CONTRIBUTIONS TO MODELING AND ANALYSIS OF METHOD COMPARISON DATA

by

Lak Nilusha Kotinkaduwe Rankothgedara

APPROVED BY SUPERVISORY COMMITTEE:

___

Pankaj K. Choudhary, Chair

___

Swati Biswas

___

Sam Efromovich

___

Frank Konietschke

To my beloved parents,

wonderful husband,

and

lovely little daughter.

CONTRIBUTIONS TO MODELING AND ANALYSIS OF

METHOD COMPARISON DATA

by

LAK NILUSHA KOTINKADUWE RANKOTHGEDARA, BS, MS

DISSERTATION

Presented to the Faculty of

The University of Texas at Dallas

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY IN

STATISTICS

THE UNIVERSITY OF TEXAS AT DALLAS

December 2018

ACKNOWLEDGMENTS

First, I would like to thank my advisor, Professor Pankaj Choudhary, for all his guidance, generosity with his time, and for offering me his mentorship, fatherly love, and care. Under his guidance, I successfully overcame many difficulties and learnt a lot. His deep knowledge of statistics, enthusiasm for the subject, and patience were invaluable in writing this dissertation. I could not have imagined having a better advisor for my Ph.D. study.

I am also very grateful to the members of my committee, namely, Professor Sam Efromovich, Professor Swati Biswas, and Professor Frank Konietschke, for their valuable advice and suggestions. I would also like to express my gratitude to the Mathematical Sciences Department at UTD for providing me an opportunity to pursue my Ph.D. by offering me admission and financial support. My sincere thanks go to everyone at the department for the warm and friendly atmosphere they provided, especially during and after my pregnancy. Thank you, to all my friends for their enormous support and being there whenever I needed a friend.

My deepest gratitude goes to my parents for all the support and love they have given me throughout my life. I am also grateful to my sisters and in-laws for their support during my Ph.D. study.

Finally, I thank my husband, Jagath, and our daughter, Leona, for their love and patience. Thank you, Jagath, for always being there with me and encouraging me in every step of this journey. I would not be here without you.

November 2018

CONTRIBUTIONS TO MODELING AND ANALYSIS OF

METHOD COMPARISON DATA

Lak Nilusha Kotinkaduwe Rankothgedara, PhD
The University of Texas at Dallas, 2018

Supervising Professor: Pankaj K. Choudhary, Chair

Method comparison studies compare a new method of measuring a continuous variable with an established method that serves as a reference. Both methods have the same unit of measurement and none of them is considered error free. The major goals in these studies are to quantify the degree of similarity and agreement between the two methods. The motivation behind the comparison is that if two methods agree well, the cheaper, simpler, or the less invasive among them can be preferred or both can be used interchangeably. Such studies are common in biomedical sciences with medical devices, assays, measurement protocols, or clinical observers serving as methods. The most popular design for conducting these studies is the paired measurements design, which leads to one measurement by each method on every subject. These paired measurements method comparison data are often analyzed by modeling them using the classical measurement error model or a special case of it, a mixed-effects model. Motivated by real applications, this dissertation makes two contributions toward modeling and analysis of these data.

First, we develop a segmented measurement error model assuming equal error variances. This model extends the classical measurement error model to allow a piecewise linear relationship between the measurements. The changepoint at which the transition takes place is treated as an unknown parameter in the model. We provide an expectation conditional maximization

(ECM) algorithm to fit the model and propose segmented-specific evaluation of similarity and agreement using appropriate extensions of the existing measures. Bootstrapping and large-sample theory of maximum likelihood estimators are used to perform the relevant inferences. We are also able to obtain an explicit expression for the Hessian matrix that is needed for this purpose. The proposed methodology is evaluated by simulation and is illustrated by analyzing a dataset containing measurements of digoxin concentration. This work is also generalized to allow unequal error variances in the segmented model.

Second, we develop a Bayesian approach that uses informative priors for error variances within a mixed-effect model framework. This approach allows taking advantage of information about error variances that may be available from previous studies, potentially leading to their improved estimation. Half-normal and hierarchical half-normal distributions are used as prior distributions for error variances and data from previous studies are used to estimate the hyperparameters of these distributions. We discuss strategies for posterior simulation to estimate the model parameters and their functions. The proposed methodology is compared with its likelihood-based counterpart in a simulation study. It is illustrated by analyzing a dataset containing oxygen saturation measurements.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Chapter organization

This dissertation focuses on paired measurements data generated in a method comparison study and makes two contributions toward their modeling and analysis. The first contribution concerns the development of a segmented measurement error model that allows a piecewise linear relationship between the measurement methods under comparison. This model is an extension of the classical measurement error model that assumes a linear relationship over the entire measurement range. In Chapter 2, we present the proposed segmented model and study the distribution theory and inference under this model assuming equal error variances for the measurement methods. The same issues are discussed in Chapter 3 but under a more general segmented model that does not assume equality of error variances. The second contribution concerns the development of a Bayesian approach for analyzing data which takes into account the information about error variances of the measurement methods that may be available from previous studies by using informative priors for error variances. This methodology is the subject of Chapter 4. In both Chapters 2 and 4, we also use Monte Carlo simulation to study the properties of the proposed methodologies and illustrate their application by analyzing datasets from the literature. We conclude in Chapter 5 by mentioning some ongoing and future work.

1.2 Preliminaries

1.2.1 Notation

Let $\phi(x)$ and $\Phi(x)$ respectively denote the probability density function and the cumulative distribution function of a $N(0,1)$ distribution. For ease of exposition, we define the following quantities that will be used in Chapters 2 and 3 of this dissertation.

$$
\begin{aligned}
g_1(x, \mu, \sigma) &= \frac{x - \mu}{\sigma}, \\
g_2(x, \mu, \sigma) &= \frac{\phi\left(g_1(x, \mu, \sigma)\right)}{1 - \Phi\left(g_1(x, \mu, \sigma)\right)}, \\
g_3(x, \mu, \sigma) &= \frac{\phi\left(g_1(x, \mu, \sigma)\right)}{\Phi\left(g_1(x, \mu, \sigma)\right)}, \\
g_4(\beta, \sigma_1, \sigma_2) &= \left(\frac{1}{\sigma_1^2} + \frac{(\beta_1 + \beta)^2}{\sigma_2^2} + \frac{1}{\sigma_b^2}\right)^{-1}, \\
g_5(\beta, \sigma_1, \sigma_2) &= \left(\frac{y_1}{\sigma_1^2} + \frac{(y_2 - \beta_0 + \beta\psi)(\beta_1 + \beta)}{\sigma_2^2} + \frac{\mu_b}{\sigma_b^2}\right), \\
g_6(\beta, \sigma_1, \sigma_2) &= g_4(\beta, \sigma_1, \sigma_2) g_5(\beta, \sigma_1, \sigma_2), \\
g_7(\beta, \mu, \sigma^2, \sigma_1, \sigma_2) &= \frac{\sqrt{2\pi\sigma^2}}{\sqrt{\sigma_1^2\sigma_2^2\sigma_b^2}} \phi\left(\frac{y_1}{\sigma_1}\right) \phi\left(\frac{y_2 - \beta_0 + \beta\psi}{\sigma_{e2}}\right) \phi\left(\frac{\mu_b}{\sigma_b}\right) \exp\left(\frac{\mu^2}{2\sigma^2}\right).
\end{aligned}
\tag{1.1}
$$

The functions g_4 to g_7 depend on other quantities as well in addition to those explicitly specified as the arguments. However, this dependence is suppressed for notational convenience.

Next, we review some distributions and their properties that will be used throughout this dissertation.

1.2.2 Truncated normal distribution

Let the random variable Y follow a $N(\mu, \sigma^2)$ distribution. When Y is restricted to lie within the interval (a_1, a_2), $-\infty < a_1 \leq a_2 < \infty$, it follows a truncated normal distribution with density (Forbes et al., 2011)

$$
f(y|a_1 < y \leq a_2) = \begin{cases} \frac{\frac{1}{\sigma}\phi\left(\frac{y-\mu}{\sigma}\right)}{\Phi\left(\frac{a_2-\mu}{\sigma}\right) - \Phi\left(\frac{a_1-\mu}{\sigma}\right)}, & \text{if } a_1 < y \leq a_2, \\ 0, & \text{otherwise.} \end{cases}
\tag{1.2}
$$

Using the notation (1.1), the first and second moments of this distribution are

$$
E(Y|a_1 < Y \leq a_2) = \mu - \sigma\left(\frac{\phi\left(g_1(a_2, \mu, \sigma)\right) - \phi\left(g_1(a_1, \mu, \sigma)\right)}{\Phi\left(g_1(a_2, \mu, \sigma)\right) - \Phi\left(g_1(a_1, \mu, \sigma)\right)}\right)
\tag{1.3}
$$

and

$$E(Y^2|a_1 < Y \le a_2) = \mu^2 + \sigma^2 \left(1 + \frac{\phi'\left(g_1(a_2,\mu,\sigma)\right) - \phi'\left(g_1(a_1,\mu,\sigma)\right)}{\Phi\left(g_1(a_2,\mu,\sigma)\right) - \Phi\left(g_1(a_1,\mu,\sigma)\right)}\right) - 2\mu\sigma \left(\frac{\phi\left(g_1(a_2,\mu,\sigma)\right) - \phi\left(g_1(a_1,\mu,\sigma)\right)}{\Phi\left(g_1(a_2,\mu,\sigma)\right) - \Phi\left(g_1(a_1,\mu,\sigma)\right)}\right), \tag{1.4}$$

respectively. Here $\phi'(x) = -x\phi(x)$. The variance of the distribution is

$$\mathrm{var}(Y|a_1 < Y \le a_2) = \sigma^2 \left\{1 + \left(\frac{-g_1(a_2,\mu,\sigma)\phi\left(g_1(a_2,\mu,\sigma)\right) + g_1(a_1,\mu,\sigma)\phi\left(g_1(a_1,\mu,\sigma)\right)}{\Phi\left(g_1(a_2,\mu,\sigma)\right) - \Phi\left(g_1(a_1,\mu,\sigma)\right)}\right) - \left(\frac{\phi\left(g_1(a_2,\mu,\sigma)\right) - \phi\left(g_1(a_1,\mu,\sigma)\right)}{\Phi\left(g_1(a_2,\mu,\sigma)\right) - \Phi\left(g_1(a_1,\mu,\sigma)\right)}\right)^2\right\}. \tag{1.5}$$

So far Y was truncated from both sides. But it may also be truncated only on one side, e.g., $Y > a_1$ or $Y \le a_2$. In these cases, we can get the analogs of (1.3), (1.4), and (1.5) by taking limit $a_2 \to \infty$ and $a_1 \to -\infty$. Doing so yields,

$$E(Y|Y > a_1) = \mu + \sigma g_2(a_1,\mu,\sigma),$$

$$E(Y^2|Y > a_1) = \mu^2 + \sigma^2 \left\{1 + g_1(a_1,\mu,\sigma)g_2(a_1,\mu,\sigma)\right\} + 2\mu\sigma g_2(a_1,\mu,\sigma),$$

$$\mathrm{var}(Y|Y > a_1) = \sigma^2 \left\{1 + g_1(a_1,\mu,\sigma)g_2(a_1,\mu,\sigma) - g_2^2(a_1,\mu,\sigma)\right\}, \tag{1.6}$$

and

$$E(Y|Y \le a_2) = \mu - \sigma g_3(a_2,\mu,\sigma),$$

$$E(Y^2|Y \le a_2) = \mu^2 + \sigma^2 \left\{1 - g_1(a_2,\mu,\sigma)g_3(a_2,\mu,\sigma)\right\} - 2\mu\sigma g_3(a_2,\mu,\sigma),$$

$$\mathrm{var}(Y|Y \le a_2) = \sigma^2 \left\{1 - g_1(a_2,\mu,\sigma)g_3(a_2,\mu,\sigma) - g_3^2(a_2,\mu,\sigma)\right\}. \tag{1.7}$$

Next, we use these one-sided moments to find the first and second moments of $YI(Y > a_1)$ and $YI(Y \le a_2)$, where $I(A)$ is the indicator function of event A. These will be needed in

Chapters 2 and 3. For the first moment, we have

$$\begin{aligned}
E(YI(Y > a_1)) &= \int_{a_1}^{\infty} yf(y)dy = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{a_1}^{\infty} ye^{-\frac{(y-\mu)^2}{2\sigma^2}} dy \\
&= \frac{1-\Phi\left(\frac{a_1-\mu}{\sigma}\right)}{\sqrt{2\pi\sigma^2}} \int_{a_1}^{\infty} y \frac{e^{-\frac{(y-\mu)^2}{2\sigma^2}}}{1-\Phi(\frac{a_1-\mu}{\sigma})} dy \\
&= \{1-\Phi\left(g_1(a_1,\mu,\sigma)\right)\} E(Y|Y > a_1) \\
&= \{1-\Phi\left(g_1(a_1,\mu,\sigma)\right)\} \mu + \sigma\phi\left(g_1(a_1,\mu,\sigma)\right) \qquad (1.8)
\end{aligned}$$

upon using (1.6) and

$$\begin{aligned}
E(YI(Y \leq a_2)) &= \int_{-\infty}^{a_2} yf(y)dy = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{a_2} ye^{-\frac{(y-\mu)^2}{2\sigma^2}} dy \\
&= \frac{\Phi\left(\frac{a_2-\mu}{\sigma}\right)}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{a_2} y \frac{e^{-\frac{(y-\mu)^2}{2\sigma^2}}}{\Phi(\frac{a_2-\mu}{\sigma})} dy \\
&= \Phi\left(g_1(a_2,\mu,\sigma)\right) E(Y|Y \leq a_2) \\
&= \Phi\left(g_1(a_2,\mu,\sigma)\right) \mu - \sigma\phi\left(g_1(a_2,\mu,\sigma)\right) \qquad (1.9)
\end{aligned}$$

upon using (1.7). Likewise, for the second moment, we have

$$\begin{aligned}
E(Y^2I(Y > a_1)) &= \{1-\Phi\left(g_1(a_1,\mu,\sigma)\right)\} E(Y^2|Y > a_1) \\
&= \{1-\Phi\left(g_1(a_1,\mu,\sigma)\right)\} \left[\mu^2 + \sigma^2\{1 + g_1(a_1,\mu,\sigma)g_2(a_1,\mu,\sigma)\}\right. \\
&\quad \left. + 2\mu\sigma g_2(a_1,\mu,\sigma)\right], \qquad (1.10)
\end{aligned}$$

and

$$\begin{aligned}
E(Y^2I(Y \leq a_2)) &= \Phi\left(g_1(a_2,\mu,\sigma)\right) E(Y^2|Y \leq a_2) \\
&= \Phi\left(g_1(a_2,\mu,\sigma)\right) \left[\mu^2 + \sigma^2\{1 - g_1(a_2,\mu,\sigma)g_3(a_2,\mu,\sigma)\}\right. \\
&\quad \left. - 2\mu\sigma g_3(a_2,\mu,\sigma)\right]. \qquad (1.11)
\end{aligned}$$

1.2.3 Half-normal distribution

Let X follow a $N(0, \sigma^2)$ distribution. Then, $Y = |X|$ follows a half-normal (HN) distribution. Its probability density function is

$$f(y) = \begin{cases} \frac{\sqrt{2}}{\sigma\sqrt{\pi}} \exp\left(-\frac{y^2}{2\sigma^2}\right), & \text{if } y > 0, \\ 0, & \text{otherwise.} \end{cases}$$

Its expected value and variance are

$$E(Y) = \sigma\sqrt{\frac{2}{\pi}}, \qquad \text{var}(Y) = \sigma^2\left(1 - \frac{2}{\pi}\right).$$

This distribution will be used as a prior distribution in Chapter 4. Let $X_1, X_2, ..., X_n$ denote a random sample from this distribution. The maximum likelihood (ML) estimator of parameter σ^2 based on these data is $\frac{1}{n}\sum_{i=1}^{n} X_i^2$.

1.2.4 Inverse gamma distribution

Let Y follow an inverse gamma (IG) distribution with shape parameter $\alpha\,(> 0)$ and scale parameter $\beta\,(> 0)$. The probability density function of Y is

$$f(y) = \begin{cases} \frac{\beta^\alpha}{\Gamma(\alpha)} y^{-\alpha-1} \exp\left(-\frac{\beta}{y}\right), & \text{if } y > 0, \\ 0, & \text{otherwise.} \end{cases}$$

The mean and variance of Y are

$$E(Y) = \frac{\beta}{\alpha - 1}, \ \alpha > 1, \ \ \text{var}(Y) = \frac{\beta^2}{(\alpha-1)^2(\alpha-2)}, \ \alpha > 2.$$

This distribution will be used as a prior distribution in Chapter 4.

1.2.5 Hierarchical half-normal distribution

Let $Y|\alpha^2$ follow a HN(α) distribution where $\alpha^2 \sim \text{IG}(a, b)$. Then, marginally Y is a said to follow a hierarchical half-normal (HHN) distribution with parameters a and b. We have:

$$f(y|\alpha^2) = \frac{\sqrt{2}}{\alpha\sqrt{\pi}} \exp\left(-\frac{y^2}{2\alpha^2}\right),$$

$$f(\alpha^2) = \frac{b^a}{\Gamma(a)}(\alpha^2)^{-a-1}\exp\left(\frac{-b}{\alpha^2}\right),$$

where y, α, a, and $b > 0$. It follows that the marginal density of Y is

$$\begin{aligned}
p(y) &= \int_0^\infty \frac{\sqrt{2}}{\alpha\sqrt{\pi}}\exp\left(-\frac{y^2}{2\alpha^2}\right)\frac{b^a}{\Gamma(a)}(\alpha^2)^{-a-1}\exp\left(\frac{-b}{\alpha^2}\right)d\alpha^2 \\
&= \sqrt{\frac{2}{\pi}}\frac{\Gamma(a+\frac{1}{2})}{\Gamma(a)}\frac{b^a}{\left(\frac{y^2+2b}{2}\right)^{a+\frac{1}{2}}}, \quad y > 0.
\end{aligned}$$

To find $E(Y)$, we can write

$$E(Y) = E\left[E(Y|\alpha^2)\right] = \sqrt{\frac{2}{\pi}}E(\alpha),$$

where

$$E(\alpha) = b\frac{\Gamma(a-\frac{1}{2})}{\Gamma(a)}, \; a > \frac{1}{2}.$$

We also have:

$$\text{var}(\alpha) = E(\alpha^2) - [E(\alpha)]^2 = \frac{b}{a-1} - b\left(\frac{\Gamma(a-\frac{1}{2})}{\Gamma(a)}\right)^2.$$

Next, to find $\text{var}(y)$, we can write

$$\begin{aligned}
\text{var}(y) &= E\left[\text{var}(y|\alpha^2)\right] + \text{var}\left[E(y|\alpha^2)\right] \\
&= \left(1 - \frac{2}{\pi}\right)\frac{b}{a-1} + \frac{2}{\pi}\text{var}(\alpha) \\
&= \frac{b}{a-1} - \frac{2b}{\pi}\left(\frac{\Gamma(a-\frac{1}{2})}{\Gamma(a)}\right)^2, \; a > 1.
\end{aligned}$$

This distribution will be used as a prior distribution in Chapter 4.

CHAPTER 2

A SEGMENTED MEASUREMENT ERROR MODEL FOR METHOD COMPARISON DATA

2.1 Introduction

Method comparison studies are generally concerned with evaluation of similarity and agreement of two methods of measuring a continuous variable to determine if they can be used interchangeably (Dunn, 2004; Carstensen, 2010; Lin et al., 2011; Choudhary and Nagaraja, 2017). The methods are assumed to have the same unit of measurement and none of them is considered error-free. Such studies are common in biomedical literature. For example, the article (Bland and Altman, 1986) that proposed the *limits of agreement* approach for analysis of method comparison data has over 25,000 citations.

The motivation for this chapter comes from a study comparing two assays, labelled 1 and 2, for measuring concentration of digoxin (Hawkins, 2002). The data consist of natural logarithm of concentrations of digoxin measured by the assays on $n = 134$ specimens. Figure 2.1 presents a scatterplot of measurements of assay 2 (Y_2) against those of assay 1 (Y_1). A plot of difference ($D = Y_2 - Y_1$) against the average of these measurements — popularly known as the *Bland-Altman plot* (Bland and Altman, 1986) — is also presented in Figure 2.2. A similar plot is presented as Figure 6 in (Hawkins, 2002). The scatterplot shows that the underlying trend is piecewise linear. The initial linear trend appears to undergo a change in slope around $y_1 = -0.5$. It is ao clear that the assays behave quite differently in the left and right segments formed by the *changepoint*. In particular, they seem to have higher agreement in the right segment than the left segment. They are also more highly correlated in the right segment than the left. The change in behavior can be seen more clearly in the Bland-Altman plot where a downward linear trend is followed by a flattening of the trend, with points on the right centered near zero. In (Hawkins, 2002), it is concluded from this

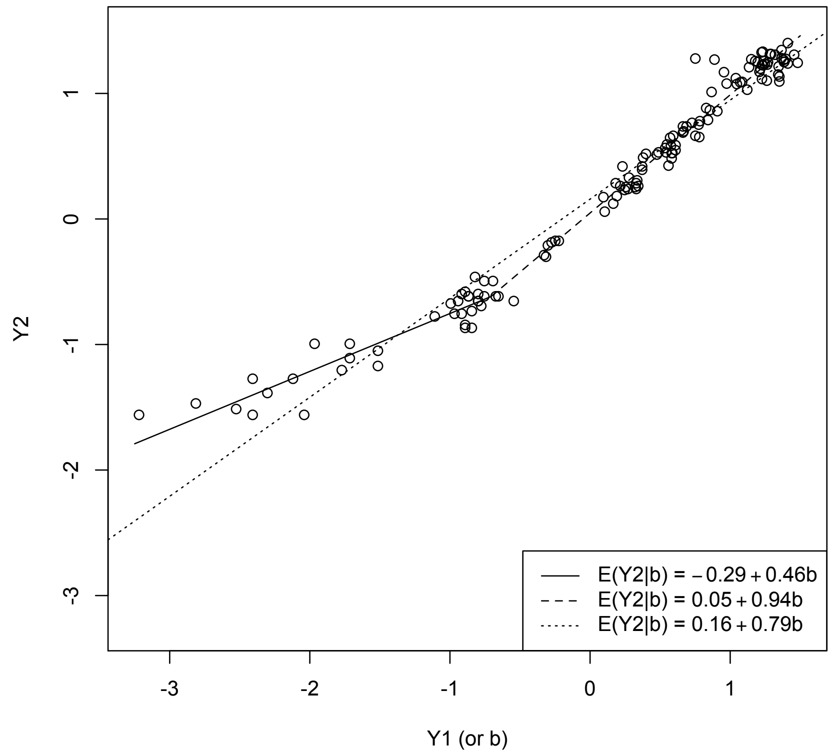

Figure 2.1. Scatterplot for log-scale digoxin data superimposed with the estimated straight line for $E(Y_2|b)$ under the classical model (dotted line) given by (2.2) and the piecewise straight line (solid and broken lines) under the segmented model given by (2.6).

plot that "the two assays are not comparable at low analyte levels" but "may be equivalent above some cut-off level."

The digoxin data are an example of paired method comparison data. Such data are often analyzed by modeling them using a classical *measurement error model* — also known as *errors-in-variables* model — or its variants, see, e.g., (Linnet, 1999; Dunn and Roberts, 1999; Hawkins, 2002; Dunn, 2004, 2007; Alanen, 2010). The literature on measurement error models, especially in the context of regression analysis, is vast and the books (Fuller, 1987; Cheng and Van Ness, 1999; Carroll et al., 2006) may be consulted for an introduction to this topic. The readers specifically interested in measurement error models for method compari-

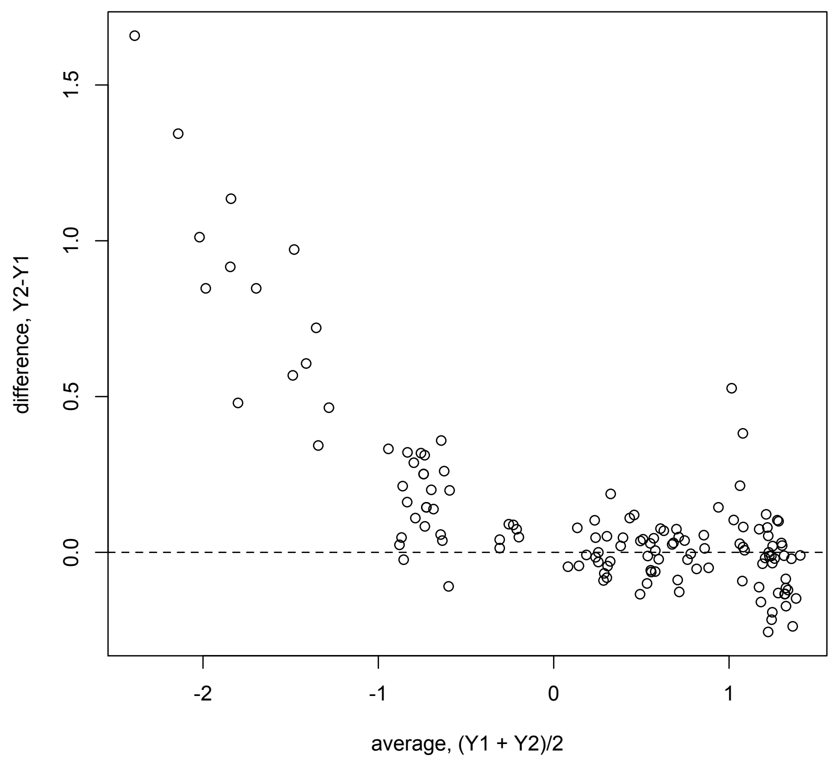

Figure 2.2. Bland-Altman plot for log-scale digoxin data superimposed with a horizontal line at zero.

son studies may begin with (Dunn, 2004). The classical measurement error model assumes a linear trend over the entire measurement range. However, it is clear from Figure 2.1 that this model must be extended to allow incorporating the change of slope that is observed in the digoxin data. What is necessary is a *segmented* measurement error model that allows a piecewise linear trend over two segments of the measurement range and treats the change-point at which the two trend lines join as an unknown parameter in the model. Studying this extension is the goal of this chapter.

Segmented models, also known as *multiphase* or *piecewise* models, for linear regression where the regression function either has different forms or involves different parameters over different segments of the domain of the covariate have been studied at least since

(Quandt, 1958), with early contributions including (Sprent, 1961; Hudson, 1966; Hinkley, 1969, 1971). These articles studied inference for piecewise simple linear regression model over two segments. The changepoint, also known as *join point*, *break point*, or *knot*, represents a value of the covariate at which the regression functions join and it may be known or unknown. The overall regression function may be continuous or discontinuous at the changepoint. A survey of the early literature on segmented regression models can be found in (Seber and Wild, 1989, Chapter 9). Segmented models in the context of logistic and other regressions have been studied by (Pastor-Barriuso et al., 2003; Muggeo, 2003; Vexler and Gurevich, 2009; Goodman et al., 2011; Fong et al., 2015). See (Piepho and Ogutu, 2003; Muggeo et al., 2014; Lai and Albert, 2014; Tang et al., 2017) for segmented models for longitudinal data.

When the covariate in regression is measured with error, giving rise to segmented measurement error models, the changepoint represents a value of the error-free covariate. An early article on this topic is (Küchenhoff and Carroll, 1997) which considered a *threshold model* — a special case of a segmented model where it is assumed that the exposure has no relation with the response up to a threshold — and focussed on estimation of the threshold (i.e., the changepoint) in the context of linear and logistic regression. It found that ignoring the measurement error leads to asymptotically biased estimator of the threshold. Moreover, the standard methods for correcting for such bias in measurement error models, namely, regression calibration, simulation extrapolation, and maximum likelihood (ML), behave quite differently when the model is segmented. In particular, if the assumed model is correct or at least plausible, the ML estimator outperforms the other two estimators in terms of bias and variance. Other articles on segmented measurement error models include (Carroll et al., 1999; Gössl and Küchenhoff, 2001; Staudenmayer and Spiegelman, 2002). They respectively consider probit regression with mixture of normals as the error distribution, logistic regression from a Bayesian perspective, and making use of external and internal validation data in addition to the main data. A primary focus of these articles is bias in the estimated changepoint.

To our knowledge, none of the existing segmented measurement error models allow the kind of piecewise linear trend over two segments that is needed for digoxin data. This model is also simple enough to allow closed-form expressions for the likelihood function. In addition to studying this model, another novel contribution of this work is that we apply the proposed model to analyze method comparison data. As indicated previously, the eventual goals in analysis of these data are not to perform inference regarding regression coefficients or the changepoint, which is typically the case in regression, but to evaluate similarity and agreement of the measurement methods (Choudhary and Nagaraja, 2017, Chapters 1 and 2). Similarity is evaluated by comparing biases and precisions of the methods and agreement is evaluated by performing inference on agreement measures such as concordance correlation coefficient (Lin, 1989) and total deviation index (Lin, 2000; Choudhary and Nagaraja, 2007; Escaramís et al., 2010; Perez-Jaume and Carrasco, 2015) (see Section 2.4).

The rest of this chapter is organized as follows. In Section 2.2, we present the classical measurement model and its proposed extension for method comparison data. Section 2.3 describes an expectation-maximization (EM) type algorithm (Meng and Rubin, 1993; McLachlan and Krishnan, 2007) for fitting the model using the ML method and presents a test for changepoint. Section 2.4 discusses evaluation of similarity and agreement under the new model. The results of a simulation study to evaluate the proposed methodology are presented in Section 2.5. We revisit the digoxin data and analyze them in Section 2.6.

2.2 Modeling of method comparison data

Let (Y_{i1}, Y_{i2}), $i = 1, \ldots, n$ denote paired measurements data collected in a method comparison study. These data are assumed to be a random sample from the distribution of (Y_1, Y_2), where Y_j represents the measurement by the jth method, $j = 1, 2$, on a randomly selected subject from the underlying population. Here method 1 is assumed to be the standard method that serves as a reference and method 2 is the test method in the comparison.

2.2.1 Classical measurement error model

The classical measurement error model for the paired measurements (Y_1, Y_2) is (Hawkins, 2002; Dunn, 2004)

$$Y_1 = b + e_1,\ Y_2 = \beta_0 + \beta_1 b + e_2, \tag{2.1}$$

where β_0 and β_1 are fixed regression coefficients, b is a random quantity representing the underlying true unobservable measurement, and e_1 and e_2 are random errors associated with the two measurement methods. The true value b is measured with error by method 1 as Y_1 and by method 2 as Y_2. The conditional means of the two methods — $E(Y_1|b) = b$ and $E(Y_2|b) = \beta_0 + \beta_1 b$ — represent their error-free values. These are linearly related. It is assumed that $b \sim N(\mu_b, {\sigma_b}^2)$, $e_1 \sim N(0, \sigma_{e1}^2)$, and $e_2 \sim N(0, \sigma_{e2}^2)$; and b, e_1, and e_2 are mutually independent. It follows that (Y_1, Y_2) have a bivariate normal distribution with parameters

$$E(Y_1) = \mu_b, \mathrm{var}(Y_1) = \sigma_b^2 + \sigma_{e1}^2, E(Y_2) = \beta_0 + \beta_1\mu_b, \mathrm{var}(Y_2) = \beta_1^2\sigma_b^2 + \sigma_{e2}^2, \mathrm{cov}(Y_1, Y_2) = \beta_1\sigma_b^2.$$

The intercept β_0 and the slope β_1 are respectively known as the *fixed bias* and *proportional bias* of method 2 (Choudhary and Nagaraja, 2017, Chapter 1). If $\beta_1 = 1$, the methods are said to have the same *scale*. If the methods have unequal scales, their precisions are measured by their *squared sensitivity* — $1/\sigma_{e1}^2$ for method 1 and β_1^2/σ_{e2}^2 for method 2 (Mandel and Stiehler, 1954; Mandel, 1978; Shyr and Gleser, 1986). They are compared using the *squared sensitivity ratio*, $\beta_1^2(\sigma_{e1}^2/\sigma_{e2}^2)$ (Choudhary and Nagaraja, 2017, Chapter 1). If $(\beta_0, \beta_1) = (0, 1)$, the two methods have the same fixed and proportional biases and hence same mean. If $(\beta_1, \sigma_{e2}^2/\sigma_{e1}^2) = (1, 1)$, they have the same precision. The comparison of precisions is unaffected by the intercept β_0.

It follows from (2.1) that the model for the observed data (Y_{i1}, Y_{i2}), $i = 1, \ldots, n$ is

$$Y_{i1} = b_i + e_{i1},\ Y_{i2} = \beta_0 + \beta_1 b_i + e_{i2}, \tag{2.2}$$

where b_i, e_{i1}, and e_{i2} are mutually independent draws from the respective distributions of b, e_1, and e_2. This model is not identifiable on the basis of paired measurements data and one constraint must be imposed on its parameters to make it identifiable. Although a number of possibilities exist, see, e.g., (Cheng and Van Ness, 1999, Chapter 1), three are common in method comparison studies. One is $\beta_1 = 1$, in which case the model becomes a mixed-effects model and is often known as the *Grubbs' model* after (Grubbs, 1948). Another is $\sigma_{e1}^2 = \sigma_{e2}^2$, see, e.g., (Hawkins, 2002). The third is that the ratio $\sigma_{e1}^2/\sigma_{e2}^2$ is known, in which case *Deming regression* (Deming, 1943; Kummell, 1879; Finney, 1996) is a widely known procedure to fit the model, see, e.g., (Dunn, 2007). See also (Dunn, 2004, Chapter 3) and (Hawkins, 2002) for a discussion of relative merits and demerits of the three approaches. Here we work under the equal error variance assumption and denote the common value by σ_e^2.

2.2.2 Segmented measurement error model

The proposed extension of the classical measurement error model (2.1) for (Y_1, Y_2) is a segmented model,

$$Y_1 = b + e_1, \quad Y_2 = \beta_0 + \beta_1 b + \beta_2 (b - \psi)_+ + e_2, \tag{2.3}$$

where $(b - \psi)_+ = \max\{0, b - \psi\}$. We can write $(b - \psi)_+ = (b - \psi) I(b > \psi)$ with $I(A)$ denoting the indicator function of A. Here ψ is the changepoint. It follows from (2.3) that the conditional mean $E(Y_2|b)$ of method 2 follows the *broken-stick model*,

$$E(Y_2|b) = \begin{cases} \beta_0 + \beta_1 b, & b \leq \psi, \\ (\beta_0 - \beta_2 \psi) + (\beta_1 + \beta_2) b, & b > \psi. \end{cases} \tag{2.4}$$

Thus, $E(Y_2|b)$ undergoes a change in slope from β_2 to $\beta_1 + \beta_2$ at the changepoint $b = \psi$ and is continuous in b at the changepoint. The latter necessitates a change in intercept also — from β_0 to $\beta_0 - \beta_2\psi$ at the changepoint.

Just like the classical model (2.1), for the distributions of the random terms in (2.3), it is assumed that

$$b \sim N(\mu_b, {\sigma_b}^2),\ e_1 \sim N(0, \sigma_e^2),\ e_2 \sim N(0, \sigma_e^2), \tag{2.5}$$

and b, e_1, and e_2 are mutually independent. Thus, the segmented model for the observed data is

$$Y_{i1} = b_i + e_{i1},\ Y_{i2} = \beta_0 + \beta_1 b_i + \beta_2 (b_i - \psi)_+ + e_{i2},\ i = 1, \ldots, n, \tag{2.6}$$

where b_i, e_{i1}, and e_{i2} are mutually independent draws from the respective distributions of b, e_1, and e_2 given in (2.5). The classical model (2.2) becomes a special case of the segmented model (2.6) when $\beta_2 = 0$ or in the limit as $\psi \to \infty$.

2.2.3 Distribution theory under the segmented model

In this section, we present some distributional results under the segmented model (2.3) that will be useful in later sections. Their proofs are provided in Chapter 3 under the more general setting of unequal error variances.

Proposition 2.1. *Consider* (Y_1, Y_2) *following the model* (2.3). *The mean and variance of* Y_1 *and* Y_2 *and their covariance are as follows:*

(a) $E(Y_1) = \mu_b$ *and* $var(Y_1) = \sigma_b^2 + \sigma_e^2$,

(b) $E(Y_2) = \beta_0 + \beta_1 \mu_b + \beta_2 m_1$ *and* $var(Y_2) = \beta_1^2 \sigma_b^2 + \beta_2^2 m_3 + 2\beta_1 \beta_2 m_4 + \sigma_e^2$,

(c) $cov(Y_1, Y_2) = \beta_1 \sigma_b^2 + \beta_2 m_4$,

where

$$
\begin{aligned}
m_1 &= E[(b-\psi)_+] = \{1-\Phi\left(g_1(\psi,\mu_b,\sigma_b)\right)\}(\mu_b-\psi) + \phi\left(g_1(\psi,\mu_b,\sigma_b)\right)\sigma_b,\\
m_2 &= E\left[\{(b-\psi)_+\}^2\right] = \{1-\Phi\left(g_1(\psi,\mu_b,\sigma_b)\right)\}\left\{(\mu_b-\psi)^2+\sigma_b^2\right\}\\
&\qquad + \left(g_1(\psi,\mu_b,\sigma_b)\sigma_b^2 + 2\mu_b\sigma_b - 2\psi\sigma_b\right)\phi\left(g_1(\psi,\mu_b,\sigma_b)\right),\\
m_3 &= var\left[(b-\psi)_+\right] = m_2 - m_1^2,\\
m_4 &= cov\left[b,(b-\psi)_+\right] = \{1-\Phi\left(g_1(\psi,\mu_b,\sigma_b)\right)\}\sigma_b^2\\
&\qquad + \left\{g_1(\psi,\mu_b,\sigma_b)\sigma_b^2 + \mu_b\sigma_b - \psi\sigma_b\right\}\phi\left(g_1(\psi,\mu_b,\sigma_b)\right),
\end{aligned}
\tag{2.7}
$$

and the functions g_1, g_2, and g_3 are given by (1.1).

