2005.0906-7590.04112.x

White, I. R., Royston, P., & Wood, A. M. (2011). Multiple imputation using chained equations: Issues and guidance for practice. *Statistics in Medicine*, *30*(4), 377–399. DOI: 10.1002/sim.4067

Zhang, Z., & Rockette, H. E. (2005). On maximum likelihood estimation in parametric regression with missing covariates. *Journal of Statistical Planning and Inference*, *134*(1), 206–223. DOI: 10.1016/j.jspi.2004.04.006

APPENDIX

FIGURES AND TABLES

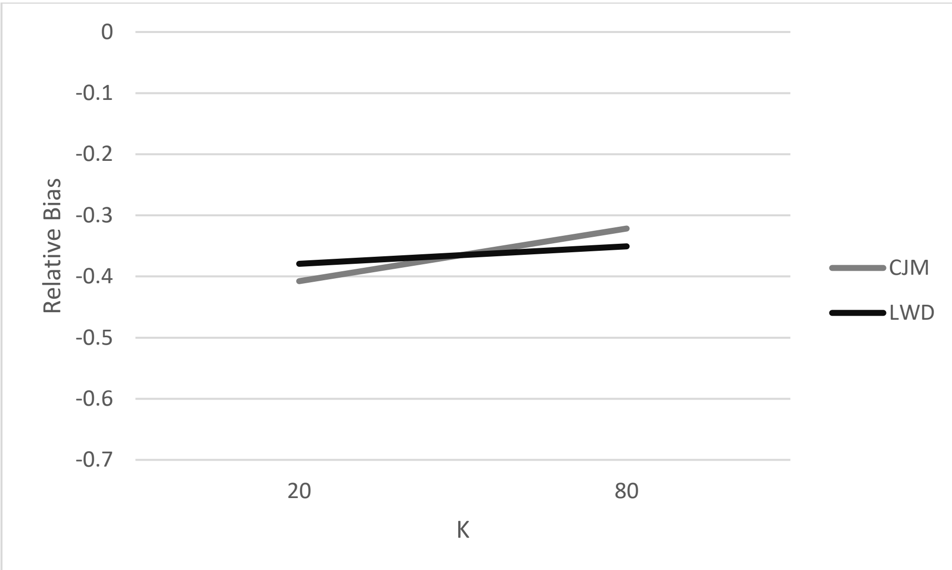

Figure 1. Two-way interaction of the effects of *K* and missing data method on relative bias for $\boldsymbol{\beta_2}$.

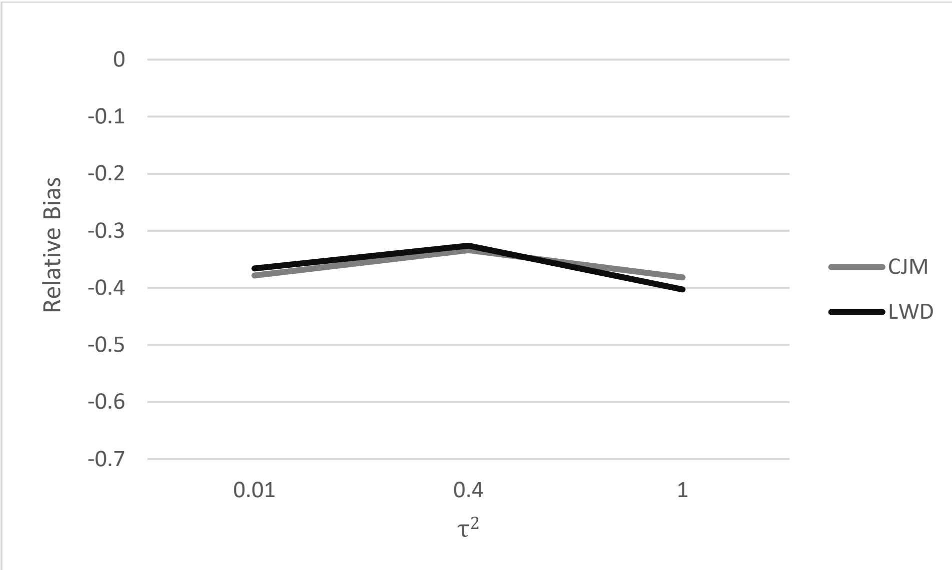

Figure 2. Two-way interaction of the effects of τ^2 and missing data method on relative bias for $\boldsymbol{\beta_2}$.

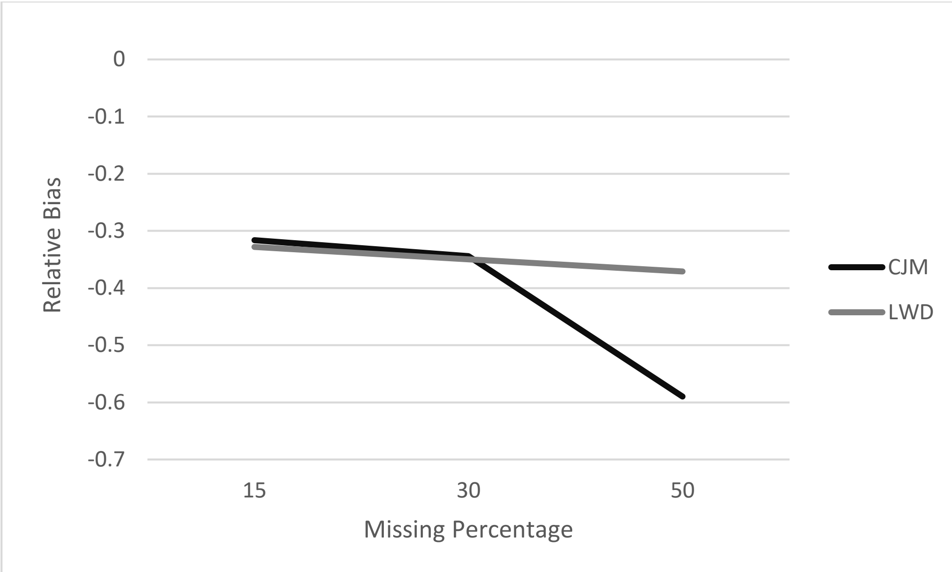

Figure 3. Two-way interaction of the effects of missing percentage and missing data method on relative bias for $\boldsymbol{\beta_2}$.

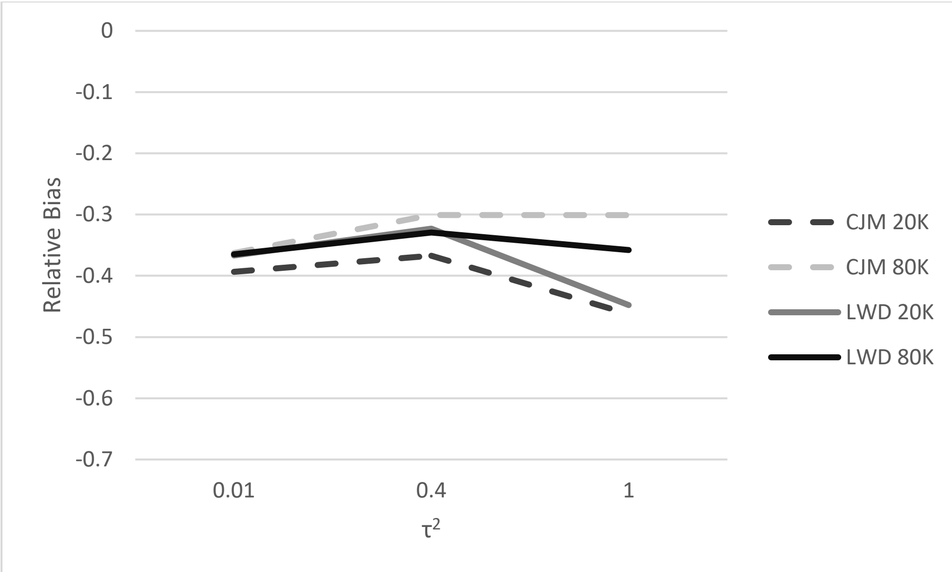

Figure 4. Three-way interaction of the effects of $\boldsymbol{\tau^2}$, *K*, and missing data method on relative bias for $\boldsymbol{\beta_2}$.

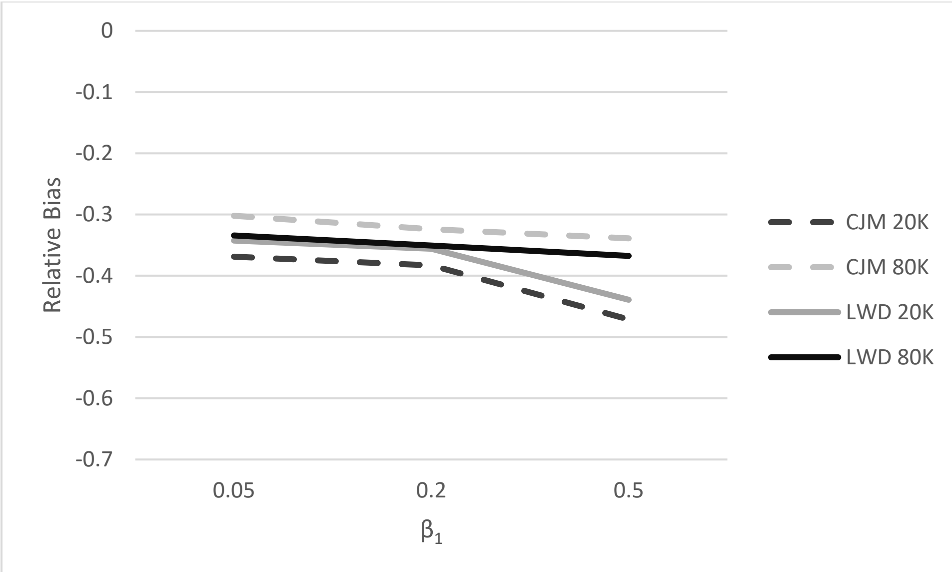

Figure 5. Three-way interaction of the effects of $\boldsymbol{\beta_1}$, *K*, and missing data method on relative bias for $\boldsymbol{\beta_2}$.

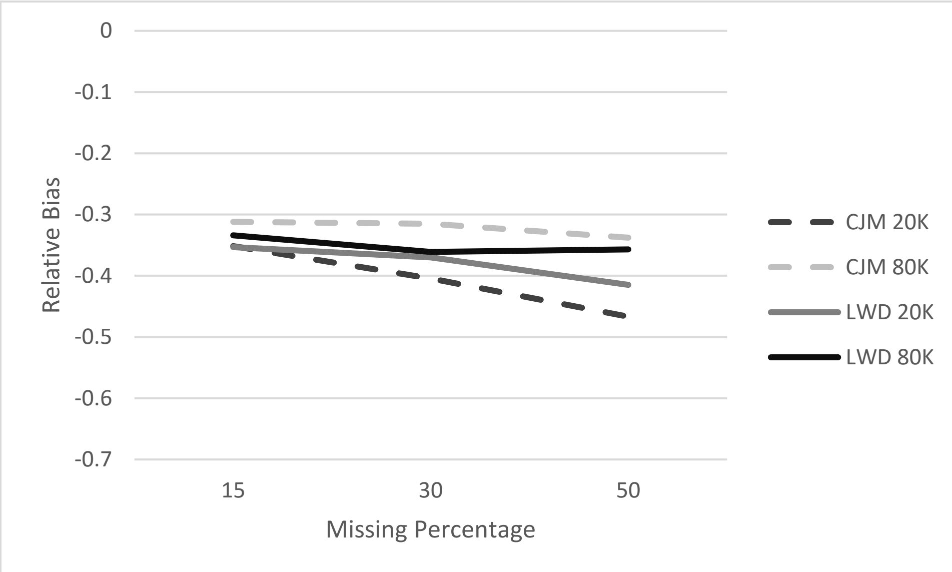

Figure 6. Three-way interaction of the effects of missing percentage, *K*, and missing data method on relative bias for $\boldsymbol{\beta_2}$.