Proposition 2.2. *Consider (Y_1, Y_2) following the model* (2.3)*. The mean and variance of Y_1 and Y_2 and their covariance when b is truncated to be either $b \le \psi$ or $b > \psi$ are as follows:*

(a) $E(Y_1|b \le \psi) = E(b|b \le \psi)$ *and* $var(Y_1|b \le \psi) = var(b|b \le \psi) + \sigma_e^2$,

(b) $E(Y_2|b \le \psi) = \beta_0 + \beta_1 E(b|b \le \psi)$ *and* $var(Y_2|b \le \psi) = \beta_1^2 var(b|b \le \psi) + \sigma_e^2$,

(c) $cov(Y_1, Y_2|b \le \psi) = \beta_1 var(b|b \le \psi)$,

(d) $E(Y_1|b > \psi) = E(b|b > \psi)$ *and* $var(Y_1|b > \psi) = var(b|b > \psi) + \sigma_e^2$,

(e) $E(Y_2|b > \psi) = (\beta_0 - \beta_2\psi) + (\beta_1 + \beta_2)E(b|b > \psi)$ *and* $var(Y_2|b > \psi) = (\beta_1 + \beta_2)^2 var(b|b > \psi) + \sigma_e^2$,

(f) $cov(Y_1, Y_2|b > \psi) = (\beta_1 + \beta_2) var(b|b > \psi)$,

where

$$
\begin{aligned}
E(b|b \le \psi) &= \mu_b - \sigma_b g_3(\psi,\mu_b,\sigma_b),\\
var(b|b \le \psi) &= \sigma_b^2\left\{1 - g_1(\psi,\mu_b,\sigma_b)g_3(\psi,\mu_b,\sigma_b) - g_3^2(\psi,\mu_b,\sigma_b)\right\}.\\
E(b|b > \psi) &= \mu_b + \sigma_b g_2(\psi,\mu_b,\sigma_b),\\
var(b|b > \psi) &= \sigma_b^2\left\{1 + g_1(\psi,\mu_b,\sigma_b)g_2(\psi,\mu_b,\sigma_b) - g_2^2(a_1,\mu_b,\sigma_b)\right\},
\end{aligned}
\tag{2.8}
$$

and the functions g_1, g_2, and g_3 are given by (1.1).

Proposition 2.3. *The joint probability density function of* (Y_1, Y_2) *following the model* (2.3) *is*

$$f(y_1, y_2) = f_1(y_1, y_2) + f_2(y_1, y_2), \tag{2.9}$$

where

$$\begin{aligned} f_1(y_1, y_2) &= \int_{-\infty}^{\psi} f(b, y_1, y_2) db = g_7(0, g_6(0, \sigma_e, \sigma_e), g_4(0, \sigma_e, \sigma_e), \sigma_e, \sigma_e) \\ &\quad \times \Phi\left(g_1(\psi, g_6(0, \sigma_e, \sigma_e), \sqrt{g_4(0, \sigma_e, \sigma_e)})\right) \\ f_2(y_1, y_2) &= \int_{\psi}^{\infty} f(b, y_1, y_2) db = g_7(\beta_2, g_6(\beta_2, \sigma_e, \sigma_e), g_4(\beta_2, \sigma_e, \sigma_e), \sigma_e, \sigma_e) \\ &\quad \times \left[1 - \Phi\left(g_1\left(\psi, g_6(\beta_2, \sigma_e, \sigma_e), \sqrt{g_4(\beta_2, \sigma_e, \sigma_e)}\right)\right)\right], \end{aligned} \tag{2.10}$$

and the functions g_1 through g_7 are given by (1.1).

Proposition 2.4. *Consider* (Y_1, Y_2) *following the model* (2.3)*. The probability density function of $D = Y_1 - Y_2$ is*

$$h(d) = h_1(d) + h_2(d), \tag{2.11}$$

where

$$\begin{aligned} h_1(d) &= \int_{-\infty}^{\psi} f(d, b) db = \frac{\sqrt{2\pi c_1}}{\sqrt{\sigma^2 \sigma_b^2}} \phi\left(\frac{d - \beta_0}{\sigma}\right) \phi\left(\frac{\mu_b}{\sigma_b}\right) \exp\left(\frac{c_2^2}{2c_1}\right) \Phi(g_1(\psi, c_2, \sqrt{c_1}), \\ h_2(d) &= \int_{\psi}^{\infty} f(d, b) db = \frac{\sqrt{2\pi c_3}}{\sqrt{\sigma^2 \sigma_b^2}} \phi\left(\frac{d - \beta_0 + \beta_2 \psi}{\sigma}\right) \phi\left(\frac{\mu_b}{\sigma_b}\right) \exp\left(\frac{c_4^2}{2c_3}\right) \\ &\quad \times \left[1 - \Phi(g_1(\psi, c_4, \sqrt{c_3}))\right], \end{aligned} \tag{2.12}$$

and

$$\sigma^2 = 2\sigma_e^2,$$
$$c_1 = \left(\frac{(\beta_1-1)^2}{\sigma^2} + \frac{1}{\sigma_b^2}\right)^{-1},$$
$$c_2 = \left(\frac{(d-\beta_0)(\beta_1-1)}{\sigma^2} + \frac{\mu_b}{\sigma_b^2}\right) c_1$$
$$c_3 = \left(\frac{(\beta_1+\beta_2-1)^2}{\sigma^2} + \frac{1}{\sigma_b^2}\right)^{-1}$$
$$c_4 = \left(\frac{(d-\beta_0+\beta_2\psi)(\beta_1+\beta_2-1)}{\sigma^2} + \frac{\mu_b}{\sigma_b^2}\right) c_3. \tag{2.13}$$

Proposition 2.5. *Consider* (Y_1, Y_2) *following the model* (2.3)*. The probability density function of* $D = Y_1 - Y_2$ *when* b *is truncated to be either* $b \leq \psi$ *or* $b > \psi$ *are as follows:*

$$h(d|b \leq \psi) = \frac{h_1(d)}{\Phi(g_1(\psi, \mu_b, \sigma_b))}, \; h(d|b > \psi) = \frac{h_2(d)}{1 - \Phi(g_1(\psi, \mu_b, \sigma_b))}, \tag{2.14}$$

where $h_1(d)$ *and* $h_2(d)$ *are given by* (2.12) *in Proposition 2.4.*

Proposition 2.6. *Consider* (Y_1, Y_2) *the model* (2.3)*. The first and second moments of* b *and* $(b-\psi)_+$ *conditional on* $(Y_1, Y_2) = (y_1, y_2)$ *are as follows:*

(a) $E[b|y_1, y_2] = (A_1 + A_2)/f(y_1, y_2)$,

(b) $E[b^2|y_1, y_2] = (A_3 + A_4)/f(y_1, y_2)$,

(c) $E\left[(b-\psi)_+|y_1, y_2\right] = (A_2 - \psi f_2(y_1, y_2))/f(y_1, y_2)$,

(d) $E\left[\{(b-\psi)_+\}^2|y_1, y_2\right] = (A_4 - 2\psi A_2 + \psi^2 f_2(y_1, y_2))/f(y_1, y_2)$,

(e) $E\left[b(b-\psi)_+|y_1, y_2\right] = (A_4 - \psi A_2)/f(y_1, y_2)$,

where

$$A_1 = \int_{-\infty}^{\psi} bf(b, y_1, y_2)db = f_1(y_1, y_2)\left\{g_6(0, \sigma_e, \sigma_e) - \sqrt{g_4(0, \sigma_e, \sigma_e)} \times g_3\left(\psi, g_6(0, \sigma_e, \sigma_e), \sqrt{g_4(0, \sigma_e, \sigma_e)}\right)\right\}$$

$$A_2 = \int_{\psi}^{\infty} bf(b, y_1, y_2)db = f_2(y_1, y_2)\left\{g_6(\beta_2, \sigma_e, \sigma_e) + \sqrt{g_4(\beta_2, \sigma_e, \sigma_e)} \times g_2\left(\psi, g_6(\beta_2, \sigma_e, \sigma_e), \sqrt{g_4(\beta_2, \sigma_e, \sigma_e)}\right)\right\}$$

$$A_3 = \int_{-\infty}^{\psi} b^2 f(b, y_1, y_2)db = f_1(y_1, y_2)\left\{g_4(0, \sigma_e, \sigma_e) + g_6^2(0, \sigma_e, \sigma_e) - g_1(\psi, g_6(0, \sigma_e, \sigma_e), \sqrt{g_4(0, \sigma_e, \sigma_e)}) \times g_4(0, \sigma_e, \sigma_e)g_3\left(\psi, g_6(0, \sigma_e, \sigma_e), \sqrt{g_4(0, \sigma_e, \sigma_e)}\right) - 2\sqrt{g_4(0, \sigma_e, \sigma_e)}g_6(0, \sigma_e, \sigma_e)g_3\left(\psi, g_6(0, \sigma_e, \sigma_e), \sqrt{g_4(0, \sigma_e, \sigma_e)}\right)\right\}$$

$$A_4 = \int_{\psi}^{\infty} b^2 f(b, y_1, y_2)db = f_2(y_1, y_2)\left\{g_4(\beta_2, \sigma_e, \sigma_e) + g_6^2(\beta_2, \sigma_e, \sigma_e) + g_1(\psi, g_6(\beta_2, \sigma_e, \sigma_e), \sqrt{g_4(\beta_2, \sigma_e, \sigma_e)}) \times g_4(\beta_2, \sigma_e, \sigma_e)g_2\left(\psi, g_6(\beta_2, \sigma_e, \sigma_e), \sqrt{g_4(\beta_2, \sigma_e, \sigma_e)}\right) + 2\sqrt{g_4(\beta_2, \sigma_e, \sigma_e)}g_6(\beta_2, \sigma_e, \sigma_e)g_2\left(\psi, g_6(\beta_2, \sigma_e, \sigma_e), \sqrt{g_4(\beta_2, \sigma_e, \sigma_e)}\right)\right\}, \tag{2.15}$$

and f_1*,* f_2*, and* f *are given by 2.9 and* (2.10) *in Proposition 2.3.*

Proposition 2.7. *Consider the model* (2.3)*. The best linear predictor of* b *using* (Y_1, Y_2) *is*

$$\hat{b} = \mu_b + [\sigma_b^2, \beta_1\sigma_b^2 + \beta_2 m_4]\left(var\left(\left[\begin{smallmatrix} Y_1 \\ Y_2 \end{smallmatrix}\right]\right)\right)^{-1}\left[\begin{smallmatrix} Y_1 - \mu_b \\ Y_2 - E(Y_2) \end{smallmatrix}\right], \tag{2.16}$$

where m_4 *is given by* (2.7) *and the moments involved are given by Proposition 2.1.*

2.3 Fitting the segmented model

2.3.1 Parameter estimation

The segmented model (2.6) has 7 unknown parameters, $(\mu_b, \beta_0, \beta_1, \beta_2, \sigma_b^2, \sigma_e^2, \psi)$. Let $\boldsymbol{\theta}$ be the 7×1 vector of these parameters. Its log-likelihood function can be written as

$$L(\boldsymbol{\theta}) = \sum_{i=1}^{n} \log\{f(y_{i1}, y_{i2}|\boldsymbol{\theta})\}, \tag{2.17}$$

where the density f is given by (2.9) in Proposition 2.3 and now we have explicitly included $\boldsymbol{\theta}$ in its notation. We may numerically maximize $L(\boldsymbol{\theta})$ to get ML estimate $\hat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$. However, in our experience, a direct numerical maximization of this function is sensitive to starting points and often leads to unstable estimates. Therefore, we consider an alternative approach to get the ML estimate. It is a variant of the EM algorithm, specifically the ECM algorithm (Meng and Rubin, 1993), in which the M-step of EM is replaced by a sequence of computationally simpler constrained maximization (CM) steps. Each iteration of ECM increases the likelihood function and the algorithm often converges to a local or global maxima (McLachlan and Krishnan, 2007, Chapter 5).

To develop the ECM algorithm, we take b_i as the *missing data* and (b_i, Y_{i1}, Y_{i2}) as the *complete data* for the ith subject. The logarithm of the joint density $f(b, y_1, y_2|\boldsymbol{\theta})$ can be written as

$$\begin{aligned}
\log\{f(b, y_1, y_2|\boldsymbol{\theta})\} &= \log f\{(y_1|b, \boldsymbol{\theta})\} + \log\{f(y_2|b, \boldsymbol{\theta})\} + \log\{f(b|\boldsymbol{\theta})\} \\
&= -\frac{3}{2}\log(2\pi) - \frac{1}{2}\log(\sigma_e^2) - \frac{1}{2}\log(\sigma_e^2) - \frac{1}{2}\log(\sigma_b^2) - \frac{1}{2\sigma_e^2}(y_1 - b)^2 \\
&\quad - \frac{1}{2\sigma_e^2}\left\{y_2 - \beta_0 - \beta_1 b - \beta_2 (b - \psi)_+\right\}^2 - \frac{1}{2\sigma_b^2}(b - \mu_b)^2 \\
&= -\frac{3}{2}\log(2\pi) - \frac{1}{2}\log(\sigma_e^2) - \frac{1}{2}\log(\sigma_e^2) - \frac{1}{2}\log(\sigma_b^2) - \frac{1}{2\sigma_e^2}\left(y_1^2 - 2y_1 b + b^2\right) \\
&\quad - \frac{1}{2\sigma_e^2}\left\{(y_2 - \beta_0)^2 - 2(y_2 - \beta_0)(\beta_1 b + \beta_2 (b - \psi)_+)\right.
\end{aligned}$$

$$
\begin{aligned}
&+ \left(\beta_1 b + \beta_2 (b-\psi)_+\right)^2 \Big\} \\
&- \frac{1}{2\sigma_b^2}\left(b^2 - 2\mu_b b + \mu_b^2\right) \\
= c &- \frac{1}{2\sigma_e^2}\left(-2y_1 b + b^2\right) - \frac{1}{2\sigma_e^2}\left\{-2(y_2-\beta_0)(\beta_1 b + \beta_2 (b-\psi)_+)\right. \\
&\left. + \left(\beta_1^2 b^2 + 2\beta_1\beta_2 b (b-\psi)_+\right) + \beta_2^2 (b-\psi)_+^2\right\} - \frac{1}{2\sigma_b^2}\left(b^2 - 2\mu_b b\right),
\end{aligned}
\tag{2.18}
$$

where c consists of terms that do not involve b and is given as

$$
c = -\frac{3}{2}\log(2\pi) - \frac{1}{2}\log(\sigma_e^2) - \frac{1}{2}\log(\sigma_e^2) - \frac{1}{2}\log(\sigma_b^2) - \frac{1}{2\sigma_e^2}y_1^2 - \frac{1}{2\sigma_e^2}(y_2-\beta_0)^2 - \frac{1}{2\sigma_b^2}\mu_b^2. \tag{2.19}
$$

It follows that the complete data log-likelihood function is

$$
\begin{aligned}
&\sum_{i=1}^{n} \log\{f(b_i, y_{i1}, y_{i2}|\boldsymbol{\theta})\} \\
&= \sum_{i=1}^{n}\left\{c_i - \frac{1}{2\sigma_e^2}\left(-2y_{i1}b_i + b_i^2\right) - \frac{1}{2\sigma_e^2}\left\{-2(y_{i2}-\beta_0)(\beta_1 b_i + \beta_2 (b_i-\psi)_+)\right.\right. \\
&\quad \left.\left. + \left(\beta_1^2 b_i^2 + 2\beta_1\beta_2 b_i (b_i-\psi)_+\right) + \beta_2^2 (b_i-\psi)_+^2\right\} - \frac{1}{2\sigma_b^2}\left(b_i^2 - 2\mu_b b_i\right)\right\},
\end{aligned}
\tag{2.20}
$$

where c_i is the value of c given by (2.19) evaluated for the ith subject.

In the rth ECM iteration, let $\boldsymbol{\theta}^{(r)}$ be the value of $\boldsymbol{\theta}$ and $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(r)})$ be the expectation of the complete data log-likelihood (2.20) with respect to the conditional distribution of $b_i|y_{i1}, y_{i2}$ evaluated at $\boldsymbol{\theta}^{(r)}$, i.e., $E\left[\sum_{i=1}^{n}\log\{f(b_i, y_{i1}, y_{i2}|\boldsymbol{\theta})\}|y_{i1}, y_{i2}, \boldsymbol{\theta}^{(r)}\right]$. Letting $E^{(r)}$ denote the expectation over the conditional distribution of $b_i|y_{i1}, y_{i2}$ evaluated at $\boldsymbol{\theta}^{(r)}$, we can write

$$
\begin{aligned}
Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(r)}) = \sum_{i=1}^{n}\Big\{ & c_i - \frac{1}{2\sigma_e^2}\left(-2y_{i1}E^{(r)}[b_i] + E^{(r)}[b_i^2]\right) - \frac{1}{2\sigma_e^2}\{-2(y_{i2}-\beta_0) \\
& \left(\beta_1 E^{(r)}[b_i] + \beta_2 E^{(r)}\left[(b_i-\psi)_+\right]\right) + \beta_1^2 E^{(r)}[b_i^2] \\
& + 2\beta_1\beta_2 E^{(r)}\left[b_i (b_i-\psi)_+\right] + \beta_2^2 E^{(r)}\left[(b_i-\psi)_+^2\right]\} \\
& - \frac{1}{2\sigma_b^2}\left(E^{(r)}[b_i^2] - 2\mu_b E^{(r)}[b_i]\right)\Big\}.
\end{aligned}
\tag{2.21}
$$

Next, we find derivative of $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(r)})$ with respect to $\boldsymbol{\theta}$ so that we can perform the CM steps. For this, let $E_{i1}^{(r)}$ and $E_{i2}^{(r)}$ respectively denote the values of $E^{(r)}[b_i]$ and $E^{(r)}[b_i^2]$ and $A_{i1}^{(r)}, \ldots, A_{i4}^{(r)}$ respectively denote the values of $A_1, \ldots, A_4$ given by (2.15) in Proposition 2.6, evaluated at $(\boldsymbol{\theta}, y_1, y_2) = (\boldsymbol{\theta}^{(r)}, y_{i1}, y_{i2})$. Further, let $f^{(r)}(y_{i1}, y_{i2})$ denote the value of $f(y_1, y_2)$ given by (2.9) in Proposition 2.3, also evaluated at $(\boldsymbol{\theta}, y_1, y_2) = (\boldsymbol{\theta}^{(r)}, y_{i1}, y_{i2})$. The derivatives of $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(r)})$ with respect to the elements of $\boldsymbol{\theta}$ are as follows:

$$
\begin{aligned}
\frac{\partial Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(r)})}{\partial \mu_b} &= -\frac{1}{2\sigma_b^2}\sum_{i=1}^{n} 2\left\{-E_{i1}^{(r)} + \mu_b\right\} \\
\frac{\partial Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(r)})}{\partial \beta_0} &= -\frac{1}{2\sigma_e^2}\sum_{i=1}^{n} 2\left\{\beta_1 E_{i1}^{(r)} + \frac{\beta_2\left(A_{i2}^{(r)} - \psi f_2^{(r)}(y_{i1}, y_{i2})\right)}{f^{(r)}(y_{i1}, y_{i2})} - (y_{i2} - \beta_0)\right\} \\
\frac{\partial Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(r)})}{\partial \beta_1} &= -\frac{1}{2\sigma_e^2}\sum_{i=1}^{n} 2\left\{-(y_{i2} - \beta_0)E_{i1}^{(r)} + \beta_1 E_{i2}^{(r)} + \frac{\beta_2(A_{i4}^{(r)} - \psi A_{i2}^{(r)})}{f^{(r)}(y_{i1}, y_{i2})}\right\} \\
\frac{\partial Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(r)})}{\partial \beta_2} &= -\frac{1}{2\sigma_e^2}\sum_{i=1}^{n} 2\left\{-\frac{(y_{i2} - \beta_0)\left(A_{i2}^{(r)} - \psi f_2^{(r)}(y_{i1}, y_{i2})\right) + \beta_1(A_{i4}^{(r)} - \psi A_{i2}^{(r)})}{f^{(r)}(y_{i1}, y_{i2})}\right. \\
&\qquad \left. + \frac{\beta_2(A_{i4}^{(r)} - 2\psi A_{i2}^{(r)} + \psi^2 f_2^{(r)}(y_{i1}, y_{i2}))}{f^{(r)}(y_{i1}, y_{i2})}\right\} \\
\frac{\partial Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(r)})}{\partial \sigma_b^2} &= -\frac{n}{2\sigma_b^2} + \frac{1}{2\sigma_b^4}\sum_{i=1}^{n}\left\{E_{i2}^{(r)} - 2\mu_b E_{i1}^{(r)} + \mu_b^2\right\} \\
\frac{\partial Q(\mathbf{b, Y_1, Y_2})}{\partial \sigma_e^2} &= -\frac{n}{\sigma_e^2} + \frac{1}{2\sigma_e^4}\sum_{i=1}^{n}\left\{-2(Y_{i2} - \beta_0)\left[\beta_1 E_{i1}^{(r)} + \frac{\beta_2(A_{i2}^{(r)} - \psi f_2^{(r)}(Y_{i1}, Y_{i2}))}{f^{(r)}(Y_{i1}, Y_{i2})}\right]\right. \\
&\qquad + \beta_1^2 E_{i2}^{(r)} + \frac{2\beta_1\beta_2(A_{i4}^{(r)} - \psi A_{i2}^{(r)}) + \beta_2^2(A_{i4}^{(r)} - 2\psi A_{i2}^{(r)} + \psi^2 f_2^{(r)}(Y_{i1}, Y_{i2}))}{f^{(r)}(Y_{i1}, Y_{i2})} \\
&\qquad \left. + (Y_{i2} - \beta_0)^2 - 2Y_{i1}E_{i1}^{(r)} + E_{i2}^{(r)} + Y_{i1}^2\right\} \\
\frac{\partial Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(r)})}{\partial \psi} &= -\frac{1}{2\sigma_e^2}\sum_{i=1}^{n} 2\left\{\frac{(y_{i2} - \beta_0)\beta_2 f_2^{(r)}(y_{i1}, y_{i2}) - \beta_1\beta_2 A_{i2}^{(r)} + \beta_2^2(-A_{i2}^{(r)} + \psi f_2^{(r)}(y_{i1}, y_{i2}))}{f^{(r)}(y_{i1}, y_{i2})}\right\}
\end{aligned} \tag{2.22}
$$

By setting each of the derivatives in (2.22) equal to zero and solving for the corresponding parameter, we get:

$$
\begin{aligned}
\mu_b &= \frac{\sum_{i=1}^n E_{i1}^{(r)}}{n} \\
\beta_0 &= \frac{\sum_{i=1}^n \left\{ y_{i2} - \beta_1 E_{i1}^{(r)} - \frac{\beta_2\left(A_{i2}^{(r)} - \psi f_2^{(r)}(y_{i1},y_{i2})\right)}{f^{(r)}(y_{i1},y_{i2})} \right\}}{n} \\
\beta_1 &= \frac{\sum_{i=1}^n \left\{ (y_{i2}-\beta_0) E_{i1}^{(r)} - \frac{\beta_2\left(A_{i4}^{(r)} - \psi A_{i2}^{(r)}\right)}{f^{(r)}(y_{i1},y_{i2})} \right\}}{\sum_{i=1}^n E_{i2}^{(r)}} \\
\beta_2 &= \frac{\sum_{i=1}^n \frac{\left\{(y_{i2}-\beta_0)\left(A_{i2}^{(r)} - \psi f_2^{(r)}(y_{i1},y_{i2})\right) - \beta_1\left(A_{i4}^{(r)} - \psi A_{i2}^{(r)}\right)\right\}}{f^{(r)}(y_{i1},y_{i2})}}{\sum_{i=1}^n \frac{A_{i4}^{(r)} - 2\psi A_{i2}^{(r)} + \psi^2 f_2^{(r)}(y_{i1},y_{i2})}{f^{(r)}(y_{i1},y_{i2})}} \\
\sigma_b^2 &= \frac{\sum_{i=1}^n \left\{ E_{i2}^{(r)} - 2\mu_b E_{i1}^{(r)} + \mu_b^2 \right\}}{n} \\
\sigma_e^2 &= \frac{1}{2n}\sum_{i=1}^n \Bigg\{ y_{i1}^2 + (y_{i2}-\beta_0)^2 - 2y_{i1}E_{i1}^{(r)} + E_{i2}^{(r)} - 2(y_{i2}-\beta_0)\Bigg[\beta_1 E_{i1}^{(r)} \\
&\quad + \frac{\beta_2(A_{i2}^{(r)} - \psi f_2^{(r)}(y_{i1},y_{i2}))}{f^{(r)}(y_{i1},y_{i2})}\Bigg] + \beta_1^2 E_{i2}^{(r)} \\
&\quad + \frac{2\beta_1\beta_2(A_{i4}^{(r)} - \psi A_{i2}^{(r)}) + \beta_2^2(A_{i4}^{(r)} - 2\psi A_{i2}^{(r)} + \psi^2 f_2^{(r)}(y_{i1},y_{i2}))}{f^{(r)}(y_{i1},y_{i2})} \Bigg\} \\
\psi &= \frac{\sum_{i=1}^n \left\{ \frac{(\beta_1+\beta_2)A_{i2}^{(r)} - (y_{i2}-\beta_0) f_2^{(r)}(y_{i1},y_{i2})}{f^{(r)}(y_{i1},y_{i2})} \right\}}{\beta_2 \sum_{i=1}^n \frac{f_2^{(r)}(y_{i1},y_{i2})}{f^{(r)}(y_{i1},y_{i2}))}}.
\end{aligned}
\tag{2.23}
$$

It may be noted that these values do not provide a solution for the simultaneous equations $\partial Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(r)})/\partial\boldsymbol{\theta} = \mathbf{0}$. Taken together, the E and CM steps in the rth iteration of our ECM algorithm are as follows:

E-step: Compute $A_{i1}^{(r)}, \ldots, A_{i4}^{(r)}$ and hence the conditional expectations in $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(r)})$.

CM-step 1: Update μ_b by maximizing $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(r)})$ with respect to μ_b while holding all other parameters at their current values. This yields $\mu_b^{(r+1)}$ as the value of μ_b given in (2.23).

CM-step 2: Update β_0 by maximizing $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(r)})$ with respect to β_0 while holding μ_b at the updated value and all other parameters at their current values. This yields $\beta_0^{(r+1)}$ as the value of β_0 given in (2.23).

CM-step 3: Update β_1 by maximizing $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(r)})$ with respect to β_1 while holding μ_b and β_0 at their updated values and all other parameters at their current values. This yields $\beta_1^{(r+1)}$ as the value of β_1 given in (2.23).

CM-step 4: Update β_2 by maximizing $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(r)})$ with respect to β_2 while holding μ_b, β_0, and β_1 at their updated values and all other parameters at their current values. This yields $\beta_2^{(r+1)}$ as the value of β_2 given in (2.23).

CM-step 5: Update σ_b^2 by maximizing $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(r)})$ with respect to σ_b^2 while holding μ_b, β_0, β_1, and β_2 at their updated values and all other parameters at their current values. This yields $\sigma_b^{2,(r+1)}$ as the value of σ_b^2 given in (2.23).

CM-step 6: Update σ_e^2 by maximizing $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(r)})$ with respect to σ_e^2 while holding μ_b, β_0, β_1, β_2, and σ_b^2 at their updated values and all other parameters at their current values. This yields $\sigma_e^{2,(r+1)}$ as the value of σ_e^2 given in (2.23).

CM-step 7: Update ψ by maximizing $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(r)})$ with respect to ψ while holding all other parameters at their updated values. This yields $\psi^{(r+1)}$ as the value of ψ given in (2.23).

This algorithm begins with a starting point and is iterated until convergence. Moreover, as is true with any EM-type algorithm, one needs to try a number of starting points to have some assurance that the algorithm converges to a global maxima $\hat{\boldsymbol{\theta}}$. Next, let $\mathbf{I} = -\partial^2 \log\{L(\boldsymbol{\theta})\}/\partial\boldsymbol{\theta}^2|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}$ denote the Hessian matrix of the negative of the log-likelihood function $L(\boldsymbol{\theta})$ evaluated at the MLE, where $L(\boldsymbol{\theta})$ is given by (2.17). This matrix is also known as the *observed information matrix*. It can be computed by numerical differentiation. We are also able to derive analytical expressions for the elements of this matrix. (See the appendix at the end of this chapter). When n is large, it follows from the large-sample theory of

ML estimators (Lehmann, 1998, Chapter 7) that the distribution of $\hat{\boldsymbol{\theta}}$ can be approximated by a $N(\boldsymbol{\theta}, \mathbf{I}^{-1})$ distribution. This result is used to perform inference on $\boldsymbol{\theta}$.

2.3.2 Fitted values

Once the model (2.6) is fit to the data, we can replace the unknown parameters in (2.16) with their ML estimates to get the estimated best linear predictor $\hat{b}_i$ of b_i. Then, the fitted values can be computed as

$$\hat{Y}_{i1} = \hat{b}_i, \; \hat{Y}_{i2} = \hat{\beta}_0 + \hat{\beta}_1 \hat{b}_i + \hat{\beta}_2 (\hat{b}_i - \hat{\psi})_+.$$

2.3.3 Testing for changepoint

In the segmented model (2.6), we are particularly interested in checking whether or not there is any need to include the changepoint. Without the changepoint, the model reduces to the classical model (2.2). Thus, the null hypothesis H_0 of interest is that the data follow (2.2) against the alternative H_1 that the data follow (2.6). We consider a likelihood ratio test for this. The test statistic is $\Lambda = -2(L_{\text{reduced}} - L_{\text{full}})$, where L_{full} and L_{reduced} respectively represent the log-likelihood functions under the full model model (2.6) and the reduced model (2.2) evaluated at the ML estimates. We use the expressions given in (Fuller, 1987, Chapter 1) to obtain ML estimates of the classical model. Let Λ_{obs} be the observed value of the test statistic. Oftentimes, when n is large, the null distribution of a likelihood ratio statistic can be approximated by a χ^2-distribution with degrees of freedom equal to the number of free parameters in the full model that are fixed to get the reduced model. However, in our case, the reduced model (2.2) can be obtained from full model (2.6) by setting either $\beta_2 = 0$ or taking limit $\psi \to \infty$, the standard χ^2 approximation does not apply. A similar issue, although in the context of segmented linear regression model, was encountered by (Hinkley, 1971) and the authors used a χ^2 approximation with 3 degrees of freedom. We may also follow along the lines of (Hinkley, 1971) and use a χ^2 approximation with degree of freedom equal to either 2, 3 or something in between such as 2.5. A better alternative may be to use bootstrap (Davison and Hinkley, 1997, Chapter 4) to approximate the null distribution of Λ and compute the p-value. Alternatively, we may also use model selection criteria such as Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) to compare the two models. The steps in the calculation of p-value by bootstrap are as follows:

Step 1: Simulate n independent pairs of observations (Y_{i1}^*, Y_{i2}^*), $i = 1, \ldots, n$ following the model (2.2) with $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$. These data represent a parametric resample of the original sample.

Step 2: Fit the full and reduced models to the resample from Step 1 and compute the test statistic Λ.

Step 3: Repeat Steps 1 and 2 a large number of times, say, B, resulting in B values $\Lambda_1^*, \ldots, \Lambda_B^*$ for the test statistic. Take the proportion of these values that are greater than or equal to Λ_{obs} as the approximate p-value for the test.

2.4 Evaluation of similarity and agreement

In a method comparison study, evaluation of similarity is a comparison of marginal characteristics of the measurement methods such as their biases and precisions. On the other hand, evaluation of agreement is essentially an examination of how close the bivariate distribution of the methods is to being degenerate on the 45^o line. In this case, the methods have perfect agreement because their measurements are identical or equivalent, their difference is zero with probability one, making them interchangeable. Evaluation of both similarity and agreement are needed in method comparison studies (Choudhary and Nagaraja, 2017, Chapter 1). In case of the classical model (2.1), similarity is evaluated by examining biases using intercept β_0 and slope β_1 and precisions using squared sensitivity ratio $\beta_1^2(\sigma_{e1}^2/\sigma_{e2}^2)$. If $(\beta_0, \beta_1, \sigma_{e1}^2/\sigma_{e2}^2) = (0, 1, 1)$, the methods have the same fixed and proportional biases and precisions, and hence are similar. In this case, the measurement methods have the same marginal distributions. Agreement between the methods is evaluated using measures of agreement such as CCC and TDI (Lin, 1989, 2000). These are defined as follows:

$$\text{CCC} = \frac{2\text{cov}(Y_1, Y_2)}{\{E(Y_1) - E(Y_2)\}^2 + \text{var}(Y_1) + \text{var}(Y_2)},$$
$$\text{TDI}(p) = p\text{th quantile of } |D|, \tag{2.24}$$

where p is a large probability specified by the practitioner. Usually $p \in \{0.85, 0.90, 0.95\}$ is used in application. By definition, $|\text{CCC}| \leq 1$ and $\text{TDI}(p) \geq 0$. Good agreement is implied by a large value for CCC or a small value for TDI. Agreement is perfect in the limiting case when CCC $= 1$ or TDI $= 0$. To evaluate similarity and agreement, we generally use two-sided confidence intervals for the measures of similarity and one-sided confidence intervals for the measures of agreement.

To evaluate similarity and agreement in case of the segmented model (2.3), we can proceed in exactly the same manner as described above but due to the change in the relationship between the methods over the two segments, this evaluation must be done separately for the two segments using segment-specific versions of the measures. Thus, for similarity evaluation, it follows from (2.4) that fixed and proportional biases can be compared using β_0 and β_1 in the left segment and $\beta_0 - \beta_2\psi$ and $\beta_1 + \beta_2$ in the right segment. Although the precisions can be compared using squared sensitivity ratio but, under the equal error variance assumption in (2.3), it equals β_1^2 in the left segment and $(\beta_1 + \beta_2)^2$ in the right segment, whose square-roots are already examined as part of the bias evaluation. Thus, the methods can be considered similar over a segment if the intercept and slope over that segment equal zero and one, respectively.

The segment-specific versions of CCC and TDI are obtained by evaluating (2.24) under the marginal distribution of (Y_1, Y_2) when b is truncated to be either $b \leq \psi$ (left segment) or $b > \psi$ (right segment). For CCC, this amounts to substituting in (2.24) the relevant conditional moments from Proposition 2.2. However, such closed-form expressions are not available in case of TDI. These are obtained numerically by solving $p = P(|D| \leq q | b \leq \psi)$ and $p = P(|D| \leq q | b > \psi)$ for q where the probabilities are obtained by integrating the relevant marginal density of D from Proposition 2.5. We may also compute a single CCC and TDI for the entire measurement range by using the marginal moments of (Y_1, Y_2) given by Proposition 2.1 and the marginal density of D given by Proposition 2.4. But, unless there is no changepoint in which case the segmented model reduces to the classical model, these overall measures may be misleading.

The various measures needed for evaluation of similarity and agreement are functions of the model parameter vector $\boldsymbol{\theta}$. As in the classical model, they are estimated by replacing $\boldsymbol{\theta}$ in their definitions with its ML estimate $\hat{\boldsymbol{\theta}}$ and their one- and two-sided confidence intervals are computed using the multivariate delta method (Lehmann, 1998, Chapter 5). To improve accuracy for parameters or parameter functions whose range is not the entire real line, the confidence intervals are computed after applying a normalizing transformation and the results are inverted back to the original scale. Specifically, a log transformation is applied to σ_e^2, σ_b^2, and TDI and the Fisher's z-transformation is applied to CCC. Alternatively, the confidence intervals can also be computed using the percentile bootstrap method (Davison and Hinkley, 1997, Chapter 5) in which case no transformation is needed.

2.5 A simulation study

A simulation study is performed to evaluate the performance of point and interval estimators of parameters in the proposed model and the test for changepoint. We are specifically interested in examining the following: (a) biases and mean squared errors (MSEs) for ML estimators; (b) accuracy of confidence intervals provided by standard large-sample theory where the Hessian matrix provided at the end of this chapter is used to compute the estimated covariance matrix of the ML estimate; (c) accuracy of confidence intervals provided by percentile bootstrap method; and (d) accuracy of the test for changepoint where p-value is computed by bootstrap (with $B = 500$ replications) and by approximating the null distribution of the test statistic by a χ^2 distribution with 2 and 3 degrees of freedom. In (a), (b), and (c), ML estimates obtained by both the ECM algorithm and direct maximization are compared. For direct maximization, we use the `optim` function in `R`.

To evaluate the accuracy of confidence intervals, we consider the following five settings for parameters of the segmented model (2.6) that are motivated by the point estimates and 95% confidence intervals obtained from the analysis of digoxin data.

- Setting 1a: Set parameters equal to their estimates obtained for the digoxin data, i.e., $(\mu_b, \beta_0, \beta_1, \beta_2, \log\sigma_e^2, \log\sigma_b^2, \psi) = (0.180, -0.292, 0.461, 0.480, -4.871, 0.172, -0.707)$.

- Setting 2a: Replace mean related parameters $(\mu_b, \beta_0, \beta_1, \beta_2)$ in Setting 1 by their lower confidence limits and keep the others as in Setting 1, i.e., $(\mu_b, \beta_0, \beta_1, \beta_2, \log\sigma_e^2, \log\sigma_b^2, \psi)$ $= (-0.005, -0.387, 0.406, 0.417, -4.871, 0.172, -0.707)$.

- Setting 3a: Replace mean related parameters $(\mu_b, \beta_0, \beta_1, \beta_2)$ in Setting 1 by their upper confidence limits and keep the others as in Setting 1, i.e., $(\mu_b, \beta_0, \beta_1, \beta_2, \log\sigma_e^2, \log\sigma_b^2, \psi)$ $= (0.365, -0.197, 0.516, 0.543, -4.871, 0.172, -0.707)$.

- Setting 4a: Replace variance related parameters $(\log\sigma_e^2, \log\sigma_b^2)$ in Setting 1 by their lower confidence limits keep the others as in Setting 1, i.e., $(\mu_b, \beta_0, \beta_1, \beta_2, \log\sigma_e^2, \log\sigma_b^2,$ $\psi) = (0.180, -0.292, 0.461, 0.480, -5.110, -0.069, -0.707)$.

- Setting 5a: Replace variance related parameters $(\log\sigma_e^2, \log\sigma_b^2)$ in Setting 1 by their upper confidence limits and keep the others as in Setting 1, i.e., $(\mu_b, \beta_0, \beta_1, \beta_2, \log\sigma_e^2,$ $\log\sigma_b^2, \psi) = (0.180, -0.292, 0.461, 0.480, -4.632, 0.413, -0.707)$.

The number of subjects is set to $n \in (30, 50, 100, 200)$. We also choose 0.95 as the nominal confidence level for the confidence intervals and 0.05 as the nominal level for the test of changepoint. Thus, altogether we investigate a total of $5 \times 4 = 20$ combinations of n and parameter settings. For a given combination, the data are simulated from the model (2.6) and the estimates are computed and the tests are performed as described in Section 2.3. Then, this process is repeated 500 times and the results are used to compute estimated biases and MSEs of the point estimators, coverage probabilities of the confidence intervals, and type I error probability for the tests. The results for point and interval estimates for the four values of n under parameter setting 1 are summarized in Tables 2.1 through 2.4 summarize. The ratio of MSEs of ML estimates obtained by ECM and direct maximization, with former in the denominator, are also presented in these tables. The results for other parameter settings are omitted as they qualitatively similar. The estimated type I error probabilities for tests for all the parameter settings are presented in Table 2.5.

We can conclude the following regarding point and interval estimates from Tables 2.1-2.4:

- With a few notable exceptions in case of $n = 30$, both ECM and direct maximization lead to nearly identical MSEs of the estimates. When the exceptions occur, ECM leads to slightly smaller MSE than direct maximization. Although, in principle, both ECM and direct maximization are expected to provide identical ML estimates, but in practice, direct maximization is more sensitive to starting points than ECM, especially when n is not large, and this explains the difference in the two sets of estimates.

- As expected, both bias and MSE for estimators of all parameters decrease as n increases.

- The coverage probabilities of the standard large-sample confidence intervals are generally less than the nominal level unless $n = 200$ in which case they are quite close for the nominal level. However, for values of $n = 30, 50$, these intervals cannot be considered accurate. With a few notable exceptions, the bootstrap confidence intervals are generally more accurate than the standard intervals. On the whole, unless n is 100 or more, bootstrap should be preferred over the standard intervals.

Table 2.1. Bias, MSE, and ratio of MSEs of estimates obtained by ECM and direct maximization (DIR) and coverage probabilities for 95% standard and bootstrap confidence intervals in case of $n = 30$.

	Bias		MSE		MSE	Coverage Probability			
						Standard		Bootstrap	
Parameter	ECM	DIR	ECM	DIR	Ratio	ECM	DIR	ECM	DIR
μ_b	-0.007	-0.007	0.036	0.036	1.000	0.954	0.954	0.942	0.942
β_0	-0.054	-0.056	0.046	0.054	0.848	0.89	0.878	0.926	0.978
β_1	-0.039	-0.041	0.025	0.033	0.770	0.908	0.9	0.922	0.972
β_2	0.045	0.048	0.025	0.032	0.769	0.928	0.922	0.922	0.986
$\log \sigma_e^2$	-0.117	-0.118	0.089	0.089	0.997	0.906	0.906	0.788	0.96
$\log \sigma_b^2$	-0.133	-0.133	0.077	0.077	1.000	0.93	0.93	0.878	0.878
ψ	0.005	0.009	0.051	0.052	0.995	0.84	0.834	0.922	0.96

Table 2.2. Bias, MSE, and ratio of MSEs of estimates obtained by ECM and direct maximization (DIR) and coverage probabilities for 95% standard and bootstrap confidence intervals in case of $n = 50$.

	Bias		MSE		MSE	Coverage Probability			
						Standard		Bootstrap	
Parameter	ECM	DIR	ECM	DIR	Ratio	ECM	DIR	ECM	DIR
μ_b	-0.013	-0.013	0.022	0.022	1.000	0.956	0.956	0.952	0.954
β_0	-0.034	-0.033	0.024	0.025	0.981	0.894	0.89	0.952	0.974
β_1	-0.023	-0.022	0.011	0.012	0.971	0.914	0.91	0.948	0.97
β_2	0.026	0.025	0.011	0.011	0.974	0.934	0.934	0.95	0.98
$\log \sigma_e^2$	-0.058	-0.058	0.047	0.047	1.001	0.924	0.924	0.86	0.976
$\log \sigma_b^2$	-0.065	-0.065	0.041	0.041	1.000	0.952	0.952	0.92	0.918
ψ	-0.008	-0.007	0.032	0.033	0.989	0.854	0.85	0.934	0.956

Table 2.3. Bias, MSE, and ratio of MSEs of estimates obtained by ECM and direct maximization (DIR) and coverage probabilities for 95% standard and bootstrap confidence intervals in case of $n = 100$.

	Bias		MSE		MSE	Coverage Probability			
						Standard		Bootstrap	
Parameter	ECM	DIR	ECM	DIR	Ratio	ECM	DIR	ECM	DIR
μ_b	-0.006	-0.006	0.012	0.012	1.000	0.932	0.932	0.938	0.94
β_0	-0.011	-0.011	0.007	0.007	0.985	0.914	0.914	0.948	0.962
β_1	-0.006	-0.006	0.003	0.003	0.982	0.922	0.922	0.95	0.96
β_2	0.005	0.005	0.003	0.003	0.981	0.94	0.94	0.952	0.962
$\log \sigma_e^2$	-0.045	-0.045	0.027	0.027	1.001	0.92	0.92	0.868	0.99
$\log \sigma_b^2$	-0.026	-0.026	0.022	0.022	1.000	0.932	0.932	0.918	0.92
ψ	-0.008	-0.008	0.011	0.011	1.006	0.914	0.914	0.944	0.964

Table 2.4. Bias, MSE, and ratio of MSEs of estimates obtained by ECM and direct maximization (DIR) and coverage probabilities for 95% standard and bootstrap confidence intervals in case of $n = 200$.

	Bias		MSE		MSE	Coverage Probability			
						Standard		Bootstrap	
Parameter	ECM	DIR	ECM	DIR	Ratio	ECM	DIR	ECM	DIR
μ_b	0.006	0.006	0.006	0.006	1.000	0.942	0.942	0.938	0.938
β_0	0.001	0.001	0.002	0.002	0.987	0.948	0.948	0.966	0.970
β_1	0.000	0.000	0.001	0.001	0.986	0.962	0.960	0.968	0.980
β_2	0.001	0.001	0.001	0.001	0.987	0.956	0.956	0.968	0.982
$\log \sigma_e^2$	-0.029	-0.028	0.011	0.011	0.999	0.936	0.936	0.902	0.982
$\log \sigma_b^2$	-0.010	-0.010	0.010	0.010	1.000	0.952	0.952	0.938	0.940
ψ	0.010	0.010	0.005	0.005	0.993	0.946	0.946	0.956	0.964

We evaluate the accuracy of the test for changepoint by examining how close the estimated type I error probabilities of the changepoint tests are to the nominal level of $\alpha = 0.05$. For this, data are simulated under the null hypothesis from the classical model (2.2) with parameter values as in settings 1a through 5a except that the parameters β_2 and ψ are omitted. We refer to these settings as 1b through 5b, respectively. The estimated type I error probabilities presented in Table 2.5 show that the χ_3^2 approximation does not work well as it is quite liberal. The χ_2^2 approximation is generally conservative for $n = 30, 50$ but seems quite accurate for $n = 100$. These conclusions hold for both ECM and direction maximization methods. In case of ECM, the bootstrap approximation also works well and it may be considered the best of the three approximations. However, this is not the case for bootstrap when direct maximization is used. On the whole, these results suggest that both χ_2^2 and bootstrap approximations can be used to perform the test of changepoint.

2.6 Illustration: Analysis of digoxin data

We now return to the digoxin data introduced in Section 2.1 and analyze them using the methodology proposed in this chapter. First, we fit the model (2.6) by ML using both ECM algorithm and direct maximization of the log-likelihood function. Table 2.6 presents point

Table 2.5. Estimated type I error probabilities for the test of changepoint for when the null distribution of the likelihood ratio test statistic is approximated by χ^2 distributions with 2 and 3 degrees of freedom and bootstrap.

		ECM			DIR		
Setting	n	χ^2_2	χ^2_3	Bootstrap	χ^2_2	χ^2_3	Bootstrap
1b	30	0.066	0.024	0.052	0.064	0.024	0.052
	50	0.068	0.028	0.056	0.054	0.018	0.060
	100	0.044	0.026	0.044	0.036	0.014	0.042
2b	30	0.074	0.038	0.056	0.060	0.026	0.054
	50	0.076	0.036	0.064	0.048	0.018	0.052
	100	0.046	0.012	0.034	0.016	0.008	0.020
3b	30	0.064	0.032	0.054	0.048	0.022	0.060
	50	0.040	0.024	0.036	0.028	0.012	0.050
	100	0.048	0.018	0.050	0.020	0.012	0.054
4b	30	0.052	0.026	0.050	0.044	0.028	0.044
	50	0.044	0.024	0.044	0.030	0.018	0.036
	100	0.054	0.026	0.058	0.040	0.018	0.056
5b	30	0.052	0.024	0.046	0.040	0.022	0.046
	50	0.052	0.020	0.050	0.050	0.028	0.068
	100	0.040	0.022	0.036	0.024	0.008	0.030

estimates obtained by ECM, standard errors of estimates obtained by numerical computing the Hessian matrix, and 95% large-sample confidence intervals for model parameters. Estimates produced by direct maximization are nearly identical and hence are omitted. None of the standard errors appears unusually high. We see that the interval $(0.42, 0.54)$ for β_2 does not contain zero, suggesting the need for a changepoint. Next, we test the null hypothesis that the data follow the classical model (2.2) using the likelihood ratio test. The observed value of the test statistic is 118.12 and the p-value, computing using bootstrap with $B = 500$ replications, is zero, confirming the need for the changepoint. Even AIC and BIC prefer the segmented model (2.6) over the classical model (2.2) as their respective values for the two models are 215.6 and 329.7 (AIC) and 235.8 and 344.2 (BIC).

The estimated changepoint is $\hat{\psi} = -0.71$. Substituting the unknowns in (2.4) with their estimates gives the fitted piecewise straight line for $E(Y_2|b)$. The resulting line is $-0.29+0.46b$ in the left segment ($b \leq -0.71$) and $0.05 + 0.94b$ in the right segment ($b > -0.71$). They

Table 2.6. ML estimates, their standard errors (SE), and 95% confidence intervals for parameters of the segmented model (2.6) for digoxin data.

Parameter	Estimate	SE	95% Interval
μ_b	0.18	0.09	$(-0.01, 0.37)$
β_0	-0.29	0.05	$(-0.39, -0.20)$
β_1	0.46	0.03	$(0.41, 0.52)$
β_2	0.48	0.03	$(0.42, 0.54)$
$\log(\sigma_e^2)$	-4.87	0.12	$(-5.11, -4.63)$
$\log(\sigma_b^2)$	0.17	0.12	$(-0.07, 0.41)$
ψ	-0.71	0.13	$(-0.96, -0.46)$

are superimposed on the scatterplot in Figure 2.1. They seem to provide a good fit to the underlying trend in the data. For comparison, also superimposed on the plot is the line $0.16 + 0.79b$ obtained by fitting the classical model. It clearly does not fit the data well.

Under the fitted segmented model (2.6), the bivariate distribution of (Y_1, Y_2) has mean $(-1.32, -0.90)$, variance $(0.27, 0.06)$, and correlation 0.92 in the left segment; and mean $(0.57, 0.59)$, variance $(0.69, 0.61)$, and correlation 0.99 in the right segment. These are obtained by replacing the unknowns in the moments given by Proposition 2.2 with their estimates. It follows that the fitted distribution of $D = Y_2 - Y_1$ has mean 0.42 and variance 0.09 in the left segment and 0.01 and 0.02 in the right segment. Thus, assay 2 has higher mean and lower variance than assay 1 in both segments but the difference is much smaller in the right segment than the left segment. The correlation between the assays is also higher in the right segment. These findings are consistent with what we saw in Figure 2.1.

Next, we consider evaluation of similarity of the assays. Table 2.7 presents estimates and 95% confidence intervals for the segment-specific similarity measures considered in Section 2.4. None of the intervals appears unusually wide. In the left segment, the intervals for the intercept β_0 and slope β_1 are $(-0.39, -0.20)$ and $(0.41, 0.52)$, respectively. In the right segment, they are $(0.00, 0.09)$ and $(0.89, 0.99)$ for the intercept $\beta_0 - \beta_2\psi$ and slope $\beta_1 + \beta_2$, respectively. None of the intercept intervals covers zero although zero is on the left boundary of the second interval. Likewise, none of the slope intervals covers one although one is near

Table 2.7. Summary of estimates for segment-specific measures of similarity and agreement. Two-sided confidence intervals are presented for similarity measures and one-sided confidence intervals are presented for agreement measures. The intercept and slope respectively refer to β_0 and β_1 in the left segment and $\beta_0 - \beta_2\psi$ and $\beta_1 + \beta_2$ in the right segment. Also, z(CCC) refers to Fisher's z-transformation of CCC.

	Left Segment			Right Segment		
Measure	Estimate	SE	95% Interval	Estimate	SE	95% Interval
Intercept	-0.29	0.05	$(-0.39, -0.20)$	0.05	0.02	$(0.00, 0.09)$
Slope	0.46	0.03	$(0.41, 0.52)$	0.94	0.03	$(0.89, 0.99)$
z(CCC)	0.52	0.08	$(0.39, \infty)$	2.49	0.09	$(2.34, \infty)$
$\log\{\text{TDI}(0.90)\}$	-0.18	0.09	$(-\infty, -0.03)$	-1.51	0.09	$(-\infty, -1.37)$

the right boundary of the second interval. Thus, because their fixed and proportional biases are not equal, the assays cannot be regarded as similar in either segment. That said, the biases differ considerably in the left segment and only moderately so in the right segment.

Our next task is to evaluate agreement between the assays. Estimates and 95% one-sided confidence bounds for the segment-specific agreement measures considered in Section 2.4 are also presented in Table 2.7 on transformed scale. These estimates and bounds can be transformed back to the original scale by applying the inverse transformation. In the left segment, the lower bound for CCC is 0.37 and it is 0.98 in the right segment. Further, the upper bounds for TDI (0.90) in the left and right segments are 0.97 and 0.25, respectively. Thus, on the basis of both measures, the assays exhibit much higher agreement in the right segment than in the left segment. This is consistent with what we expect on the basis of Figure 2.1. Focusing on the right segment, we see that the CCC estimate and lower bound are nearly one, indicating potentially excellent agreement between the assays. But this conclusion may be misleading considering that, in these data, $\hat{\sigma}_e^2 = \exp(-4.87) \approx 0.01$ is much smaller than $\hat{\sigma}_b^2 = \exp(0.17) \approx 1.19$. In such a case, the CCC estimate tends to be high despite how close the actual measurements of the two methods are (Atkinson and Nevill, 1997). A better picture of agreement is given by the TDI upper bound which indicates that 90% of ratios of measurements by the two assays are estimated to fall within

$\exp(-0.25) \approx 0.78$ to $\exp(0.25) \approx 1.28$ with 95% confidence. Given how wide this range is around one, the extent of agreement between the assays cannot be considered strong.

Thus, altogether we see that the two digoxin assays exhibit considerably more similarity and agreement at higher analyte levels than at low levels. However, the above findings show that even at high analyte levels, the assays cannot be regarded as similar in the sense of having equal fixed and proportional biases or having well enough agreement for interchangeable use.

2.7 Appendix: Hessian matrix

Let $d_{AB} = \frac{dA}{dB}$, $d^2_{AB} = \frac{d^2A}{dB^2}$, and $d^2_{ABC} = \frac{d^2A}{dBdC}$. The first derivatives of the functions $g_4(\beta, \sigma_1, \sigma_2)$ and $g_6(\beta, \sigma_1, \sigma_2)$, given by (1.1), with respect to parameters of the model (2.3) are as follows.

$$d_{g_4(\beta,\sigma_e^2,\sigma_e^2)\mu_b} = \frac{dg_4(\beta,\sigma_e^2,\sigma_e^2)}{d\mu_b} = 0$$

$$d_{g_4(\beta,\sigma_e^2,\sigma_e^2)\beta_0} = \frac{dg_4(\beta,\sigma_e^2,\sigma_e^2)}{d\beta_0} = 0$$

$$d_{g_4(\beta,\sigma_e^2,\sigma_e^2)\beta_1} = \frac{dg_4(\beta,\sigma_e^2,\sigma_e^2)}{d\beta_1} = \frac{-2(\beta_1+\beta)\left[g_4(\beta,\sigma_e^2,\sigma_e^2)\right]^2}{\sigma_e^2}$$

$$d_{g_4(\beta,\sigma_e^2,\sigma_e^2)\beta_2} = \frac{dg_4(\beta,\sigma_e^2,\sigma_e^2)}{d\beta_2} = \begin{cases} \frac{-2(\beta_1+\beta_2)\left[g_4(\beta_2,\sigma_e^2,\sigma_e^2)\right]^2}{\sigma_e^2} & ;\text{if} \quad \beta=\beta_2 \\ 0 \quad ; \quad \text{if} \quad \beta = 0 \end{cases}$$

$$d_{g_4(\beta,\sigma_e^2,\sigma_e^2)\sigma_b^2} = \frac{dg_4(\beta,\sigma_e^2,\sigma_e^2)}{d\sigma_b^2} = \frac{\left[g_4(\beta,\sigma_e^2,\sigma_e^2)\right]^2}{\sigma_b^4}$$

$$d_{g_4(\beta,\sigma_e^2,\sigma_e^2)\sigma_e^2} = \frac{dg_4(\beta,\sigma_e^2,\sigma_e^2)}{d\sigma_e^2} = \frac{\left[g_4(\beta,\sigma_e^2,\sigma_e^2)\right]^2(1+(\beta_1+\beta_2)^2)}{\sigma_e^4}$$

$$d_{g_4(\beta,\sigma_e^2,\sigma_e^2)\psi} = \frac{dg_4(\beta,\sigma_e^2,\sigma_e^2)}{d\psi} = 0$$

$$
\begin{aligned}
d_{g_6(\beta,\sigma_e^2,\sigma_e^2)\mu_b} &= \frac{dg_6(\beta,\sigma_e^2,\sigma_e^2)}{d\mu_b} = g_6(\beta,\sigma_e^2,\sigma_e^2)\frac{1}{\sigma_b^2} \\
d_{g_6(\beta,\sigma_e^2,\sigma_e^2)\beta_0} &= \frac{dg_6(\beta,\sigma_e^2,\sigma_e^2)}{d\beta_0} = -g_6(\beta,\sigma_e^2,\sigma_e^2)\frac{\beta_1+\beta}{\sigma_e^2} \\
d_{g_6(\beta,\sigma_e^2,\sigma_e^2)\beta_1} &= \frac{dg_6(\beta,\sigma_e^2,\sigma_e^2)}{d\beta_1} = \frac{\frac{1}{g_6(\beta,\sigma_e^2,\sigma_e^2)}(y_2-\beta_0+\beta\psi) - 2g_5(\beta,\sigma_e^2,\sigma_e^2)(\beta_1+\beta)}{\frac{\sigma_e^2}{[g_6(\beta,\sigma_e^2,\sigma_e^2)]^2}} \\
d_{g_6(\beta,\sigma_e^2,\sigma_e^2)\beta_2} &= \frac{dg_6(\beta,\sigma_e^2,\sigma_e^2)}{d\beta_2} \\
&= \begin{cases} \frac{\frac{1}{g_6(\beta_2,\sigma_e^2,\sigma_e^2)}(y_2-\beta_0+\beta_1\psi+2\beta_2\psi)-2g_5(\beta_2,\sigma_e^2,\sigma_e^2)(\beta_1+\beta_2)}{\frac{\sigma_e^2}{[g_6(\beta_2,\sigma_e^2,\sigma_e^2)]^2}} & ; \quad \text{if} \quad \beta=\beta_2 \\ 0 \quad ; \quad \text{if} \quad \beta = 0 \end{cases} \\
d_{g_6(\beta,\sigma_e^2,\sigma_e^2)\sigma_b^2} &= \frac{dg_6(\beta,\sigma_e^2,\sigma_e^2)}{d\sigma_b^2} = \frac{\frac{-\mu_b}{g_6(\beta,\sigma_e^2,\sigma_e^2)} + g_5(\beta,\sigma_e^2,\sigma_e^2)}{\frac{\sigma_b^4}{[g_6(\beta,\sigma_e^2,\sigma_e^2)]^2}} \\
d_{g_6(\beta,\sigma_e^2,\sigma_e^2)\sigma_e^2} &= \frac{dg_6(\beta,\sigma_e^2,\sigma_e^2)}{d\sigma_e^2} \\
&= \frac{\frac{1}{g_6(\beta,\sigma_e^2,\sigma_e^2)}\left\{y_1+(\beta_1+\beta)(y_2-\beta_0+\beta\psi)+g_5(\beta,\sigma_e^2,\sigma_e^2)[1+(\beta_1+\beta)^2]\right\}}{\frac{\sigma_e^4}{[g_6(\beta,\sigma_e^2,\sigma_e^2)]^2}} \\
d_{g_6(\beta,\sigma_e^2,\sigma_e^2)\psi} &= \frac{dg_6(\beta,\sigma_e^2,\sigma_e^2)}{d\psi} = \frac{g_6(\beta,\sigma_e^2,\sigma_e^2)\beta(\beta_1+\beta)}{\sigma_e^2}
\end{aligned}
$$

Next, the log likelihood function for the model (2.3) can be expressed as follows.

$$
\begin{aligned}
l(y_1,y_2) &= \log(f(y_1,y_2)) \\
&= \log\left\{\exp\left\{a_1+a_2+a_3+a_4+a_5+a_6+a_7\right\} + \exp\left\{a_1+a_2+a_4+a_8+a_9+a_{10}+a_{11}\right\}\right\} \\
&= \log\{\exp(t_1)+\exp(t_2)\}, \qquad (2.25)
\end{aligned}
$$

where $t_1 = a_1+a_2+a_3+a_4+a_5+a_6+a_7$, $t_2 = a_1+a_2+a_4+a_8+a_9+a_{10}+a_{11}$, and

$$
\begin{aligned}
a_1 &= -\frac{1}{2}\left[2\log\sigma_e^2 + \log\sigma_b^2\right] \\
a_2 &= -\frac{1}{2}\log(2\pi) - \frac{y_1^2}{2\sigma_e^2}
\end{aligned}
$$

$$
\begin{aligned}
a_3 &= -\frac{1}{2}\log(2\pi) - \frac{(y_1 - \beta_0 + \beta_2\psi)^2}{2\sigma_e^2} \\
a_4 &= -\frac{1}{2}\log(2\pi) - \frac{\mu_b^2}{2\sigma_b^2} \\
a_5 &= -\frac{1}{2}\log(2\pi) + \frac{1}{2}\log g_4(\beta_2, \sigma_e, \sigma_e) \\
a_6 &= \frac{g_6^2(\beta_2, \sigma_e, \sigma_e)}{2g_4(\beta_2, \sigma_e, \sigma_e)} \\
a_7 &= \log\left[1 - \Phi(g_1(\psi, g_6(\beta_2, \sigma_e, \sigma_e), \sqrt{g_4(\beta_2, \sigma_e, \sigma_e)}))\right] = \log(k_7) \\
a_8 &= -\frac{1}{2}\log(2\pi) - \frac{(y_1 - \beta_0)^2}{2\sigma_e^2} \\
a_9 &= -\frac{1}{2}\log(2\pi) + \frac{1}{2}\log g_4(0, \sigma_e, \sigma_e) \\
a_{10} &= \frac{g_6^2(0, \sigma_e, \sigma_e)}{2g_4(0, \sigma_e, \sigma_e)} \\
a_{11} &= \log\left(\Phi(g_1(\psi, g_6(0, \sigma_e, \sigma_e), \sqrt{g_4(0, \sigma_e, \sigma_e)}))\right) = \log(k_{11}).
\end{aligned}
$$

Let γ and α denote any of the model parameters. Then the first derivative of the log likelihood function with respect to parameter γ is,

$$
d_{l\gamma} = \frac{dl(y_1, y_2)}{d\gamma} = \frac{1}{f(y_1, y_2)}\left\{e^{t_1}\frac{dt_1}{d\gamma} + e^{t_2}\frac{dt_2}{d\gamma}\right\} \tag{2.26}
$$

The second derivative of the log likelihood function with respect to parameters γ and α can be written as,

$$
\begin{aligned}
d^2_{l\gamma\alpha} = \frac{d^2 l(y_1, y_2)}{d\gamma d\alpha} &= \frac{1}{f(y_1, y_2)}\left\{e^{t_1}\frac{dt_1}{d\gamma}\frac{dt_1}{d\alpha} + e^{t_1}\frac{d^2 t_1}{d\gamma d\alpha} + e^{t_2}\frac{dt_2}{d\gamma}\frac{dt_2}{d\alpha} + e^{t_2}\frac{d^2 t_2}{d\gamma d\alpha}\right\} \\
&\quad - \frac{1}{f^2(y_1, y_2)}\frac{df(y_1, y_2)}{d\alpha}\left\{e^{t_1}\frac{dt_1}{d\gamma} + e^{t_2}\frac{dt_2}{d\gamma}\right\} && (2.27) \\
&= \frac{1}{f(y_1, y_2)}\left\{e^{t_1}\frac{dt_1}{d\gamma}\frac{dt_1}{d\alpha} + e^{t_1}\frac{d^2 t_1}{d\gamma d\alpha} + e^{t_2}\frac{dt_2}{d\gamma}\frac{dt_2}{d\alpha} + e^{t_2}\frac{d^2 t_2}{d\gamma d\alpha}\right\} - \frac{dl(y_1, y_2)}{d\gamma}\frac{dl(y_1, y_2)}{d\alpha} && (2.28)
\end{aligned}
$$

The first derivative of t_1 and t_2 with respect to model parameters are given by the (2.29) and (2.30) below. We can plug in these expressions in (2.26) to obtain the first derivative of log likelihood function with respect to the model parameters.

$$
\begin{aligned}
\frac{dt_1}{d\mu_b} &= -\frac{\mu_b}{\sigma_b^2} + d_{a_5\mu_b} + d_{a_6\mu_b} + d_{a_7\mu_b} \\
\frac{dt_1}{d\beta_0} &= \frac{y_1 - \beta_0 + \beta_2\psi}{\sigma_e^2} + d_{a_5\beta_0} + d_{a_6\beta_0} + d_{a_7\beta_0} \\
\frac{dt_1}{d\beta_1} &= d_{a_5\beta_1} + d_{a_6\beta_1} + d_{a_7\beta_1} \\
\frac{dt_1}{d\beta_2} &= -\frac{(y_1 - \beta_0 + \beta_2\psi)\psi}{\sigma_e^2} + d_{a_5\beta_2} + d_{a_6\beta_2} + d_{a_7\beta_2} \\
\frac{dt_1}{d\sigma_b^2} &= -\frac{1}{2\sigma_b^2} + \frac{\mu_b^2}{2\sigma_b^4} + d_{a_5\sigma_b^2} + d_{a_6\sigma_b^2} + d_{a_7\sigma_b^2} \\
\frac{dt_1}{d\sigma_e^2} &= -\frac{1}{\sigma_e^2} + \frac{y_1^2}{2\sigma_e^4} + \frac{(y_1 - \beta_0 + \beta_2\psi)^2}{2\sigma_e^4} + d_{a_5\sigma_e^2} + d_{a_6\sigma_e^2} + d_{a_7\sigma_e^2} \\
\frac{dt_1}{d\psi} &= -\frac{(y_1 - \beta_0 + \beta_2\psi)\beta_2}{\sigma_e^2} + d_{a_5\psi} + d_{a_6\psi} + d_{a_7\psi}
\end{aligned}
\tag{2.29}
$$

$$
\begin{aligned}
\frac{dt_2}{d\mu_b} &= -\frac{\mu_b}{\sigma_b^2} + d_{a_9\mu_b} + d_{a_{10}\mu_b} + d_{a_{11}\mu_b} \\
\frac{dt_2}{d\beta_0} &= \frac{y_2 - \beta_0}{\sigma_e^2} + d_{a_9\beta_0} + d_{a_{10}\beta_0} + d_{a_{11}\beta_0} \\
\frac{dt_2}{d\beta_1} &= d_{a_9\beta_1} + d_{a_{10}\beta_1} + d_{a_{11}\beta_1} \\
\frac{dt_2}{d\beta_2} &= d_{a_9\beta_2} + d_{a_{10}\beta_2} + d_{a_{11}\beta_2} \\
\frac{dt_2}{d\sigma_b^2} &= -\frac{1}{2\sigma_b^2} + \frac{\mu_b^2}{2\sigma_b^4} + d_{a_9\sigma_b^2} + d_{a_10\sigma_b^2} + d_{a_{11}\sigma_b^2} \\
\frac{dt_2}{d\sigma_e^2} &= -\frac{1}{\sigma_e^2} + \frac{y_1^2}{2\sigma_e^4} + \frac{(y_2 - \beta_0)^2}{2\sigma_e^4} + d_{a_9\sigma_e^2} + d_{a_10\sigma_e^2} + d_{a_{11}\sigma_e^2} \\
\frac{dt_2}{d\psi} &= d_{a_9\psi} + d_{a_{10}\psi} + d_{a_{11}\psi}
\end{aligned}
\tag{2.30}
$$

Here,

$$
\begin{aligned}
d_{a_5\gamma} &= \frac{da_5}{d\gamma} = \frac{1}{2g_4(\beta_2, \sigma_e^2, \sigma_e^2)} d_{g_4(\beta_2,\sigma_e^2,\sigma_e^2)\gamma} \\
d_{a_6\gamma} &= \frac{da_6}{d\gamma} = \frac{1}{2}\left[2\frac{g_6(\beta_2, \sigma_e, \sigma_e)}{g_4(\beta_2, \sigma_e, \sigma_e)} - \frac{1}{g_4^2(\beta_2, \sigma_e, \sigma_e)} d_{g_4(\beta_2,\sigma_e,\sigma_e)\gamma} g_6^2(\beta_2, \sigma_e, \sigma_e)\right] \\
d_{a_7\gamma} &= \frac{da_7}{d\gamma} = \frac{1}{k_7}\frac{dk_7}{d\gamma} \\
d_{a_9\gamma} &= \frac{da_9}{d\gamma} = \frac{1}{2g_4(0, \sigma_e^2, \sigma_e^2)} d_{g_4(0,\sigma_e^2,\sigma_e^2)\gamma} \\
d_{a_10\gamma} &= \frac{da_10}{d\gamma} = \frac{1}{2}\left[2\frac{g_6(0, \sigma_e, \sigma_e)}{g_4(0, \sigma_e, \sigma_e)} - \frac{1}{g_4^2(0, \sigma_e, \sigma_e)} d_{g_4(0,\sigma_e,\sigma_e)\gamma} g_6^2(0, \sigma_e, \sigma_e)\right] \\
d_{a_{11}\gamma} &= \frac{da_{11}}{d\gamma} = \frac{1}{k_{11}}\frac{dk_{11}}{d\gamma}
\end{aligned}
$$

The expressions in (2.31) and (2.32) provide second derivative of the terms t_1 and t_2 with respect to the model parameters.

$$
\begin{aligned}
\frac{d^2t_1}{d\mu_b^2} &= -\frac{1}{\sigma_b^2} + d^2_{a_5\mu_b} + d^2_{a_6\mu_b} + d^2_{a_7\mu_b} \\
\frac{d^2t_1}{d\beta_0^2} &= -\frac{1}{\sigma_e^2} + d^2_{a_5\beta_0} + d^2_{a_6\beta_0} + d^2_{a_7\beta_0} \\
\frac{d^2t_1}{d\beta_1^2} &= d^2_{a_5\beta_1} + d^2_{a_6\beta_1} + d^2_{a_7\beta_1} \\
\frac{d^2t_1}{d\beta_2^2} &= -\frac{\psi^2}{\sigma_e^2} + d^2_{a_5\beta_2} + d^2_{a_6\beta_2} + d^2_{a_7\beta_2} \\
\frac{d^2t_1}{d(\sigma_b^2)^2} &= \frac{1}{2\sigma_b^4} - \frac{\mu_b^2}{\sigma_b^6} + d^2_{a_5\sigma_b^2} + d^2_{a_6\sigma_b^2} + d^2_{a_7\sigma_b^2} \\
\frac{d^2t_1}{d(\sigma_e^2)^2} &= \frac{1}{\sigma_e^4} - \frac{y_1^2}{\sigma_e^6} - \frac{(y_1 - \beta_0 + \beta_2\psi)^2}{\sigma_e^6} + d^2_{a_5\sigma_e^2} + d^2_{a_6\sigma_e^2} + d^2_{a_7\sigma_e^2} \\
\frac{d^2t_1}{d\psi^2} &= -\frac{\beta_2^2}{\sigma_e^2} + d^2_{a_5\psi} + d^2_{a_6\psi} + d^2_{a_7\psi}
\end{aligned} \tag{2.31}
$$

$$\frac{d^2 t_2}{d\mu_b^2} = -\frac{1}{\sigma_b^2} + d^2_{a_9\mu_b} + d^2_{a_{10}\mu_b} + d^2_{a_{11}\mu_b}$$
$$\frac{dt_2}{d\beta_0^2} = -\frac{1}{\sigma_e^2} + d^2_{a_9\beta_0} + d^2_{a_{10}\beta_0} + d^2_{a_{11}\beta_0}$$
$$\frac{d^2 t_2}{d\beta_1^2} = d^2_{a_9\beta_1} + d^2_{a_{10}\beta_1} + d^2_{a_{11}\beta_1}$$
$$\frac{d^2 t_2}{d\beta_2^2} = d^2_{a_9\beta_2} + d^2_{a_{10}\beta_2} + d^2_{a_{11}\beta_2}$$
$$\frac{d^2 t_2}{d(\sigma_b^2)^2} = \frac{1}{2\sigma_b^4} - \frac{\mu_b^2}{\sigma_b^6} + d^2_{a_9\sigma_b^2} + d^2_{a_{10}\sigma_b^2} + d^2_{a_{11}\sigma_b^2}$$
$$\frac{d^2 t_2}{d(\sigma_e^2)^2} = \frac{1}{\sigma_e^4} - \frac{y_1^2}{\sigma_e^6} - \frac{(y_2 - \beta_0)^2}{\sigma_e^6} + d^2_{a_9\sigma_e^2} + d^2_{a_{10}\sigma_e^2} + d^2_{a_{11}\sigma_e^2}$$
$$\frac{d^2 t_2}{d\psi^2} = d^2_{a_9\psi} + d^2_{a_{10}\psi} + d^2_{a_{11}\psi} \tag{2.32}$$

The following expressions give the partial derivatives of t_1 and t_2 with respect to the model parameters.