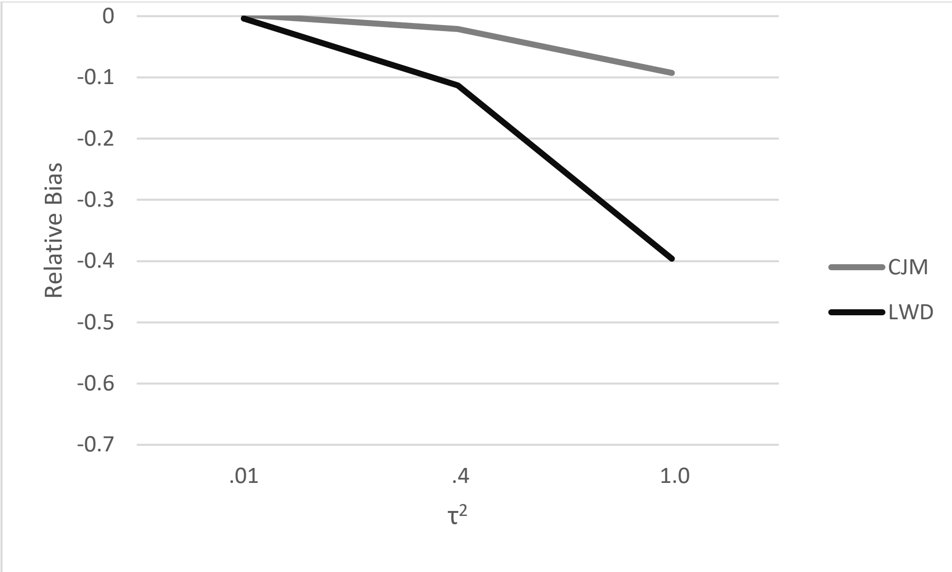

Figure 7. Two-way interaction of the effects of $\boldsymbol{\tau^2}$ and missing data method on relative bias for $\boldsymbol{g}$.

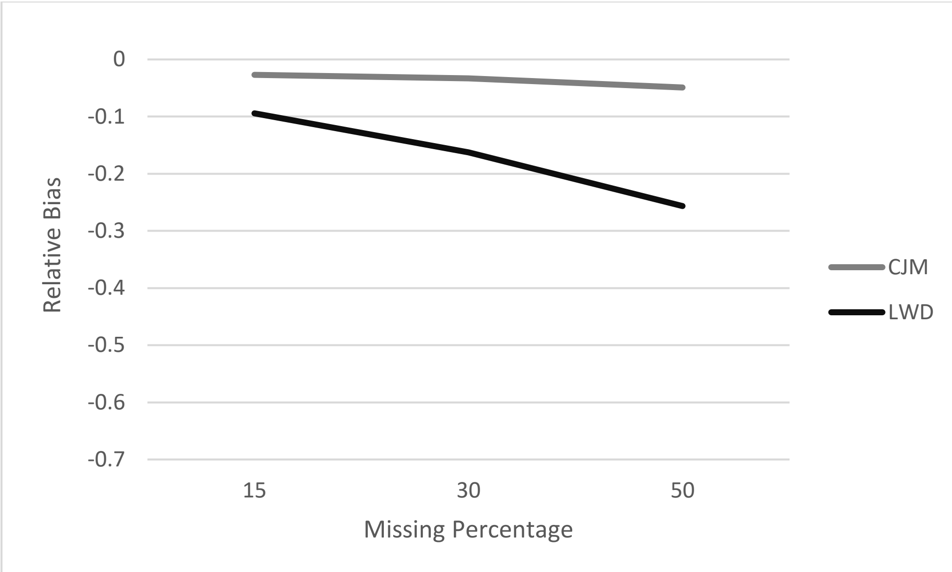

Figure 8. Two-way interaction of the effects of missing percentage and missing data method on relative bias for $\boldsymbol{g}$.

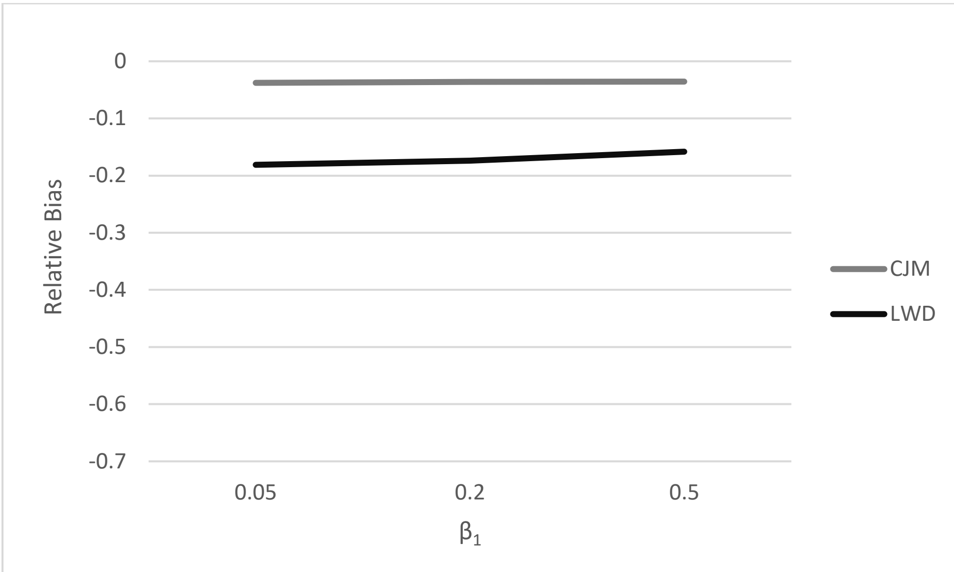

Figure 9. Two-way interaction of the effects of $\boldsymbol{\beta_1}$ and missing data method on relative bias for $\boldsymbol{g}$.

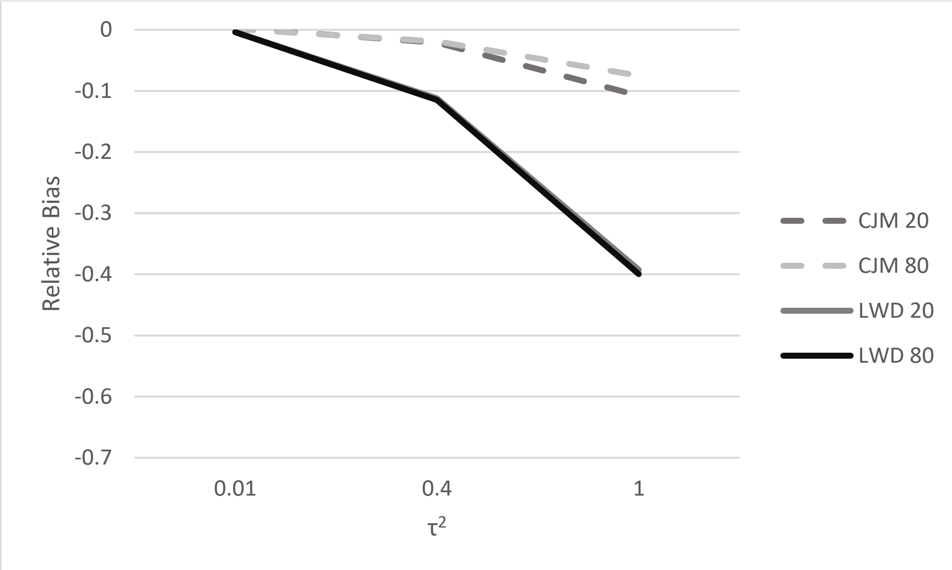

Figure 10. Three-way interaction of the effects of $\boldsymbol{\tau^2}$, *K*, and missing data method on relative bias for $\boldsymbol{g}$.

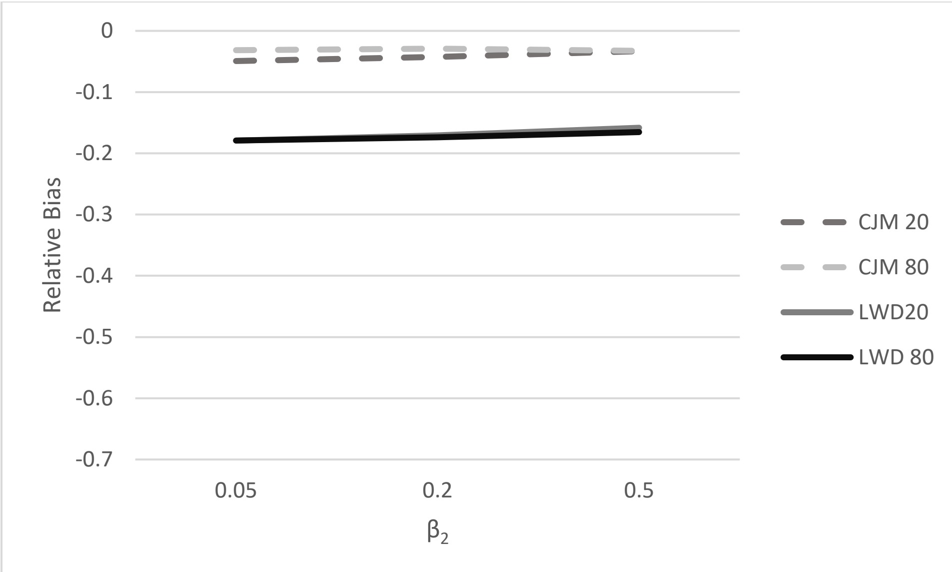

Figure 11. Three-way interaction of the effects of $\boldsymbol{\beta_2}$, *K*, and missing data method on relative bias for $\boldsymbol{g}$.

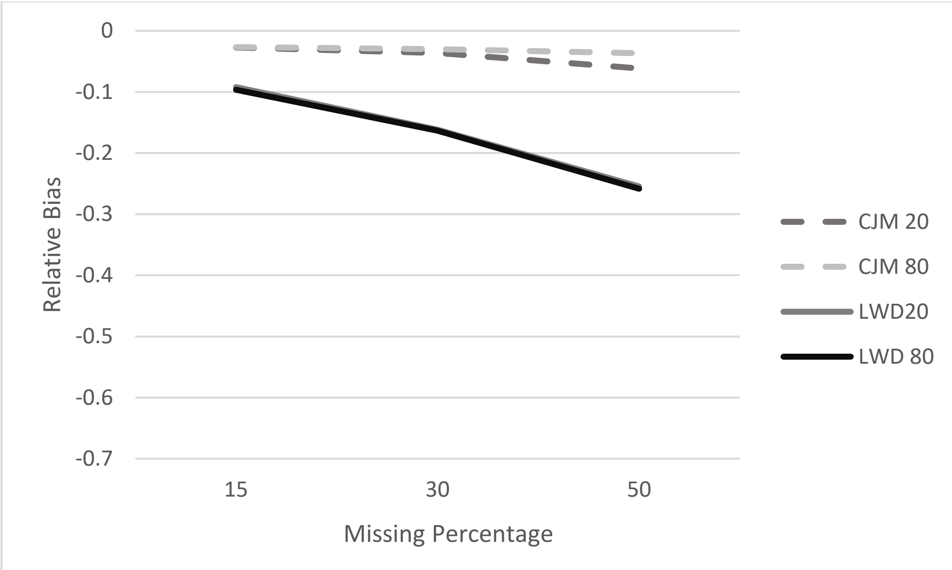

Figure 12. Three-way interaction of the effects of missing percentage, K, and missing data method on relative bias for $\boldsymbol{g}$.

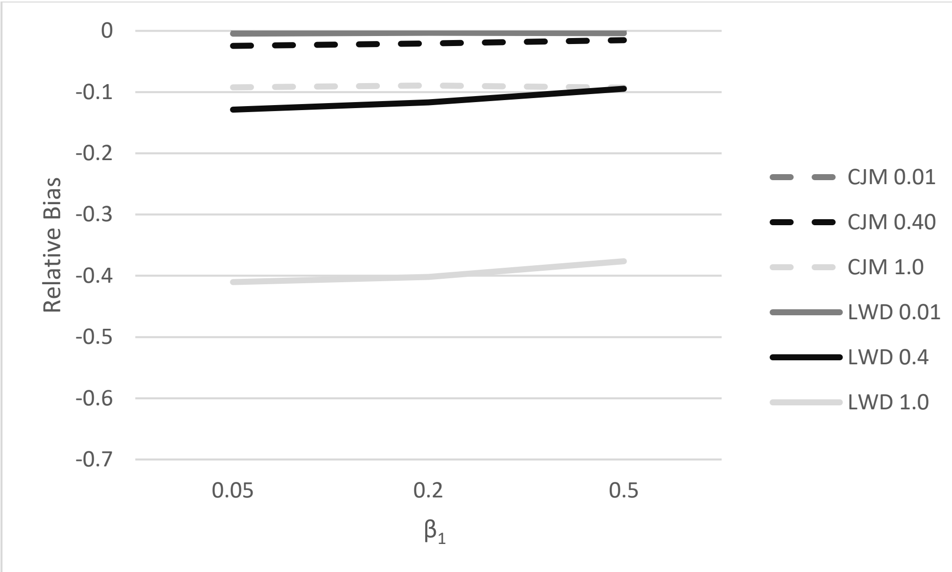

Figure 13. Three-way interaction of the effects of $\boldsymbol{\beta_1}$, $\boldsymbol{\tau^2}$, and missing data method on relative bias for $\boldsymbol{g}$.