$$\frac{d^2 t_1}{d\mu_b d\beta_0} = d^2_{a_5\mu_b\beta_0} + d^2_{a_6\mu_b\beta_0} + d^2_{a_7\mu_b\beta_0}$$
$$\frac{d^2 t_1}{d\mu_b d\beta_1} = d^2_{a_5\mu_b\beta_1} + d^2_{a_6\mu_b\beta_1} + d^2_{a_7\mu_b\beta_1}$$
$$\frac{d^2 t_1}{d\mu_b d\beta_2} = d^2_{a_5\mu_b\beta_2} + d^2_{a_6\mu_b\beta_2} + d^2_{a_7\mu_b\beta_2}$$
$$\frac{d^2 t_1}{d\mu_b d\sigma_b^2} = \frac{\mu_b}{\sigma_b^4} + d^2_{a_5\mu_b\sigma_b^2} + d^2_{a_6\mu_b\sigma_b^2} + d^2_{a_7\mu_b\sigma_b^2}$$
$$\frac{d^2 t_1}{d\mu_b d\sigma_e^2} = d^2_{a_5\mu_b\sigma_e^2} + d^2_{a_6\mu_b\sigma_e^2} + d^2_{a_7\mu_b\sigma_e^2}$$
$$\frac{d^2 t_1}{d\mu_b . d\psi} = d^2_{a_5\mu_b\psi} + d^2_{a_6\mu_b\psi} + d^2_{a_7\mu_b\psi}$$
$$\frac{d^2 t_1}{d\beta_0 d\beta_1} = d^2_{a_5\beta_0\beta_1} + d^2_{a_6\beta_0\beta_1} + d^2_{a_7\beta_0\beta_1}$$
$$\frac{d^2 t_1}{d\beta_0 d\beta_2} = \frac{\psi}{\sigma_e^2} + d^2_{a_5\beta_0\beta_2} + d^2_{a_6\beta_0\beta_2} + d^2_{a_7\beta_0\beta_2}$$
$$\frac{d^2 t_1}{d\beta_0 d\sigma_b^2} = d^2_{a_5\beta_0\sigma_b^2} + d^2_{a_6\beta_0\sigma_b^2} + d^2_{a_7\beta_0\sigma_b^2}$$

$$
\begin{aligned}
\frac{d^2t_1}{d\beta_0 d\sigma_e^2} &= -\frac{(y_1-\beta_0+\beta_2\psi)}{\sigma_e^4} + d^2_{a_5\beta_0\sigma_e^2} + d^2_{a_6\beta_0\sigma_e^2} + d^2_{a_7\beta_0\sigma_e^2} \\
\frac{d^2t_1}{\beta_0 d\psi} &= \frac{\beta_2}{\sigma_e^2} + d^2_{a_5\beta_0\psi} + d^2_{a_6\beta_0\psi} + d^2_{a_7\beta_0\psi} \\
\frac{d^2t_1}{d\beta_1 d\beta_2} &= d^2_{a_5\beta_1\beta_2} + d^2_{a_6\beta_1\beta_2} + d^2_{a_7\beta_1\beta_2} \\
\frac{d^2t_1}{d\beta_1 d\sigma_b^2} &= d^2_{a_5\beta_1\sigma_b^2} + d^2_{a_6\beta_1\sigma_b^2} + d^2_{a_7\beta_1\sigma_b^2} \\
\frac{d^2t_1}{d\beta_1 d\sigma_e^2} &= d^2_{a_5\beta_1\sigma_e^2} + d^2_{a_6\beta_1\sigma_e^2} + d^2_{a_7\beta_1\sigma_e^2} \\
\frac{d^2t_1}{d\beta_1 d\psi} &= d^2_{a_5\beta_1\psi} + d^2_{a_6\beta_1\psi} + d^2_{a_7\beta_1\psi} \\
\frac{d^2t_1}{d\beta_2 d\sigma_b^2} &= d^2_{a_5\beta_2\sigma_b^2} + d^2_{a_6\beta_2\sigma_b^2} + d^2_{a_7\beta_2\sigma_b^2} \\
\frac{d^2t_1}{d\beta_2 d\sigma_e^2} &= \frac{(y_1-\beta_0+\beta_2\psi)\psi}{\sigma_e^4} + d^2_{a_5\beta_2\sigma_e^2} + d^2_{a_6\beta_2\sigma_e^2} + d^2_{a_7\beta_2\sigma_e^2} \\
\frac{d^2t_1}{d\beta_2 d\psi} &= -\frac{(y_1-\beta_0+2\beta_2\psi)}{\sigma_e^2} - \frac{\beta_2\psi}{\sigma_e^4} + d^2_{a_5\beta_2\psi} + d^2_{a_6\beta_2\psi} + d^2_{a_7\beta_2\psi} \\
\frac{d^2t_1}{d\sigma_b^2 d\sigma_e^2} &= d^2_{a_5\sigma_b^2\sigma_e^2} + d^2_{a_6\sigma_b^2\sigma_e^2} + d^2_{a_7\sigma_b^2\sigma_e^2} \\
\frac{d^2t_1}{d\sigma_b^2 d\psi} &= d^2_{a_5\sigma_b^2\psi} + d^2_{a_6\sigma_b^2\psi} + d^2_{a_7\sigma_b^2\psi} \\
\frac{d^2t_1}{d\sigma_e^2 d\psi} &= \frac{(y_1-\beta_0+\beta_2\psi)\beta_2}{\sigma_e^4} + d^2_{a_5\sigma_e^2\psi} + d^2_{a_6\sigma_e^2\psi} + d^2_{a_7\sigma_e^2\psi}
\end{aligned}
\tag{2.33}
$$

$$
\begin{aligned}
\frac{d^2t_2}{d\mu_b d\beta_0} &= d^2_{a_9\mu_b\beta_0} + d^2_{a_{10}\mu_b\beta_0} + d^2_{a_{11}\mu_b\beta_0} \\
\frac{d^2t_2}{d\mu_b d\beta_1} &= d^2_{a_9\mu_b\beta_1} + d^2_{a_{10}\mu_b\beta_1} + d^2_{a_{11}\mu_b\beta_1} \\
\frac{d^2t_2}{d\mu_b d\beta_2} &= d^2_{a_9\mu_b\beta_2} + d^2_{a_{10}\mu_b\beta_2} + d^2_{a_{11}\mu_b\beta_2} \\
\frac{d^2t_2}{d\mu_b d\sigma_b^2} &= \frac{\mu_b}{\sigma_b^4} + d^2_{a_9\mu_b\sigma_b^2} + d^2_{a_{10}\mu_b\sigma_b^2} + d^2_{a_{11}\mu_b\sigma_b^2} \\
\frac{d^2t_2}{d\mu_b d\sigma_e^2} &= d^2_{a_9\mu_b\sigma_e^2} + d^2_{a_{10}\mu_b\sigma_e^2} + d^2_{a_{11}\mu_b\sigma_e^2} \\
\frac{d^2t_2}{d\mu_b d\psi} &= d^2_{a_9\mu_b\psi} + d^2_{a_{10}\mu_b\psi} + d^2_{a_{11}\mu_b\psi}
\end{aligned}
$$

$$
\begin{aligned}
\frac{d^2 t_2}{d\beta_0 d\beta_1} &= d^2_{a_9\beta_0\beta_1} + d^2_{a_{10}\beta_0\beta_1} + d^2_{a_{11}\beta_0\beta_1}\\
\frac{d^2 t_2}{d\beta_0 d\beta_2} &= d^2_{a_9\beta_0\beta_2} + d^2_{a_{10}\beta_0\beta_2} + d^2_{a_{11}\beta_0\beta_2}\\
\frac{d^2 t_2}{d\beta_0 d\sigma_b^2} &= d^2_{a_9\beta_0\sigma_b^2} + d^2_{a_{10}\beta_0\sigma_b^2} + d^2_{a_{11}\beta_0\sigma_b^2}\\
\frac{d^2 t_2}{d\beta_0 d\sigma_e^2} &= -\frac{(y_2-\beta_0)}{\sigma_e^4} + d^2_{a_9\beta_0\sigma_e^2} + d^2_{a_{10}\beta_0\sigma_e^2} + d^2_{a_{11}\beta_0\sigma_e^2}\\
\frac{d^2 t_2}{\beta_0 d\psi} &= d^2_{a_9\beta_0\psi} + d^2_{a_{10}\beta_0\psi} + d^2_{a_{11}\beta_0\psi}\\
\frac{d^2 t_2}{d\beta_1 d\beta_2} &= d^2_{a_9\beta_1\beta_2} + d^2_{a_{10}\beta_1\beta_2} + d^2_{a_{11}\beta_1\beta_2}\\
\frac{d^2 t_2}{d\beta_1 d\sigma_b^2} &= d^2_{a_9\beta_1\sigma_b^2} + d^2_{a_{10}\beta_1\sigma_b^2} + d^2_{a_{11}\beta_1\sigma_b^2}\\
\frac{d^2 t_2}{d\beta_1 d\sigma_e^2} &= d^2_{a_9\beta_1\sigma_e^2} + d^2_{a_{10}\beta_1\sigma_e^2} + d^2_{a_{11}\beta_1\sigma_e^2}\\
\frac{d^2 t_2}{d\beta_1 d\psi} &= d^2_{a_9\beta_1\psi} + d^2_{a_{10}\beta_1\psi} + d^2_{a_{11}\beta_1\psi}\\
\frac{d^2 t_2}{d\beta_2 d\sigma_b^2} &= d^2_{a_9\beta_2\sigma_b^2} + d^2_{a_{10}\beta_2\sigma_b^2} + d^2_{a_{11}\beta_2\sigma_b^2}\\
\frac{d^2 t_2}{d\beta_2 d\sigma_e^2} &= d^2_{a_9\beta_2\sigma_e^2} + d^2_{a_{10}\beta_2\sigma_e^2} + d^2_{a_{11}\beta_2\sigma_e^2}\\
\frac{d^2 t_2}{d\beta_2 d\psi} &= d^2_{a_9\beta_2\psi} + d^2_{a_{10}\beta_2\psi} + d^2_{a_{11}\beta_2\psi}\\
\frac{d^2 t_2}{d\sigma_b^2 d\sigma_e^2} &= d^2_{a_9\sigma_b^2\sigma_e^2} + d^2_{a_{10}\sigma_b^2\sigma_e^2} + d^2_{a_{11}\sigma_b^2\sigma_e^2}\\
\frac{d^2 t_2}{d\sigma_b^2 d\psi} &= d^2_{a_9\sigma_b^2\psi} + d^2_{a_{10}\sigma_b^2\psi} + d^2_{a11\sigma_b^2\psi}\\
\frac{d^2 t_2}{d\sigma_e^2 d\psi} &= d^2_{a_9\sigma_e^2\psi} + d^2_{a_{10}\sigma_e^2\psi} + d^2_{a_{11}\sigma_e^2\psi} \qquad (2.34)
\end{aligned}
$$

Here,

$$
\begin{aligned}
d^2_{a_5\gamma\alpha} = \frac{d^2 a_5}{d\gamma d\alpha} &= -\frac{1}{2}\left[\frac{1}{g_4(\beta_2,\sigma_e,\sigma_e)} d^2_{g_4(\beta_2,\sigma_e,\sigma_e)\gamma\alpha} - \frac{1}{g_4^2(\beta_2,\sigma_e,\sigma_e)} d_{g_4(\beta_2,\sigma_e,\sigma_e)\gamma} d_{g_4(\beta_2,\sigma_e,\sigma_e)\alpha}\right]\\
d^2_{a_6\gamma\alpha} = \frac{d^2 a_6}{d\gamma d\alpha} &= \frac{1}{2}\Bigg\{2\Bigg(\frac{1}{g_4(\beta_2,\sigma_e,\sigma_e)} d_{g_6(\beta_2,\sigma_e,\sigma_e)\gamma} d_{g_6(\beta_2,\sigma_e,\sigma_e)\alpha}\\
&\quad -\frac{g_6(\beta_2,\sigma_e,\sigma_e)}{g_4^2(\beta_2,\sigma_e,\sigma_e)} d_{g_6(\beta_2,\sigma_e,\sigma_e)\gamma} d_{g_4(\beta_2,\sigma_e,\sigma_e)\alpha} + \frac{g_6(\beta_2,\sigma_e,\sigma_e)}{g_4(\beta_2,\sigma_e,\sigma_e)} d^2_{g_6(\beta_2,\sigma_e,\sigma_e)\gamma\alpha}\Bigg)
\end{aligned}
$$

$$
- \left(\frac{2g_6(\beta_2, \sigma_e, \sigma_e)}{g_4^2(\beta_2, \sigma_e, \sigma_e)} d_{g_4(\beta_2,\sigma_e,\sigma_e)\gamma} d_{g_6(\beta_2,\sigma_e,\sigma_e)\alpha} + \frac{g_6^2(\beta_2, \sigma_e, \sigma_e)}{g_4^2(\beta_2, \sigma_e, \sigma_e)} d^2_{g_4(\beta_2,\sigma_e,\sigma_e)\gamma\alpha} \right.
$$

$$
\left. - \frac{2g_6^2(\beta_2, \sigma_e, \sigma_e)}{g_4^3(\beta_2, \sigma_e, \sigma_e)} d_{g_4(\beta_2,\sigma_e,\sigma_e)\gamma} d_{g_4(\beta_2,\sigma_e,\sigma_e)\alpha} \right) \Bigg\}
$$

$$
d^2_{a_7\gamma\alpha} = \frac{d^2 a_7}{d\gamma d\alpha} = \frac{1}{k_7} d^2_{k_7\gamma\alpha} - \frac{1}{k_7^2} d_{k_7\gamma} d_{k_7\alpha}
$$

$$
d^2_{a_9\gamma\alpha} = \frac{d^2 a_9}{d\gamma d\alpha} = -\frac{1}{2} \left[\frac{1}{g_4(0, \sigma_e, \sigma_e)} d^2_{g_4(0,\sigma_e,\sigma_e)\gamma\alpha} - \frac{1}{g_4^2(0, \sigma_e, \sigma_e)} d_{g_4(0,\sigma_e,\sigma_e)\gamma} d_{g_4(0,\sigma_e,\sigma_e)\alpha} \right]
$$

$$
d^2_{a_{10}\gamma\alpha} = \frac{d^2 a_{10}}{d\gamma d\alpha} = \frac{1}{2} \Bigg\{ 2 \left(\frac{1}{g_4(0, \sigma_e, \sigma_e)} d_{g_6(0,\sigma_e,\sigma_e)\gamma} d_{g_6(0,\sigma_e,\sigma_e)\alpha} \right.
$$

$$
\left. - \frac{g_6(0, \sigma_e, \sigma_e)}{g_4^2(0, \sigma_e, \sigma_e)} d_{g_6(0,\sigma_e,\sigma_e)\gamma} d_{g_4(0,\sigma_e,\sigma_e)\alpha} + \frac{g_6(0, \sigma_e, \sigma_e)}{g_4(0, \sigma_e, \sigma_e)} d^2_{g_6(0,\sigma_e,\sigma_e)\gamma\alpha} \right)
$$

$$
- \left(\frac{2g_6(0, \sigma_e, \sigma_e)}{g_4^2(0, \sigma_e, \sigma_e)} d_{g_4(0,\sigma_e,\sigma_e)\gamma} d_{g_6(0,\sigma_e,\sigma_e)\alpha} + \frac{g_6^2(0, \sigma_e, \sigma_e)}{g_4^2(0, \sigma_e, \sigma_e)} d^2_{g_4(0,\sigma_e,\sigma_e)\gamma\alpha} \right.
$$

$$
\left. - \frac{2g_6^2(0, \sigma_e, \sigma_e)}{g_4^3(0, \sigma_e, \sigma_e)} d_{g_4(0,\sigma_e,\sigma_e)\gamma} d_{g_4(0,\sigma_e,\sigma_e)\alpha} \right) \Bigg\}
$$

$$
d^2_{a_{11}\gamma\alpha} = \frac{d^2 a_{11}}{d\gamma d\alpha} = \frac{1}{k_{11}} d^2_{k_{11}\gamma\alpha} - \frac{1}{k_{11}^2} d_{k_{11}\gamma} d_{k_{11}\alpha}
$$

The elements of the Hessian matrix can be obtained by substituting the expressions from the (2.31), (2.32) , (2.33) and (2.34) in (2.27).

CHAPTER 3

A GENERALIZED SEGMENTED MEASUREMENT ERROR MODEL

3.1 Introduction

In this chapter, we consider a generalization of the segmented measurement error model (2.3) from Chapter 2 that does not assume equal error variances for the two methods. Thus, we study the model

$$Y_1 = b + e_1,\ Y_2 = \beta_0 + \beta_1 b + \beta_2 (b - \psi)_+ + e_2, \tag{3.1}$$

where

$$b \sim N(\mu_b, \sigma_b^2),\ e_1 \sim N(0, \sigma_{e1}^2),\ e_2 \sim N(0, \sigma_{e2}^2), \tag{3.2}$$

and b, e_1, and e_2 are mutually independent. The model (2.3) is a special case of this model when $\sigma_{e1}^2 = \sigma_{e2}^2 = \sigma_e^2$. Our goal in this chapter is to parallel the development in Chapter 2 to derive the distribution and estimation theory under the model (3.1). In this process, we also provide the proofs of Propositions 2.1 to 2.7 under the more general setup of this chapter.

3.2 Distribution theory

Proposition 3.1. *Consider* (Y_1, Y_2) *following the model* (3.1). *The mean and variance of* Y_1 *and* Y_2 *and their covariance are as follows:*

(a) $E(Y_1) = \mu_b$ *and* $var(Y_1) = \sigma_b^2 + \sigma_{e1}^2$,

(b) $E(Y_2) = \beta_0 + \beta_1 \mu_b + \beta_2 m_1$ *and* $var(Y_2) = \beta_1^2 \sigma_b^2 + \beta_2^2 m_3 + 2\beta_1 \beta_2 m_4 + \sigma_{e2}^2$,

(c) $cov(Y_1, Y_2) = \beta_1 \sigma_b^2 + \beta_2 m_4$,

where

$$
\begin{aligned}
m_1 &= E[(b-\psi)_+] = \{1-\Phi\left(g_1(\psi,\mu_b,\sigma_b)\right)\}\left(\mu_b-\psi\right)+\phi\left(g_1(\psi,\mu_b,\sigma_b)\right)\sigma_b,\\
m_2 &= E\left[\{(b-\psi)_+\}^2\right] = \{1-\Phi\left(g_1(\psi,\mu_b,\sigma_b)\right)\}\left\{(\mu_b-\psi)^2+\sigma_b^2\right\}\\
&\qquad\qquad + \left(g_1(\psi,\mu_b,\sigma_b)\sigma_b^2+2\mu_b\sigma_b-2\psi\sigma_b\right)\phi\left(g_1(\psi,\mu_b,\sigma_b)\right),\\
m_3 &= var\left[(b-\psi)_+\right] = m_2-m_1^2,\\
m_4 &= cov\left[b,(b-\psi)_+\right] = \{1-\Phi\left(g_1(\psi,\mu_b,\sigma_b)\right)\}\sigma_b^2\\
&\qquad\qquad + \left\{g_1(\psi,\mu_b,\sigma_b)\sigma_b^2+\mu_b\sigma_b-\psi\sigma_b\right\}\phi\left(g_1(\psi,\mu_b,\sigma_b)\right),
\end{aligned}
\tag{3.3}
$$

and the functions g_1, g_2, and g_3 are given by (1.1).

Proof. Let us first consider the expressions for the moments in (3.3). By definition,

$$
\begin{aligned}
m_1 &= E[(b-\psi)_+] = E[(b-\psi)I(b>\psi)] = E[bI(b>\psi)]-\psi E[I(b>\psi)],\\
m_2 &= E\left[\{(b-\psi)_+\}^2\right] = E[(b^2-2b\psi+\psi^2)I(b>\psi)]\\
&= E[b^2I(b>\psi)]-2\psi E[bI(b>\psi)]+\psi^2E[I(b>\psi)],\\
m_4 &= \mathrm{cov}\left[b,(b-\psi)_+\right] = E\left[b(b-\psi)_+\right]-E(b)E[(b-\psi)_+]\\
&= E[b^2I(b>\psi)]-\psi E[bI(b>\psi)]-\mu_bE[I(b>\psi)].
\end{aligned}
$$

Their expressions now follow from an application of (1.8) and (1.10) and simplification. The expression for m_3 is just the definition of variance. Next, to obtain the moments of Y_1 and Y_2, we use their definitions in (3.1) and apply (3.3). □

Proposition 3.2. *Consider (Y_1, Y_2) following the model* (3.1)*. The mean and variance of Y_1 and Y_2 and their covariance when b is truncated to be either $b \le \psi$ or $b > \psi$ are as follows:*

(a) $E(Y_1|b\le\psi) = E(b|b\le\psi)$ *and* $var(Y_1|b\le\psi) = var(b|b\le\psi)+\sigma_{e1}^2$*,*

(b) $E(Y_2|b\le\psi) = \beta_0+\beta_1E(b|b\le\psi)$ *and* $var(Y_2|b\le\psi) = \beta_1^2var(b|b\le\psi)+\sigma_{e2}^2$*,*

(c) $cov(Y_1, Y_2|b \le \psi) = \beta_1 var(b|b \le \psi)$,

(d) $E(Y_1|b > \psi) = E(b|b > \psi)$ *and* $var(Y_1|b > \psi) = var(b|b > \psi) + \sigma_{e1}^2$,

(e) $E(Y_2|b > \psi) = (\beta_0 - \beta_2\psi) + (\beta_1 + \beta_2)E(b|b > \psi)$ *and* $var(Y_2|b > \psi) = (\beta_1 + \beta_2)^2 var(b|b > \psi) + \sigma_{e2}^2$,

(f) $cov(Y_1, Y_2|b > \psi) = (\beta_1 + \beta_2) var(b|b > \psi)$,

where

$$\begin{aligned} E(b|b \le \psi) &= \mu_b - \sigma_b g_3(\psi, \mu_b, \sigma_b), \\ var(b|b \le \psi) &= \sigma_b^2 \left\{1 - g_1(\psi, \mu_b, \sigma_b) g_3(\psi, \mu_b, \sigma_b) - g_3^2(\psi, \mu_b, \sigma_b)\right\}. \\ E(b|b > \psi) &= \mu_b + \sigma_b g_2(\psi, \mu_b, \sigma_b), \\ var(b|b > \psi) &= \sigma_b^2 \left\{1 + g_1(\psi, \mu_b, \sigma_b) g_2(\psi, \mu_b, \sigma_b) - g_2^2(a_1, \mu_b, \sigma_b)\right\}, \end{aligned} \tag{3.4}$$

and the functions g_1*,* g_2*, and* g_3 *are given by* (1.1).

Proof. From the definition of model (3.1), we know that

$$Y_2 = \begin{cases} \beta_0 + \beta_1 b + e_2, & b \le \psi, \\ (\beta_0 - \beta_2\psi) + (\beta_1 + \beta_2) b + e_2, & b > \psi. \end{cases}$$

The expressions in (a)-(f) are now easily derived using the properties of mean, variance, and covariance operators. The mean and variance of truncated distribution b in (3.4) follow from moments of truncated normal random variables given in (1.6) and (1.7). □

Proposition 3.3. *The joint probability density function of* (Y_1, Y_2) *following the model* (3.1) *is*

$$f(y_1, y_2) = f_1(y_1, y_2) + f_2(y_1, y_2), \tag{3.5}$$

where

$$f_1(y_1,y_2) = \int_{-\infty}^{\psi} f(b,y_1,y_2)db = g_7(0, g_6(0,\sigma_{e1},\sigma_{e2}), g_4(0,\sigma_{e1},\sigma_{e2}),\sigma_{e1},\sigma_{e2})$$
$$\times \Phi\left(g_1(\psi, g_6(0,\sigma_{e1},\sigma_{e2}), \sqrt{g_4(0,\sigma_{e1},\sigma_{e2})})\right)$$
$$f_2(y_1,y_2) = \int_{\psi}^{\infty} f(b,y_1,y_2)db = g_7(\beta_2, g_6(\beta_2,\sigma_{e1},\sigma_{e2}), g_4(\beta_2,\sigma_{e1},\sigma_{e2}),\sigma_{e1},\sigma_{e2})$$
$$\times \left[1-\Phi\left(g_1\left(\psi, g_6(\beta_2,\sigma_{e1},\sigma_{e2}), \sqrt{g_4(\beta_2,\sigma_{e1},\sigma_{e2})}\right)\right)\right], \qquad (3.6)$$

and the functions g_1 *through* g_7 *are given by* (1.1).

Proof. To begin, we note that $Y_1|b \sim N(b,\sigma_{e1}^2)$, $Y_2|b \sim N(\beta_0+\beta_1 b,\sigma_{e2}^2)$ if $b \le \psi$, and $Y_2|b \sim N(\beta_0-\beta_2\psi+(\beta_1+\beta_2)b,\sigma_{e2}^2)$ if $b \ge \psi$. Next, we write

$$f(y_1,y_2) = f_1(y_1,y_2) + f_2(y_1,y_2)$$
$$\int_{-\infty}^{\psi} f(y_1|b)f(y_2|b)f(b)db + \int_{\psi}^{\infty} f(y_1|b)f(y_2|b)f(b)db. \qquad (3.7)$$

It now suffices to obtain the expressions in (3.6). For $f_1(y_1,y_2)$, we can write

$$\int_{-\infty}^{\psi} \frac{1}{\sqrt{2\pi\sigma_{e1}^2}} \exp\left\{\frac{-(y_1-b)^2}{2\sigma_{e1}^2}\right\} \frac{1}{\sqrt{2\pi\sigma_{e2}^2}} \exp\left\{\frac{-(y_2-\beta_0-\beta_1 b)^2}{2\sigma_{e2}^2}\right\}$$
$$\frac{1}{\sqrt{2\pi\sigma_b^2}} \exp\left\{\frac{-(b-\mu_b)^2}{2\sigma_b^2}\right\} db$$
$$= \frac{1}{(2\pi)^{3/2}\sqrt{\sigma_{e1}^2\sigma_{e2}^2\sigma_b^2}} \int_{-\infty}^{\psi} \exp\left\{-\frac{1}{2}\left(\frac{(y_1-b)^2}{\sigma_{e1}^2} + \frac{(y_2-\beta_0-\beta_1 b)^2}{\sigma_{e2}^2} + \frac{(b-\mu_b)^2}{\sigma_b^2}\right)\right\} db$$
$$= \frac{1}{(2\pi)^{3/2}\sqrt{\sigma_{e1}^2\sigma_{e2}^2\sigma_b^2}} \int_{-\infty}^{\psi} \exp\left\{-\frac{1}{2}\left(\frac{(y_1^2-2y_1b+b^2)}{\sigma_{e1}^2}\right.\right.$$
$$\left.\left.+\frac{((y_2-\beta_0)^2-2(y_2-\beta_0)\beta_1 b+\beta_1^2b^2)}{\sigma_{e2}^2} + \frac{(b^2-2\mu_b b+\mu_b^2)}{\sigma_b^2}\right)\right\} db$$
$$= \frac{1}{(2\pi)^{3/2}\sqrt{\sigma_{e1}^2\sigma_{e2}^2\sigma_b^2}} \exp\left\{-\frac{1}{2}\left(\frac{y_1^2}{\sigma_{e1}^2} + \frac{(y_2-\beta_0)^2}{\sigma_{e2}^2} + \frac{\mu_b^2}{\sigma_b^2}\right)\right\}$$
$$\int_{-\infty}^{\psi} \exp\left\{-\frac{1}{2}\left[\left(\frac{1}{\sigma_{e1}^2}+\frac{\beta_1^2}{\sigma_{e2}^2}+\frac{1}{\sigma_b^2}\right)b^2 - 2\left(\frac{y_1}{\sigma_{e1}^2}+\frac{(y_2-\beta_0)\beta_1}{\sigma_{e2}^2}+\frac{\mu_b}{\sigma_b^2}\right)b\right]\right\} db$$
$$= \frac{1}{\sqrt{\sigma_{e1}^2\sigma_{e2}^2\sigma_b^2}} \phi\left(\frac{y_1}{\sigma_{e1}}\right)\phi\left(\frac{y_2-\beta_0}{\sigma_{e2}}\right)\phi\left(\frac{\mu_b}{\sigma_b}\right)$$

$$\int_{-\infty}^{\psi} \exp\left\{-\frac{1}{2}\left[\frac{b^2}{g_4(0,\sigma_{e1},\sigma_{e2})} - 2g_5(0,\sigma_{e1},\sigma_{e2})b\right]\right\} db$$

$$= \frac{1}{\sqrt{\sigma_{e1}^2\sigma_{e2}^2\sigma_b^2}}\phi\left(\frac{y_1}{\sigma_{e1}}\right)\phi\left(\frac{y_2-\beta_0}{\sigma_{e2}}\right)\phi\left(\frac{\mu_b}{\sigma_b}\right)\int_{-\infty}^{\psi}\exp\left\{-\frac{1}{2g_4(0,\sigma_{e1},\sigma_{e2})}\left[b^2\right.\right.$$

$$\left.\left. -2g_4(0,\sigma_{e1},\sigma_{e2})g_5(0,\sigma_{e1},\sigma_{e2})b\right]\right\} db$$

$$= \frac{1}{\sqrt{\sigma_{e1}^2\sigma_{e2}^2\sigma_b^2}}\phi\left(\frac{y_1}{\sigma_{e1}}\right)\phi\left(\frac{y_2-\beta_0}{\sigma_{e2}}\right)\phi\left(\frac{\mu_b}{\sigma_b}\right)$$

$$\int_{-\infty}^{\psi}\exp\left\{-\frac{1}{2g_4(0,\sigma_{e1},\sigma_{e2})}\left[b^2 - 2g_6(0,\sigma_{e1},\sigma_{e2})b\right]\right\} db$$

$$= \frac{1}{\sqrt{\sigma_{e1}^2\sigma_{e2}^2\sigma_b^2}}\phi\left(\frac{y_1}{\sigma_{e1}}\right)\phi\left(\frac{y_2-\beta_0}{\sigma_{e2}}\right)\phi\left(\frac{\mu_b}{\sigma_b}\right)$$

$$\int_{-\infty}^{\psi}\exp\left\{-\frac{1}{2g_4(0,\sigma_{e1},\sigma_{e2})}\left[(b-g_6(0,\sigma_{e1},\sigma_{e2}))^2 - g_6^2(0,\sigma_{e1},\sigma_{e2})\right]\right\} db$$

$$= \frac{1}{\sqrt{\sigma_{e1}^2\sigma_{e2}^2\sigma_b^2}}\phi\left(\frac{y_1}{\sigma_{e1}}\right)\phi\left(\frac{y_2-\beta_0}{\sigma_{e2}}\right)\phi\left(\frac{\mu_b}{\sigma_b}\right)\exp\left\{\frac{g_6^2(0,\sigma_{e1},\sigma_{e2})}{2g_4(0,\sigma_{e1},\sigma_{e2})}\right\}$$

$$\frac{\sqrt{2\pi g_4(0,\sigma_{e1},\sigma_{e2})}}{\sqrt{2\pi g_4(0,\sigma_{e1},\sigma_{e2})}}\int_{-\infty}^{\psi}\exp\left\{-\frac{(b-g_6(0,\sigma_{e1},\sigma_{e2}))^2}{2g_4(0,\sigma_{e1},\sigma_{e2})}\right\} db$$

$$= \frac{g_7(0,g_6(0,\sigma_{e1},\sigma_{e2}),g_4(0,\sigma_{e1},\sigma_{e2}),\sigma_{e1},\sigma_{e2})}{\sqrt{2\pi g_4(0,\sigma_{e1},\sigma_{e2}))}}\int_{-\infty}^{\psi}\exp\left\{-\frac{(b-g_6(0,\sigma_{e1},\sigma_{e2}))^2}{2g_4(0,\sigma_{e1},\sigma_{e2})}\right\} db$$

$$= g_7(0,g_6(0,\sigma_{e1},\sigma_{e2}),g_4(0,\sigma_{e1},\sigma_{e2}),\sigma_{e1},\sigma_{e2})$$

$$\times\,\Phi\left(g_1(\psi,g_6(0,\sigma_{e1},\sigma_{e2}),\sqrt{g_4(0,\sigma_{e1},\sigma_{e2})})\right), \tag{3.8}$$

which is as required. For $f_2(y_1,y_2)$, we can write

$$\int_{\psi}^{\infty}\frac{1}{\sqrt{2\pi\sigma_{e1}^2}}\exp\left\{\frac{-(y_1-b)^2}{2\sigma_{e1}^2}\right\}\frac{1}{\sqrt{2\pi\sigma_{e2}^2}}\exp\left\{\frac{-\left(y_2-\beta_0-\beta_1 b-\beta_2(b-\psi)\right)^2}{2\sigma_{e2}^2}\right\}$$

$$\frac{1}{\sqrt{2\pi\sigma_b^2}}\exp\left\{\frac{-(b-\mu_b)^2}{2\sigma_b^2}\right\} db$$

$$= \frac{1}{(2\pi)^{3/2}\sqrt{\sigma_{e1}^2\sigma_{e2}^2\sigma_b^2}}\int_{\psi}^{\infty}\exp\left\{-\frac{1}{2}\left(\frac{(y_1-b)^2}{\sigma_{e1}^2}+\frac{(y_2-\beta_0-\beta_1 b-\beta_2(b-\psi))^2}{\sigma_{e2}^2}\right.\right.$$

$$\left.\left.+\frac{(b-\mu_b)^2}{\sigma_b^2}\right)\right\} db$$

$$= \frac{1}{(2\pi)^{3/2}\sqrt{\sigma_{e1}^2\sigma_{e2}^2\sigma_b^2}}\int_{\psi}^{\infty}\exp\left\{-\frac{1}{2}\left(\frac{(y_1^2-2y_1b+b^2)}{\sigma_{e1}^2}+\frac{(b^2-2\mu_b b+\mu_b^2)}{\sigma_b^2}\right.\right.$$

$$
+\frac{((y_2-\beta_0+\beta_2\psi)^2-2(\beta_1+\beta_2)(y_2-\beta_0+\beta_2\psi)b+(\beta_1+\beta_2)^2b^2)}{\sigma_{e2}^2}\Bigg)\Bigg\}\,db
$$

$$
=\frac{1}{(2\pi)^{3/2}\sqrt{\sigma_{e1}^2\sigma_{e2}^2\sigma_b^2}}\exp\left\{-\frac{1}{2}\left(\frac{y_1^2}{\sigma_{e1}^2}+\frac{(y_2-\beta_0+\beta_2\psi)^2}{\sigma_{e2}^2}+\frac{\mu_b^2}{\sigma_b^2}\right)\right\}
$$

$$
\int_\psi^\infty \exp\Bigg\{-\frac{1}{2}\Bigg[\left(\frac{1}{\sigma_{e1}^2}+\frac{(\beta_1+\beta_2)^2}{\sigma_{e2}^2}+\frac{1}{\sigma_b^2}\right)b^2
$$

$$
-2\left(\frac{y_1}{\sigma_{e1}^2}+\frac{(y_2-\beta_0+\beta_2\psi)(\beta_1+\beta_2)}{\sigma_{e2}^2}+\frac{\mu_b}{\sigma_b^2}\right)b\Bigg]\Bigg\}\,db
$$

$$
=\frac{1}{\sqrt{\sigma_{e1}^2\sigma_{e2}^2\sigma_b^2}}\phi\left(\frac{y_1}{\sigma_{e1}}\right)\phi\left(\frac{y_2-\beta_0+\beta_2\psi}{\sigma_{e2}}\right)\phi\left(\frac{\mu_b}{\sigma_b}\right)\int_\psi^\infty \exp\Bigg\{-\frac{1}{2}\Bigg[\frac{b^2}{g_4(\beta_2,\sigma_{e1},\sigma_{e2})}
$$

$$
-2\left(\frac{y_1}{\sigma_{e1}^2}+\frac{(y_2-\beta_0+\beta_2\psi)(\beta_1+\beta_2)}{\sigma_{e2}^2}+\frac{\mu_b}{\sigma_b^2}\right)b\Bigg]\Bigg\}\,db
$$

$$
=\frac{1}{\sqrt{\sigma_{e1}^2\sigma_{e2}^2\sigma_b^2}}\phi\left(\frac{y_1}{\sigma_{e1}}\right)\phi\left(\frac{y_2-\beta_0+\beta_2\psi}{\sigma_{e2}}\right)\phi\left(\frac{\mu_b}{\sigma_b}\right)\int_\psi^\infty \exp\Bigg\{-\frac{1}{2g_4(\beta_2,\sigma_{e1},\sigma_{e2})}\Big[b^2
$$

$$
-2g_4(\beta_2,\sigma_{e1},\sigma_{e2})g_5(\beta_2,\sigma_{e1},\sigma_{e2})b\Big]\Bigg\}\,db
$$

$$
=\frac{1}{\sqrt{\sigma_{e1}^2\sigma_{e2}^2\sigma_b^2}}\phi\left(\frac{y_1}{\sigma_{e1}}\right)\phi\left(\frac{y_2-\beta_0+\beta_2\psi}{\sigma_{e2}}\right)\phi\left(\frac{\mu_b}{\sigma_b}\right)
$$

$$
\int_\psi^\infty \exp\left\{-\frac{1}{2g_4(\beta_2,\sigma_{e1},\sigma_{e2})}\left[b^2-2g_6(\beta_2,\sigma_{e1},\sigma_{e2})b\right]\right\}\,db
$$

$$
=\frac{1}{\sqrt{\sigma_{e1}^2\sigma_{e2}^2\sigma_b^2}}\phi\left(\frac{y_1}{\sigma_{e1}}\right)\phi\left(\frac{y_2-\beta_0+\beta_2\psi}{\sigma_{e2}}\right)\phi\left(\frac{\mu_b}{\sigma_b}\right)
$$

$$
\int_\psi^\infty \exp\left\{-\frac{1}{2g_4(\beta_2,\sigma_{e1},\sigma_{e2})}\left[(b-g_6(\beta_2,\sigma_{e1},\sigma_{e2}))^2-g_6^2(\beta_2,\sigma_{e1},\sigma_{e2})\right]\right\}\,db
$$

$$
=\frac{1}{\sqrt{\sigma_{e1}^2\sigma_{e2}^2\sigma_b^2}}\phi\left(\frac{y_1}{\sigma_{e1}}\right)\phi\left(\frac{y_2-\beta_0+\beta_2\psi}{\sigma_{e2}}\right)\phi\left(\frac{\mu_b}{\sigma_b}\right)\exp\left\{\frac{g_6^2(\beta_2,\sigma_{e1},\sigma_{e2})}{2g_4(\beta_2,\sigma_{e1},\sigma_{e2})}\right\}
$$

$$
\frac{\sqrt{2\pi g_4(\beta_2,\sigma_{e1},\sigma_{e2})}}{\sqrt{2\pi g_4(\beta_2,\sigma_{e1},\sigma_{e2})}}\int_\psi^\infty \exp\left\{-\frac{(b-g_6(\beta_2,\sigma_{e1},\sigma_{e2}))^2}{2g_4(\beta_2,\sigma_{e1},\sigma_{e2})}\right\}\,db
$$

$$
=\frac{g_7(\beta_2,g_6(\beta_2,\sigma_{e1},\sigma_{e2}),g_4(\beta_2,\sigma_{e1},\sigma_{e2}),\sigma_{e1},\sigma_{e2})}{\sqrt{2\pi g_4(\beta_2,\sigma_{e1},\sigma_{e2})})}\int_\psi^\infty \exp\left\{-\frac{(b-g_6(\beta_2,\sigma_{e1},\sigma_{e2}))^2}{2g_4(\beta_2,\sigma_{e1},\sigma_{e2})}\right\}\,db
$$

$$
=g_7(\beta_2,g_6(\beta_2,\sigma_{e1},\sigma_{e2}),g_4(\beta_2,\sigma_{e1},\sigma_{e2}),\sigma_{e1},\sigma_{e2})
$$

$$
\times\left[1-\Phi\left(g_1\left(\psi,g_6(\beta_2,\sigma_{e1},\sigma_{e2}),\sqrt{g_4(\beta_2,\sigma_{e1},\sigma_{e2})}\right)\right)\right], \tag{3.9}
$$

which is also as required, completing the proof. □

Proposition 3.4. *Consider* (Y_1, Y_2) *following the model* (3.1). *The probability density function of* $D = Y_1 - Y_2$ *is*

$$h(d) = h_1(d) + h_2(d), \tag{3.10}$$

where

$$\begin{aligned} h_1(d) &= \int_{-\infty}^{\psi} f(d,b)db = \frac{\sqrt{2\pi c_1}}{\sqrt{\sigma^2\sigma_b^2}}\phi\left(\frac{d-\beta_0}{\sigma}\right)\phi\left(\frac{\mu_b}{\sigma_b}\right)\exp\left(\frac{c_2^2}{2c_1}\right)\Phi(g_1(\psi, c_2, \sqrt{c_1})), \\ h_2(d) &= \int_{\psi}^{\infty} f(d,b)db = \frac{\sqrt{2\pi c_3}}{\sqrt{\sigma^2\sigma_b^2}}\phi\left(\frac{d-\beta_0+\beta_2\psi}{\sigma}\right)\phi\left(\frac{\mu_b}{\sigma_b}\right)\exp\left(\frac{c_4^2}{2c_3}\right) \\ &\qquad \times\left[1-\Phi(g_1(\psi, c_4, \sqrt{c_3}))\right], \end{aligned} \tag{3.11}$$

and

$$\begin{aligned} \sigma^2 &= \sigma_{e1}^2 + \sigma_{e2}^2, \\ c_1 &= \left(\frac{(\beta_1-1)^2}{\sigma^2} + \frac{1}{\sigma_b^2}\right)^{-1}, \\ c_2 &= \left(\frac{(d-\beta_0)(\beta_1-1)}{\sigma^2} + \frac{\mu_b}{\sigma_b^2}\right)c_1 \\ c_3 &= \left(\frac{(\beta_1+\beta_2-1)^2}{\sigma^2} + \frac{1}{\sigma_b^2}\right)^{-1} \\ c_4 &= \left(\frac{(d-\beta_0+\beta_2\psi)(\beta_1+\beta_2-1)}{\sigma^2} + \frac{\mu_b}{\sigma_b^2}\right)c_3. \end{aligned} \tag{3.12}$$

Proof. By definition,

$$D|b \sim \begin{cases} N(\beta_0 + \beta_1 b + \beta_2(b-\psi) - b, \sigma^2), & b \geq \psi \\ N(\beta_0 + \beta_1 b - b, \sigma^2), & b \leq \psi. \end{cases}$$

To find the density of D, we can write

$$h(d) = \int_{-\infty}^{\infty} f(d,b)db = h_1(d) + h_2(d), \tag{3.13}$$

where $h_1(d) = \int_{-\infty}^{\psi} f(d|b)f(b)db$ and $h_2(d) = \int_{\psi}^{\infty} f(d|b)f(b)db$. It suffices now to obtain the expressions for $h_1(d)$ and $h_2(d)$ given in (3.11). For $h_1(d)$, we can write:

$$\int_{-\infty}^{\psi} f(d|b)f(b)db$$

$$= \int_{-\infty}^{\psi} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{[d-(\beta_0+\beta_1 b - b)]^2}{2\sigma^2}\right\} \frac{1}{\sqrt{2\pi\sigma_b^2}} \exp\left\{-\frac{(b-\mu_b)^2}{2\sigma_b^2}\right\} db$$

$$= \frac{1}{2\pi\sqrt{\sigma^2\sigma_b^2}} \int_{-\infty}^{\psi} \exp\left\{-\frac{1}{2}\left(\frac{[(d-\beta_0)-(\beta_1-1)b]^2}{\sigma^2} + \frac{(b-\mu_b)^2}{\sigma_b^2}\right)\right\} db$$

$$= \frac{1}{2\pi\sqrt{\sigma^2\sigma_b^2}} \int_{-\infty}^{\psi} \exp\left\{-\frac{1}{2}\frac{[(d-\beta_0)^2 - 2(d-\beta_0)(\beta_1-1)b + (\beta_1-1)^2 b^2]}{\sigma^2}\right.$$

$$\left. + \frac{(b^2 - 2\mu_b b + \mu_b^2)}{\sigma_b^2}\right\} db$$

$$= \frac{1}{2\pi\sqrt{\sigma^2\sigma_b^2}} \exp\left\{-\frac{(d-\beta_0)^2}{2\sigma^2} - \frac{\mu_b^2}{2\sigma_b^2}\right\}$$

$$\int_{-\infty}^{\psi} \exp\left\{-\frac{1}{2}\left[-2\left(\frac{(d-\beta_0)(\beta_1-1)}{\sigma^2} + \frac{\mu_b}{\sigma_b^2}\right) b + \left(\frac{(\beta_1-1)^2}{\sigma^2} + \frac{1}{\sigma_b^2}\right) b^2\right]\right\}$$

$$= \frac{1}{\sqrt{\sigma^2\sigma_b^2}} \phi\left(\frac{d-\beta_0}{\sigma}\right) \phi\left(\frac{\mu_b}{\sigma_b}\right)$$

$$\int_{-\infty}^{\psi} \exp\left\{-\frac{\left(\frac{(\beta_1-1)^2}{\sigma^2} + \frac{1}{\sigma_b^2}\right)}{2}\left[b^2 - 2\frac{\left(\frac{(d-\beta_0)(\beta_1-1)}{\sigma^2} + \frac{\mu_b}{\sigma_b^2}\right)}{\left(\frac{(\beta_1-1)^2}{\sigma^2} + \frac{1}{\sigma_b^2}\right)} b\right]\right\} db$$

$$= \frac{1}{\sqrt{\sigma^2\sigma_b^2}} \phi\left(\frac{d-\beta_0}{\sigma}\right) \phi\left(\frac{\mu_b}{\sigma_b}\right) \int_{-\infty}^{\psi} \exp\left\{-\frac{1}{2c_1}\left(b^2 - 2c_2 b + c_2^2 - c_2^2\right)\right\} db$$

$$= \frac{1}{\sqrt{\sigma^2\sigma_b^2}} \phi\left(\frac{d-\beta_0}{\sigma}\right) \phi\left(\frac{\mu_b}{\sigma_b}\right) \int_{-\infty}^{\psi} \exp\left\{-\frac{1}{2c_1}\left((b-c_2)^2 - c_2^2\right)\right\} db$$

$$= \frac{1}{\sqrt{\sigma^2\sigma_b^2}} \phi\left(\frac{d-\beta_0}{\sigma}\right) \phi\left(\frac{\mu_b}{\sigma_b}\right) \exp\left(\frac{c_2^2}{2c_1}\right) \int_{-\infty}^{\psi} \exp\left\{-\frac{1}{2c_1}\left((b-c_2)^2\right)\right\} db,$$

$$= \frac{\sqrt{2\pi c_1}}{\sqrt{\sigma^2\sigma_b^2}} \phi\left(\frac{d-\beta_0}{\sigma}\right) \phi\left(\frac{\mu_b}{\sigma_b}\right) \exp\left(\frac{c_2^2}{2c_1}\right) \Phi(g_1(\psi, c_2, \sqrt{c_1})), \tag{3.14}$$

which is the expression for $h_1(d)$ given in (3.11). For $h_2(d)$, we can write as before:

$$\int_{-\infty}^{\psi} f(d|b)f(b)db$$

$$= \int_{\psi}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{[d-(\beta_0+\beta_1 b+\beta_2 b-\beta_2\psi-b)]^2}{2\sigma^2}\right\}$$

$$\frac{1}{\sqrt{2\pi\sigma_b^2}} \exp\left\{-\frac{(b-\mu_b)^2}{2\sigma_b^2}\right\} db$$

$$= \frac{1}{2\pi\sqrt{\sigma^2\sigma_b^2}} \int_\psi^\infty \exp\left\{-\frac{1}{2}\left(\frac{[(d-\beta_0+\beta_2\psi)-(\beta_1+\beta_2-1)b]^2}{\sigma^2} + \frac{(b-\mu_b)^2}{\sigma_b^2}\right)\right\} db$$

$$= \frac{1}{2\pi\sqrt{\sigma^2\sigma_b^2}} \int_\psi^\infty \exp\left\{-\frac{1}{2}\left(\frac{1}{\sigma^2}\left[(d-\beta_0+\beta_2\psi)^2 - 2(d-\beta_0+\beta_2\psi)(\beta_1+\beta_2-1)b\right.\right.\right.$$

$$\left.\left.\left. +(\beta_1+\beta_2-1)^2\right] + \frac{(b^2-2\mu_b b+\mu_b^2)}{\sigma_b^2}\right)\right\} db$$

$$= \frac{1}{2\pi\sqrt{\sigma^2\sigma_b^2}} \exp\left\{-\frac{(d-\beta_0+\beta_2\psi)^2}{2\sigma^2} - \frac{\mu_b^2}{2\sigma_b^2}\right\}$$

$$\int_\psi^\infty \exp\left\{-\frac{1}{2}\left[-2\left(\frac{(d-\beta_0+\beta_2\psi)(\beta_1+\beta_2-1)}{\sigma^2} + \frac{\mu_b}{\sigma_b^2}\right)b + \left(\frac{(\beta_1+\beta_2-1)^2}{\sigma^2} + \frac{1}{\sigma_b^2}\right)b^2\right]\right\}$$

$$= \frac{1}{\sqrt{\sigma^2\sigma_b^2}}\phi\left(\frac{d-\beta_0+\beta_2\psi}{\sigma}\right)\phi\left(\frac{\mu_b}{\sigma_b}\right)$$

$$\int_\psi^\infty \exp\left\{-\frac{\left(\frac{(\beta_1+\beta_2-1)^2}{\sigma^2}+\frac{1}{\sigma_b^2}\right)}{2}\left[b^2 - 2\frac{\left(\frac{(d-\beta_0+\beta_2\psi)(\beta_1+\beta_2-1)}{\sigma^2}+\frac{\mu_b}{\sigma_b^2}\right)}{\left(\frac{(\beta_1+\beta_2-1)^2}{\sigma^2}+\frac{1}{\sigma_b^2}\right)}b\right]\right\} db$$

$$= \frac{1}{\sqrt{\sigma^2\sigma_b^2}}\phi\left(\frac{d-\beta_0+\beta_2\psi}{\sigma}\right)\phi\left(\frac{\mu_b}{\sigma_b}\right)\int_\psi^\infty \exp\left\{-\frac{1}{2c_3}\left((b-c_4)^2-c_4^2\right)\right\} db$$

$$= \frac{\sqrt{2\pi c_3}}{\sqrt{\sigma^2\sigma_b^2}}\phi\left(\frac{d-\beta_0+\beta_2\psi}{\sigma}\right)\phi\left(\frac{\mu_b}{\sigma_b}\right)\exp\left(\frac{c_4^2}{2c_3}\right)\left[1-\Phi(g_1(\psi, c_4, \sqrt{c_3}))\right], \tag{3.15}$$

which is the expression for $h_2(d)$ given in (3.11). □

Proposition 3.5. *Consider* (Y_1, Y_2) *following the model* (3.1). *The probability density function of* $D = Y_1 - Y_2$ *when* b *is truncated to be either* $b \le \psi$ *or* $b > \psi$ *are as follows:*

$$h(d|b \le \psi) = \frac{h_1(d)}{\Phi(g_1(\psi, \mu_b, \sigma_b))},\ h(d|b > \psi) = \frac{h_2(d)}{1-\Phi(g_1(\psi, \mu_b, \sigma_b))}, \tag{3.16}$$

where $h_1(d)$ *and* $h_2(d)$ *are given by* (3.11) *in Proposition 3.4.*

Proof. By definition,

$$h(d|b \le \psi) = \int_{-\infty}^{\psi} f(d,b)db/P(b \le \psi) = h_1(d)/\Phi(g_1(\psi, \mu_b, \sigma_b)).$$

Similarly,

$$h(d|b > \psi) = \int_{\psi}^{\infty} f(d,b)db/P(b > \psi) = h_1(d)/\{1 - \Phi(g_1(\psi, \mu_b, \sigma_b))\},$$

completing the proof. □

Proposition 3.6. *Consider the model* (3.1). *The first and second moments of* b *and* $(b-\psi)_+$ *conditional on* $(Y_1, Y_2) = (y_1, y_2)$ *are as follows:*

(a) $E[b|y_1, y_2] = (A_1 + A_2)/f(y_1, y_2)$,

(b) $E[b^2|y_1, y_2] = (A_3 + A_4)/f(y_1, y_2)$,

(c) $E\left[(b-\psi)_+|y_1, y_2\right] = (A_2 - \psi f_2(y_1, y_2))/f(y_1, y_2)$,

(d) $E\left[\{(b-\psi)_+\}^2|y_1, y_2\right] = (A_4 - 2\psi A_2 + \psi^2 f_2(y_1, y_2))/f(y_1, y_2)$,

(e) $E\left[b(b-\psi)_+|y_1, y_2\right] = (A_4 - \psi A_2)/f(y_1, y_2)$,

where

$$\begin{aligned}
A_1 &= \int_{-\infty}^{\psi} bf(b, y_1, y_2)db = f_1(y_1, y_2)\Big\{g_6(0, \sigma_{e1}, \sigma_{e2}) - \sqrt{g_4(0, \sigma_{e1}, \sigma_{e2})} \\
&\qquad \times g_3\left(\psi, g_6(0, \sigma_{e1}, \sigma_{e2}), \sqrt{g_4(0, \sigma_{e1}, \sigma_{e2})}\right)\Big\} \\
A_2 &= \int_{\psi}^{\infty} bf(b, y_1, y_2)db = f_2(y_1, y_2)\Big\{g_6(\beta_2, \sigma_{e1}, \sigma_{e2}) + \sqrt{g_4(\beta_2, \sigma_{e1}, \sigma_{e2})} \\
&\qquad \times g_2\left(\psi, g_6(\beta_2, \sigma_{e1}, \sigma_{e2}), \sqrt{g_4(\beta_2, \sigma_{e1}, \sigma_{e2})}\right)\Big\} \\
A_3 &= \int_{-\infty}^{\psi} b^2 f(b, y_1, y_2)db = f_1(y_1, y_2)\big\{g_4(0, \sigma_{e1}, \sigma_{e2}) + g_6^2(0, \sigma_{e1}, \sigma_{e2}) \\
&\qquad - g_1(\psi, g_6(0, \sigma_{e1}, \sigma_{e2}), \sqrt{g_4(0, \sigma_{e1}, \sigma_{e2})}) \\
&\qquad \times g_4(0, \sigma_{e1}, \sigma_{e2}) g_3\left(\psi, g_6(0, \sigma_{e1}, \sigma_{e2}), \sqrt{g_4(0, \sigma_{e1}, \sigma_{e2})}\right) \\
&\qquad - 2\sqrt{g_4(0, \sigma_{e1}, \sigma_{e2})} g_6(0, \sigma_{e1}, \sigma_{e2}) g_3\left(\psi, g_6(0, \sigma_{e1}, \sigma_{e2}), \sqrt{g_4(0, \sigma_{e1}, \sigma_{e2})}\right)\Big\} \\
A_4 &= \int_{\psi}^{\infty} b^2 f(b, y_1, y_2)db = f_2(y_1, y_2)\big\{g_4(\beta_2, \sigma_{e1}, \sigma_{e2}) + g_6^2(\beta_2, \sigma_{e1}, \sigma_{e2})
\end{aligned}$$

$$
\begin{aligned}
&+g_1(\psi, g_6(\beta_2,\sigma_{e1},\sigma_{e2}), \sqrt{g_4(\beta_2,\sigma_{e1},\sigma_{e2})}) \\
&\times g_4(\beta_2,\sigma_{e1},\sigma_{e2}) g_2\left(\psi, g_6(\beta_2,\sigma_{e1},\sigma_{e2}), \sqrt{g_4(\beta_2,\sigma_{e1},\sigma_{e2})}\right) \\
&+2\sqrt{g_4(\beta_2,\sigma_{e1},\sigma_{e2})} g_6(\beta_2,\sigma_{e1},\sigma_{e2}) g_2\left(\psi, g_6(\beta_2,\sigma_{e1},\sigma_{e2}), \sqrt{g_4(\beta_2,\sigma_{e1},\sigma_{e2})}\right)\Bigg\},
\end{aligned}
\tag{3.17}
$$

and f_1, f_2, and f are given by Proposition 3.3.

Proof. Let us first obtain the integrals in (3.17). Upon proceeding as in the proof of Proposition 3.3, we can write

$$
\begin{aligned}
A_1 &= \int_{-\infty}^{\psi} b f(b, y_1, y_2) db \\
&= \frac{g_7(0, g_6(0,\sigma_{e1},\sigma_{e2}), g_4(0,\sigma_{e1},\sigma_{e2}), \sigma_{e1}, \sigma_{e2})}{\sqrt{2\pi g_4(0,\sigma_{e1},\sigma_{e2})})} \int_{-\infty}^{\psi} b \exp\left\{-\frac{(b - g_6(0,\sigma_{e1},\sigma_{e2}))^2}{2g_4(0,\sigma_{e1},\sigma_{e2})}\right\} db \\
&= \frac{g_7(0, g_6(0,\sigma_{e1},\sigma_{e2}), g_4(0,\sigma_{e1},\sigma_{e2}), \sigma_{e1}, \sigma_{e2}) \Phi\left(g_1(\psi, g_6(0,\sigma_{e1},\sigma_{e2}), \sqrt{g_4(0,\sigma_{e1},\sigma_{e2})})\right)}{\sqrt{2\pi g_4(0,\sigma_{e1},\sigma_{e2})})} \\
&\quad \int_{-\infty}^{\psi} b \frac{\exp\left\{-\frac{(b-g_6(0,\sigma_{e1},\sigma_{e2}))^2}{2g_4(0,\sigma_{e1},\sigma_{e2})}\right\}}{\Phi\left(g_1(\psi, g_6(0,\sigma_{e1},\sigma_{e2}), \sqrt{g_4(0,\sigma_{e1},\sigma_{e2})})\right)} \\
&= f_1(y_1,y_2)\left\{g_6(0,\sigma_{e1},\sigma_{e2}) - \sqrt{g_4(0,\sigma_{e1},\sigma_{e2})} g_3\left(\psi, g_6(0,\sigma_{e1},\sigma_{e2}), \sqrt{g_4(0,\sigma_{e1},\sigma_{e2})}\right)\right\},
\end{aligned}
$$

where the last equality follows from (1.7) and (3.8). Likewise,

$$
\begin{aligned}
A_2 &= \int_{\psi}^{\infty} b f(b, y_1, y_2) db \\
&= \frac{g_7(\beta_2, g_6(\beta_2,\sigma_{e1},\sigma_{e2}), g_4(\beta_2,\sigma_{e1},\sigma_{e2}), \sigma_{e1}, \sigma_{e2})}{\sqrt{2\pi g_4(\beta_2,\sigma_{e1},\sigma_{e2})})} \\
&\quad \int_{\psi}^{\infty} b \exp\left\{-\frac{(b - g_6(\beta_2,\sigma_{e1},\sigma_{e2}))^2}{2g_4(\beta_2,\sigma_{e1},\sigma_{e2})}\right\} db \\
&= \frac{\begin{array}{c} g_7(\beta_2, g_6(\beta_2,\sigma_{e1},\sigma_{e2}), g_4(\beta_2,\sigma_{e1},\sigma_{e2}), \sigma_{e1}, \sigma_{e2}) \\ \left[1 - \Phi\left(g_1(\psi, g_6(\beta_2,\sigma_{e1},\sigma_{e2}), \sqrt{g_4(\beta_2,\sigma_{e1},\sigma_{e2})})\right)\right] \end{array}}{\sqrt{2\pi g_4(\beta_2,\sigma_{e1},\sigma_{e2})})}
\end{aligned}
$$

$$\int_{\psi}^{\infty} b \frac{\exp\left\{-\frac{(b-g_6(\beta_2,\sigma_{e1},\sigma_{e2}))^2}{2g_4(\beta_2,\sigma_{e1},\sigma_{e2})}\right\}}{\left[1-\Phi\left(g_1(\psi, g_6(\beta_2,\sigma_{e1},\sigma_{e2}), \sqrt{g_4(\beta_2,\sigma_{e1},\sigma_{e2})})\right)\right]} db$$

$$= f_2(y_1,y_2)\left\{g_6(\beta_2,\sigma_{e1},\sigma_{e2}) + \sqrt{g_4(\beta_2,\sigma_{e1},\sigma_{e2})}\, g_2\left(\psi, g_6(\beta_2,\sigma_{e1},\sigma_{e2}), \sqrt{g_4(\beta_2,\sigma_{e1},\sigma_{e2})}\right)\right\},$$

where the last equality follows from (1.6) and (3.9). In a similar manner, upon using (1.10) and (1.11), we can write

$$\begin{aligned} A_3 &= \int_{-\infty}^{\psi} b^2 f(b,y_1,y_2) db \\ &= \frac{g_7(0, g_6(0,\sigma_{e1},\sigma_{e2}), g_4(0,\sigma_{e1},\sigma_{e2}), \sigma_{e1},\sigma_{e2})}{\sqrt{2\pi g_4(0,\sigma_{e1},\sigma_{e2})}} \int_{-\infty}^{\psi} b^2 \exp\left\{-\frac{(b-g_6(0,\sigma_{e1},\sigma_{e2}))^2}{2g_4(0,\sigma_{e1},\sigma_{e2})}\right\} db \\ &= \frac{g_7(0, g_4(0,\sigma_{e1},\sigma_{e2}), g_6(0,\sigma_{e1},\sigma_{e2}), \sigma_{e1},\sigma_{e2})\Phi\left(g_1(\psi, g_6(0,\sigma_{e1},\sigma_{e2}), \sqrt{g_4(0,\sigma_{e1},\sigma_{e2})})\right)}{\sqrt{2\pi g_4(0,\sigma_{e1},\sigma_{e2})})} \\ &\quad \int_{-\infty}^{\psi} b^2 \frac{\exp\left\{-\frac{(b-g_6(0,\sigma_{e1},\sigma_{e2}))^2}{2g_4(0,\sigma_{e1},\sigma_{e2})}\right\}}{\Phi\left(g_1(\psi, g_6(0,\sigma_{e1},\sigma_{e2}), \sqrt{g_4(0,\sigma_{e1},\sigma_{e2})})\right)} \\ &= f_1(y_1,y_2)\left\{g_4(0,\sigma_{e1},\sigma_{e2}) - g_1(\psi, g_6(0,\sigma_{e1},\sigma_{e2}), \sqrt{g_4(0,\sigma_{e1},\sigma_{e2})}) g_4(0,\sigma_{e1},\sigma_{e2}) \right. \\ &\quad g_3\left(\psi, g_6(0,\sigma_{e1},\sigma_{e2}), \sqrt{g_4(0,\sigma_{e1},\sigma_{e2})}\right) + g_6^2(0,\sigma_{e1},\sigma_{e2}) \\ &\quad \left. -2\sqrt{g_4(0,\sigma_{e1},\sigma_{e2})}\, g_6(0,\sigma_{e1},\sigma_{e2}) g_3\left(\psi, g_6(0,\sigma_{e1},\sigma_{e2}), \sqrt{g_4(0,\sigma_{e1},\sigma_{e2})}\right)\right\}, \end{aligned}$$

and

$$\begin{aligned} A_4 &= \int_{\psi}^{\infty} b^2 f(b,y_1,y_2) db \\ &= \frac{g_7(\beta_2, g_6(\beta_2,\sigma_{e1},\sigma_{e2}), g_4(\beta_2,\sigma_{e1},\sigma_{e2}), \sigma_{e1},\sigma_{e2})}{\sqrt{2\pi g_4(\beta_2,\sigma_{e1},\sigma_{e2})})} \\ &\quad \int_{-\infty}^{\psi} b^2 \exp\left\{-\frac{(b-g_6(\beta_2,\sigma_{e1},\sigma_{e2}))^2}{2g_4(\beta_2,\sigma_{e1},\sigma_{e2})}\right\} db \\ &= \frac{g_7(\beta_2, g_6(\beta_2,\sigma_{e1},\sigma_{e2}), g_4(\beta_2,\sigma_{e1},\sigma_{e2}), \sigma_{e1},\sigma_{e2})}{\sqrt{2\pi g_4(\beta_2,\sigma_{e1},\sigma_{e2})})} \\ &\qquad \times \left[1-\Phi\left(g_1(\psi, g_6(\beta_2,\sigma_{e1},\sigma_{e2}), \sqrt{g_4(\beta_2,\sigma_{e1},\sigma_{e2})})\right)\right] \end{aligned}$$

$$\times \int_{-\infty}^{\psi} b^2 \frac{\exp\left\{-\frac{(b-g_6(\beta_2,\sigma_{e1},\sigma_{e2}))^2}{2g_4(\beta_2,\sigma_{e1},\sigma_{e2})}\right\}}{\left[1-\Phi\left(g_1(\psi, g_6(\beta_2,\sigma_{e1},\sigma_{e2}), \sqrt{g_4(\beta_2,\sigma_{e1},\sigma_{e2})})\right)\right]} db$$

$$= f_2(y_1,y_2)\Big\{g_4(\beta_2,\sigma_{e1},\sigma_{e2}) + g_1(\psi, g_6(\beta_2,\sigma_{e1},\sigma_{e2}), \sqrt{g_4(\beta_2,\sigma_{e1},\sigma_{e2})})g_4(\beta_2,\sigma_{e1},\sigma_{e2})$$
$$g_2\left(\psi, g_6(\beta_2,\sigma_{e1},\sigma_{e2}), \sqrt{g_4(\beta_2,\sigma_{e1},\sigma_{e2})}\right) + g_6^2(\beta_2,\sigma_{e1},\sigma_{e2})$$
$$+2\sqrt{g_4(\beta_2,\sigma_{e1},\sigma_{e2})}g_6(\beta_2,\sigma_{e1},\sigma_{e2})g_2\left(\psi, g_6(\beta_2,\sigma_{e1},\sigma_{e2}), \sqrt{g_4(\beta_2,\sigma_{e1},\sigma_{e2})}\right)\Big\}$$

Now the moments in (a) and (b) are readily obtained. For the rest, we first note that $(b-\psi)_+ = (b-\psi)I(b>\psi)$. Therefore,

$$\begin{aligned}\int_{-\infty}^{\infty} (b-\psi)_+ f(b,y_1,y_2)db &= \int_{-\infty}^{\infty} (b-\psi)\, I(b>\psi) f(b,y_1,y_2)db\\ &= \int_{\psi}^{\infty} bf(b,y_1,y_2)db - \psi\int_{\psi}^{\infty} f(b,y_1,y_2)db\\ &= A_2 - \psi f_2(y_1,y_2),\end{aligned}$$

$$\begin{aligned}\int_{-\infty}^{\infty} b\,(b-\psi)_+ f(b,y_1,y_2)db &= \int_{-\infty}^{\infty} b\,(b-\psi)\, I(b>\psi) f(b,y_1,y_2)db\\ &= \int_{\psi}^{\infty} b^2 f(b,y_1,y_2)db - \psi\int_{\psi}^{\infty} bf(b,y_1,y_2)db\\ &= A_4 - \psi A_2,\end{aligned}$$

and

$$\int_{-\infty}^{\infty}\{(b-\psi)_+\}^2 f(b,y_1,y_2)db = \int_{-\infty}^{\infty} (b-\psi)^2\, I(b>\psi) f(b,y_1,y_2)db$$
$$= \int_{\psi}^{\infty} b^2 f(b,y_1,y_2)db - 2\psi\int_{\psi}^{\infty} bf(b,y_1,y_2)db + \psi^2\int_{\psi}^{\infty} f(b,y_1,y_2)db$$
$$= A_4 - 2\psi A_2 + \psi^2 f_2(y_1,y_2),$$

The moments in (c), (d), and (e) are now obtained by dividing these integrals by $f(y_1,y_2)$. □

Proposition 3.7. *Consider the model* (3.1)*. The best linear predictor of* b *using* (Y_1,Y_2) *is*

$$\hat{b} = \mu_b + [\sigma_b^2, \beta_1\sigma_b^2 + \beta_2 m_4]\left(var\left(\left[\begin{smallmatrix}Y_1\\Y_2\end{smallmatrix}\right]\right)\right)^{-1}\left[\begin{smallmatrix}Y_1-\mu_b\\Y_2-E(Y_2)\end{smallmatrix}\right], \tag{3.18}$$

where m_4 *is given by* (3.3) *and the moments involved are given by Proposition 3.1.*

Proof. The proof follows from an application of a well-known result (McCulloch et al., 2008, Chapter 13), implying that

$$\hat{b} = E(b) + [\text{cov}(b, Y_1), \text{cov}(b, Y_2)] \left(\text{var}\left(\left[\begin{smallmatrix} Y_1 \\ Y_2 \end{smallmatrix}\right]\right)\right)^{-1} \left[\begin{smallmatrix} Y_1 - E(Y_1) \\ Y_2 - E(Y_2) \end{smallmatrix}\right].$$

□

3.3 Estimation of model parameters

As in Chapter 2, we use $\boldsymbol{\theta}$ to denote the 8×1 vector of parameters $(\mu_b, \beta_0, \beta_1, \beta_2, \sigma_b^2, \sigma_{e1}^2, \sigma_{e2}^2, \psi)$ in the model (3.1) and develop an ECM algorithm for its estimation based on n pairs of observations (Y_{i1}, Y_{i2}), $i = 1, \ldots, n$ of (Y_1, Y_2) collected from n randomly selected subjects from the population that follow the model (3.1). Thus, the data are modeled as

$$Y_{i1} = b_i + e_{i1},\ Y_{i2} = \beta_0 + \beta_1 b_i + \beta_2 (b_i - \psi)_+ + e_{i2}, \tag{3.19}$$

where b_i and e_{ij} are mutually independent draws from the respective distributions of b and e_j given in (3.2), $j = 1, 2$, $i = 1, \ldots, n$. The log-likelihood function can be written as

$$L(\boldsymbol{\theta}) = \sum_{i=1}^{n} \log\{f(y_{i1}, y_{i2}|\boldsymbol{\theta})\}, \tag{3.20}$$

where the density f is given by (3.5) in Proposition 3.3. To develop the ECM algorithm, we take b_i as the *missing data* and (b_i, Y_{i1}, Y_{i2}) as the *complete data* for the ith subject. The logarithm of the joint density $f(b, y_1, y_2|\boldsymbol{\theta})$ can be written as

$$\begin{aligned}
\log\{f(b, y_1, y_2|\boldsymbol{\theta})\} &= \log f\{(y_1|b, \boldsymbol{\theta})\} + \log\{f(y_2|b, \boldsymbol{\theta})\} + \log\{f(b|\boldsymbol{\theta})\} \\
&= -\frac{3}{2}\log(2\pi) - \frac{1}{2}\log(\sigma_{e1}^2) - \frac{1}{2}\log(\sigma_{e2}^2) - \frac{1}{2}\log(\sigma_b^2) - \frac{1}{2\sigma_{e1}^2}(y_1 - b)^2 \\
&\quad - \frac{1}{2\sigma_{e2}^2}\left\{y_2 - \beta_0 - \beta_1 b - \beta_2 (b - \psi)_+\right\}^2 - \frac{1}{2\sigma_b^2}(b - \mu_b)^2 \\
&= -\frac{3}{2}\log(2\pi) - \frac{1}{2}\log(\sigma_{e1}^2) - \frac{1}{2}\log(\sigma_{e2}^2) - \frac{1}{2}\log(\sigma_b^2)
\end{aligned}$$

$$
\begin{aligned}
&-\frac{1}{2\sigma_{e1}^2}\left(y_1^2-2y_1b+b^2\right)-\frac{1}{2\sigma_{e2}^2}\left\{(y_2-\beta_0)^2\right.\\
&\left.-2(y_2-\beta_0)(\beta_1 b+\beta_2\left(b-\psi\right)_+)+\left(\beta_1 b+\beta_2\left(b-\psi\right)_+\right)^2\right\}\\
&-\frac{1}{2\sigma_b^2}\left(b^2-2\mu_b b+\mu_b^2\right)\\
=\ &c-\frac{1}{2\sigma_{e1}^2}\left(-2y_1b+b^2\right)-\frac{1}{2\sigma_{e2}^2}\left\{-2(y_2-\beta_0)(\beta_1 b+\beta_2\left(b-\psi\right)_+)\right.\\
&\left.+\left(\beta_1^2b^2+2\beta_1\beta_2 b\left(b-\psi\right)_+\right)+\beta_2^2\left(b-\psi\right)_+^2\right\}-\frac{1}{2\sigma_b^2}\left(b^2-2\mu_b b\right),
\end{aligned}
\tag{3.21}
$$

where c consists of terms that do not involve b and is given as

$$
c=-\frac{3}{2}\log(2\pi)-\frac{1}{2}\log(\sigma_{e1}^2)-\frac{1}{2}\log(\sigma_{e2}^2)-\frac{1}{2}\log(\sigma_b^2)-\frac{1}{2\sigma_{e1}^2}y_1^2-\frac{1}{2\sigma_{e2}^2}\left(y_2-\beta_0\right)^2-\frac{1}{2\sigma_b^2}\mu_b^2. \tag{3.22}
$$

It follows that the complete data log-likelihood function is

$$
\begin{aligned}
&\sum_{i=1}^n \log\{f(b_i,y_{i1},y_{i2}|\boldsymbol{\theta})\}\\
&=\sum_{i=1}^n\left\{c_i-\frac{1}{2\sigma_{e1}^2}\left(-2y_{i1}b_i+b_i^2\right)-\frac{1}{2\sigma_{e2}^2}\left\{-2(y_{i2}-\beta_0)(\beta_1 b_i+\beta_2\left(b_i-\psi\right)_+)\right.\right.\\
&\qquad\left.\left.+\left(\beta_1^2b_i^2+2\beta_1\beta_2 b_i\left(b_i-\psi\right)_+\right)+\beta_2^2\left(b_i-\psi\right)_+^2\right\}-\frac{1}{2\sigma_b^2}\left(b_i^2-2\mu_b b_i\right)\right\},
\end{aligned}
\tag{3.23}
$$

where c_i is the value of c given by (3.22) evaluated for the ith subject.