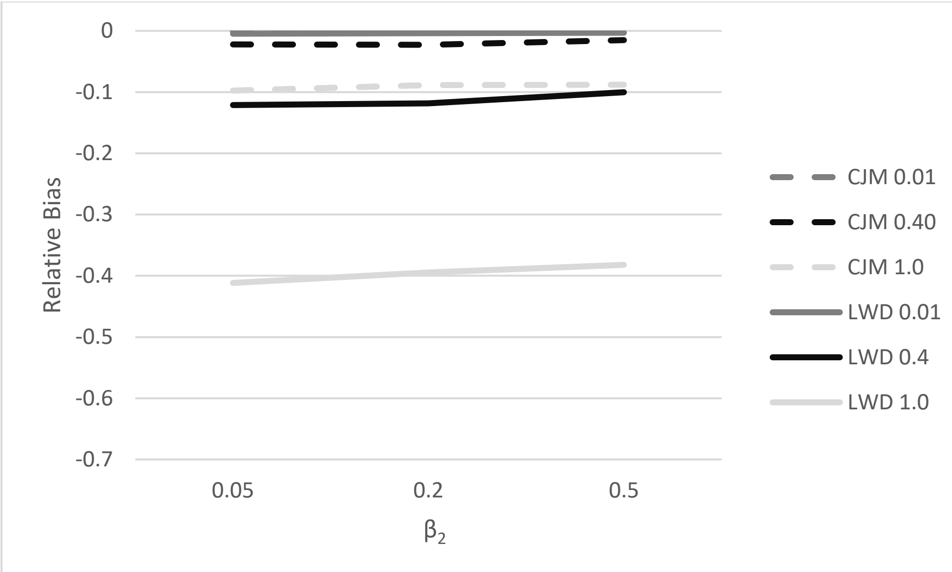

Figure 14. Three-way interaction of the effects of $\boldsymbol{\beta_2}$, $\boldsymbol{\tau^2}$, and missing data method on relative bias for $\boldsymbol{g}$.

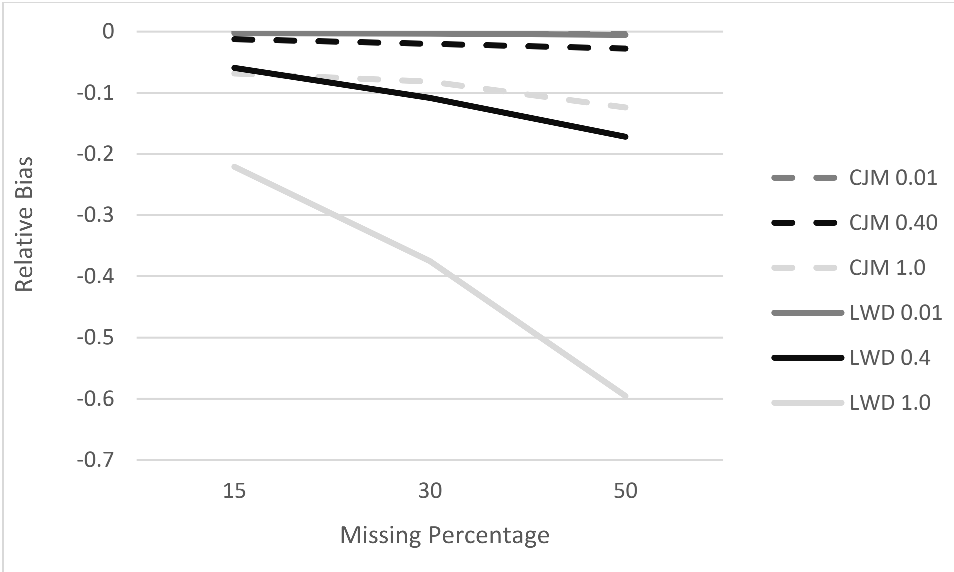

Figure 15. Three-way interaction of the effects of missing percentage, $\boldsymbol{\tau^2}$, and missing data method on relative bias for $\boldsymbol{g}$.

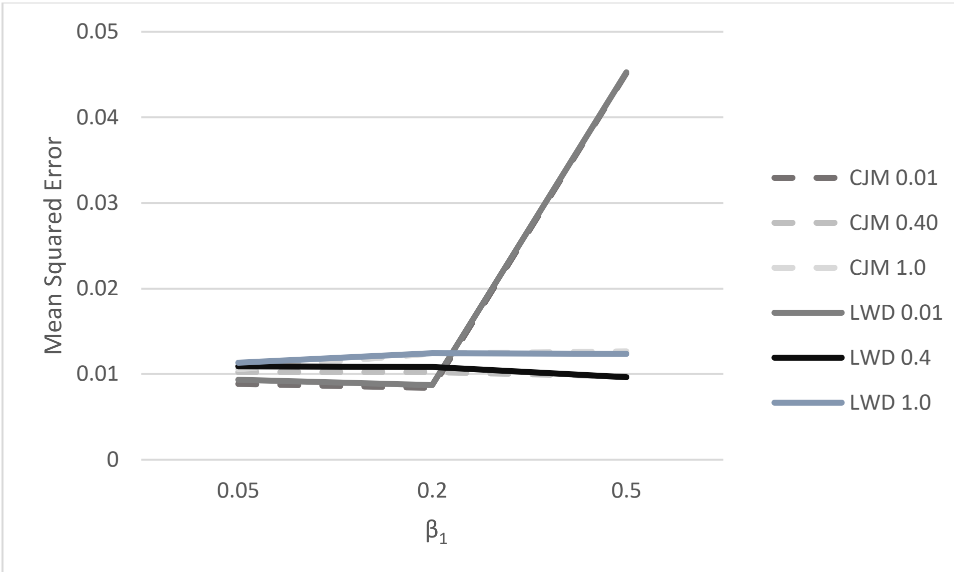

Figure 16. Three-way interaction of the effects of $\boldsymbol{\beta_1}$, $\boldsymbol{\tau^2}$, and missing data method on mean squared error for $\boldsymbol{\beta_2}$.

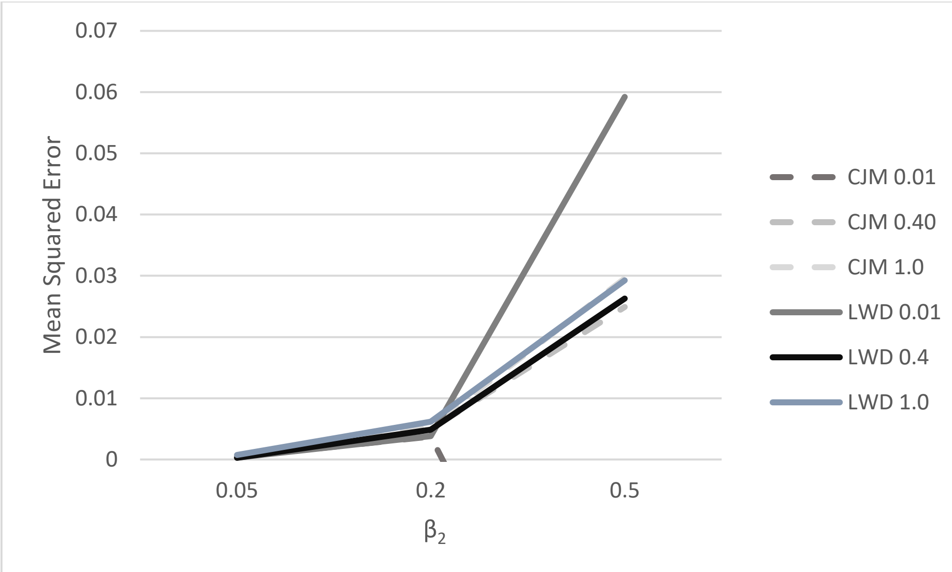

Figure 17. Three-way interaction of the effects of $\boldsymbol{\beta_2}$, $\boldsymbol{\tau^2}$, and missing data method on mean squared error for $\boldsymbol{\beta_2}$.

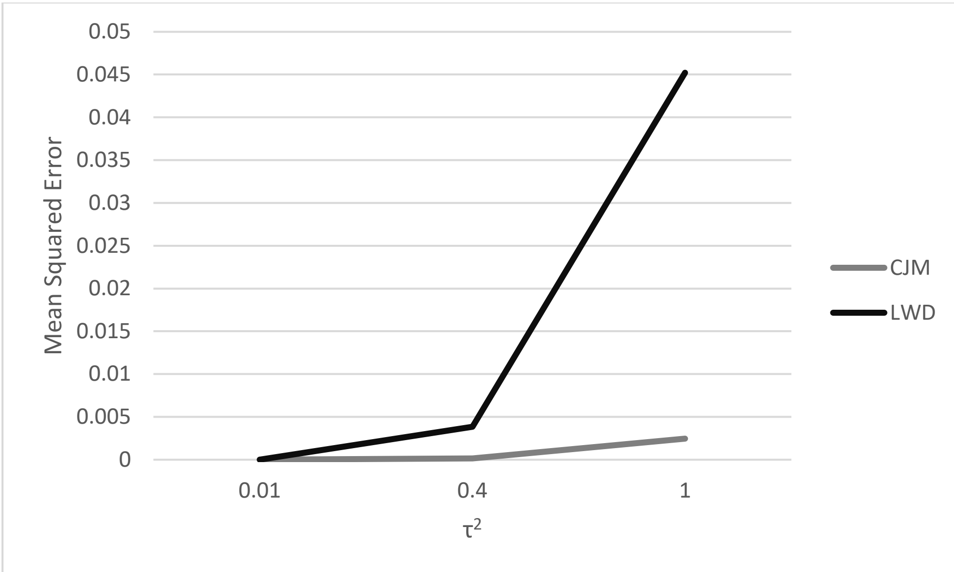

Figure 18. Two-way interaction of the effects of $\boldsymbol{\tau^2}$ and missing data method on mean squared error for $\boldsymbol{g}$.

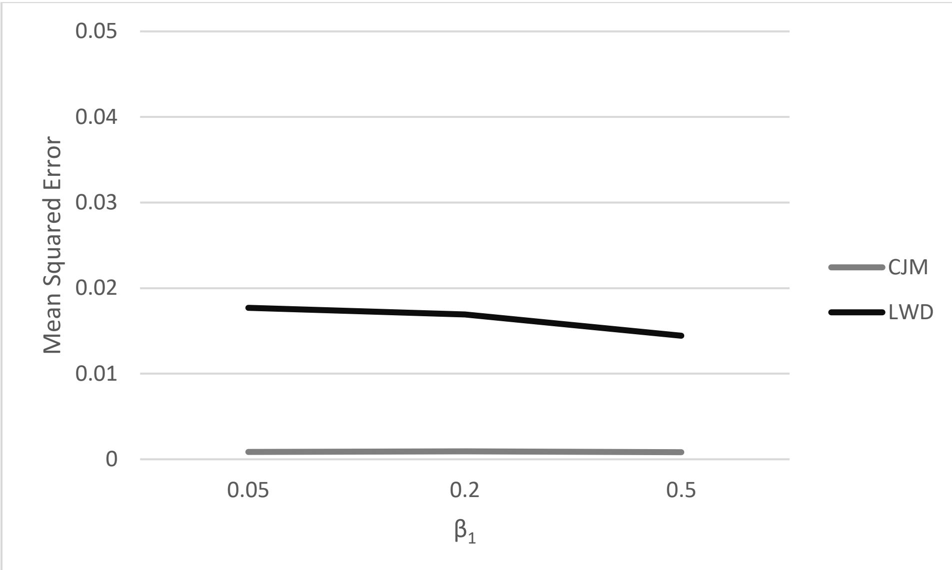

Figure 19. Two-way interaction of the effects of $\boldsymbol{\beta_1}$ and missing data method on mean squared error for $\boldsymbol{g}$.

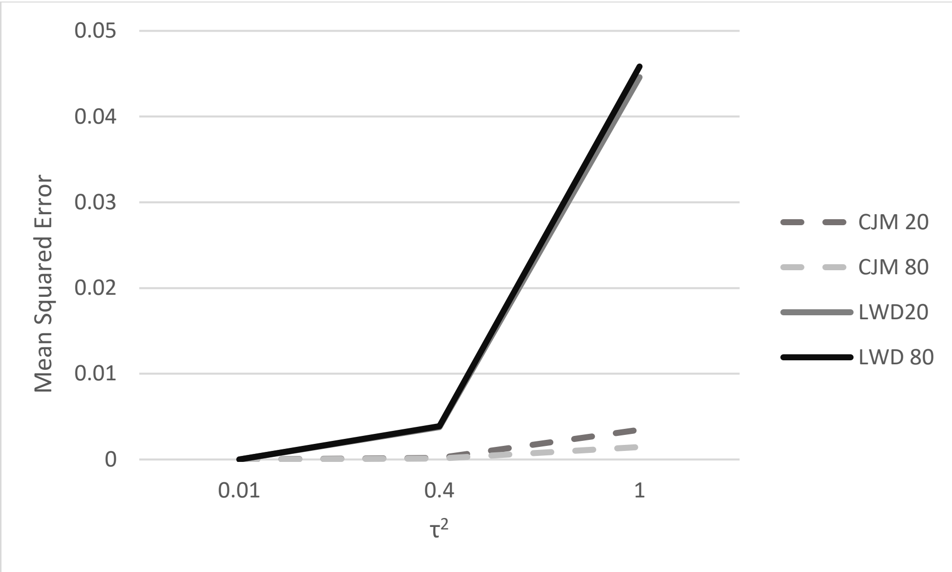

Figure 20. Three-way interaction of the effects of $\boldsymbol{\tau^2}$, *K*, and missing data method on mean squared error for $\boldsymbol{g}$.

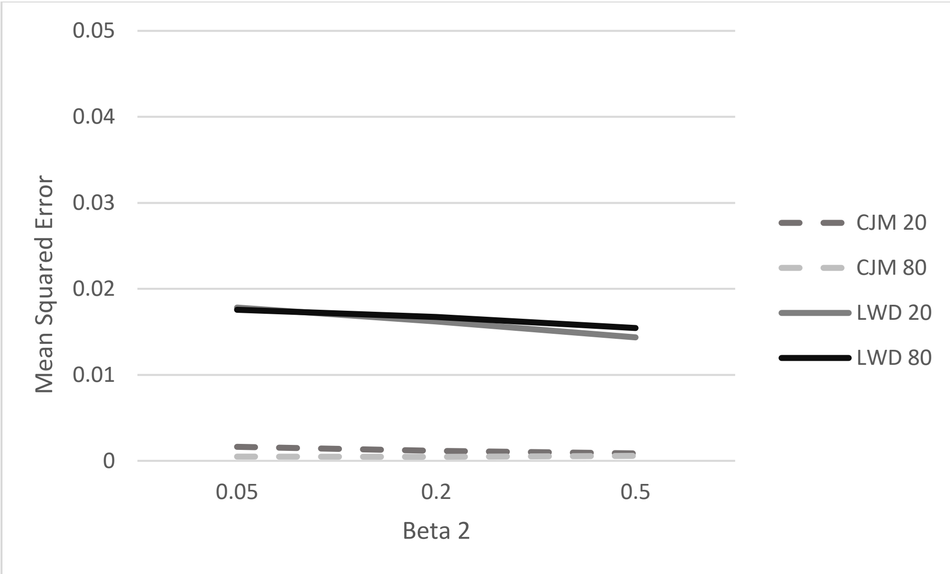

Figure 21. Three-way interaction of the effects of $\boldsymbol{\beta_2}$, *K*, and missing data method on mean squared error for $\boldsymbol{g}$.

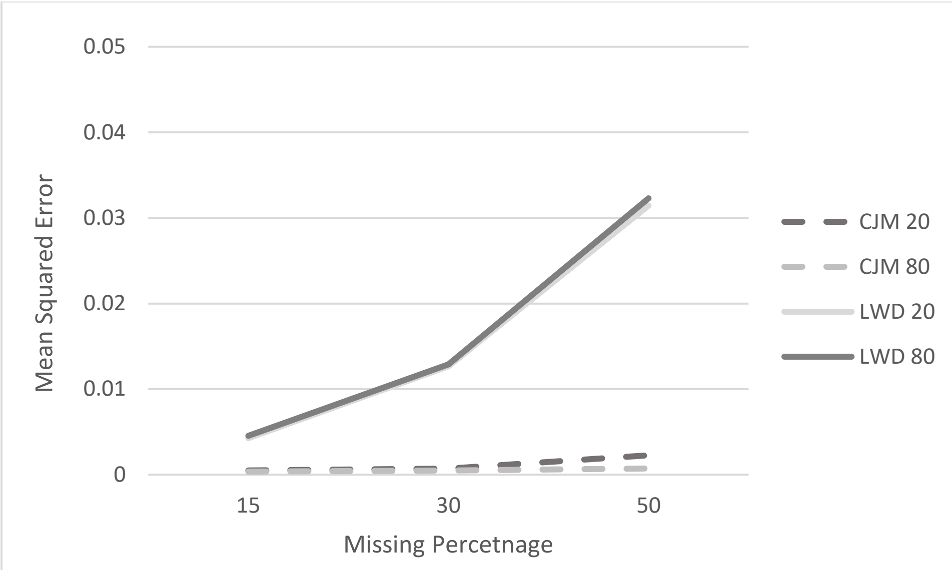

Figure 22. Three-way interaction of the effects of missing percentage, K, and missing data method on mean squared error for $\boldsymbol{g}$.

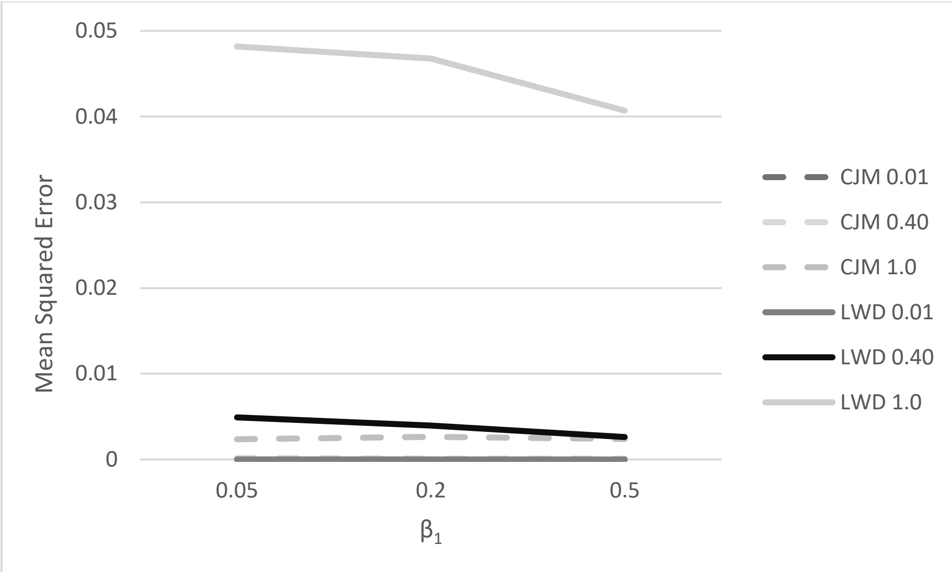

Figure 23. Three-way interaction of the effects of $\boldsymbol{\beta_1}$, $\boldsymbol{\tau^2}$, and missing data method on mean squared error for $\boldsymbol{g}$.

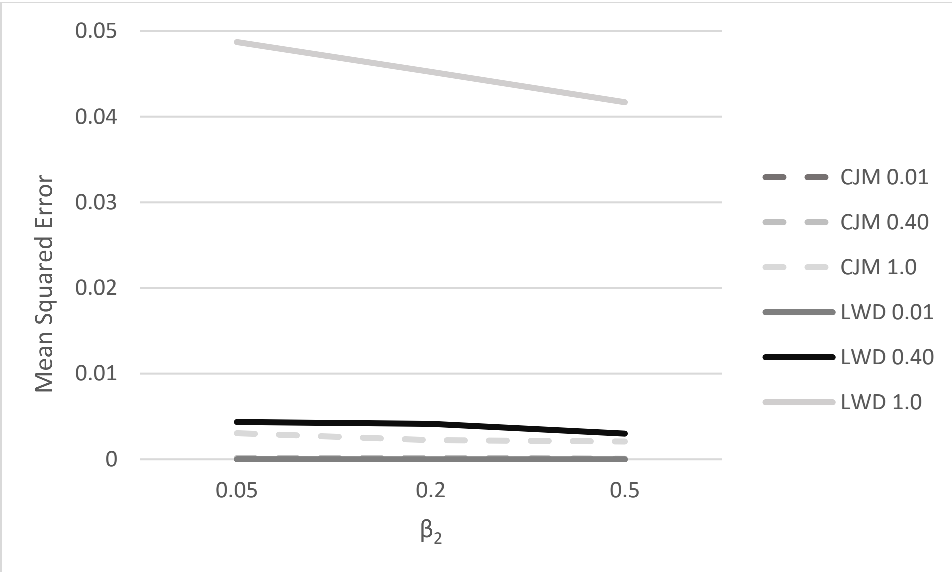

Figure 24. Three-way interaction of the effects of $\boldsymbol{\beta_2}$, $\boldsymbol{\tau^2}$, and missing data method on mean squared error for $\boldsymbol{g}$.

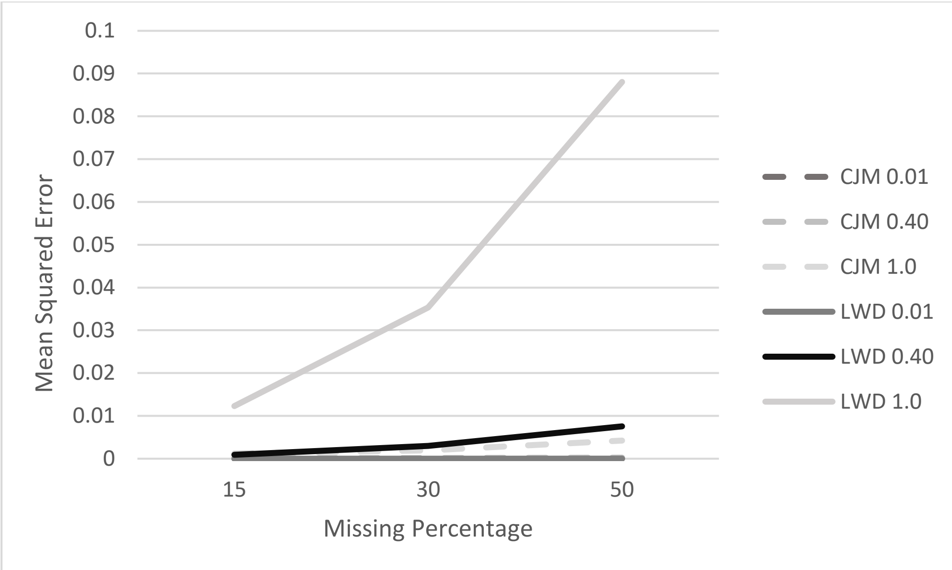

Figure 25. Three-way interaction of the effects of missing percentage, $\boldsymbol{\tau^2}$, and missing data method on mean squared error for $\boldsymbol{g}$.

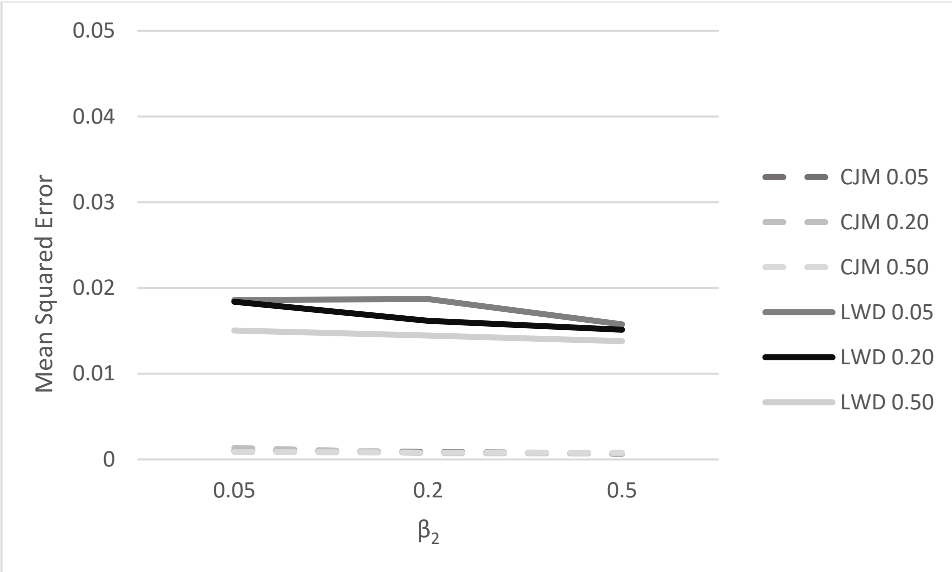

Figure 26. Three-way interaction of the effects of $\boldsymbol{\beta_1}$, $\boldsymbol{\beta_2}$ and missing data method on mean squared error for $\boldsymbol{g}$.