In the rth ECM iteration, let $\boldsymbol{\theta}^{(r)}$ be the value of $\boldsymbol{\theta}$ and $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(r)})$ be the expectation of the complete data log-likelihood (3.23), i.e., $E\left[\sum_{i=1}^n \log\{f(b_i,y_{i1},y_{i2}|\boldsymbol{\theta})\}|y_{i1},y_{i2},\boldsymbol{\theta}^{(r)}\right]$. Letting $E^{(r)}$ denote the expectation over the conditional distribution of $b_i|y_{i1},y_{i2}$ evaluated at $\boldsymbol{\theta}^{(r)}$, we can write

$$
\begin{aligned}
Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(r)})=\sum_{i=1}^n&\left\{c_i-\frac{1}{2\sigma_{e1}^2}\left(-2y_{i1}E^{(r)}[b_i]+E^{(r)}[b_i^2]\right)-\frac{1}{2\sigma_{e2}^2}\left\{-2(y_{i2}-\beta_0)\right.\right.\\
&\left(\beta_1E^{(r)}[b_i]+\beta_2E^{(r)}\left[(b_i-\psi)_+\right]\right)+\beta_1^2E^{(r)}[b_i^2]\\
&\left.+2\beta_1\beta_2E^{(r)}\left[b_i\left(b_i-\psi\right)_+\right]+\beta_2^2E^{(r)}\left[(b_i-\psi)_+^2\right]\right\}\\
&\left.-\frac{1}{2\sigma_b^2}\left(E^{(r)}[b_i^2]-2\mu_bE^{(r)}[b_i]\right)\right\}.
\end{aligned}
\tag{3.24}
$$

Next, we find derivative of $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(r)})$ with respect to $\boldsymbol{\theta}$ so that we can perform the CM steps. For this, let $E_{i1}^{(r)}$ and $E_{i2}^{(r)}$ respectively denote the values of $E^{(r)}[b_i]$ and $E^{(r)}[b_i^2]$ and $A_{i1}^{(r)}, \ldots, A_{i4}^{(r)}$ respectively denote the values of $A_1, \ldots, A_4$ given by (3.17) in Proposition 3.6 and evaluated at $(\boldsymbol{\theta}, y_1, y_2) = (\boldsymbol{\theta}^{(r)}, y_{i1}, y_{i2})$. Further, let $f^{(r)}(y_{i1}, y_{i2})$ denote the value of $f(y_1, y_2)$ given by (3.5) in Proposition 3.3 and also evaluated at $(\boldsymbol{\theta}, y_1, y_2) = (\boldsymbol{\theta}^{(r)}, y_{i1}, y_{i2})$. The derivatives of $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(r)})$ with respect to the elements of $\boldsymbol{\theta}$ are as follows:

$$
\begin{aligned}
\frac{\partial Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(r)})}{\partial \mu_b} &= -\frac{1}{2\sigma_b^2}\sum_{i=1}^{n} 2\left\{-E_{i1}^{(r)} + \mu_b\right\} \\
\frac{\partial Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(r)})}{\partial \beta_0} &= -\frac{1}{2\sigma_{e2}^2}\sum_{i=1}^{n} 2\left\{\beta_1 E_{i1}^{(r)} + \frac{\beta_2\left(A_{i2}^{(r)} - \psi f_2^{(r)}(y_{i1}, y_{i2})\right)}{f^{(r)}(y_{i1}, y_{i2})} - (y_{i2} - \beta_0)\right\} \\
\frac{\partial Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(r)})}{\partial \beta_1} &= -\frac{1}{2\sigma_{e2}^2}\sum_{i=1}^{n} 2\left\{-(y_{i2} - \beta_0)E_{i1}^{(r)} + \beta_1 E_{i2}^{(r)} + \frac{\beta_2(A_{i4}^{(r)} - \psi A_{i2}^{(r)})}{f^{(r)}(y_{i1}, y_{i2})}\right\} \\
\frac{\partial Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(r)})}{\partial \beta_2} &= -\frac{1}{2\sigma_{e2}^2}\sum_{i=1}^{n} 2\left\{-\frac{(y_{i2} - \beta_0)\left(A_{i2}^{(r)} - \psi f_2^{(r)}(y_{i1}, y_{i2})\right) + \beta_1(A_{i4}^{(r)} - \psi A_{i2}^{(r)})}{f^{(r)}(y_{i1}, y_{i2})}\right. \\
&\qquad \left. + \frac{\beta_2(A_{i4}^{(r)} - 2\psi A_{i2}^{(r)} + \psi^2 f_2^{(r)}(y_{i1}, y_{i2})}{f^{(r)}(y_{i1}, y_{i2})}\right\} \\
\frac{\partial Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(r)})}{\partial \sigma_b^2} &= -\frac{n}{2\sigma_b^2} + \frac{1}{2\sigma_b^4}\sum_{i=1}^{n}\left\{E_{i2}^{(r)} - 2\mu_b E_{i1}^{(r)} + \mu_b^2\right\} \\
\frac{\partial Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(r)})}{\partial \sigma_{e1}^2} &= -\frac{n}{2\sigma_{e1}^2} + \frac{1}{2\sigma_{e1}^4}\sum_{i=1}^{n}\left\{-2y_{i1}E_{i1}^{(r)} + E_{i2}^{(r)} + y_{i1}^2\right\} \\
\frac{\partial Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(r)})}{\partial \sigma_{e2}^2} &= -\frac{n}{2\sigma_{e2}^2} + \frac{1}{2\sigma_{e2}^4}\sum_{i=1}^{n}\left\{-2(y_{i2} - \beta_0)\left[\beta_1 E_{i1}^{(r)} + \frac{\beta_2(A_{i2}^{(r)} - \psi f_2^{(r)}(y_{i1}, y_{i2}))}{f^{(r)}(y_{i1}, y_{i2})}\right]\right. \\
&\qquad + \beta_1^2 E_{i2}^{(r)} + \frac{2\beta_1\beta_2(A_{i4}^{(r)} - \psi A_{i2}^{(r)}) + \beta_2^2(A_{i4}^{(r)} - 2\psi A_{i2}^{(r)} + \psi^2 f_2^{(r)}(y_{i1}, y_{i2}))}{f^{(r)}(y_{i1}, y_{i2})} \\
&\qquad \left. + (y_{i2} - \beta_0)^2\right\} \\
\frac{\partial Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(r)})}{\partial \psi} &= -\frac{1}{2\sigma_{e2}^2}\sum_{i=1}^{n} 2\left\{\frac{\begin{array}{c}(y_{i2} - \beta_0)\beta_2 f_2^{(r)}(y_{i1}, y_{i2}) - \beta_1\beta_2 A_{i2}^{(r)} \\ + \beta_2^2(-A_{i2}^{(r)} + \psi f_2^{(r)}(y_{i1}, y_{i2}))\end{array}}{f^{(r)}(y_{i1}, y_{i2})}\right\}
\end{aligned} \tag{3.25}
$$

By setting each of the derivatives in (3.25) equal to zero and solving for the corresponding parameter, we get:

$$
\begin{aligned}
\mu_b &= \frac{\sum_{i=1}^n E_{i1}^{(r)}}{n} \\
\beta_0 &= \frac{\sum_{i=1}^n \left\{ y_{i2} - \beta_1 E_{i1}^{(r)} - \frac{\beta_2\left(A_{i2}^{(r)} - \psi f_2^{(r)}(y_{i1},y_{i2})\right)}{f^{(r)}(y_{i1},y_{i2})} \right\}}{n} \\
\beta_1 &= \frac{\sum_{i=1}^n \left\{ (y_{i2} - \beta_0) E_{i1}^{(r)} - \frac{\beta_2\left(A_{i4}^{(r)} - \psi A_{i2}^{(r)}\right)}{f^{(r)}(y_{i1},y_{i2})} \right\}}{\sum_{i=1}^n E_{i2}^{(r)}} \\
\beta_2 &= \frac{\sum_{i=1}^n \frac{\left\{(y_{i2}-\beta_0)\left(A_{i2}^{(r)} - \psi f_2^{(r)}(y_{i1},y_{i2})\right) - \beta_1\left(A_{i4}^{(r)} - \psi A_{i2}^{(r)}\right)\right\}}{f^{(r)}(y_{i1},y_{i2})}}{\sum_{i=1}^n \frac{A_{i4}^{(r)} - 2\psi A_{i2}^{(r)} + \psi^2 f_2^{(r)}(y_{i1},y_{i2})}{f^{(r)}(y_{i1},y_{i2})}} \\
\sigma_b^2 &= \frac{\sum_{i=1}^n \left\{ E_{i2}^{(r)} - 2\mu_b E_{i1}^{(r)} + \mu_b^2 \right\}}{n} \\
\sigma_{e1}^2 &= \frac{\sum_{i=1}^n \left\{ (y_{i1} - E_{i1}^{(r)})^2 + E_{i2}^{(r)} - (E_{i1}^{(r)})^2 \right\}}{n} \\
\sigma_{e2}^2 &= \frac{1}{n} \sum_{i=1}^n \Bigg\{ (y_{i2} - \beta_0)^2 - 2(y_{i2} - \beta_0) \left[\beta_1 E_{i1}^{(r)} + \frac{\beta_2 (A_{i2}^{(r)} - \psi f_2^{(r)}(y_{i1}, y_{i2}))}{f^{(r)}(y_{i1}, y_{i2})} \right] \\
&\qquad \beta_1^2 E_{i2}^{(r)} + \frac{2\beta_1\beta_2 (A_{i4}^{(r)} - \psi A_{i2}^{(r)}) + \beta_2^2 (A_{i4}^{(r)} - 2\psi A_{i2}^{(r)} + \psi^2 f_2^{(r)}(y_{i1}, y_{i2}))}{f^{(r)}(y_{i1}, y_{i2})} \Bigg\} \\
\psi &= \frac{\sum_{i=1}^n \left\{ \frac{(\beta_1+\beta_2) A_{i2}^{(r)} - (y_{i2}-\beta_0) f_2^{(r)}(y_{i1},y_{i2})}{f^{(r)}(y_{i1},y_{i2})} \right\}}{\beta_2 \sum_{i=1}^n \frac{f_2^{(r)}(y_{i1},y_{i2})}{f^{(r)}(y_{i1},y_{i2}))}}.
\end{aligned}
\tag{3.26}
$$

Taken together, the E and CM steps in the rth iteration of our ECM algorithm are as follows:

E-step: Compute $A_{i1}^{(r)}, \ldots, A_{i4}^{(r)}$ and hence the conditional expectations in $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(r)})$.

CM-step 1: Update μ_b by maximizing $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(r)})$ with respect to μ_b while holding all other parameters at their current values. This yields $\mu_b^{(r+1)}$ as the value of μ_b given in (3.26).

CM-step 2: Update β_0 by maximizing $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(r)})$ with respect to β_0 while holding μ_b at the updated value and all other parameters at their current values. This yields $\beta_0^{(r+1)}$ as the value of β_0 given in (3.26).

CM-step 3: Update β_1 by maximizing $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(r)})$ with respect to β_1 while holding μ_b and β_0 at their updated values and all other parameters at their current values. This yields $\beta_1^{(r+1)}$ as the value of β_1 given in (3.26).

CM-step 4: Update β_2 by maximizing $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(r)})$ with respect to β_2 while holding μ_b, β_0, and β_1 at their updated values and all other parameters at their current values. This yields $\beta_2^{(r+1)}$ as the value of β_2 given in (3.26).

CM-step 5: Update σ_b^2 by maximizing $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(r)})$ with respect to σ_b^2 while holding μ_b, β_0, β_1, and β_2 at their updated values and all other parameters at their current values. This yields $\sigma_b^{2,(r+1)}$ as the value of σ_b^2 given in (3.26).

CM-step 6: Update σ_{e1}^2 by maximizing $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(r)})$ with respect to σ_{e1}^2 while holding μ_b, β_0, β_1, β_2, and σ_b^2 at their updated values and all other parameters at their current values. This yields $\sigma_{e1}^{2,(r+1)}$ as the value of σ_{e1}^2 given in (3.26).

CM-step 7: Update σ_{e2}^2 by maximizing $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(r)})$ with respect to σ_{e2}^2 while holding μ_b, β_0, β_1, β_2, σ_b^2, and σ_{e1}^2 at their updated values and ψ^2 at its current value. This yields $\sigma_{e2}^{2,(r+1)}$ as the value of σ_{e2}^2 given in (3.26).

CM-step 8: Update ψ by maximizing $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(r)})$ with respect to ψ while holding all other parameters at their updated values. This yields $\psi^{(r+1)}$ as the value of ψ given in (3.26).

This algorithm is essentially a slight modification of the ECM algorithm presented in Chapter 2 that reflects that the equal error variance is not being made. Once $\boldsymbol{\theta}$ is estimated, further inference on it proceeds as described in Chapter 2. We can also perform model evaluation using residuals as in Chapter 2.

3.4 Discussion

In this chapter, we developed the distribution theory and presented an ECM algorithm for parameter estimation under the segmented measurement error model (3.1) that allows the error variances to be unequal. These results generalize their analogs presented in Chapter 2 that make the equal error variance assumption. A study of the properties of estimators under this model, including the impact of equal of error variance assumption, using both simulated and real data is planned in the near future.

CHAPTER 4

A BAYESIAN APPROACH FOR ANALYSIS OF METHOD COMPARISON DATA WITH INFORMATIVE PRIORS

4.1 Introduction

Let (Y_1, Y_2) be paired measurements by two methods on a randomly selected subject from a population. Consider the paired measurements data (Y_{i1}, Y_{i2}), $i = 1, \ldots, n$ collected in a method comparison study. Here Y_{ij} is the measurement made by the jth method on the ith subject, $j = 1, 2$, $i = 1, \ldots, n$. These data are assumed to be a random sample from the distribution of (Y_1, Y_2). They are often modeled using a mixed-effects model (Choudhary and Nagaraja, 2017, Chapters 1 and 3)

$$Y_{ij} = \beta_j + b_i + \epsilon_{ij},\ i = 1, \ldots, n,\ j = 1, 2, \tag{4.1}$$

where β_j is the fixed effect of method j, b_i is the random effect of subject i, and ϵ_{ij} is the within-subject random error. It is assumed that, for $i = 1, \ldots, n$, the random effects b_i are independently distributed as $N_1(0, \psi^2)$, the errors e_{ij} are independently distributed as $N_1(0, \sigma_{ej}^2)$, and the random effects and the errors are mutually independent. This model is a special case of the classical measurement error model (2.2) when the slope is assumed to be one. It is also known as the Grubbs' model after (Grubbs, 1948). Under (4.1), it follows that

$$\begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} \sim N_2 \left(\begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}, \begin{bmatrix} \psi^2 + \sigma_{e1}^2 & \psi^2 \\ \psi^2 & \psi^2 + \sigma_{e2}^2 \end{bmatrix} \right).$$

Although, in principle, this model is identifiable for method comparison data, but in practice, it suffers from a near non-identifiability problem. In particular, when the model is fit to the data, the error standard deviations (SDs) σ_{e1} and σ_{e2} are generally not estimated well in that at least one of their estimates has a rather large standard error (SE).

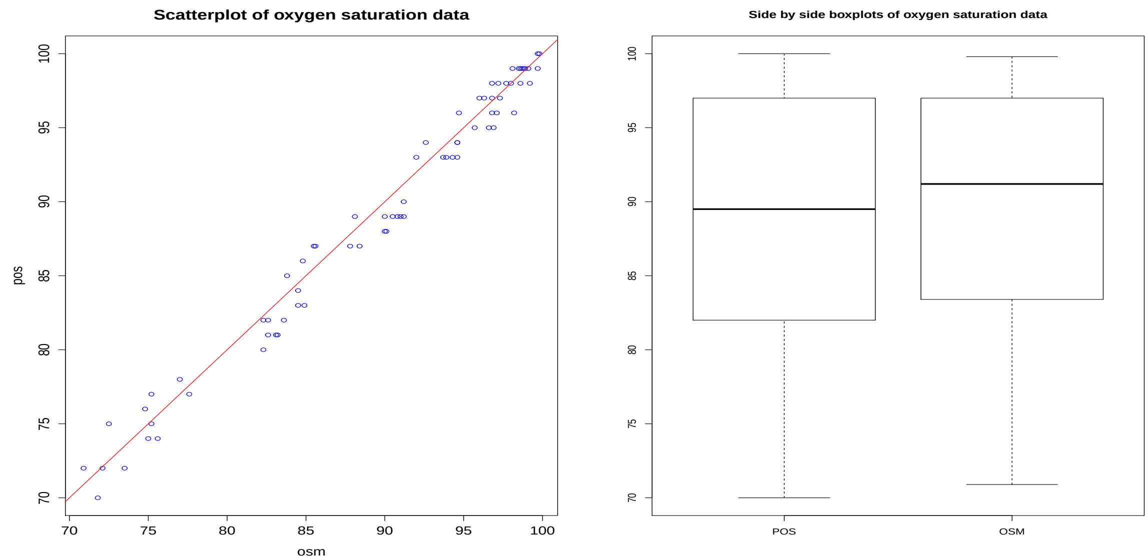

Figure 4.1. Scatterplot and side-by-side boxplots for oxygen saturation data.

As a specific example, consider the oxygen saturation data. These data consist of paired measurements of percent saturation of hemoglobin with oxygen from $n = 72$ subjects measured using two methods — pulse oximetry saturation (POS, method 1) and oxygen saturation monitor (OSM, method 2). Figure 4.1 displays the scatterplot and the side-by-side boxplots of the data. The points in the scatterplot seem tightly clustered around the 45° line. However, there are more points below the line than above the line, suggesting that the center of the distribution of OSM measurements is shifted to the right of that for POS measurements. This is confirmed by the boxplots where the median of OSM measurements is clearly greater than that for POS. The boxplots also show that the variability in POS may be slightly higher than that of OSM. On the whole, this exploratory data analysis indicates that OSM has higher mean than POS but the two may be considered to have equal variances or it may be that POS has a slightly higher variance than OSM.

Table 4.1 presents ML estimates of parameters, their SEs, and 95% confidence intervals when the model (4.1) is fit to these data. The ML estimates are obtained by numerically maximizing the likelihood function using `optim` function in `R`. The SEs and confidence in-

Table 4.1. ML estimates of parameters of model (4.1) without and with the equal error variance assumption. Method 1 refers to POS and method 2 refers to OSM.

	Unequal error variances			Equal error variances		
Parameter	Estimate	SE	95% CI	Estimate	SE	95% CI
β_1	89.08	1.02	$(87.07, 91.09)$	89.08	1.02	$(87.08, 91.09)$
β_2	89.50	1.02	$(87.50, 91.49)$	89.50	1.02	$(87.49, 91.50)$
$\log(\psi)$	2.16	0.08	$(1.99, 2.32)$	2.16	0.08	(1.99, 2.32)
$\log(\sigma_{e1})$	0.13	2.85	$(-5.46, 5.71)$	-	-	-
$\log(\sigma_{e2})$	-0.95	0.42	$(-1.78, -0.12)$	-	-	-
$\log(\sigma_e)$	-	-	-	-0.16	0.08	$(-0.32, 0.00)$

tervals are computed using the standard large-sample theory (Lehmann, 1998; Choudhary and Nagaraja, 2017). The Hessian matrix needed for this calculation is numerically computed. The SD parameters are transformed on the log scale for model fitting. We see that the estimated mean of OSM is slightly higher than that of POS. Moreover, using the estimates reported for ψ, σ_{e1}, and σ_{e2}, the estimates of marginal variances of POS and OSM are $\exp(4.310) + \exp(0.260) \approx 75.74$ and $\exp(4.310) + \exp(-1.904) \approx 74.59$, respectively, which may be considered similar. These findings are consistent with what we expect from the exploratory data analysis. However, there is a stark difference in the estimates of the error SDs — $\exp(-0.95) \approx 0.39$ versus $\exp(0.13) \approx 1.14$, with OSM appear much more precise than POS — and the SE of the POS estimate on log-scale is unusually large, indicating that the estimate is not reliable. Table 4.1 also presents the parameter estimates under the equal error variance assumption, i.e., $\sigma_{e1}^2 = \sigma_{e2}^2 = \sigma_e^2$, with σ_e^2 representing the common error variance. The estimate of σ_e is $\exp(-0.16) \approx 0.85$ and the SE of the estimate on log-scale is not unusually large. These observations confirm the aforementioned near non-identifiability problem — the data do not have enough information to estimate two error variances reliably. We return to these data in Section 4.5.

Upon a little reflection, we find that the reason for the difficulty in estimating two separate error variances is that we are trying to do so without having replicated measurements on each subject by every method. This matter is made worse by the fact that the paired measurements in method comparison studies tend to be highly correlated, often with estimated correlation in excess of 0.99. Indeed, the sample correlation in the oxygen saturation data is 0.99. The bottomline is that the paired measurements data do not have enough information to estimate the method-specific error variances in method comparison studies. This motivated us to ask the following question: "Can we get reliable estimates of the error variances if we bring in external information in the form of informative prior distributions for the error SDs and use a Bayesian approach for inference?" The work in this chapter is devoted to answering this question.

As the primary motivation behind a Bayesian approach is the potential for reliable estimation of error variances, we let the other parameters, namely, the fixed effects vector $\boldsymbol{\beta} = [\beta_1, \beta_2]'$ and the random effect variance ψ^2, follow the standard noninformative prior distributions from the literature (Gelman et al., 2013, Chapter 2). Specifically, we assume

$$\boldsymbol{\beta} \sim N_2(\mathbf{0}, \mathbf{S}), \quad \mathbf{S} = \text{diag}\{V_1^2, V_2^2\}, \quad \psi^2 \sim \text{IG}(A_\psi, B_\psi), \tag{4.2}$$

where the hyperparamaters V_1^2, V_2^2, A_ψ, and B_ψ are known. Their values will be specified to ensure that the resulting priors are noninformative.

It is convenient to write the model (4.1) with priors (4.2) in matrix notation. For this, define the $2n \times 1$ vector of observations $\mathbf{Y} = [Y_{11}, \ldots, Y_{n1}, Y_{12}, \ldots, Y_{n2}]'$, the corresponding $2n \times 1$ vector of random errors $\boldsymbol{\epsilon} = [\epsilon_{11}, \ldots, \epsilon_{n1}, \epsilon_{12}, \ldots, \epsilon_{n2}]'$, its $2n \times 2n$ covariance matrix $\mathbf{R} = \text{diag}\{\sigma_{e1}^2 \mathbf{I}_{n\times n}, \sigma_{e2}^2 \mathbf{I}_{n\times n}\}$, the $(2+n)$ vector of fixed and random effects $\boldsymbol{\gamma} = [\beta_1, \beta_2, b_1, \ldots, b_n]'$, and the associated $2n \times (2+n)$ design matrix and the $(2+n) \times (2+n)$ covariance matrix

$$\mathbf{C} = \begin{bmatrix} \mathbf{1}_{n\times 1} & \mathbf{0}_{n\times 1} & \mathbf{I}_{n\times n} \\ \mathbf{0}_{n\times 1} & \mathbf{1}_{n\times 1} & \mathbf{I}_{n\times n} \end{bmatrix}, \ \mathbf{G} = \text{diag}\{V_1^2, V_2^2, \psi^2, \ldots, \psi^2\}.$$

Here $\mathbf{R}^{-1} = \text{diag}\{(1/\sigma_{e1}^2)\mathbf{I}_{n\times n}, (1/\sigma_{e2}^2)\mathbf{I}_{n\times n}\}$ and $\mathbf{G}^{-1} = \text{diag}\{1/V_1^2, 1/V_2^2, 1/\psi^2, \ldots, 1/\psi^2\}$. The model (4.1) with priors (4.2) can be written as

$$\mathbf{Y} = \mathbf{C}\boldsymbol{\gamma} + \boldsymbol{\epsilon},\ \boldsymbol{\epsilon}|\sigma_{e1}^2, \sigma_{e2}^2 \sim N_{2n}(\mathbf{0}_{2n\times 1}, \mathbf{R}),\ \boldsymbol{\gamma}|\psi^2 \sim N_{2+n}(\mathbf{0}_{(2+n)\times 1}, \mathbf{G}),\ \psi^2 \sim \text{IG}(A_\psi, B_\psi). \tag{4.3}$$

It follows that

$$\mathbf{Y}|\boldsymbol{\gamma}, \sigma_{e1}^2, \sigma_{e2}^2 \sim N_{2n}(\mathbf{C}\boldsymbol{\gamma}, \mathbf{R}),\ \boldsymbol{\gamma}|\psi^2 \sim N_{2+n}(\mathbf{0}_{(2+n)\times 1}, \mathbf{G}). \tag{4.4}$$

For ease of exposition, we begin in Section 4.2 by describing the Bayesian approach assuming equal error variances. This assumption is dropped in Section 4.3.

4.2 Model fitting assuming equal error variances

In this section, we assume $\sigma_{e1}^2 = \sigma_{e2}^2 = \sigma_e^2$. This implies $\mathbf{R} = \sigma_e^2\mathbf{I}_{2n\times 2n}$ in (4.3) and (4.4), and $\mathbf{R}^{-1} = (1/\sigma_e^2)\mathbf{I}_{2n\times 2n}$. We consider two choices for priors for σ_e. One is a half-normal prior, i.e., $\sigma_e \sim \text{HN}(\alpha)$, where the hyperparameter α is specified, and the other is a hierarchical half-normal prior, i.e., $\sigma_e|\alpha^2 \sim \text{HN}(\alpha)$ and $\alpha^2 \sim \text{IG}(A_\alpha, B_\alpha)$, where the hyperparameters A_α and B_α are specified. We denote this distribution as $\text{HHN}(A_\alpha, B_\alpha)$ distribution. The motivation for a half-normal prior comes from (Spiegelhalter et al., 2004) where several choices of priors for SD are explored and the use of half-normal is encouraged for the case when prior information is strong and important. The hyperparameter of the half-normal is estimated from external data. We also explore a hierarchical version of this prior — the HHN prior — to take into account the uncertainty in this estimation. The issue of posterior simulation under each of HN and HHN priors is discussed in the next two subsections.

4.2.1 Half-normal prior

Assuming $\sigma_e \sim \text{HN}(\alpha)$, the posterior distribution of model parameters — $(\boldsymbol{\gamma}, \psi^2, \sigma_e)$ — can be written as:

$$\begin{aligned} p(\boldsymbol{\gamma}, \psi^2, \sigma_e|\mathbf{y}) \propto p(\mathbf{y}|\boldsymbol{\gamma}, \psi^2, \sigma_e)p(\boldsymbol{\gamma}, \psi^2, \sigma_e) &= p(\mathbf{y}|\boldsymbol{\gamma}, \sigma_e)p(\boldsymbol{\gamma}|\psi^2)p(\psi^2)p(\sigma_e) \\ &\propto \frac{1}{|\mathbf{R}|^{1/2}} \exp\left\{-\frac{1}{2}(\mathbf{y}-\mathbf{C}\boldsymbol{\gamma})'\mathbf{R}^{-1}(\mathbf{y}-\mathbf{C}\boldsymbol{\gamma})\right\} \frac{1}{|\mathbf{G}|^{1/2}} \exp\left\{-\frac{1}{2}\boldsymbol{\gamma}'\mathbf{G}^{-1}\boldsymbol{\gamma}\right\} \\ &\quad \times (\psi^2)^{-A_\psi-1} \exp\left\{-\frac{B_\psi}{\psi^2}\right\} \exp\left\{-\frac{\sigma_e^2}{2\alpha^2}\right\}. \end{aligned} \tag{4.5}$$

This distribution is not available in a closed form. Therefore, we consider a Markov chain Monte Carlo (MCMC) method for posterior simulation. Specifically, we take a Gibbs sampler approach and find the *full conditional distribution* — the posterior distribution of a parameter conditional on the others — for each of the parameters. We begin with the full conditional of $\boldsymbol{\gamma}$. This density is

$$\begin{aligned} p(\boldsymbol{\gamma}|\psi^2, \sigma_e, \mathbf{y}) &\propto \exp\left\{-\frac{1}{2}(\mathbf{y}-\mathbf{C}\boldsymbol{\gamma})'\mathbf{R}^{-1}(\mathbf{y}-\mathbf{C}\boldsymbol{\gamma})\right\} \exp\left\{-\frac{1}{2}\boldsymbol{\gamma}'\mathbf{G}^{-1}\boldsymbol{\gamma}\right\} \\ &\propto \exp\left\{-\frac{1}{2}\left[\left(\boldsymbol{\gamma} - \left(\mathbf{C}'\mathbf{R}^{-1}\mathbf{C}+\mathbf{G}^{-1}\right)^{-1}\mathbf{C}'\mathbf{R}^{-1}\mathbf{y}\right)' (\mathbf{C}'\mathbf{R}^{-1}\mathbf{C}+\mathbf{G}^{-1}) \right.\right. \\ &\quad \left.\left. \times \left(\boldsymbol{\gamma} - \left(\mathbf{C}'\mathbf{R}^{-1}\mathbf{C}+\mathbf{G}^{-1}\right)^{-1}\mathbf{C}'\mathbf{R}^{-1}\mathbf{y}\right)\right]\right\}, \end{aligned}$$

implying that

$$\boldsymbol{\gamma}|\psi^2, \sigma_e, \mathbf{y} \sim N_{2+n}\left(\left(\mathbf{C}'\mathbf{R}^{-1}\mathbf{C}+\mathbf{G}^{-1}\right)^{-1}\mathbf{C}'\mathbf{R}^{-1}\mathbf{y}, (\mathbf{C}'\mathbf{R}^{-1}\mathbf{C}+\mathbf{G}^{-1})^{-1}\right). \tag{4.6}$$

Next, we consider the full conditional of ψ^2. This density is

$$\begin{aligned} p(\psi^2|\boldsymbol{\gamma}, \sigma_e, \mathbf{y}) &\propto \frac{1}{|\mathbf{G}|^{1/2}} \exp\left\{-\frac{1}{2}\boldsymbol{\gamma}'\mathbf{G}^{-1}\boldsymbol{\gamma}\right\} (\psi^2)^{-A_\psi-1} \exp\left\{-\frac{B_\psi}{\psi^2}\right\} \\ &\propto (\psi^2)^{-A_\psi-(n/2)-1} \exp\left\{-\frac{(B_\psi + \sum_{i=1}^n b_i^2/2)}{\psi^2}\right\}, \end{aligned}$$

implying that

$$\psi^2|\boldsymbol{\gamma}, \sigma_e, \mathbf{y} \sim \text{IG}\left(A_\psi + n/2, B_\psi + \sum_{i=1}^n b_i^2/2\right). \tag{4.7}$$

Now, we consider the full conditional of σ_e. Its density is

$$p(\sigma_e|\boldsymbol{\gamma}, \psi^2, \mathbf{y}) \propto \frac{1}{|\mathbf{R}|^{1/2}} \exp\left\{-\frac{1}{2}(\mathbf{y}-\mathbf{C}\boldsymbol{\gamma})'\mathbf{R}^{-1}(\mathbf{y}-\mathbf{C}\boldsymbol{\gamma})\right\} \exp\left\{-\frac{\sigma_e^2}{2\alpha^2}\right\}$$
$$= \frac{1}{\sigma_e^{2n}} \exp\left\{-\frac{1}{2\sigma_e^2}\sum_{i=1}^{n}\sum_{j=1}^{2}(y_{ij}-b_i-\beta_j)^2\right\} \exp\left\{-\frac{\sigma_e^2}{2\alpha^2}\right\},$$

which is not available in a closed form. Therefore, to sample from this distribution, we will use a Metropolis algorithm to first draw from the full conditional of $\lambda = \log(\sigma_e)$, which is given as

$$p(\lambda|\boldsymbol{\gamma}, \psi^2, \mathbf{y}) \propto \exp\left\{(1-2n)\lambda - \frac{1}{2\exp(2\lambda)}\sum_{i=1}^{n}\sum_{j=1}^{2}(y_{ij}-b_i-\beta_j)^2 - \frac{\exp(2\lambda)}{2\alpha^2}\right\}, \quad (4.8)$$

using a normal proposal distribution, and then exponentiate the draw. Taken together, the MCMC algorithm for posterior simulation is as follows.

Algorithm 1: Begin with a starting point $(\boldsymbol{\gamma}_0, \psi_0^2, \sigma_{e,0})$ and iterate the following until convergence. For $t = 1, 2, \ldots$

1. Draw $\boldsymbol{\gamma}_t|\psi_{t-1}^2, \sigma_{e,t-1}, \mathbf{y}$ from the full conditional in (4.6).

2. Draw $\psi_t^2|\boldsymbol{\gamma}_t, \sigma_{e,t-1}, \mathbf{y}$ from the full conditional in (4.7).

3. Draw $\lambda_t|\boldsymbol{\gamma}_t, \psi_t^2, \mathbf{y}$ from the full conditional in (4.8) using a Metropolis algorithm with a $N(\lambda_{t-1}, 0.05)$ proposal distribution. This involves drawing a proposal $\lambda^* \sim N(\lambda_{t-1}, 0.05)$, calculating the ratio of densities,

$$r = \frac{p(\lambda^*|\boldsymbol{\gamma}_t, \psi_t^2, \mathbf{y})}{p(\lambda_{t-1}|\boldsymbol{\gamma}_t, \psi_t^2, \mathbf{y})}$$

   setting

$$\lambda_t = \begin{cases} \lambda^*, & \text{with probability } \min\{r, 1\}, \\ \lambda_{t-1}, & \text{otherwise,} \end{cases}$$

   and taking $\sigma_{e,t} = \exp\{\lambda_t\}$.

This algorithm requires specification of the hyperparameters $V_1^2, V_2^2, A_\psi, B_\psi$, and α and also the starting points. An illustration is provided in Section 4.5.

4.2.2 Hierarchical half-normal prior

In this section, we assume that $\sigma_e \sim \text{HHN}(A_\alpha, B_\alpha)$. This is equivalent to assuming that $\sigma_e|\alpha^2 \sim \text{HN}(\alpha)$ and $\alpha^2 \sim \text{IG}(A_\alpha, B_\alpha)$, where the hyperparameters A_α and B_α are specified. The model parameters in this case consist of $(\boldsymbol{\gamma}, \psi^2, \sigma_e, \alpha^2)$. As in (4.5), the posterior density can be expressed as

$$
\begin{aligned}
p(\boldsymbol{\gamma}, \psi^2, \sigma_e, \alpha^2|\mathbf{y}) &\propto p(\mathbf{y}|\boldsymbol{\gamma}, \psi^2, \sigma_e, \alpha^2)p(\boldsymbol{\gamma}, \psi^2, \sigma_e, \alpha^2) = p(\mathbf{y}|\boldsymbol{\gamma}, \sigma_e)p(\boldsymbol{\gamma}|\psi^2)p(\psi^2)p(\sigma_e|\alpha^2)p(\alpha^2) \\
&\propto \frac{1}{|\mathbf{R}|^{1/2}} \exp\left\{-\frac{1}{2}(\mathbf{y}-\mathbf{C}\boldsymbol{\gamma})'\mathbf{R}^{-1}(\mathbf{y}-\mathbf{C}\boldsymbol{\gamma})\right\} \frac{1}{|\mathbf{G}|^{1/2}} \exp\left\{-\frac{1}{2}\boldsymbol{\gamma}'\mathbf{G}^{-1}\boldsymbol{\gamma}\right\} \\
&\quad \times (\psi^2)^{-A_\psi-1} \exp\left\{-\frac{B_\psi}{\psi^2}\right\} \frac{1}{\alpha} \exp\left\{-\frac{\sigma_e^2}{2\alpha^2}\right\} \\
&\quad \times (\alpha^2)^{-A_\alpha-1} \exp\left\{-\frac{B_\alpha}{\alpha^2}\right\}.
\end{aligned} \tag{4.9}
$$

This distribution is also not available in a closed form. Therefore, we resort to an MCMC algorithm for posterior simulation. The full conditionals of $\boldsymbol{\gamma}$, ψ^2, and σ_e are same as in the case of HN prior. The full conditional density of the additional parameter α^2 can be written as

$$
p(\alpha^2|\boldsymbol{\gamma}, \psi^2, \sigma_e, \mathbf{y}) \propto (\alpha^2)^{-(A_\alpha+1/2)-1} \exp\left\{-\left(\frac{\sigma_e^2 + 2B_\alpha}{2\alpha^2}\right)\right\},
$$

implying that

$$
\alpha^2|\boldsymbol{\gamma}, \psi^2, \sigma_e, \mathbf{y} \sim \text{IG}\left(A_\alpha + 1/2, (\sigma_e^2 + 2B_\alpha)/2\right). \tag{4.10}
$$

For posterior simulation, we modify Algorithm 1 of the previous subsection by adding a step for sampling α^2. The resulting algorithm is:

Algorithm 2: Begin with a starting point $(\boldsymbol{\gamma}_0, \psi_0^2, \sigma_{e,0}, \alpha_0^2)$ and iterate the following until convergence. For $t = 1, 2, \ldots$

1. Draw $\boldsymbol{\gamma}_t|\psi_{t-1}^2, \sigma_{e,t-1}, \alpha_{t-1}^2, \mathbf{y}$ from the full conditional in (4.6).

2. Draw $\psi_t^2|\boldsymbol{\gamma}_t, \sigma_{e,t-1}, \alpha_{t-1}^2, \mathbf{y}$ from the full conditional in (4.7).

3. Draw $\sigma_{e,t}|\boldsymbol{\gamma}_t, \psi_t^2, \alpha_{t-1}^2, \mathbf{y}$ as in Algorithm 1.

4. Draw $\alpha_t^2|\boldsymbol{\gamma}_t, \psi_t^2, \sigma_{e,t}, \mathbf{y}$ from the full conditional in (4.10).

The algorithm requires specification of the hyperparameters $V_1^2, V_2^2, A_\psi, B_\psi, A_\alpha$, and B_α and also the starting points. An illustration is provided in Section 4.5.

4.3 Model fitting without assuming equality of error variances

In this section, we extend of the approach of Section 4.2 to deal with the case when the error variances σ_{e1}^2 and σ_{e2}^2 in (4.3) are two separate parameters instead of assumed to be equal. Essentially, we mimic the approach of Section 4.2 to develop posterior simulation algorithms under HN and HHN priors for σ_{e1} and σ_{e2}.

4.3.1 Half-normal priors

We assume that the error SDs follow independent HN priors, i.e., $\sigma_{ej} \sim \text{HN}(\alpha_j)$ for $j = 1, 2$, where the hyperparameters α_1 and α_2 are known. The priors for other parameters in the model, namely, $\boldsymbol{\gamma}$ and ψ^2, are as in (4.4). As in (4.5), the posterior density can be expressed as

$$\begin{aligned}
p(\boldsymbol{\gamma}, \psi^2, \sigma_{e1}, \sigma_{e2}|\mathbf{y}) &\propto p(\mathbf{y}|\boldsymbol{\gamma}, \sigma_{e1}, \sigma_{e2})p(\boldsymbol{\gamma}|\psi^2)p(\psi^2)p(\sigma_{e1})p(\sigma_{e2}) \\
&\propto \frac{1}{|\mathbf{R}|^{1/2}} \exp\left\{-\frac{1}{2}(\mathbf{y} - \mathbf{C}\boldsymbol{\gamma})'\mathbf{R}^{-1}(\mathbf{y} - \mathbf{C}\boldsymbol{\gamma})\right\} \frac{1}{|\mathbf{G}|^{1/2}} \exp\left\{-\frac{1}{2}\boldsymbol{\gamma}'\mathbf{G}^{-1}\boldsymbol{\gamma}\right\} \\
&\quad \times (\psi^2)^{-A_\psi - 1} \exp\left\{-\frac{B_\psi}{\psi^2}\right\} \exp\left\{-\frac{\sigma_{e1}^2}{2\alpha_1^2}\right\} \exp\left\{-\frac{\sigma_{e2}^2}{2\alpha_2^2}\right\}, \qquad (4.11)
\end{aligned}$$

where $\mathbf{R} = \text{diag}\{\sigma_{e1}^2\mathbf{I}_{n\times n}, \sigma_{e2}^2\mathbf{I}_{n\times n}\}$. Since this distribution is not in a closed form, we next derive the full conditional distributions so that we can employ an MCMC algorithm for posterior simulation.