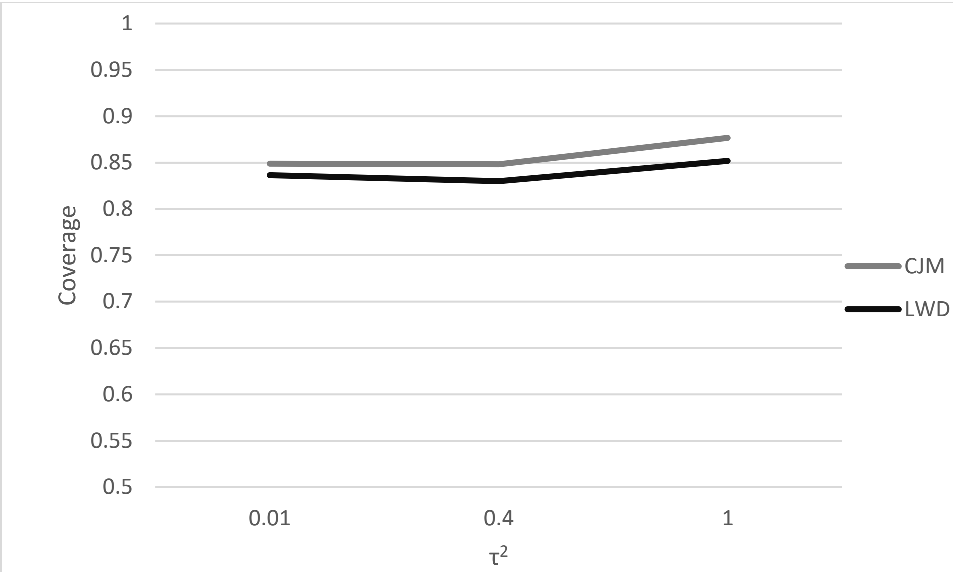

Figure 27. Two-way interaction of the effects of $\boldsymbol{\tau^2}$ and missing data method on coverage for $\boldsymbol{\beta_2}$.

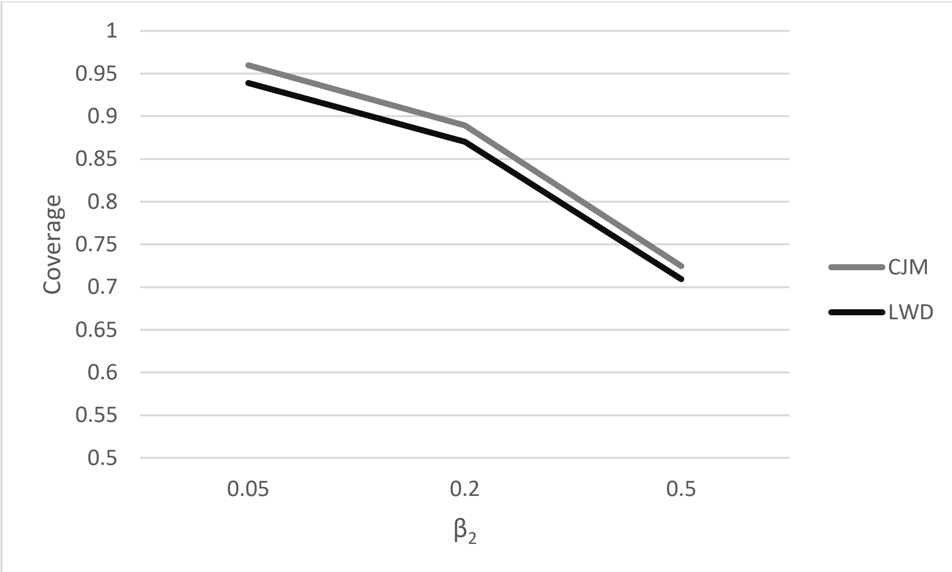

Figure 28. Two-way interaction of the effects of $\boldsymbol{\beta_2}$ and missing data method on coverage for $\boldsymbol{\beta_2}$.

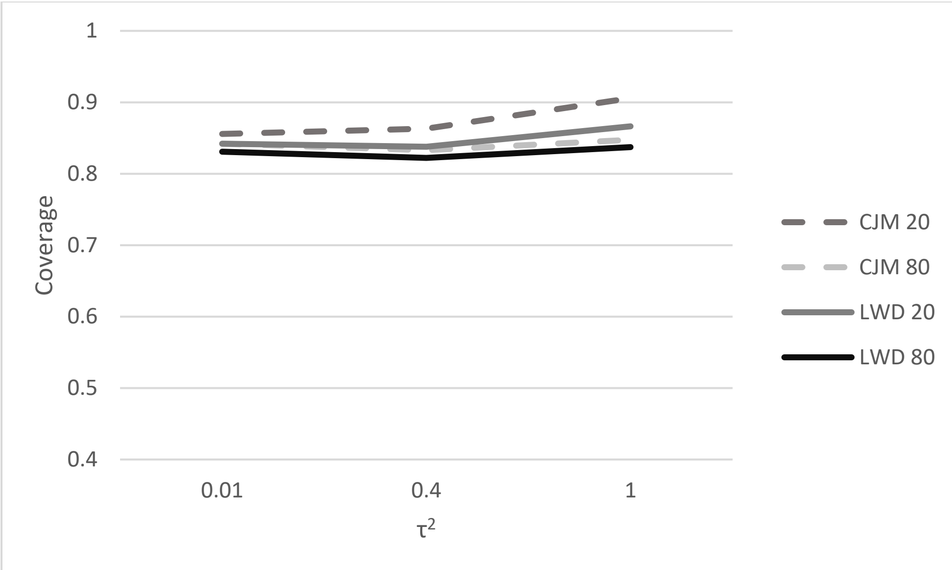

Figure 29. Three-way interaction of the effects of $\boldsymbol{\tau^2}$, *K*, and missing data method on coverage for $\boldsymbol{\beta_2}$.

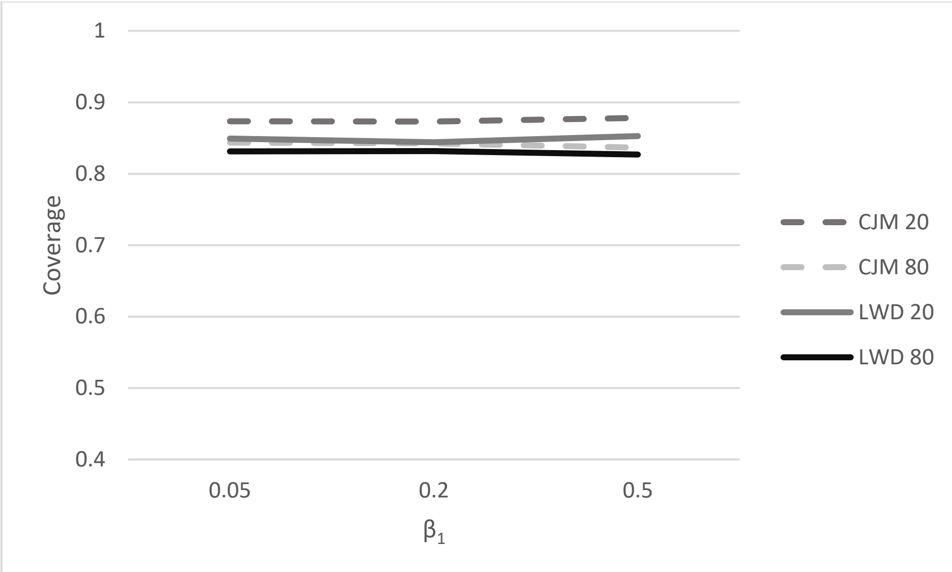

Figure 30. Three-way interaction of the effects of $\boldsymbol{\beta_1}$, *K*, and missing data method on coverage for $\boldsymbol{\beta_2}$.

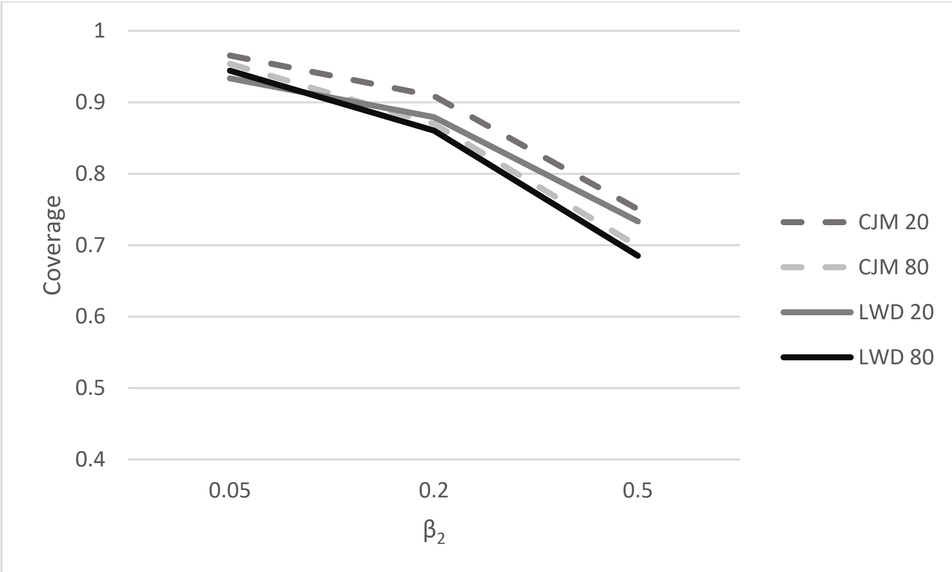

Figure 31. Three-way interaction of the effects of $\boldsymbol{\beta_2}$, *K*, and missing data method on coverage for $\boldsymbol{\beta_2}$.

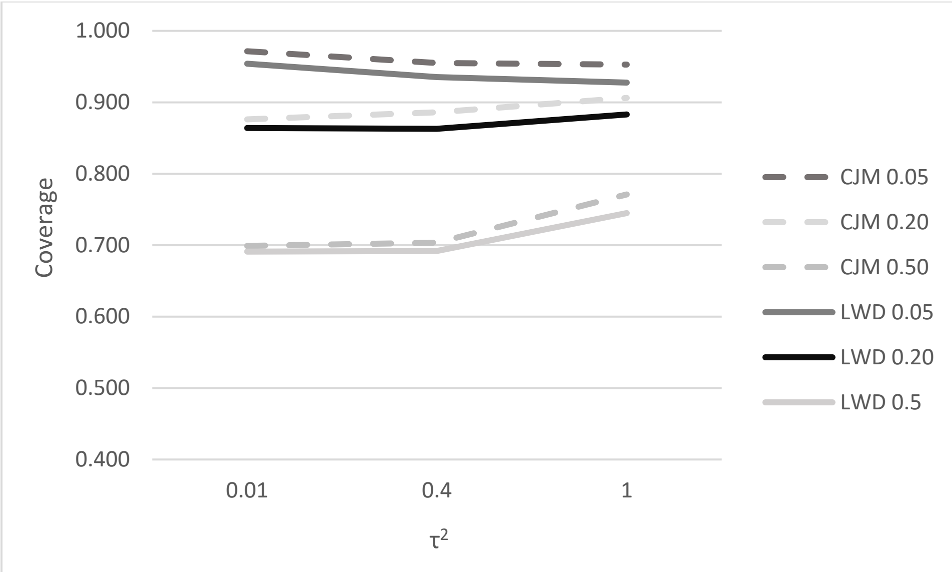

Figure 32. Three-way interaction of the effects of $\boldsymbol{\beta_2}$, $\boldsymbol{\tau^2}$, and missing data method on coverage for $\boldsymbol{\beta_2}$.

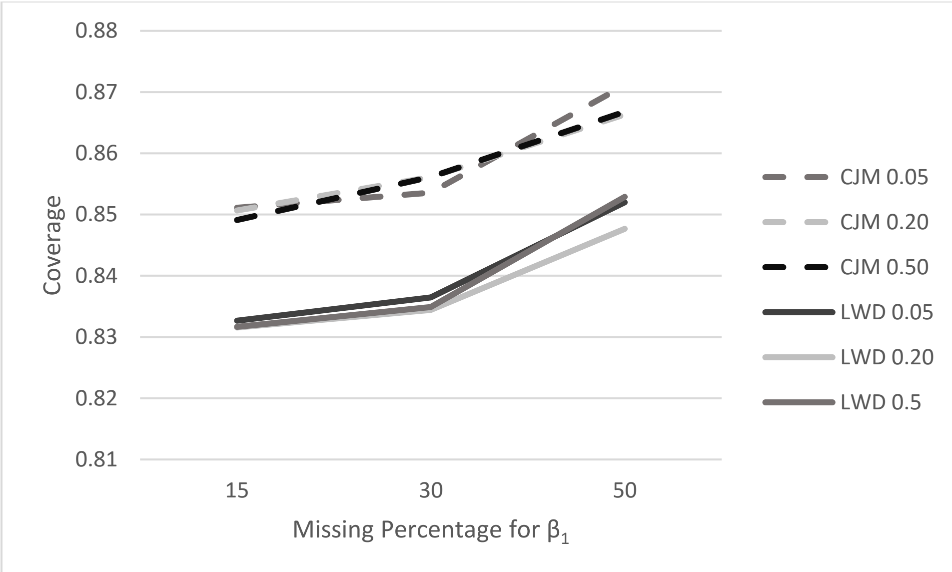

Figure 33. Three-way interaction of the effects of $\boldsymbol{\beta_1}$, missing percentage, and missing data method on coverage for $\boldsymbol{\beta_2}$.

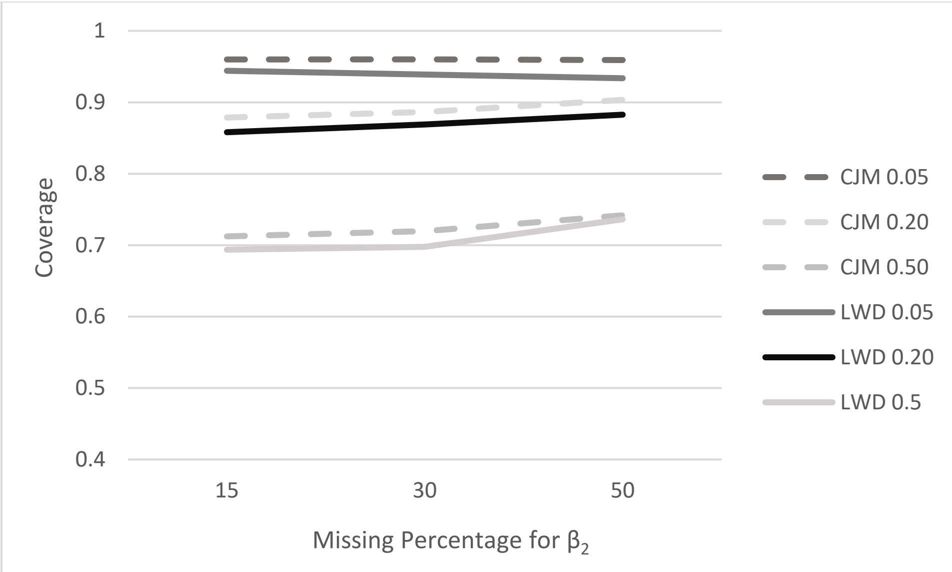

Figure 34. Three-way interaction of the effects of $\boldsymbol{\beta_2}$, missing percentage, and missing data method on coverage for $\boldsymbol{\beta_2}$.

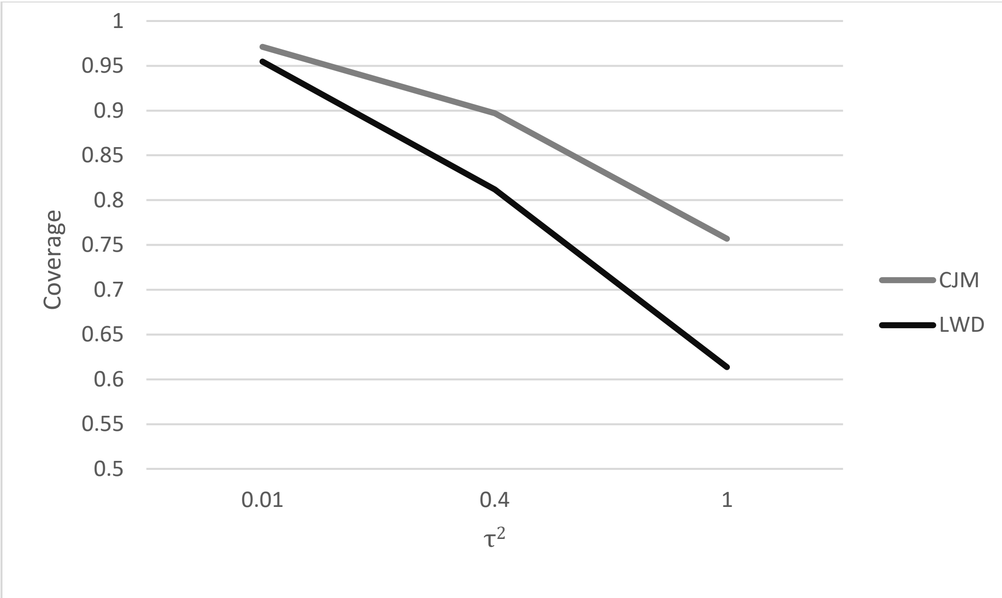

Figure 35. Two-way interaction of the effects of $\boldsymbol{\tau^2}$ and missing data method on coverage for $\boldsymbol{g}$.

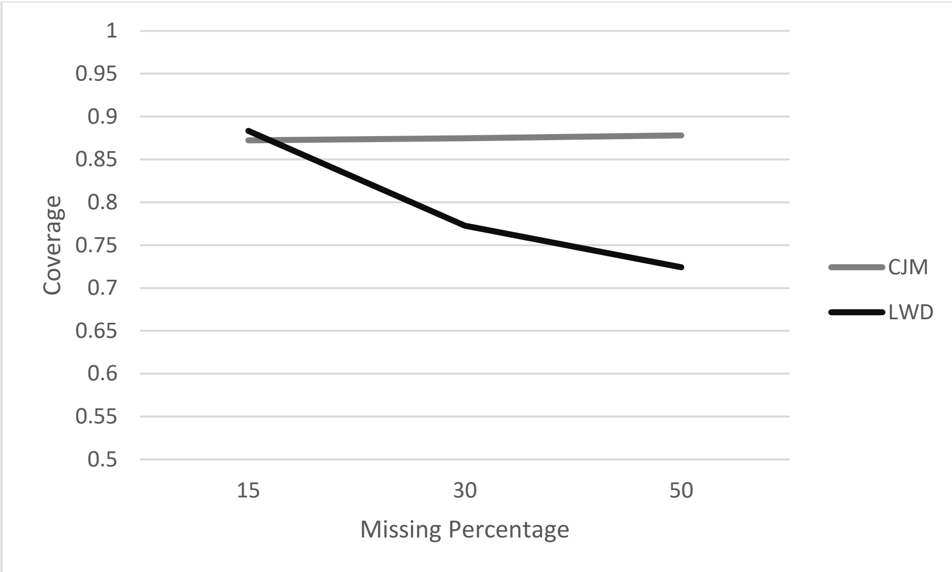

Figure 36. Two-way interaction of the effects of missing percentage and missing data method on coverage for $\boldsymbol{g}$.

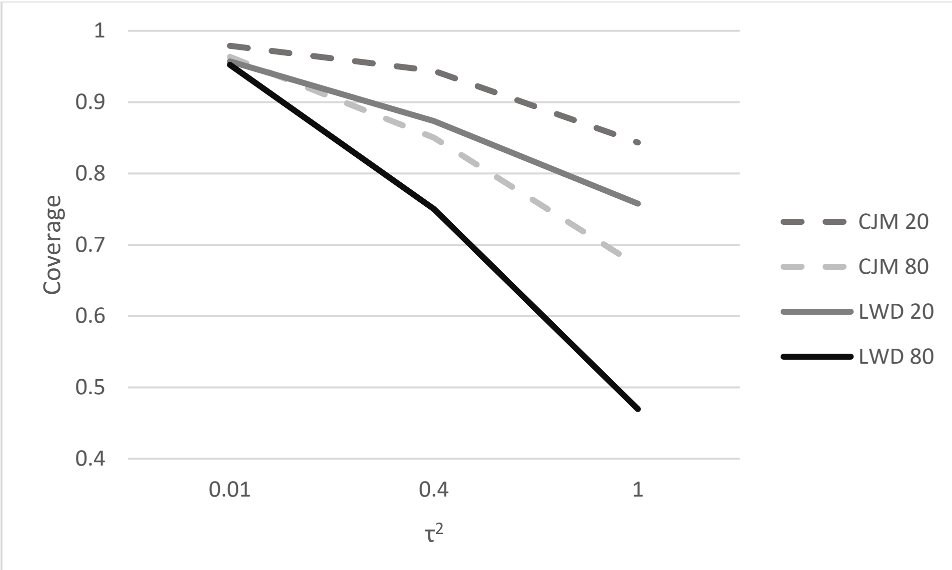

Figure 37. Three-way interaction of the effects of K, $\boldsymbol{\tau^2}$, and missing data method on coverage for $\boldsymbol{g}$.

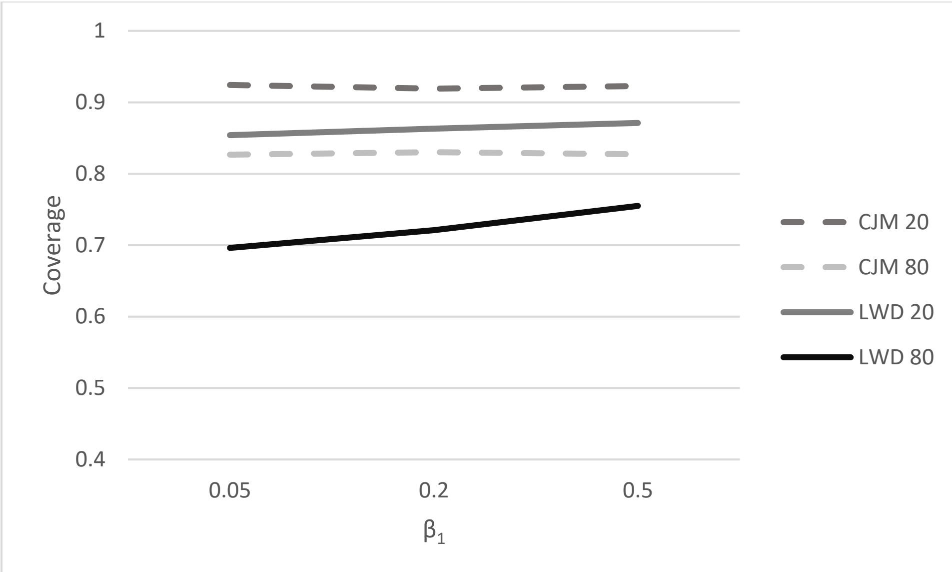

Figure 38. Three-way interaction of the effects of *K*, $\boldsymbol{\beta_1}$,and missing data method on coverage for $\boldsymbol{g}$.

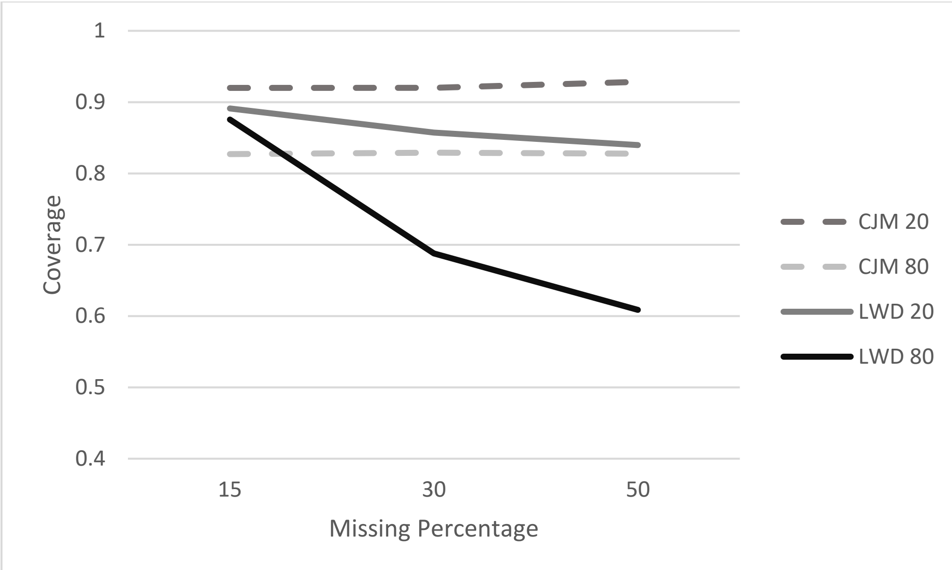

Figure 39. Three-way interaction of the effects of *K*, missing percentage, and missing data method on coverage for $\boldsymbol{g}$.