Proceeding along the lines of Section 4.2.1, we can see that the full conditionals of $\boldsymbol{\gamma}$ and ψ^2 are given by (4.6) and (4.7), respectively. Moreover, the full conditional density for $(\sigma_{e1}, \sigma_{e2})$ can be written as

$$p(\sigma_{e1}, \sigma_{e2}|\boldsymbol{\gamma}, \psi^2, \mathbf{y}) \propto \frac{1}{|\mathbf{R}|^{1/2}} \exp\left\{-\frac{1}{2}(\mathbf{y} - \mathbf{C}\boldsymbol{\gamma})'\mathbf{R}^{-1}(\mathbf{y} - \mathbf{C}\boldsymbol{\gamma})\right\} \exp\left\{-\frac{\sigma_{e1}^2}{2\alpha_1^2}\right\} \exp\left\{-\frac{\sigma_{e2}^2}{2\alpha_2^2}\right\}$$
$$= \prod_{j=1}^{2} \frac{1}{\sigma_{ej}^n} \exp\left\{-\frac{1}{2\sigma_{ej}^2} \sum_{i=1}^{n} (y_{ij} - b_i - \beta_j)^2\right\} \exp\left\{-\frac{\sigma_{ej}^2}{2\alpha_j^2}\right\}.$$

To sample from this distribution, we transform $\lambda_j = \log(\sigma_{ej})$, $j = 1, 2$, obtaining the full conditional distribution of (λ_1, λ_2) as

$$p(\lambda_1, \lambda_2|\boldsymbol{\gamma}, \psi^2, \mathbf{y}) \propto \prod_{j=1}^{2} \exp\left\{(1-n)\lambda_j - \frac{1}{2\exp(2\lambda_j)} \sum_{i=1}^{n} (y_{ij} - b_i - \beta_j)^2 - \frac{\exp(2\lambda_j)}{2\alpha_j^2}\right\}. \tag{4.12}$$

We will use a Metropolis algorithm to sample from this distribution. Thus, the algorithm for simulating from posterior of $(\boldsymbol{\gamma}, \psi^2, \sigma_{e1}, \sigma_{e2})$ is as follows.

Algorithm 3: Begin with a starting point $(\boldsymbol{\gamma}_0, \psi_0^2, \sigma_{e1,0}, \sigma_{e2,0})$ and iterate the following until convergence. For $t = 1, 2, \ldots$

1. Draw $\boldsymbol{\gamma}_t|\psi_{t-1}^2, \sigma_{e1,t-1}, \sigma_{e2,t-1}, \mathbf{y}$ from the full conditional in (4.6).

2. Draw $\psi_t^2|\boldsymbol{\gamma}_t, \sigma_{e1,t-1}, \sigma_{e2,t-1}, \mathbf{y}$ from the full conditional in (4.7).

3. Draw $(\sigma_{e1,t}, \sigma_{e2,t})|\boldsymbol{\gamma}_t, \psi_t^2, \mathbf{y}$ from the full conditional in (4.8) using a Metropolis algorithm with a bivariate normal proposal distribution. This involves drawing a proposal $(\lambda_1^*, \lambda_2^*) \sim N_2(\mathbf{0}, \text{diag}(0.05, 0.07))$, calculating the ratio of densities,

$$r = \frac{p(\lambda_1^*, \lambda_2^*|\boldsymbol{\gamma}_t, \psi_t^2, \mathbf{y})}{p(\lambda_{1,t-1}, \lambda_{2,t-1}|\boldsymbol{\gamma}_t, \psi_t^2, \mathbf{y})}$$

setting

$$(\lambda_{1,t}, \lambda_{2,t}) = \begin{cases} (\lambda_1^*, \lambda_2^*), & \text{with probability } \min\{r, 1\}, \\ (\lambda_{1,t-1}, \lambda_{2,t-1}), & \text{otherwise,} \end{cases}$$

and taking $\sigma_{j,t} = \exp\{\lambda_{j,t}\}$, $j = 1, 2$.

This algorithm requires specification of the hyperparameters $V_1^2, V_2^2, A_\psi, B_\psi, \alpha_1$, and α_2 and also the starting points. An illustration is provided in Section 4.5.

4.3.2 Hierarchical half-normal priors

In this section, we assume that the error SDs follow independent HHN priors, $\sigma_{ej} \sim$ HHN $(A_{\alpha_j}, B_{\alpha_j})$, $j = 1, 2$. This is equivalent to assuming that $\sigma_{ej}|\alpha_j^2 \sim \mathrm{HN}(\alpha_j)$, $\alpha_j^2 \sim \mathrm{IG}(A_{\alpha_j}, B_{\alpha_j})$, $j = 1, 2$. The model parameters in this case consist of $(\boldsymbol{\gamma}, \psi^2, \sigma_{e1}, \sigma_{e2}, \alpha_1^2, \alpha_2^2)$. As in (4.9), the posterior density can be expressed as

$$\begin{aligned} p(\boldsymbol{\gamma}, \psi^2, \sigma_{e1}, \sigma_{e2}, \alpha_1^2, \alpha_2^2|\mathbf{y}) &\propto p(\mathbf{y}|\boldsymbol{\gamma}, \sigma_{e1}, \sigma_{e2})p(\boldsymbol{\gamma}|\psi^2)p(\psi^2)p(\sigma_{e1}|\alpha_1^2)p(\sigma_{e2}|\alpha_2^2)p(\alpha_1^2)p(\alpha_2^2) \\ &\propto \frac{1}{|\mathbf{R}|^{1/2}} \exp\left\{-\frac{1}{2}(\mathbf{y}-\mathbf{C}\boldsymbol{\gamma})'\mathbf{R}^{-1}(\mathbf{y}-\mathbf{C}\boldsymbol{\gamma})\right\} \\ &\quad \times \frac{1}{|\mathbf{G}|^{1/2}} \exp\left\{-\frac{1}{2}\boldsymbol{\gamma}'\mathbf{G}^{-1}\boldsymbol{\gamma}\right\} (\psi^2)^{-A_\psi - 1} \exp\left\{-\frac{B_\psi}{\psi^2}\right\} \\ &\quad \times \frac{1}{\alpha_1} \exp\left\{-\frac{\sigma_{e1}^2}{2\alpha_1^2}\right\} \frac{1}{\alpha_2} \exp\left\{-\frac{\sigma_{e2}^2}{2\alpha_2^2}\right\} \\ &\quad \times (\alpha_1^2)^{-A_{\alpha_1}-1} \exp\left\{-\frac{B_{\alpha_1}}{\alpha_1^2}\right\} (\alpha_2^2)^{-A_{\alpha_2}-1} \exp\left\{-\frac{B_{\alpha_2}}{\alpha_2^2}\right\}, \end{aligned} \quad (4.13)$$

with $\mathbf{R} = \mathrm{diag}\{\sigma_{e1}^2\mathbf{I}_{n\times n}, \sigma_{e2}^2\mathbf{I}_{n\times n}\}$. Next, we derive the full conditional distributions for posterior simulation via MCMC. The full conditionals for $\boldsymbol{\gamma}$, ψ^2, σ_{e1}, and σ_{e2} are same as in the case of HN prior. The full conditional density of (α_1^2, α_2^2) can be written as

$$p(\alpha_1^2, \alpha_2^2|\boldsymbol{\gamma}, \psi^2, \sigma_{e1}, \sigma_{e2}, \mathbf{y}) \propto \prod_{j=1}^{2} (\alpha_j^2)^{-(A_{\alpha_j}+1/2)-1} \exp\left\{-\left(\frac{\sigma_{ej}^2 + 2B_{\alpha_j}}{2\alpha_j^2}\right)\right\},$$

implying that

$$\alpha_j^2|\boldsymbol{\gamma}, \psi^2, \sigma_{e1}, \sigma_{e2}, \mathbf{y} \sim \mathrm{IG}\left(A_{\alpha_j} + 1/2, (\sigma_{ej}^2 + 2B_{\alpha_j})/2\right), \; j = 1, 2, \quad (4.14)$$

and these distributions are independent. This leads to the following posterior simulation algorithm for $(\boldsymbol{\gamma}, \psi^2, \sigma_{e1}, \sigma_{e2}, \alpha_1^2, \alpha_2^2)$.

Algorithm 4: Begin with a starting point $(\boldsymbol{\gamma}_0, \psi_0^2, \sigma_{e1,0}, \sigma_{e2,0}, \alpha_{1,0}^2, \alpha_{2,0}^2)$ and iterate the following until convergence. For $t = 1, 2, \ldots$

1. Draw $\boldsymbol{\gamma}_t|\psi^2_{t-1}, \sigma_{e1,t-1}, \sigma_{e2,t-1}, \alpha^2_{1,t-1}, \alpha^2_{2,t-1}, \mathbf{y}$ from the full conditional in (4.6).

2. Draw $\psi^2_t|\boldsymbol{\gamma}_t, \sigma_{e1,t-1}, \sigma_{e2,t-1}, \alpha^2_{1,t-1}, \alpha^2_{2,t-1}, \mathbf{y}$ from the full conditional in (4.7).

3. Draw $(\sigma_{e1,t}, \sigma_{e2,t})|\boldsymbol{\gamma}_t, \psi^2_t, \alpha^2_{1,t-1}, \alpha^2_{2,t-1}, \mathbf{y}$ as in Algorithm 3.

4. Draw $(\alpha^2_{1,t}, \alpha^2_{2,t})|\boldsymbol{\gamma}_t, \psi^2_t, \sigma_{e1,t}, \sigma_{e2,t}, \mathbf{y}$ independently from full conditionals in (4.14).

This algorithm requires specification of the hyperparameters $V_1^2, V_2^2, A_\psi, B_\psi, A_{\alpha_1}$, B_{α_1}, A_{α_2}, and B_{α_2} and also the starting points. An illustration is provided in Section 4.5.

For posterior simulation, we follow the recommendations in (Gelman et al., 2013, Chapter 11) to run three MCMC chains with overdispersed starting points. The convergence is diagnosed using Gelman-Rubin potential scale reduction factors and trace plots. In case of a Metropolis algorithm, the parameters of the proposal distribution are tuned to get the acceptance rates near the recommendation in (Gelman et al., 2013, Chapter 11).

4.4 Evaluation of similarity and agreement

We use measures of similarity and agreement under the linear mixed-effects model to evaluate similarity and agreement (see also Section 2.4). Two widely accepted measure are mean difference $\beta_1 - \beta_2$ and precision ratio $\sigma^2_{e2}/\sigma^2_{e1}$. Mean difference near zero and precision ratio near one indicate that two methods have similar marginal characteristics. In the classical approach to inference, similarity is evaluated using two-sided confidence intervals for the similarity measures. In a Bayesian approach, we can use their two-sided posterior credible intervals.

To evaluate the agreement between two methods, we use two common agreement measures, CCC and TDI. They are defined as follows:

$$CCC = \frac{2Cov(Y_1, Y_2)}{\{E(Y_1 - Y_2)\}^2 + Var(Y_1) + Var(Y_2)}.$$

$$TDI(p) = p^{th} \text{ percentile of } |D| \quad \text{where} \quad D = Y_1 - Y_2$$
$$= \sigma_D \left\{ \chi^2_{1,p} \left\{ \frac{\mu_D^2}{\sigma_D^2} \right\} \right\}^{\frac{1}{2}} \quad \text{if} \qquad D \sim N(\mu_D, \sigma_D^2).$$

Here p is a large probability which need to be specified. The CCC is bounded above by one and TDI is bounded below by zero. Thus, a larger value of CCC and a small value of TDI indicate a good agreement between two methods. In the classical approach to inference, agreement is evaluated using one-sided confidence bounds for these measures, in particular, a lower confidence bound for CCC and an upper confidence bound for TDI. In a Bayesian approach, we can use their posterior percentiles.

4.5 Illustration: Analysis of oxygen saturation data

We now return to the oxygen saturation data introduced in Section 4.1 for comparing two methods — POS (method 1) and OSM (method 2) — for measuring percent saturation of hemoglobin with oxygen. One of the first tasks in analyzing them using the Bayesian methods of this chapter is to specify the hyperparameters of the prior distributions. Since we plan to use non-informative priors for β_1, β_2, and ψ^2, we set their hyperparameters $(V_1^2, V_2^2, A_\psi, B_\psi)$ equal to $(10^6, 10^6, 10^{-2}, 10^{-2})$, resulting in densities that are essentially flat in the regions where the true parameter values are likely. For the error SDs, we plan to use informative priors. We next describe their elicitation using estimates reported in the literature. Since it is difficult to find articles that study both OSM and POS methods, we decided to assume a common prior distribution for both error SDs. Of course, this issue is relevant only when the error SDs are allowed to be different in the model, not when they are assumed to be equal.

4.5.1 Elicitation of hyperparameters

Table 4.2 presents error SD estimates obtained from 9 studies. The first two studies directly report the estimate. However, the remaining 7 provide estimates of SD of difference in errors

Table 4.2. Error SD estimates obtained from the literature. The first two studies directly report the error SD estimate. However, the rest provide estimates of SD of difference in errors of two methods. The error SD estimate is imputed from them assuming that the errors are independent and have equal variances.

Source	Error SD estimate
(James et al., 2006)	0.870, 1.195, 1.061
(Phattraprayoon et al., 2012)	0.664, 0.615, 0.658, 0.658
(Nickerson et al., 1988)	2.121
(Jubran and Tobin, 1990)	0.848, 1.909
(Das et al., 2010)	0.892, 0.771, 0.726, 1.118, 1.597, 0.758, 1.016, 1.230, 1.100, 1.211, 0716, 1.211
(Hannhart et al., 1991)	0.919, 0.990, 1.061, 1.131, 1.343, 1.343, 1.484, 1.484, 1.556, 1.556, 1.626, 1.768, 1.838, 2.121, 2.263, 1.980, 1.768, 1.273, 2.333, 1.626, 1.626, 3.606, 2.121, 1.273, 2.333, 2.616, 1.838, 3.253
(Desiderio et al., 1990)	1.331, 0.718
(Al Khudhairi et al., 1990)	0.920, 1.130, 0.800, 1.090, 1.040, 0.930, 0.600, 0.900, 0.520, 0.520, 0.520, 0.550
(Praud et al., 1989)	1.061, 1.697, 3.111

of two methods. From these, we compute the error SD estimate assuming that the errors have equal variances. In other words, the error SD estimate is equal to the estimated SD of difference in the errors divided by $\sqrt{2}$. We treat the estimates in Table 4.2 as a random sample from the assumed prior distribution for error SD and use the ML method to estimate its parameters. The estimated values are taken as the hyperparameters.

Half-normal prior

Assuming that the estimates reported in Table 4.2 are a random sample from an HN(α) distribution, from Section 1.2.3, the ML estimate of α is 1.34 — the square-root of the second sample moment of the observations. Thus, the hyperparameters of the HN distributions are taken to be $\alpha = \alpha_1 = \alpha_2 = 1.34$. The density function of HN(1.34) distribution is plotted in Figure 4.2. It provides support to lower error variance values. It has a mean of 1.069.

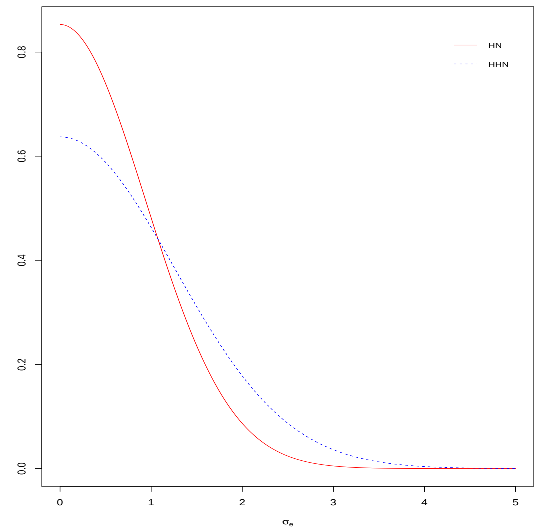

Figure 4.2. The probability density functions of the elicited prior distributions, HN(1.34) and HHN(10509.13, 16481.6).

Hierarchical half-normal prior

In case of a random sample from an HHN(A_α, B_α) distribution, the ML estimates are not available in closed-form. They have to be obtained via numerical maximization. This is the approach we use here. Specifically, assuming that the estimates reported in Table 4.2 are a random sample from this distribution, the ML estimates of (A_α, B_α) are $(10509.13, 16481.6)$. Therefore, the hyperparameters in this case are taken to be $\hat{A}_\alpha = \hat{A}_{\alpha_1} = \hat{A}_{\alpha_2} = 10509.13$ and $\hat{B}_\alpha = \hat{B}_{\alpha_1} = \hat{B}_{\alpha_2} = 16481.6$. The density function of HHN$(10509.13, 16481.6)$ distribution is also plotted in Figure 4.2. It also provides support to lower error variance values and its its mean is 0.997. Its tail is heavier than that of the half-normal distribution.

4.5.2 Model fitting

To fit a model, we run three chains with starting points as the ML estimates of the parameters involved and limits of their 95% confidence intervals given by the likelihood-based analysis. We use the `monitor` function in `R` package `RStan` (Stan Development Team, 2018) to diagnose

Table 4.3. Point and 95% interval estimates for selected model parameters computed using the ML approach and two Bayesian approaches assuming equal error variances. The intervals for CCC and TDI are one-sided. CI refers to "confidence interval" for the ML approach and "credible interval" for the Bayesian approach.

	ML approach		Bayesian approach			
			HN		HHN	
	Estimate	95% CI	Estimate	95% CI	Estimate	95% CI
β_1	89.08	$(87.08, 91.09)$	89.07	$(87.04, 91.08)$	89.08	$(87.01, 91.14)$
β_2	89.50	$(87.49, 91.50)$	89.48	$(87.44, 91.50)$	89.49	$(87.42, 91.56)$
$\log(\psi)$	2.16	(1.99, 2.32)	2.17	$(2.01, 2.34)$	2.17	$(2.01, 2.34)$
$\log(\sigma_e)$	-0.16	$(-0.32, 0.00)$	-0.15	$(-0.30, 0.02)$	-0.14	$(-0.29, 0.02)$
α	-	-	-	-	1.57	$(1.54, 1.60)$
$\beta_1 - \beta_2$	-0.42	(-0.69, -0.13)	-0.41	$(-0.70, -0.13)$	-0.41	$(-0.70, -0.13)$
CCC	0.99	(0.98, 1.00)	0.989	$(0.984, 1.000)$	0.989	$(0.984, 1.000)$
TDI	2.10	(0.00, 2.37)	2.140	$(0.000, 2.456)$	2.148	$(0.000, 2.459)$

convergence. It automatically throws away the first half of each chain and keeps only the second half.

4.5.3 Assuming equality of error variances

We first fit the model (4.1) assuming a common error variance σ_e^2 and using the HN and HHN priors for σ_e. The priors for the remaining model parameters are given by (4.2). In both cases, three chains were run for 6000 iterations. The variance of the normal proposal distribution for $\lambda = \log(\sigma_e)$ was set at 0.05 as it provided acceptance rates in line with the recommendation of (Gelman et al., 2013, Chapter 11). The values of Gelman-Rubin scale reduction factor for all parameters were under 1.01 for both priors, indicating convergence. Table 4.3 provides point and 95% interval estimates for key model parameters and their functions obtained using the classical likelihood-based approach and the two Bayesian approaches. For a Bayesian approach, the posterior mean is presented as the point estimate and the 95% central region is presented as the credible interval.

The ML estimates for model parameters in Table 4.3 are reproduced from Table 4.1. We see that the results for all three methods are remarkably similar.

The prior and posterior means for error SD using HN prior are 1.069 and $\exp(-0.15) = 0.86$ respectively. Similarly, the prior mean for error SD using HHN prior is 1.00 and posterior mean is $\exp(-0.14) = 0.87$. Thus, the prior mean and posterior mean differ from each other regardless of the choice of the prior. This indicates that the posterior distribution is dominated by information provided by the data.

4.5.4 Not assuming equality of error variances

In this section, we discuss the results when no assumption regarding equality of error variances is made. For both HN and HHN priors, three chains were run for 60,000 iterations. In the Metropolis algorithm, a bivariate normal distribution with variances (0.05, 0.07) and correlation zero was used as the proposal distribution for $(\lambda_1, \lambda_2) = (\log(\sigma_{e1}), \log(\sigma_{e2}))$. The values of Gelman-Rubin scale reduction factor for all parameters were under 1.04 in case of HN prior and under 1.09 in case of HHN prior, indicating convergence. Table 4.4 provides point and 95% interval estimates for the quantities in Table 4.3 and some additional quantities, including $\log(\sigma_{e1}/\sigma_{e2})$.

The ML estimates for model parameters in Table 4.4 are reproduced from Table 4.1. We see that the results for β_1, β_2, $\beta_2 - \beta_1$ and $\log\psi$ are similar across the three methods. Further, the two Bayesian methods give similar results for the error variances, but they differ markedly from the results for the ML method, which are known to be unrealistic and unreliable from Section 4.1. Specifically, the Bayesian estimates of $\log(\sigma_{e1})$ and $\log(\sigma_{e2})$ around -0.2 and -0.5, respectively, appear much more reasonable than the ML estimates of -0.9 and 0.1, respectively. Moreover, unlike the case of ML method, none of the Bayesian interval estimates is unrealistically wide. Thus, the Bayesian methods have succeeded in the goal of obtaining reliable estimates of error variances by bringing in additional information through informative prior distributions for the error SDs. Moreover, the fact that the two prior distributions lead to similar results indicates that the estimates are not overly sensitive to the choice of the prior distribution.

Table 4.4. Point and 95% interval estimates for selected model parameters computed using the ML approach and two Bayesian approaches with no assumption of equality of error variances. The intervals for CCC and TDI are one-sided. CI refers to "confidence interval" for the ML approach and "credible interval" for the Bayesian approach.

			Bayesian approach			
	ML approach		HN		HHN	
	Estimate	95% CI	Estimate	95% CI	Estimate	95% CI
β_1	89.08	$(87.07, 91.09)$	89.09	$(87.03, 91.14)$	89.08	$(87.04, 91.14)$
β_2	89.50	$(87.50, 91.49)$	89.50	$(87.44, 91.55)$	89.49	$(87.45, 91.55)$
$\log(\psi)$	2.16	$(1.99, 2.32)$	2.17	$(2.01, 2.34)$	2.17	$(2.01, 2.34)$
$\log(\sigma_{e1})$	-0.95	$(-1.78, -0.12)$	-0.19	$(-1.56, 0.29)$	-0.22	$(-1.65, 0.30)$
$\log(\sigma_{e2})$	0.13	$(-5.46, 5.71)$	-0.49	$(-2.29, 0.24)$	-0.49	$(-2.69, 0.26)$
α_1^2	-	-	-	-	1.57	$(1.54, 1.60)$
α_2^2	-	-	-	-	1.57	$(1.54, 1.60)$
$\beta_1 - \beta_2$	-0.42	(-0.69, -0.13)	-0.41	$(-0.70, -0.13)$	-0.41	$(-0.70, -0.13)$
$\log(\sigma_{e2}/\sigma_{e1})$	2.16	(-46.67, 51.00)	-0.58	$(-4.98, 3.50)$	-0.55	$(-5.78, 3.66)$
CCC	0.99	(0.66, 1.00)	0.99	$(0.98, 1.00)$	0.99	$(0.98, 1.00)$
TDI	2.10	(0.00, 34.12)	2.15	$(0.00, 2.47)$	2.15	$(0.00, 2.47)$

The Bayesian approach, however, is not a panacea for the lack of replicated measurements which are ideal for estimation of error variances. This is clear from Table 4.4 where we see that although the estimates of error variances are reliable but the same cannot be said for their ratio — a key measure of similarity — because the interval estimate of $\log(\sigma_{e2}/\sigma_{e1})$ ranges from zero to more than 33. Thus, on the whole, the Bayesian approach with informative priors only partly ameliorates the issue of lack of replicated measurements.

Next, we discuss the similarity and agreement evaluation. Tables 4.3 and 4.4 provide estimates and confidence intervals for mean difference and agreements measures CCC and TDI. The estimate and the confidence interval for mean difference, -0.42 and (-0.69, -0.13) are similar for both classical and Bayesian approaches. This implies that the mean for POS is smaller than the OSM and confirms the result we obtain in explanatory data analysis. Under equal error variance assumption, all three approaches provide the same CCC estimates and lower bounds 0.99 and 0.98 respectively. This indicates that there is a strong agreement between two methods. TDI estimates and upper bounds under equal error variance assumption are similar for two Bayesian approaches and slightly greater than that of ML method.

Obviously, the inference on mean difference does not depend on the equal error variance assumption. Even though, the estimates of CCC and TDI remain unchanged under unequal error variances, the corresponding lower/upper bounds differ significantly compared to the values under equal error variances. Further, ML method has much wider confidence bounds for CCC and TDI compared to two Bayesian approaches. This indicates that Bayesian approaches provide more reasonable confidence bounds for agreement measures than the ML approach.

4.6 A Simulation study

In this section, we perform a Monte Carlo simulation study to evaluate properties of the Bayesian approaches based on HN and HHN priors and compare them with the ML approach. We are specifically interested in determining bias and mean squared error (MSE) of point estimates of model parameters as well as measures of similarity and agreement using the three approaches and comparing them. We are also interested in evaluating coverage accuracy of the Bayesian interval estimates. For a Bayesian approach, the posterior mean is taken as the point estimator and the central credible region is taken as the interval estimator.

The data are generated from the model (4.1) with parameter values motivated by the analysis of oxygen saturation data. In particular, we set the fixed effects $(\beta_1, \beta_2) = (89, 90)$ and the random effect variance $\psi^2 = 78$. To study the impact of equal error variance assumption, the data are simulated assuming both equal and unequal variances and the models are also fit assuming both equal and unequal variances. The prior distributions used for model fitting are identical to those assumed for the data analysis. Thus, for HN priors, $\alpha = \alpha_1 = \alpha_2 = 1.34$ is taken and for HHN priors, $A_\alpha = A_{\alpha_1} = A_{\alpha_2} = 10509.13$ and $B_\alpha = B_{\alpha_1} = B_{\alpha_2} = 16481.6$ are taken.

We are also interested in studying what happens when there is a conflict between prior and data distributions with regard to the error SDs. To understand this issue, first consider the case of HN($\alpha = 1.34$) prior. Its density shown in Figure 4.2 shows more support for smaller values of σ_e than its larger values. Therefore, choosing a low percentile of the HN distribution as the true value for error SD in the simulation corresponds to no prior-data conflict. On the other hand, simulating data assuming a high percentile for error SD corresponds to a conflict between data and prior. Thus, to simulate data under the equal variance assumption with no prior-data conflict and prior-data conflict, we respectively choose the 10^{th} and 90^{th} percentiles of the HN(1.34) distribution as the true value for σ_e. Similarly, under the unequal error variance assumption, to simulate data with no prior-data conflict we choose the 5^{th} and 15^{th} percentiles of the HN distribution as the true values of σ_{e1} and σ_{e2}. Also, we choose the 85^{th} and 95^{th} percentiles to simulate data that conflict with prior.

Next, consider the case of HHN ($A_\alpha = 10509.13, B_\alpha = 16481.6$) prior distribution. As can be seen in Figure 4.2, HHN prior also supports smaller error SD values. To determine the true values for error SD under unequal error variances, we first generate two sets of a large number of random draws for α_1^2 and α_2^2 from IG$(10509.13, 16481.6)$. Then we draw two random sets of HN(α_1) and HN(α_2) using the generated α_1^2 and α_2^2 values. Thus, we can consider these as two random samples from the distributions of σ_{e1} and σ_{e2}. We choose 5^{th} percentile of each sample as the true value for error SDs to simulate data with prior-data conflict. Similarly, 95^{th} percentiles of each sample are considered as the true values for σ_{e1} and σ_{e2} to simulate data that has no prior data conflict. The same approach is followed to determine the true value of σ_e under equal error variance. In this case, we choose 5^{th} and 95^{th} percentiles to simulate data with and without prior-data conflict respectively.

Thus, altogether we have 4 broad scenarios — equal and unequal variance in the true model and in the fitted model, and each scenario has two cases — with and without prior-data conflict. These scenarios and the corresponding parameter values are summarized below.

- Scenario 1: Data are generated assuming equal error variances and the model is fit assuming unequal error variances. The two cases are: 1Y and 1N corresponding to with and without prior-data conflict

- Scenario 2: Data are generated and the model is fit assuming equal error variances. The two cases are: 2Y and 2N as in Scenario 1.

- Scenario 3: Data are generated and the model is fit assuming unequal error variances. The two cases are: 3Y and 3N corresponding to with and without conflict.

- Scenario 4: Data are generated assuming unequal error variances and the model is fit assuming equal error variances. The two cases are: 4Y and 4N as in Scenario 3.

Table 4.5 displays the true values for error SDs for each Scenario.

Table 4.5. True values for error SDs for various scenarios.

Scenario	HN			HHN		
	σ_e	σ_{e1}	σ_{e2}	σ_e	σ_{e1}	σ_{e2}
1Y	1.54	-	-	2.46	-	-
1N	0.12	-	-	0.09	-	-
2Y	1.54	-	-	2.46	-	-
2N	0.12	-	-	0.09	-	-
3Y	-	1.35	1.83	-	2.46	2.66
3N	-	0.06	0.18	-	0.09	0.07
4Y	-	1.35	1.83	-	2.46	2.66
4N	-	0.06	0.18	-	0.09	0.07

Note that the other parameter values, namely, $(\beta_1, \beta_2, \psi^2)$ are set at $(89, 90, 78)$ in all scenarios. For each scenario, the data are simulated with number of subjects $n \in (20, 30, 60, 100)$, and we take $1 - \alpha = 0.95$ for the interval estimate. Bias and MSE of point estimates for all three approaches and coverage probability of Bayesian interval estimates are computed using 500 Monte Carlo replications. To compare a Bayesian point estimator (the posterior mean) with the ML estimator, we compute their relative efficiency by dividing the MSE of the ML estimator by the MSE of the posterior mean. Thus, an efficiency greater than 1 indicates that the posterior mean is better. Tables 4.6 to 4.9 present results assuming the HN prior and the Tables 4.10 to 4.13 present the corresponding results

Table 4.6. Bias, MSE, and relative efficiency of ML estimators and posterior means and coverage probability of 95% credible intervals in case of Scenarios 1Y and 1N under a half-normal prior for error SD.

n	Parameter	True value	ML Bias	ML MSE	Bayes Bias	Bayes MSE	MSE ratio	Cov. prob.
				Scenario 1Y				
20	β_1	89.00	0.81	3.67	-0.10	3.23	1.14	0.92
	β_2	90.00	-1.10	4.42	-0.19	3.04	1.45	0.92
	$\log\psi$	2.18	-0.07	0.03	-0.01	0.02	1.22	0.94
	$\log\sigma_{e1}$	0.43	-2.38	> 10	-0.44	0.28	> 10	1.00
	$\log\sigma_{e2}$	0.43	-2.74	> 10	-0.38	0.24	> 10	0.98
	CCC	0.96	0.00	0.00	0.00	0.00	1.13	0.90
	TDI	3.94	-0.21	0.65	-0.02	0.52	1.24	0.86
	$\beta_1 - \beta_2$	-1.00	0.08	0.28	0.08	0.28	1.00	0.92
	$\log(\sigma_{e1}^2/\sigma_{e2}^2)$	0.00	-0.07	> 10	0.13	0.97	> 10	1.00
100	β_1	89.00	1.08	3.07	0.10	1.66	1.86	0.98
	β_2	90.00	-0.90	2.46	0.08	1.92	1.28	0.98
	$\log\psi$	2.18	-0.04	0.02	0.00	0.02	1.06	0.92
	$\log\sigma_{e1}$	0.43	-2.50	> 10	-0.36	0.27	> 10	1.00
	$\log\sigma_{e2}$	0.43	-2.37	> 10	-0.34	0.23	> 10	1.00
	CCC	0.96	0.00	0.00	0.00	0.00	1.01	0.98
	TDI	3.94	-0.05	0.07	0.00	0.06	1.06	1.00
	$\beta_1 - \beta_2$	-1.00	0.03	0.04	0.03	0.045	1.00	0.92
	$\log(\sigma_{e1}^2/\sigma_{e2}^2)$	0.00	0.17	> 10	0.06	3.74	23.24	0.96
				Scenario 1N				
20	β_1	89.00	0.93	6.23	-0.08	5.39	1.16	0.86
	β_2	90.00	-1.08	6.55	-0.08	5.38	1.22	0.86
	$\log\psi$	2.18	-0.06	0.03	0.00	0.02	1.14	0.96
	$\log\sigma_{e1}$	-2.12	-3.32	> 10	-0.24	0.09	> 10	1.00
	$\log\sigma_{e2}$	-2.12	-3.93	> 10	-0.25	0.09	> 10	1.00
	CCC	0.99	0.00	0.00	0.00	0.00	1.20	0.96
	TDI	1.22	0.00	0.00	0.03	0.00	0.65	1.00
	$\beta_1 - \beta_2$	-1.00	0.00	0.00	0.00	0.00	1.00	0.98
100	β_1	89.00	0.86	1.28	-0.14	0.55	2.34	0.96
	β_2	90.00	-1.14	1.82	-0.14	0.55	3.30	0.96
	$\log\psi$	2.18	-0.02	0.00	-0.01	0.00	1.05	0.94
	$\log\sigma_{e1}$	-2.12	-2.11	> 10	-0.38	0.20	> 10	1.00
	$\log\sigma_{e2}$	-2.12	-4.21	> 10	-0.30	0.12	> 10	1.00
	CCC	0.99	0.00	0.00	0.00	0.00	1.06	1.00
	TDI	1.22	0.00	0.00	0.00	0.00	0.98	0.96
	$\beta_1 - \beta_2$	-1.00	0.00	0.00	0.00	0.00	1.00	0.94

Table 4.7. Bias, MSE, and relative efficiency of ML estimators and posterior means and coverage probability of 95% credible intervals in case of Scenarios 2Y and 2N under a half-normal prior for error SD.

n	Parameter	True value	ML Bias	ML MSE	Bayes Bias	Bayes MSE	MSE ratio	Cov. prob.
Scenario 2Y								
20	β_1	89.00	0.81	3.67	-0.10	3.23	1.14	0.92
	β_2	90.00	-1.10	4.42	-0.19	3.04	1.45	0.92
	$\log\psi$	2.18	-0.06	0.03	-0.01	0.02	1.14	0.96
	$\log\sigma_e$	0.43	-0.09	0.06	-0.06	0.04	1.27	0.86
	CCC	0.96	0.00	0.00	-0.01	0.00	0.94	0.88
	TDI	3.94	-0.21	0.65	0.01	0.58	1.11	0.82
	$\beta_1 - \beta_2$	-1.00	0.08	0.28	0.09	0.28	1.00	0.92
100	β_1	89.00	-0.13	0.78	0.84	1.42	1.82	0.94
	β_2	90.00	-0.16	0.76	-1.13	2.05	2.71	0.94
	$\log\psi$	2.18	0.01	0.00	0.00	0.00	0.98	0.98
	$\log\sigma_e$	0.43	-0.01	0.00	-0.01	0.00	1.09	0.92
	CCC	0.96	0.00	0.00	0.00	0.00	1.00	0.98
	TDI	3.94	-0.05	0.07	0.00	0.07	1.05	1.00
	$\beta_1 - \beta_2$	-1.00	0.03	0.05	0.03	0.05	1.00	0.92
Scenario 2N								
20	β_1	89.00	0.93	6.23	-0.08	5.39	1.16	0.86
	β_2	90.00	-1.08	6.55	-0.07	5.38	1.22	0.86
	$\log\psi$	2.18	-0.05	0.03	0.00	0.02	1.13	0.94
	$\log\sigma_e$	-2.12	-0.02	0.03	0.06	0.03	0.85	0.90
	CCC	0.99	0.00	0.00	0.00	0.00	1.19	0.94
	TDI	1.22	0.00	0.00	0.02	0.00	0.73	0.96
	$\beta_1 - \beta_2$	-1.00	0.00	0.00	0.00	0.00	0.99	0.98
100	β_1	89.00	0.86	1.28	-0.14	0.55	2.34	0.98
	β_2	90.00	-1.14	1.82	-0.13	0.55	3.29	0.96
	$\log\psi$	2.18	-0.02	0.00	-0.01	0.00	1.04	0.94
	$\log\sigma_e$	-2.12	-0.02	0.00	0.00	0.00	1.05	0.96
	CCC	0.99	0.00	0.00	0.00	0.00	1.04	1.00
	TDI	1.22	0.00	0.00	0.00	0.00	1.00	0.96
	$\beta_1 - \beta_2$	-1.00	0.00	0.00	0.00	0.00	1.00	0.94

Table 4.8. Bias, MSE, and relative efficiency of ML estimators and posterior means and coverage probability of 95% credible intervals in case of Scenarios 3Y and 3N under a half-normal prior for error SD.