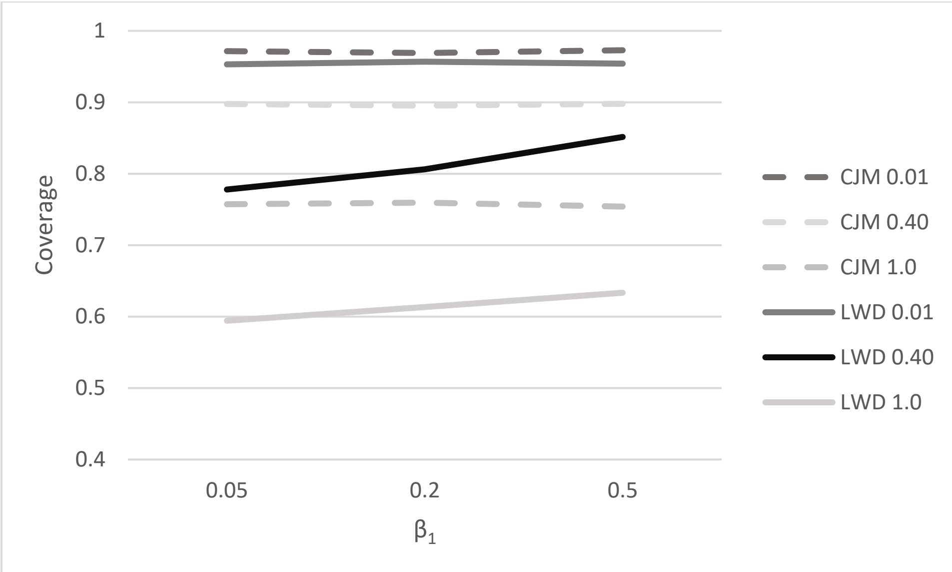

Figure 40. Three-way interaction of the effects of $\boldsymbol{\beta_1}$, $\boldsymbol{\tau^2}$, and missing data method on coverage for $\boldsymbol{g}$.

Table 1

Simulation Conditions

Parameter	# Levels	Conditions
τ^2	3	.01, .40, 10
k	2	20, 80
n*	1	500
β_0	1	.50
β_1	3	.05, .20, .50
β_2	3	.05, .20 .50
Missing Data Handling Procedure	2	Listwise Deletion Conditional Joint Modeling
Percent Missing	3	15, 30, 50

Note. *The combined within-study sample size

Table 2

Bias Performance for Missing β_2

		CJM						Listwise Deletion					
		K = 20			K = 80			K = 20			K = 80		
True β_1	True β_2	15% Missing	30% Missing	50% Missing	15% Missing	30% Missing	50% Missing	15% Missing	30% Missing	50% Missing	15% Missing	30% Missing	50% Missing
τ^2 = 0.01													
0.05	0.05	-0.355	-0.462	-0.521	-0.311	-0.328	-0.361	-0.338	-0.417	-0.376	-0.304	-0.348	-0.314
	0.2	-0.309	-0.329	-0.355	-0.279	-0.282	-0.282	-0.306	-0.329	-0.357	-0.285	-0.307	-0.287
	0.5	-0.319	-0.316	-0.304	-0.285	-0.281	-0.289	-0.322	-0.322	-0.313	-0.291	-0.303	-0.299
0.2	0.05	-0.330	-0.363	-0.437	-0.336	-0.350	-0.423	-0.297	-0.278	-0.299	-0.328	-0.336	-0.380
	0.2	-0.311	-0.302	-0.346	-0.272	-0.279	-0.271	-0.311	-0.299	-0.324	-0.276	-0.301	-0.275
	0.5	-0.306	-0.304	-0.302	-0.282	-0.278	-0.277	-0.308	-0.304	-0.305	-0.289	-0.300	-0.286
0.5	0.05	-0.288	-0.331	-0.523	-0.431	-0.378	-0.520	-0.269	-0.274	-0.371	-0.432	-0.366	-0.503
	0.2	-0.326	-0.323	-0.358	-0.284	-0.279	-0.295	-0.320	-0.318	-0.353	-0.288	-0.298	-0.302
	0.5	-0.730	-0.729	-0.743	-0.714	-0.711	-0.718	-0.728	-0.727	-0.741	-0.715	-0.719	-0.721
τ^2 = 0.4													
0.05	0.05	-0.344	-0.382	-0.274	-0.313	-0.360	-0.359	-0.330	-0.292	-0.023	-0.334	-0.373	-0.357
	0.2	-0.341	-0.369	-0.490	-0.304	-0.309	-0.326	-0.340	-0.346	-0.448	-0.333	-0.364	-0.348
	0.5	-0.308	-0.316	-0.379	-0.293	-0.298	-0.301	-0.308	-0.317	-0.356	-0.312	-0.342	-0.339
0.2	0.05	-0.211	-0.257	-0.521	-0.215	-0.336	-0.362	-0.149	-0.148	-0.520	-0.222	-0.392	-0.377
	0.2	-0.355	-0.388	-0.469	-0.284	-0.285	-0.310	-0.355	-0.384	-0.436	-0.300	-0.341	-0.327
	0.5	-0.316	-0.327	-0.383	-0.292	-0.280	-0.304	-0.322	-0.329	-0.354	-0.312	-0.320	-0.337
0.5	0.05	-0.339	-0.514	-0.506	-0.269	-0.283	-0.272	-0.305	-0.461	-0.292	-0.306	-0.395	-0.288
	0.2	-0.346	-0.391	-0.392	-0.287	-0.298	-0.322	-0.318	-0.340	-0.278	-0.303	-0.325	-0.334
	0.5	-0.309	-0.333	-0.358	-0.286	-0.279	-0.295	-0.316	-0.329	-0.324	-0.299	-0.308	-0.305
τ^2 = 1													
0.05	0.05	-0.272	-0.362	-0.599	-0.330	-0.190	-0.346	-0.412	-0.276	-0.594	-0.382	-0.307	-0.375
	0.2	-0.375	-0.380	-0.448	-0.294	-0.267	-0.273	-0.392	-0.414	-0.380	-0.355	-0.358	-0.343
	0.5	-0.297	-0.348	-0.404	-0.291	-0.289	-0.313	-0.305	-0.329	-0.306	-0.332	-0.363	-0.367
0.2	0.05	-0.381	-0.677	-0.662	-0.441	-0.436	-0.489	-0.403	-0.644	-0.528	-0.493	-0.486	-0.585
	0.2	-0.300	-0.366	-0.521	-0.325	-0.364	-0.400	-0.345	-0.410	-0.489	-0.360	-0.403	-0.434
	0.5	-0.363	-0.388	-0.454	-0.274	-0.285	-0.297	-0.355	-0.372	-0.343	-0.314	-0.343	-0.346
0.5	0.05	-0.615	-0.887	-0.957	-0.134	-0.161	-0.038	-0.607	-0.592	-1.357	-0.187	-0.287	-0.058
	0.2	-0.384	-0.406	-0.444	-0.311	-0.316	-0.358	-0.426	-0.391	-0.388	-0.346	-0.398	-0.403
	0.5	-0.362	-0.379	-0.454	-0.285	-0.306	-0.318	-0.347	-0.342	-0.344	-0.320	-0.367	-0.352

Table 3

Bias Performance for Hedges' g

		CJM						Listwise Deletion					
		K = 20			K = 80			K = 20			K = 80		
True β_1	True β_2	15% Missing	30% Missing	50% Missing	15% Missing	30% Missing	50% Missing	15% Missing	30% Missing	50% Missing	15% Missing	30% Missing	50% Missing
$\tau^2 = 0.01$													
0.05	0.05	-0.001	0.000	-0.002	-0.002	-0.001	0.000	-0.005	-0.004	-0.013	-0.005	-0.005	-0.007
	0.2	0.004	0.007	0.015	-0.001	0.001	0.005	-0.001	-0.001	-0.006	-0.005	-0.008	-0.008
	0.5	0.002	0.007	0.015	0.002	0.004	0.005	-0.002	-0.001	0.000	-0.001	-0.003	-0.006
0.2	0.05	0.000	-0.001	0.001	-0.003	-0.002	-0.002	-0.001	-0.005	-0.005	-0.005	-0.004	-0.004
	0.2	-0.001	0.004	0.005	0.001	0.004	0.006	-0.003	-0.002	-0.005	-0.001	-0.001	-0.004
	0.5	0.002	0.004	0.013	0.000	0.002	0.005	-0.002	-0.002	-0.003	-0.003	-0.004	-0.004
0.5	0.05	-0.001	-0.001	0.000	-0.002	-0.002	-0.002	-0.003	-0.005	0.001	-0.004	-0.005	-0.005
	0.2	0.001	0.004	0.004	0.000	0.002	0.005	-0.002	-0.004	-0.010	-0.002	-0.002	-0.003
	0.5	0.001	0.004	0.004	0.000	0.002	0.005	-0.002	-0.004	-0.010	-0.002	-0.002	-0.003
$\tau^2 = 0.4$													
0.05	0.05	-0.022	-0.042	-0.053	-0.011	-0.013	-0.019	-0.073	-0.143	-0.214	-0.072	-0.127	-0.211
	0.2	-0.020	-0.028	-0.054	-0.017	-0.021	-0.027	-0.069	-0.135	-0.210	-0.072	-0.129	-0.205
	0.5	-0.015	-0.015	-0.025	-0.018	-0.022	-0.023	-0.052	-0.099	-0.169	-0.058	-0.104	-0.174
0.2	0.05	-0.013	-0.023	-0.034	-0.021	-0.024	-0.029	-0.072	-0.119	-0.181	-0.074	-0.125	-0.194
	0.2	-0.022	-0.027	-0.044	-0.015	-0.018	-0.023	-0.070	-0.120	-0.173	-0.063	-0.109	-0.184
	0.5	-0.003	-0.004	-0.008	-0.018	-0.018	-0.025	-0.043	-0.097	-0.158	-0.058	-0.099	-0.161
0.5	0.05	-0.008	-0.015	-0.032	-0.013	-0.015	-0.016	-0.044	-0.087	-0.150	-0.053	-0.095	-0.147
	0.2	0.023	-0.030	-0.038	-0.016	-0.018	-0.018	-0.052	-0.095	-0.148	-0.055	-0.096	-0.148
	0.5	-0.004	-0.008	-0.004	-0.016	-0.019	-0.028	-0.039	-0.085	-0.129	-0.049	-0.089	-0.139
$\tau^2 = 1$													
0.05	0.05	-0.079	-0.099	-0.172	-0.067	-0.078	-0.093	-0.229	-0.387	-0.606	-0.238	-0.417	-0.635
	0.2	-0.075	-0.117	-0.158	-0.055	-0.065	-0.088	-0.243	-0.421	-0.622	-0.230	-0.385	-0.626
	0.5	-0.062	-0.084	-0.119	-0.068	-0.082	-0.097	-0.210	-0.365	-0.561	-0.233	-0.377	-0.601
0.2	0.05	-0.111	-0.072	-0.221	-0.048	-0.057	-0.074	-0.272	-0.424	-0.682	-0.221	-0.380	-0.613
	0.2	-0.056	-0.078	-0.151	-0.061	-0.071	-0.088	-0.183	-0.344	-0.594	-0.225	-0.377	-0.607
	0.5	-0.059	-0.083	-0.138	-0.069	-0.072	-0.095	-0.217	-0.356	-0.552	-0.225	-0.370	-0.590
0.5	0.05	-0.075	-0.097	-0.159	-0.073	-0.080	-0.099	-0.208	-0.347	-0.564	-0.226	-0.374	-0.587
	0.2	-0.074	-0.090	-0.157	-0.058	-0.068	-0.084	-0.196	-0.342	-0.562	-0.213	-0.360	-0.572
	0.5	-0.072	-0.099	-0.148	-0.070	-0.076	-0.090	-0.195	-0.368	-0.551	-0.212	-0.358	-0.537

Table 4

MSE Performance for Missing β_2

		CJM						Listwise Deletion					
		K = 20			K = 80			K = 20			K = 80		
True β_1	True β_2	15% Missing	30% Missing	50% Missing	15% Missing	30% Missing	50% Missing	15% Missing	30% Missing	50% Missing	15% Missing	30% Missing	50% Missing
τ^2 = 0.01													
0.05	0.05	0.000	0.001	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	0.2	0.004	0.004	0.005	0.003	0.003	0.003	0.004	0.004	0.005	0.003	0.004	0.003
	0.5	0.025	0.025	0.023	0.020	0.020	0.021	0.026	0.026	0.024	0.021	0.023	0.022
0.2	0.05	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	0.2	0.004	0.004	0.005	0.003	0.003	0.003	0.004	0.004	0.004	0.003	0.004	0.003
	0.5	0.023	0.023	0.023	0.020	0.019	0.019	0.024	0.023	0.023	0.021	0.023	0.020
0.5	0.05	0.000	0.000	0.001	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.001
	0.2	0.004	0.004	0.005	0.003	0.003	0.003	0.004	0.004	0.005	0.003	0.004	0.004
	0.5	0.133	0.133	0.138	0.127	0.127	0.129	0.133	0.132	0.137	0.128	0.129	0.130
τ^2 = 0.4													
0.05	0.05	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	0.2	0.005	0.005	0.010	0.004	0.004	0.004	0.005	0.005	0.008	0.004	0.005	0.005
	0.5	0.024	0.025	0.036	0.021	0.022	0.023	0.024	0.025	0.032	0.024	0.029	0.029
0.2	0.05	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000
	0.2	0.005	0.006	0.009	0.003	0.003	0.004	0.005	0.006	0.008	0.004	0.005	0.004
	0.5	0.025	0.027	0.037	0.021	0.020	0.023	0.026	0.027	0.031	0.024	0.026	0.028
0.5	0.05	0.000	0.001	0.001	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000
	0.2	0.005	0.006	0.006	0.003	0.004	0.004	0.004	0.005	0.003	0.004	0.004	0.004
	0.5	0.024	0.028	0.032	0.020	0.019	0.022	0.025	0.027	0.026	0.022	0.024	0.023
τ^2 = 1													
0.05	0.05	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000
	0.2	0.006	0.006	0.008	0.003	0.003	0.003	0.006	0.007	0.006	0.005	0.005	0.005
	0.5	0.022	0.030	0.041	0.021	0.021	0.024	0.023	0.027	0.023	0.028	0.033	0.034
0.2	0.05	0.000	0.001	0.001	0.000	0.000	0.001	0.000	0.001	0.001	0.001	0.001	0.001
	0.2	0.004	0.005	0.011	0.004	0.005	0.006	0.005	0.007	0.010	0.005	0.006	0.008
	0.5	0.033	0.038	0.052	0.019	0.020	0.022	0.032	0.035	0.029	0.025	0.029	0.030
0.5	0.05	0.001	0.002	0.002	0.000	0.000	0.000	0.001	0.001	0.005	0.000	0.000	0.000
	0.2	0.006	0.007	0.008	0.004	0.004	0.005	0.007	0.006	0.006	0.005	0.006	0.006
	0.5	0.033	0.036	0.051	0.020	0.023	0.025	0.030	0.029	0.030	0.026	0.034	0.031

Table 5

MSE Performance for Hedges' g

		CJM						Listwise Deletion					
		K = 20			*K* = 80			*K* = 20			*K* = 80		
True β_1	True β_2	15% Missing	30% Missing	50% Missing	15% Missing	30% Missing	50% Missing	15% Missing	30% Missing	50% Missing	15% Missing	30% Missing	50% Missing
τ^2 = 0.01													
0.05	0.05	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	0.2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	0.5	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0.2	0.05	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	0.2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	0.5	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0.5	0.05	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	0.2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	0.5	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
τ^2 = 0 .4													
0.05	0.05	0.000	0.000	0.001	0.000	0.000	0.000	0.001	0.005	0.011	0.001	0.004	0.011
	0.2	0.000	0.000	0.001	0.000	0.000	0.000	0.001	0.005	0.011	0.001	0.004	0.011
	0.5	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.002	0.007	0.001	0.003	0.008
0.2	0.05	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.004	0.008	0.001	0.004	0.009
	0.2	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.004	0.007	0.001	0.003	0.008
	0.5	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.006	0.001	0.002	0.006
0.5	0.05	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.006	0.001	0.002	0.005
	0.2	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.002	0.005	0.001	0.002	0.005
	0.5	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.004	0.001	0.002	0.005
τ^2 = 1													
0.05	0.05	0.002	0.002	0.007	0.001	0.002	0.002	0.013	0.037	0.092	0.014	0.043	0.101
	0.2	0.001	0.003	0.006	0.001	0.001	0.002	0.015	0.044	0.097	0.013	0.037	0.098
	0.5	0.001	0.002	0.004	0.001	0.002	0.002	0.011	0.033	0.079	0.014	0.036	0.090
0.2	0.05	0.003	0.005	0.012	0.001	0.001	0.001	0.019	0.045	0.116	0.012	0.036	0.094
	0.2	0.001	0.002	0.006	0.001	0.001	0.002	0.008	0.030	0.088	0.013	0.036	0.092
	0.5	0.001	0.002	0.005	0.001	0.001	0.002	0.012	0.032	0.076	0.013	0.034	0.087
0.5	0.05	0.001	0.002	0.006	0.001	0.002	0.002	0.011	0.030	0.080	0.013	0.035	0.086
	0.2	0.001	0.002	0.006	0.001	0.001	0.002	0.010	0.029	0.079	0.011	0.032	0.082
	0.5	0.001	0.002	0.006	0.001	0.001	0.002	0.010	0.034	0.076	0.011	0.032	0.072

Table 6

Coverage Performance for Missing β_2

		CJM						Listwise Deletion					
		K = 20			K = 80			K = 20			K = 80		
True β_1	True β_2	15% Missing	30% Missing	50% Missing	15% Missing	30% Missing	50% Missing	15% Missing	30% Missing	50% Missing	15% Missing	30% Missing	50% Missing
τ^2 = 0.01													
0.05	0.05	0.976	0.976	0.982	0.978	0.958	0.95	0.954	0.958	0.954	0.964	0.946	0.94
	0.2	0.864	0.87	0.922	0.84	0.858	0.908	0.856	0.852	0.908	0.84	0.848	0.902
	0.5	0.692	0.69	0.716	0.7	0.698	0.714	0.664	0.676	0.716	0.686	0.674	0.712
0.2	0.05	0.97	0.974	0.982	0.962	0.96	0.962	0.956	0.942	0.966	0.948	0.95	0.946
	0.2	0.878	0.888	0.898	0.848	0.866	0.876	0.85	0.864	0.912	0.834	0.86	0.89
	0.5	0.686	0.7	0.722	0.706	0.69	0.714	0.682	0.686	0.74	0.686	0.666	0.712
0.5	0.05	0.99	0.994	0.99	0.966	0.956	0.96	0.974	0.976	0.972	0.942	0.946	0.938
	0.2	0.864	0.882	0.9	0.848	0.878	0.882	0.84	0.852	0.872	0.842	0.852	0.876
	0.5	0.688	0.698	0.712	0.686	0.68	0.69	0.666	0.69	0.754	0.68	0.658	0.69
τ^2 = 0.4													
0.05	0.05	0.95	0.96	0.976	0.944	0.952	0.938	0.93	0.918	0.936	0.934	0.944	0.934
	0.2	0.898	0.914	0.898	0.85	0.854	0.874	0.876	0.902	0.856	0.828	0.834	0.866
	0.5	0.7	0.706	0.744	0.694	0.686	0.702	0.692	0.694	0.74	0.666	0.654	0.692
0.2	0.05	0.958	0.968	0.96	0.958	0.94	0.946	0.932	0.95	0.916	0.956	0.944	0.93
	0.2	0.908	0.912	0.91	0.864	0.862	0.86	0.878	0.874	0.854	0.85	0.85	0.854
	0.5	0.684	0.714	0.726	0.674	0.684	0.692	0.676	0.676	0.704	0.668	0.672	0.674
0.5	0.05	0.948	0.964	0.948	0.96	0.952	0.966	0.93	0.912	0.902	0.958	0.952	0.954
	0.2	0.912	0.924	0.934	0.85	0.848	0.876	0.888	0.916	0.884	0.83	0.834	0.858
	0.5	0.712	0.726	0.756	0.686	0.684	0.694	0.704	0.716	0.764	0.676	0.664	0.722
τ^2= 1													
0.05	0.05	0.954	0.968	0.956	0.964	0.966	0.96	0.924	0.942	0.902	0.96	0.962	0.954
	0.2	0.93	0.914	0.932	0.872	0.872	0.926	0.89	0.884	0.884	0.866	0.878	0.916
	0.5	0.816	0.824	0.86	0.698	0.698	0.722	0.782	0.818	0.822	0.676	0.672	0.702
0.2	0.05	0.964	0.95	0.952	0.956	0.95	0.956	0.932	0.906	0.89	0.956	0.924	0.944
	0.2	0.912	0.924	0.948	0.874	0.886	0.886	0.866	0.9	0.896	0.86	0.878	0.88
	0.5	0.806	0.826	0.852	0.704	0.716	0.752	0.742	0.794	0.806	0.696	0.684	0.744
0.5	0.05	0.952	0.956	0.95	0.928	0.938	0.93	0.926	0.9	0.904	0.918	0.928	0.924
	0.2	0.936	0.934	0.936	0.868	0.866	0.894	0.894	0.896	0.896	0.858	0.868	0.884
	0.5	0.812	0.834	0.858	0.678	0.696	0.726	0.782	0.794	0.82	0.662	0.674	0.738

Table 7

Coverage Performance for Hedges' g

		CJM						Listwise Deletion					
		K = 20			K = 80			K = 20			K = 80		
True β_1	True β_2	15% Missing	30% Missing	50% Missing	15% Missing	30% Missing	50% Missing	15% Missing	30% Missing	50% Missing	15% Missing	30% Missing	50% Missing
τ^2 = 0.01													
0.05	0.05	0.992	0.992	0.988	0.970	0.970	0.964	0.947	0.963	0.964	0.947	0.950	0.958
	0.2	0.968	0.974	0.980	0.958	0.958	0.960	0.933	0.950	0.970	0.967	0.937	0.956
	0.5	0.970	0.974	0.986	0.960	0.972	0.954	0.940	0.973	0.970	0.933	0.960	0.938
0.2	0.05	0.972	0.974	0.982	0.964	0.964	0.970	0.987	0.967	0.964	0.940	0.963	0.932
	0.2	0.978	0.980	0.980	0.958	0.958	0.946	0.947	0.947	0.954	0.953	0.967	0.950
	0.5	0.972	0.966	0.984	0.962	0.968	0.964	0.967	0.947	0.968	0.947	0.963	0.962
0.5	0.05	0.990	0.990	0.988	0.954	0.954	0.952	0.940	0.963	0.972	0.933	0.927	0.960
	0.2	0.968	0.986	0.972	0.972	0.974	0.962	0.953	0.960	0.960	0.960	0.960	0.948
	0.5	0.960	0.976	0.990	0.972	0.976	0.976	0.913	0.953	0.970	0.967	0.963	0.972
τ^2 = 0.4													
0.05	0.05	0.942	0.946	0.944	0.838	0.850	0.844	0.927	0.817	0.808	0.833	0.633	0.506
	0.2	0.934	0.934	0.946	0.850	0.846	0.862	0.887	0.850	0.812	0.853	0.670	0.556
	0.5	0.950	0.954	0.972	0.846	0.852	0.844	0.860	0.883	0.840	0.880	0.733	0.654
0.2	0.05	0.936	0.932	0.946	0.844	0.844	0.866	0.853	0.897	0.826	0.793	0.677	0.600
	0.2	0.924	0.924	0.940	0.858	0.852	0.854	0.893	0.873	0.864	0.853	0.720	0.612
	0.5	0.944	0.944	0.942	0.858	0.852	0.856	0.933	0.860	0.842	0.933	0.760	0.718
0.5	0.05	0.950	0.952	0.952	0.872	0.872	0.874	0.927	0.907	0.848	0.940	0.800	0.736
	0.2	0.944	0.942	0.936	0.840	0.848	0.842	0.913	0.887	0.840	0.887	0.740	0.706
	0.5	0.942	0.940	0.968	0.818	0.836	0.834	0.960	0.890	0.886	0.933	0.763	0.764
τ^2 = 1													
0.05	0.05	0.824	0.826	0.830	0.670	0.676	0.666	0.827	0.717	0.700	0.773	0.310	0.172
	0.2	0.830	0.848	0.846	0.676	0.676	0.666	0.767	0.723	0.680	0.773	0.370	0.164
	0.5	0.858	0.864	0.882	0.660	0.670	0.662	0.840	0.767	0.744	0.780	0.373	0.216
0.2	0.05	0.844	0.840	0.856	0.688	0.688	0.690	0.827	0.750	0.666	0.827	0.397	0.202
	0.2	0.838	0.838	0.854	0.658	0.658	0.660	0.807	0.733	0.720	0.767	0.357	0.214
	0.5	0.830	0.852	0.844	0.682	0.674	0.678	0.840	0.767	0.706	0.787	0.430	0.244
0.5	0.05	0.842	0.838	0.846	0.666	0.664	0.662	0.787	0.723	0.726	0.827	0.413	0.234
	0.2	0.816	0.818	0.832	0.672	0.672	0.668	0.867	0.717	0.738	0.813	0.423	0.274
	0.5	0.854	0.840	0.880	0.666	0.662	0.674	0.820	0.763	0.736	0.840	0.413	0.286