			ML		Bayes			
n	Parameter	True value	Bias	MSE	Bias	MSE	MSE ratio	Cov. prob.
			Scenario 3Y					
20	β_1	89.00	0.83	3.68	-0.03	3.13	1.18	0.96
	β_2	90.00	-1.04	4.20	-0.17	3.02	1.39	0.96
	$\log\psi$	2.18	-0.09	0.03	-0.03	0.03	1.31	0.96
	$\log\sigma_{e1}$	0.30	-2.04	> 10	-0.21	0.16	> 10	1.00
	$\log\sigma_{e2}$	0.60	-2.75	> 10	-0.52	0.36	> 10	1.00
	CCC	0.96	-0.01	0.00	-0.01	0.00	1.01	0.98
	TDI	4.08	-0.20	0.31	-0.02	0.23	1.36	0.94
	$\beta_1 - \beta_2$	-1.00	0.13	0.22	0.13	0.22	1.00	0.96
	$\log(\sigma_{e1}^2/\sigma_{e2}^2)$	0.61	-1.42	180.90	-0.62	1.76	> 10	1.00
100	β_1	89.00	1.06	1.97	0.05	0.77	2.54	0.96
	β_2	90.00	-0.95	1.67	0.06	0.84	1.98	0.96
	$\log\psi$	2.18	-0.02	0.00	-0.01	0.00	1.10	0.96
	$\log\sigma_{e1}$	0.30	0.11	1.50	-0.41	0.40	3.77	1.00
	$\log\sigma_{e2}$	0.60	-2.73	> 10	-0.20	0.17	> 10	0.98
	CCC	0.96	0.00	0.00	0.00	0.00	0.99	0.90
	TDI	4.08	0.03	0.07	0.07	0.07	0.97	0.92
	$\beta_1 - \beta_2$	-1.00	-0.01	0.06	-0.01	0.06	1.00	0.96
	$\log(\sigma_{e1}^2/\sigma_{e2}^2)$	0.61	-5.67	> 10	0.41	2.65	> 10	1.00
			Scenario 3N					
20	β_1	89.00	1.09	5.41	0.09	4.20	1.29	0.96
	β_2	90.00	-0.91	5.02	0.09	4.23	1.19	0.96
	$\log\psi$	2.18	-0.05	0.03	0.01	0.03	1.07	0.92
	$\log\sigma_{e1}$	-2.81	-2.52	> 10	0.52	0.30	> 10	1.00
	$\log\sigma_{e2}$	-1.71	-4.26	> 10	-0.57	0.35	> 10	1.00
	CCC	0.99	0.00	0.00	0.00	0.00	1.15	0.96
	TDI	1.24	-0.01	0.00	0.02	0.00	0.77	1.00
	$\beta_1 - \beta_2$	-1.00	0.00	0.00	0.00	0.00	1.00	0.96
	$\log(\sigma_{e1}^2/\sigma_{e2}^2)$	2.20	-3.50	> 10	-2.18	4.86	> 10	1.00
100	β_1	89.00	0.90	1.67	-0.10	0.87	1.92	0.96
	β_2	90.00	-1.10	2.07	-0.10	0.87	2.38	0.96
	$\log\psi$	2.18	-0.01	0.01	0.00	0.01	1.03	0.94
	$\log\sigma_{e1}$	-2.81	-1.44	> 10	0.50	0.29	> 10	1.00
	$\log\sigma_{e2}$	-1.71	-3.85	> 10	-0.59	0.38	> 10	1.00
	CCC	0.99	0.00	0.00	0.00	0.00	1.04	0.94
	TDI	1.24	0.00	0.00	0.01	0.00	0.89	0.96
	$\beta_1 - \beta_2$	-1.00	0.00	0.00	0.00	0.00	1.00	0.98
	$\log(\sigma_{e1}^2/\sigma_{e2}^2)$	2.20	-4.82	> 10	-2.19	5.19	> 10	1.00

Table 4.9. Bias, MSE, and relative efficiency of ML estimators and posterior means and coverage probability of 95% credible intervals in case of Scenarios 4Y and 4N under a half-normal prior for error SD.

			ML		Bayes			
n	Parameter	True value	Bias	MSE	Bias	MSE	MSE ratio	Cov. prob.
			Scenario 4Y					
20	β_1	89.00	0.83	3.68	-0.03	3.13	1.18	0.96
	β_2	90.00	-1.04	4.20	-0.17	3.02	1.39	0.96
	$\log\psi$	2.18	-0.08	0.03	-0.03	0.02	1.20	0.96
	$\log\sigma_{e1}$	0.30	0.11	0.04	0.15	0.04	0.84	0.84
	$\log\sigma_{e2}$	0.60	-0.20	0.06	-0.16	0.05	1.35	0.86
	CCC	0.96	0.00	0.00	-0.01	0.00	0.80	0.98
	TDI	4.08	-0.20	0.31	0.03	0.27	1.17	0.94
	$\beta_1-\beta_2$	-1.00	0.13	0.22	0.13	0.22	1.00	0.96
100	β_1	89.00	1.06	1.97	0.05	0.77	2.55	0.96
	β_2	90.00	-0.95	1.67	0.06	0.84	1.99	0.96
	$\log\psi$	2.18	-0.02	0.00	-0.01	0.00	1.07	0.96
	$\log\sigma_{e1}$	0.30	0.18	0.04	0.18	0.04	0.93	0.22
	$\log\sigma_{e2}$	0.60	-0.13	0.02	-0.12	0.02	1.11	0.62
	CCC	0.96	0.00	0.00	0.00	0.00	0.91	0.90
	TDI	4.08	0.03	0.07	0.08	0.08	0.92	0.92
	$\beta_1-\beta_2$	-1.00	-0.01	0.06	-0.01	0.06	1.00	0.96
			Scenario 4N					
20	β_1	89.00	1.09	5.41	0.09	4.22	1.28	0.96
	β_2	90.00	-0.91	5.02	0.09	4.24	1.18	0.96
	$\log\psi$	2.18	-0.04	0.03	0.01	0.03	1.07	0.92
	$\log\sigma_{e1}$	-2.81	0.76	0.59	0.84	0.72	0.82	0.00
	$\log\sigma_{e2}$	-1.71	-0.34	0.13	-0.26	0.08	1.59	0.74
	CCC	0.99	0.00	0.00	0.00	0.00	1.15	0.96
	TDI	1.24	-0.01	0.00	0.02	0.00	0.85	1.00
	$\beta_1-\beta_2$	-1.00	0.00	0.00	0.00	0.00	1.00	0.94
100	β_1	89.00	0.90	1.67	-0.10	0.86	1.93	0.96
	β_2	90.00	-1.10	2.07	-0.10	0.87	2.39	0.96
	$\log\psi$	2.18	-0.01	0.01	0.00	0.01	1.03	0.94
	$\log\sigma_{e1}$	-2.81	0.82	0.67	0.83	0.70	0.96	0.00
	$\log\sigma_{e2}$	-1.71	-0.28	0.08	-0.26	0.07	1.11	0.08
	CCC	0.99	0.00	0.00	0.00	0.00	1.04	0.94
	TDI	1.24	0.00	0.00	0.01	0.00	0.90	0.96
	$\beta_1-\beta_2$	-1.00	0.00	0.00	0.00	0.00	1.00	0.98

for the HHN prior. In each table, the results are presented for $n = 20, 100$. The results for intermediate values of n, i.e., $n = 30, 60$, are omitted.

The key conclusions from simulation results can be summarized as follows. The Bayesian estimators for $\log \sigma_{e1}$, $\log \sigma_{e2}$ and $\log(\sigma_{e1}/\sigma_{e2})$ are better than ML estimators regardless of the settings for n under the unequal variance assumption. Moreover, in Scenario 2N, the ML estimator for $\log \sigma_e$ is better than the Bayesian estimator for small n and is comparable for large n. Furthermore, if data are generated and model is fit assuming unequal error variances, then Bayesian estimators for $\log \sigma_{e1}$, $\log \sigma_{e2}$ and $\log(\sigma_{e1}/\sigma_{e2})$ are better than ML estimators regardless of the prior-data conflict. Bayesian estimators and ML estimators for TDI are comparable when there is prior data conflict. However, when there is no such conflict, ML estimator for TDI is better than Bayesian estimator for small n and comparable for large n. Further, Bayesian estimators and ML estimators for CCC are comparable for all settings. We also see that, in most cases, the coverage probabilities of the Bayesian credible intervals are not close to 0.95 and in many cases the probabilities are one. This indicates that the credible intervals cannot be interpreted as confidence intervals at the same level of confidence. Similar conclusions hold for almost all scenarios with HHN prior.

4.7 Discussion

Drawing inferences on error standard deviation and precision ratio of error standard deviations in mixed effect model by ML method seems to be a challenging problem. Thus, we propose a Bayesian methodology that incorporates informative priors for error SDs. The methodology involves representing model in Bayesian framework where model parameters have specific prior distributions. We use informative priors for error SDs and non-informative priors for the rest of the model parameters. First, we gather error SD data from past studies and then estimate hyperparameters of prior distributions for error SDs by ML method. We consider two different priors for error SDs — HN and HHN. The model parameters are

Table 4.10. Bias, MSE, and relative efficiency of ML estimators and posterior means and coverage probability of 95% credible intervals in case of Scenarios 1Y and 1N under a hierarchical half-normal prior for error SD.

n	Parameter	True value	ML Bias	ML MSE	Bayes Bias	Bayes MSE	MSE ratio	Cov. prob.
				Scenario 1Y				
	β_1	89.00	-0.01	4.64	-0.01	4.64	1.00	0.96
	β_2	90.00	-0.05	4.50	-0.05	4.50	1.00	0.96
	$\log\psi$	2.18	0.00	0.03	0.06	0.04	0.88	0.90
	$\log\sigma_{e1}$	0.90	-1.43	> 10	-0.53	0.42	> 10	1.00
20	$\log\sigma_{e2}$	0.90	-2.64	> 10	-0.36	0.27	> 10	1.00
	CCC	0.92	-0.01	0.00	0.00	0.00	1.46	0.86
	TDI	5.95	-0.03	0.80	-0.25	0.56	1.42	0.90
	$\beta_1 - \beta_2$	-1.00	0.04	0.42	0.04	0.42	1.00	0.98
	$\log(\sigma_{e1}^2/\sigma_{e2}^2)$	0.00	-2.41	> 10	0.33	2.09	> 10	1.00
	β_1	89.00	-0.07	0.61	-0.07	0.61	1.00	1.00
	β_2	90.00	-0.09	0.61	-0.09	0.61	1.00	1.00
	$\log\psi$	2.18	-0.01	0.00	0.01	0.00	1.03	0.96
	$\log\sigma_{e1}$	0.90	-0.39	2.58	-0.23	0.22	> 10	0.94
100	$\log\sigma_{e2}$	0.90	-0.20	1.51	-0.29	0.26	> 10	0.96
	CCC	0.92	0.00	0.00	0.00	0.00	1.21	0.98
	TDI	5.95	0.00	0.11	-0.06	0.10	1.11	0.96
	$\beta_1 - \beta_2$	-1.00	0.03	0.08	0.03	0.08	1.00	1.00
	$\log(\sigma_{e1}^2/\sigma_{e2}^2)$	0.00	0.39	18.00	-0.12	2.41	7.47	0.94
				Scenario 1N				
	β_1	89.00	-0.19	2.66	-0.20	2.66	1.00	0.98
	β_2	90.00	-0.20	2.64	-0.20	2.64	1.00	1.00
	$\log\psi$	2.18	-0.03	0.02	0.02	0.02	1.03	0.94
	$\log\sigma_{e1}$	-2.41	-3.51	> 10	-0.25	0.10	> 10	1.00
20	$\log\sigma_{e2}$	-2.41	-3.55	> 10	-0.26	0.09	> 10	1.00
	CCC	0.99	0.00	0.00	0.00	0.00	1.19	0.92
	TDI	1.16	0.00	0.00	0.02	0.00	0.80	0.96
	$\beta_1 - \beta_2$	-1.00	0.00	0.00	0.00	0.00	1.00	0.90
	β_1	89.00	-0.10	0.94	-0.10	0.94	1.00	0.94
	β_2	90.00	-0.10	0.94	-0.10	0.94	1.00	0.94
	$\log\psi$	2.18	-0.02	0.01	-0.01	0.01	1.05	0.90
	$\log\sigma_{e1}$	-2.41	-3.45	> 10	-0.33	0.15	> 10	1.00
100	$\log\sigma_{e2}$	-2.41	-2.99	> 10	-0.32	0.14	> 10	1.00
	CCC	0.99	0.00	0.00	0.00	0.00	1.06	0.92
	TDI	1.16	0.00	0.00	0.00	0.00	0.96	0.98
	$\beta_1 - \beta_2$	-1.00	0.00	0.00	0.00	0.00	1.00	0.96

Table 4.11. Bias, MSE, and relative efficiency of ML estimators and posterior means and coverage probability of 95% credible intervals in case of Scenarios 2Y and 2N under a hierarchical half-normal prior for error SD.

			ML		Bayes			
n	Parameter	True value	Bias	MSE	Bias	MSE	MSE ratio	Cov. prob.
			Scenario 2Y					
20	β_1	89.00	-0.01	4.64	0.00	4.62	1.00	0.96
	β_2	90.00	-0.05	4.50	-0.04	4.48	1.00	0.98
	$\log \psi$	2.18	0.01	0.03	0.06	0.04	0.90	0.90
	$\log \sigma_e$	0.90	-0.03	0.03	-0.05	0.02	1.25	0.98
	CCC	0.92	0.00	0.00	0.00	0.00	1.22	0.90
	TDI	5.95	-0.03	0.80	0.02	0.63	1.28	0.92
	$\beta_1 - \beta_2$	-1.00	0.04	0.42	0.04	0.42	1.00	0.98
100	β_1	89.00	-0.07	0.61	-0.07	0.61	1.00	1.00
	β_2	90.00	-0.09	0.61	-0.09	0.61	1.00	1.00
	$\log \psi$	2.18	-0.01	0.00	0.00	0.00	1.03	1.00
	$\log \sigma_e$	0.90	0.00	0.00	-0.01	0.00	1.05	1.00
	CCC	0.92	0.00	0.00	0.00	0.00	1.03	0.98
	TDI	5.95	0.00	0.11	0.02	0.11	1.04	1.00
	$\beta_1 - \beta_2$	-1.00	0.03	0.08	0.03	0.08	1.00	1.00
			Scenario 2N					
20	β_1	89.00	-0.19	2.66	-0.20	2.65	1.00	1.00
	β_2	90.00	-0.20	2.64	-0.20	2.63	1.00	1.00
	$\log \psi$	2.18	-0.03	0.02	0.02	0.02	1.02	0.94
	$\log \sigma_e$	-2.41	-0.03	0.03	0.04	0.03	0.97	0.96
	CCC	0.99	0.00	0.00	0.00	0.00	1.17	0.92
	TDI	1.16	0.00	0.00	0.01	0.00	0.87	0.96
	$\beta_1 - \beta_2$	-1.00	0.00	0.00	0.00	0.00	1.00	0.88
100	β_1	89.00	-0.1	0.94	-0.1	0.94	1.01	0.94
	β_2	90.00	-0.1	0.94	-0.1	0.94	1.01	0.94
	$\log \psi$	2.18	-0.02	0.01	-0.01	0.01	1.04	0.90
	$\log \sigma_e$	-2.41	-0.01	0.00	0.01	0.00	0.99	0.98
	CCC	0.99	0.00	0.00	0.00	0.00	1.05	0.92
	TDI	1.16	0.00	0.00	0.00	0.00	0.98	0.98
	$\beta_1 - \beta_2$	-1.00	0.00	0.00	0.00	0.00	1.00	0.96

Table 4.12. Bias, MSE, and relative efficiency of ML estimators and posterior means and coverage probability of 95% credible intervals in case of Scenarios 3Y and 3N under a hierarchical half-normal prior for error SD.

			ML		Bayes			
n	Parameter	True value	Bias	MSE	Bias	MSE	MSE ratio	Cov. prob.
			Scenario 3Y					
20	β_1	89.00	-0.01	4.64	-0.01	4.64	1.00	0.96
	β_2	90.00	-0.05	4.54	-0.05	4.54	1.00	0.96
	$\log\psi$	2.18	0.00	0.03	0.06	0.04	0.88	0.90
	$\log\sigma_{e1}$	0.90	-0.99	9.40	-0.53	0.45	> 10	1.00
	$\log\sigma_{e2}$	0.98	-2.83	> 10	-0.38	0.29	> 10	1.00
	CCC	0.92	-0.01	0.00	0.01	0.00	1.48	0.84
	TDI	6.18	-0.03	0.87	-0.31	0.63	1.39	0.86
	$\beta_1 - \beta_2$	-1.00	0.04	0.45	0.04	0.45	1.00	0.98
	$\log(\sigma_{e1}^2/\sigma_{e2}^2)$	0.16	-3.68	> 10	0.30	2.22	> 10	1.00
100	β_1	89.00	-0.07	0.61	-0.07	0.61	1.00	1.00
	β_2	90.00	-0.10	0.62	-0.10	0.62	1.00	1.00
	$\log\psi$	2.18	-0.01	0.00	0.01	0.00	1.04	0.96
	$\log\sigma_{e1}$	0.90	-0.15	1.20	-0.21	0.21	5.58	0.92
	$\log\sigma_{e2}$	0.98	-0.28	1.58	-0.29	0.26	5.98	0.96
	CCC	0.92	0.00	0.00	0.00	0.00	1.21	0.98
	TDI	6.18	0.01	0.12	-0.07	0.11	1.10	0.96
	$\beta_1 - \beta_2$	-1.00	0.03	0.09	0.03	0.09	1.00	1.00
	$\log(\sigma_{e1}^2/\sigma_{e2}^2)$	0.16	-0.25	> 10	-0.16	2.42	5.09	0.92
			Scenario 3N					
20	β_1	89.00	-0.20	3.02	-0.21	3.02	1.00	1.00
	β_2	90.00	-0.20	3.03	-0.20	3.03	1.00	1.00
	$\log\psi$	2.18	-0.03	0.03	0.03	0.03	1.00	0.94
	$\log\sigma_{e1}$	-2.41	-3.60	> 10	-0.40	0.19	> 10	1.00
	$\log\sigma_{e2}$	-2.66	-3.58	> 10	-0.13	0.04	> 10	1.00
	CCC	0.99	0.00	0.00	0.00	0.00	1.16	0.92
	TDI	1.14	0.00	0.00	0.02	0.00	0.75	0.98
	$\beta_1 - \beta_2$	-1.00	0.00	0.00	0.00	0.00	1.00	0.98
	$\log(\sigma_{e1}^2/\sigma_{e2}^2)$	-0.50	0.04	> 10	0.54	0.36	> 10	1.00
100	β_1	89.00	-0.17	1.00	-0.17	1.00	1.00	0.94
	β_2	90.00	-0.17	1.00	-0.17	1.00	1.00	0.94
	$\log\psi$	2.18	-0.02	0.01	-0.01	0.01	1.05	0.90
	$\log\sigma_{e1}$	-2.41	-3.40	> 10	-0.47	0.27	> 10	1.00
	$\log\sigma_{e2}$	-2.65	-2.51	> 10	-0.14	0.06	> 10	1.00
	CCC	0.99	0.00	0.00	0.00	0.00	1.06	0.94
	TDI	1.14	0.00	0.00	0.00	0.00	0.96	0.94
	$\beta_1 - \beta_2$	-1.00	0.00	0.00	0.00	0.00	1.00	0.94
	$\log(\sigma_{e1}^2/\sigma_{e2}^2)$	-0.50	1.80	> 10	0.65	0.93	> 10	1.00

Table 4.13. Bias, MSE, and relative efficiency of ML estimators and posterior means and coverage probability of 95% credible intervals in case of Scenarios 4Y and 4N under a hierarchical half-normal prior for error SD.

			ML		Bayes			
n	Parameter	True value	Bias	MSE	Bias	MSE	MSE ratio	Cov. prob.
				Scenario 4Y				
	β_1	89.00	-0.01	4.64	0.00	4.62	1.00	0.96
	β_2	90.00	-0.05	4.54	-0.04	4.52	1.00	0.98
	$\log\psi$	2.18	0.01	0.03	0.06	0.04	0.90	0.90
	$\log\sigma_{e1}$	0.90	0.01	0.03	-0.01	0.02	1.38	0.98
20	$\log\sigma_{e2}$	0.98	-0.07	0.03	-0.09	0.03	1.12	0.94
	CCC	0.92	0.00	0.00	0.00	0.00	1.25	0.90
	TDI	6.18	-0.03	0.87	-0.01	0.66	1.31	0.90
	$\beta_1 - \beta_2$	-1.00	0.04	0.45	0.04	0.45	1.00	0.98
	β_1	89.00	-0.07	0.61	-0.07	0.61	1.00	1.00
	β_2	90.00	-0.10	0.62	-0.10	0.62	1.00	1.00
	$\log\psi$	2.18	-0.01	0.00	0.00	0.00	1.03	1.00
	$\log\sigma_{e1}$	0.90	0.04	0.00	0.03	0.00	1.16	0.96
100	$\log\sigma_{e2}$	0.98	-0.04	0.00	-0.04	0.00	0.94	0.98
	CCC	0.92	0.00	0.00	0.00	0.00	1.04	0.98
	TDI	6.18	0.01	0.12	0.01	0.12	1.05	1.00
	$\beta_1 - \beta_2$	-1.00	0.03	0.09	0.03	0.09	1.00	1.00
				Scenario 4N				
	β_1	89.00	-0.20	3.02	-0.20	3.02	1.00	0.98
	β_2	90.00	-0.20	3.03	-0.20	3.03	1.00	0.98
	$\log\psi$	2.18	-0.03	0.03	0.03	0.03	1.00	0.94
	$\log\sigma_{e1}$	-2.41	-0.15	0.05	-0.07	0.03	1.52	0.96
20	$\log\sigma_{e2}$	-2.66	0.10	0.04	0.18	0.06	0.63	0.74
	CCC	0.99	0.00	0.00	0.00	0.00	1.15	0.94
	TDI	1.14	0.00	0.00	0.01	0.00	0.82	0.96
	$\beta_1 - \beta_2$	-1.00	0.00	0.00	0.00	0.00	1.00	0.98
	β_1	89.00	-0.17	1.00	-0.17	1.00	1.00	0.94
	β_2	90.00	-0.17	1.00	-0.17	1.00	1.00	0.94
	$\log\psi$	2.18	-0.02	0.01	-0.01	0.01	1.04	0.90
	$\log\sigma_{e1}$	-2.41	-0.11	0.02	-0.10	0.02	1.20	0.72
100	$\log\sigma_{e2}$	-2.65	0.14	0.02	0.15	0.03	0.84	0.42
	CCC	0.99	0.00	0.00	0.00	0.00	1.05	0.94
	TDI	1.14	0.00	0.00	0.00	0.00	0.96	0.94
	$\beta_1 - \beta_2$	-1.00	0.00	0.00	0.00	0.00	1.00	0.94

estimated by posterior simulation. Monte Carlo method is used to estimate biases, MSE and coverage probabilities. Here, we compare the results of Bayesian method with those of the ML method. Simulation results demonstrate that Bayesian method works better on estimating error SDs than ML method unequal error variance assumption and the two are comparable under equal variance assumption. Thus, there is an advantage of bringing in external information from the literature. The results also show that the estimates are robust to the choice of the prior distribution — HN or HHN. However, even with the Bayesian approach, we are unable to obtain reliable estimates for the ratio of error variances. This is one of the limitations of our approach. We speculate that this is due to lack of replicated measurements in the data. Another limitation is that the credible intervals for many of the parameters cannot be interpreted as confidence intervals as their coverage probabilities are not close to the nominal level.

CHAPTER 5

ONGOING AND FUTURE WORK

In this chapter we list some ongoing work and possible future extensions of the methodologies presented in this dissertation. For the segmented measurement error model developed in Chapter 2:

- We only presented simulation results showing accuracy of inference procedures for the model parameters. We are currently in the process of performing additional simulations to evaluate accuracy of the segmented-specific measures of similarity and agreement, which are functions of the model parameters.

- We also plan to investigate the issue of model evaluation and see how certain violations of the model assumptions, such as non-normality of the distribution of the true value, affects the results.

- To test the changepoint hypothesis, the null distribution of the likelihood ratio statistic was obtained using bootstrap and a χ^2 approximation. Although they worked well in the simulation studies, it would be of interest to find the exact asymptotic distribution of the statistic as the standard asymptotic theory of likelihood ratio tests does not apply in this situation.

- The model assumed homoscedasticity for the errors. However, there is some evidence in the digoxin data that the error variances may be changing with the true value. This calls for an extension of the methodology to deal with heteroscedastic errors.

- The methodology is designed for paired measurements data. We would like to extend it to deal with repeated measurements data where the measurement by each method on every subject is replicated multiple times.

- Here we have taken a fully parametric approach that assumes normality for distribution of errors and the true values. An extension of this approach that replaces the assumption of normality with more flexible distributions such as skew-normal and skew-t would also be of interest.

For the extended segmented measurement error model presented in Chapter 3 that lets each method have its own error variance:

- We would like to apply the methodology to real datasets and perform simulation studies to evaluate the methodology.

- We would also like to see what advantage this model offers over its simpler special case in Chapter 2.

- An extension of this model to deal with repeated measurements data and heteroscedasticity in errors is also of interest.

For the Bayesian approach with informative priors presented in Chapter 4:

- We would like to apply to methodology to other datasets to see how well its advantage over the ML approach holds up.

- We would like to extend the methodology to repeated measurements data and see if bringing in external information through informative priors for error variances help in their estimation even in the presence of replicated measurements.

REFERENCES

Al Khudhairi, D., R. Prabhu, M. El Sharkawy, and R. Burtles (1990). Evaluation of a pulse oximeter during profound hypothermia. An assessment of the Biox 3700 during induction of hypothermia before cardiac surgery in paediatric patients. *International Journal of Clinical Monitoring and Computing 7*, 217–222.

Alanen, E. (2010). Everything all right in method comparison studies? *Statistical Methods in Medical Research 21*, 297–309.

Atkinson, G. and A. Nevill (1997). Comment on the use of concordance correlation to assess the agreement between two variables. *Biometrics 53*, 775–777.

Bland, J. M. and D. Altman (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet 327*, 307–310.

Carroll, R. J., K. Roeder, and L. Wasserman (1999). Flexible parametric measurement error models. *Biometrics 55*, 44–54.

Carroll, R. J., D. Ruppert, C. M. Crainiceanu, and L. A. Stefanski (2006). *Measurement Error in Nonlinear Models: A Modern Perspective*. Chapman and Hall/CRC.

Carstensen, B. (2010). *Comparing Clinical Measurement Methods: A Practical Guide*. John Wiley.

Cheng, C.-L. and J. W. Van Ness (1999). *Statistical Regression with Measurement Error*. John Wiley.

Choudhary, P. K. and H. N. Nagaraja (2007). Tests for assessment of agreement using probability criteria. *Journal of Statistical Planning and Inference 137*, 279–290.

Choudhary, P. K. and H. N. Nagaraja (2017). *Measuring Agreement: Models, Methods, and Applications*. Wiley.

Das, J., A. Aggarwal, and N. K. Aggarwal (2010). Pulse oximeter accuracy and precision at five different sensor locations in infants and children with cyanotic heart disease. *Indian Journal of Anaesthesia 54*, 531.

Davison, A. C. and D. V. Hinkley (1997). *Bootstrap Methods and Their Application*. Cambridge University Press.

Deming, W. E. (1943). *Statistical Adjustment of Data*. John Wiley, New York.

Desiderio, D. P., G. Wong, N. K. Shah, J. Liu, C. J. Loughlin, and R. F. Bedford (1990). A clinical evaluation of pulse oximetry during thoracic surgery. *Journal of Cardiothoracic Anesthesia 4*, 30–34.

Dunn, G. (2004). *Statistical Evaluation of Measurement Errors* (2nd ed.). John Wiley.

Dunn, G. (2007). Regression models for method comparison data. *Journal of Biopharmaceutical Statistics 17*, 739–756.

Dunn, G. and C. Roberts (1999). Modelling method comparison data. *Statistical Methods in Medical Research 8*, 161–179.

Escaramís, G., C. Ascaso, and J. L. Carrasco (2010). The total deviation index estimated by tolerance intervals to evaluate the concordance of measurement devices. *BMC Medical Research Methodology 10*, 31.

Finney, D. J. (1996). A note on the history of regression. *Journal of Applied Statistics 23*, 555–557.

Fong, Y., C. Di, and S. Permar (2015). Change point testing in logistic regression models with interaction term. *Statistics in Medicine 34*, 1483–1494.

Forbes, C., M. Evans, N. Hastings, and B. Peacock (2011). *Statistical Distributions*. John Wiley & Sons.

Fuller, W. (1987). *Measurement Error Models*. John Wiley.

Gelman, A., H. S. Stern, J. B. Carlin, D. B. Dunson, A. Vehtari, and D. B. Rubin (2013). *Bayesian Data Analysis*. Chapman and Hall/CRC.

Goodman, M. S., Y. Li, and R. C. Tiwari (2011). Detecting multiple change points in piecewise constant hazard functions. *Journal of Applied Statistics 38*, 2523–2532.

Gössl, C. and H. Küchenhoff (2001). Bayesian analysis of logistic regression with an unknown change point and covariate measurement error. *Statistics in Medicine 20*, 3109–3121.

Grubbs, F. E. (1948). On estimating precision of measuring instruments and product variability. *Journal of the American Statistical Association 43*, 243–264.

Hannhart, B., J. Haberer, C. Saunier, and M. Laxenaire (1991). Accuracy and precision of fourteen pulse oximeters. *European Respiratory Journal 4*, 115–119.

Hawkins, D. M. (2002). Diagnostics for conformity of paired quantitative measurements. *Statistics in Medicine 21*, 1913–1935.

Hinkley, D. V. (1969). Inference about the intersection in two-phase regression. *Biometrika 56*, 495–504.

Hinkley, D. V. (1971). Inference in two-phase regression. *Journal of the American Statistical Association 66*, 736–743.

Hudson, D. J. (1966). Fitting segmented curves whose join points have to be estimated. *Journal of the American Statistical Association 61*, 1097–1129.

James, J., L. Tiwari, P. Upadhyay, V. Sreenivas, V. Bhambhani, and J. M. Puliyel (2006). Evaluation of pulse-oximetry oxygen saturation taken through skin protective covering. *BMC Pediatrics 6*, 14.

Jubran, A. and M. J. Tobin (1990). Reliability of pulse oximetry in titrating supplemental oxygen therapy in ventilator-dependent patients. *Chest 97*, 1420–1425.

Küchenhoff, H. and R. Carroll (1997). Segmented regression with errors in predictors: Semi-parametric and parametric methods. *Statistics in Medicine 16*, 169–188.

Kummell, C. H. (1879). Reduction of observation equations which contain more than one observed quantity. *The Analyst 6*, 97–105.

Lai, Y. and P. S. Albert (2014). Identifying multiple change points in a linear mixed effects model. *Statistics in Medicine 33*, 1015–1028.

Lehmann, E. L. (1998). *Elements of Large-Sample Theory*. Springer.

Lin, L. I. (1989). A concordance correlation coefficient to evaluate reproducibility. *Biometrics 45*, 255–268. Corrections: 2000, **56**, 324-325.

Lin, L. I. (2000). Total deviation index for measuring individual agreement with applications in laboratory performance and bioequivalence. *Statistics in Medicine 19*, 255–270.

Lin, L. I., A. S. Hedayat, and W. Wu (2011). *Statistical Tools for Measuring Agreement*. Springer.

Linnet, K. (1999). Necessary sample size for method comparison studies based on regression analysis. *Clinical Chemistry 45*, 882–894.

Mandel, J. (1978). Accuracy and precision: Evaluation and interpretation of analytical results. In *Treatise on Analytical Chemistry, Part I, Theory and Practice*, 2nd edition, volume 1, pp. 243-298, Kolthoff, I. M. and Elving, P. J. (Editors), John Wiley, New York.

Mandel, J. and R. D. Stiehler (1954). Sensitivity – a criterion for the comparison of methods of test. *Journal of Research of the National Bureau of Standards 53*, 155–159.

McCulloch, C. E., S. R. Searle, and J. M. Neuhaus (2008). *Generalized, Linear, and Mixed Models* (2nd ed.). John Wiley.

McLachlan, G. and T. Krishnan (2007). *The EM Algorithm and Extensions*, Volume 382. John Wiley & Sons.

Meng, X.-L. and D. B. Rubin (1993). Maximum likelihood estimation via the ECM algorithm: A general framework. *Biometrika 80*, 267–278.

Muggeo, V. M. (2003). Estimating regression models with unknown break-points. *Statistics in Medicine 22*, 3055–3071.

Muggeo, V. M., D. C. Atkins, R. J. Gallop, and S. Dimidjian (2014). Segmented mixed models with random changepoints: A maximum likelihood approach with application to treatment for depression study. *Statistical Modelling 14*, 293–313.

Nickerson, B. G., C. Sarkisian, and K. Tremper (1988). Bias and precision of pulse oximeters and arterial oximeters. *Chest 93*, 515–517.

Pastor-Barriuso, R., E. Guallar, and J. Coresh (2003). Transition models for change-point estimation in logistic regression. *Statistics in Medicine 22*, 1141–1162.

Perez-Jaume, S. and J. L. Carrasco (2015). A non-parametric approach to estimate the total deviation index for non-normal data. *Statistics in Medicine 34*, 3318–3335.

Phattraprayoon, N., S. Sardesai, M. Durand, and R. Ramanathan (2012). Accuracy of pulse oximeter readings from probe placement on newborn wrist and ankle. *Journal of Perinatology 32*, 276.

Piepho, H. and J. Ogutu (2003). Inference for the break point in segmented regression with application to longitudinal data. *Biometrical Journa 45*, 591–601.

Praud, J.-P., C. Gaultier, A. Carofilis, F. Lacaille, M. Dehan, and F. Bridey (1989). Accuracy of two wavelength pulse oximetry in neonates and infants. *Pediatric Pulmonology 6*, 180–182.

Quandt, R. E. (1958). The estimation of the parameters of a linear regression system obeying two separate regimes. *Journal of the American Statistical Association 53*, 873–880.

Seber, G. A. F. and C. J. Wild (1989). *Nonlinear Regression.* Wiley.

Shyr, J. Y. and L. J. Gleser (1986). Inference about comparative precision in linear structural relationships. *Journal of Statistical Planning and Inference 14*, 339–358.

Spiegelhalter, D. J., K. R. Abrams, and J. P. Myles (2004). *Bayesian Approaches to Clinical Trials and Health-Care Evaluation.* John Wiley & Sons.

Sprent, P. (1961). Some hypotheses concerning two phase regression lines. *Biometrics 17*, 634–645.

Stan Development Team (2018). *RStan: the R interface to Stan.* R package version 2.17.4.

Staudenmayer, J. and D. Spiegelman (2002). Segmented regression in the presence of covariate measurement error in main study/validation study designs. *Biometrics 58*, 871–877.

Tang, X., M. I. Miller, and L. Younes (2017). Biomarker change-point estimation with right censoring in longitudinal studies. *The Annals of Applied Statistics 11*, 1738–1762.

Vexler, A. and G. Gurevich (2009). Average most powerful tests for a segmented regression. *Communications in Statistics—Theory and Methods 38*, 2214–2231.

BIOGRAPHICAL SKETCH

Lak Nilusha Kotinkaduwa grew up in Kandy, Sri Lanka. She received a bachelor's degree in statistics from the University of Peradeniya, Sri Lanka in 2012, with first class honors. She entered The University of Texas at Dallas to pursue her doctoral degree in statistics in Fall 2013.

CURRICULUM VITAE

Lak Nilusha Kotinakaduwa

Contact Information:

Email: laknilusha.kotinkaduwa@utdallas.edu
LinkedIn: https://www.linkedin.com/in/lak-nilusha-kotinkaduwa-6a085416a/

Educational History:

- Doctor of Philosophy (Ph.D.) in Statistics, December 2018
  The University of Texas at Dallas, Richardson, TX
- Master of Science (M.S.) in Statistics, May 2016
  The University of Texas at Dallas, Richardson, TX
- Bachelor of Science (B.S.) in Statistics , September 2011
  University of Peradeniya, Peradeniya, Sri Lanka

Employment History:

- Teaching Assistant, The University of Texas at Dallas, August 2013–Present.
- Teaching Assistant, Department of Statistics and Computer Science, University of Peradeniya, Peradeniya, Sri Lanka, October 2011 - June 2013.
- Intern, Colombo Stock Exchange, July 2010– October 2010.

Honors and Awards:

- Science, Technology, Engineering and Mathematics (STEM) Fellowship for Ph.D. study, UT Dallas, 2013-2018.
- Teaching assistantship during Ph.D. program, UT Dallas, 2013-2018.
- University Prize for Academic Excellence, University of Peradeniya, Sri Lanka, 2011

Leadership Experience:

- Member, The Institute of Applied Statistics, Sri Lanka (IASSL), September 2013 - present
- Program coordinator, Sri Lankan Student Association, University of Texas at Dallas, August 2015 - July 2016
- Program coordinator, Statistical Circle, University of Peradeniya, Sri Lanka, September 2010 - September 2011