THE UNIVERSITY OF CHICAGO

OBTAINING MORE INFORMATION FROM LOCATION SCALE MODELS FOR INTENSIVE LONGITUDINAL DATA: EFFECT SIZES, DIAGNOSTICS, AND LATENT CLASSES

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE BIOLOGICAL SCIENCES
AND THE PRITZKER SCHOOL OF MEDICINE
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

DEPARTMENT OF PUBLIC HEALTH SCIENCES

BY

XINGRUO ZHANG

CHICAGO, ILLINOIS

JUNE 2024

To my husband, Qichen.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to first thank my advisor, Donald Hedeker, for always being patient and encouraging even when I doubted myself. He is the kindest and most understanding person I have ever met. It was with his help that I managed to not only grow academically but also experience the most fantastic time of my life. I would also like to thank Professor Robert Gibbons and Professor Yuan Ji from the University of Chicago and Professor Juned Siddique from Northwestern University for serving on my dissertation committee and providing invaluable feedback.

Many thanks to my parents and my grandparents. Despite many difficulties, they have always ensured that I receive the best education possible, without asking for anything in return. I would also like to extend my gratitude and love to my dear husband. None of this would be possible without him — proofreading almost every piece of my work and helping me dry run almost every presentation. My family always provides me with endless support and unconditional love, no matter what I do.

I would also like to thank all my fellow students in the Department of Public Health Sciences, especially Yihao, Bei, and Dehua. I truly treasured their company during my time at the University of Chicago. My heartfelt appreciation also goes to my dear friends, including Krystal, Felicity, Jennifer, Dana, Jennie, Sherry, Jiajie, Rita, Iris, Juliace, Chenchen, Elinor, Liuyi, Ni, and many others, who comforted me, inspired me, and gave me really good life advice.

Special thanks to my high school math teacher, Jianbin Guo, for his belief in me. He told me that I was no worse than anyone else. I would not have believed that I could do a Ph.D. without those words.

At the end, I want to mention that this dissertation is supported financially by National Institute of Diabetes and Digestive and Kidney Diseases (grant number R01 DK125414).

ABSTRACT

Location-scale models are a particularly useful class of models for their ability to simultaneously model the intraindividual variability through the scale sub-model and model the mean outcome through the location sub-model. This thesis extended methods to enhance the interpretability of mixed-effects location scale (MELS) models and address the need for influence analysis and subgrouping techniques within location-scale models.

First, due to the relatively complicated structure of MELS models, there are challenges associated with integrating the interpretations of random effects and fixed effects. Researchers often have difficulty comparing the roles that fixed effects and random effects, which represent effects of covariates and between-subject (BS) heterogeneity, respectively, play in the models. Therefore, in the first study, we addressed this problem and enhanced the interpretation of MELS models by introducing R-squared measures as standardized effect size measures for both the location sub-model and the scale sub-model of MELS models. The methodology aligns with the increasing emphasis on reporting actual sizes of effects in addition to or even instead of p-values.

Second, as healthcare continues to shift towards personalized approaches, the demand for robust statistical tools to identify individuals with distinctive health behaviors and outcomes is on the rise. In this context, MELS models are valuable instruments for discerning differences among individuals. By developing an influence analysis method framework using MELS models, we aim to enhance our capacity to pinpoint and understand the impact of individual behavior on health conditions and outcomes, contributing to more effective personalized healthcare strategies. In the second study, the proposed approach has successfully demonstrated the use of MELS models in detecting abnormal subjects in intensive longitudinal data that cannot be detected using existing methods based on regular MRMs. We also showcased how this method facilitates a comprehensive understanding of subject heterogeneity in the association between sleep quality and learning goal achievement using data

from a published study.

Last but not least, the trajectory in intraindividual variability can be an important indicator of treatment success in many research fields, including psychology, behavioral sciences, and clinical developments, to name a few. However, there lacks available methodology to distinguish distinct groups in intraindividual variability trajectory. We proposed a Bayesian location-scale model with latent classes in both mean trajectory and intraindividual variability trajectory to fill this gap. Our developed method, powered by Stan programs, stands as a user-friendly and widely applicable tool. It seamlessly integrates with multiple programming languages, including R, Python, and Stata, making it accessible to researchers, even those without extensive experience in Bayesian modeling and coding.

Codes to implement methods described in this thesis are included in the supplementary files.

CHAPTER 1

INTRODUCTION

1.1 Location-Scale Models

When analyzing longitudinal data, the traditional approach involves modeling the mean outcome as a function of time and other covariates, often referred to as a location model. However, with intensive longitudinal data, which also capture how the outcome varies within each time period, there has been a growing focus on modeling the association between within-subject (WS) variability, termed scale, and covariates. A more general form of location-scale models, named mixed-effects location scale (MELS) models, also incorporates random effects into both components of the models to account for dependencies within the same subject and variations across different subjects. Nevertheless, there is room for improvement of existing methods related to location-scale models, and this thesis focuses on enhancing these methodologies, particularly in the areas of effect sizes, influence analysis, and latent class modeling.

1.2 R-squared Measures

The coefficient of determination, also referred to as R-squared (R^2), quantifies the proportion of variance accounted for by a specific factor, such as a covariate or random effects. Because R^2 values always fall within the range of 0 to 1, they serve as valuable standardized effect size metrics, enabling direct comparisons between different components within a model.

In the context of multilevel modeling, R^2 measures prove particularly useful for assessing the impact of fixed effects associated with time-invariant covariates, fixed effects linked to time-varying covariates, as well as random effects at each level of the model. They also facilitate comparisons of the explained variances across various levels of the model.

This dissertation outlines the application of R^2 measures to both the location and scale components of a location-scale model. Furthermore, it introduces methodologies for incorporating random scale effects and addressing different approaches of modeling between-subject variances in mean outcomes when computing these R^2 values.

1.3 Influence Analysis

The presence of subjects whose behaviors deviate markedly from others, thereby affecting statistical inference, poses a well-recognized challenge in modeling longitudinal data. Such subjects also demand special attention, as they may require individualized interventions. Existing literature reveals a gap, specifically in methods for analyzing the influence of data from each subject on multilevel models for intraindividual variability. This dissertation addresses this gap by discussing how to use MELS models as a valuable tool for performing influence analysis, encompassing both location and scale aspects, in the context of intensive longitudinal data.

1.4 Latent Class Modeling

In longitudinal data analysis, an important topic is obtaining more information on how individuals change differently in the long term [31]. Similar to random effects, latent classes provide a method for accounting for heterogeneity among subjects within a population. This approach involves categorizing subjects into several relatively homogeneous subgroups to extract additional insights from the data.

Latent class modeling holds significant promise in personalized medicine by stratifying patients into meaningful subgroups. Consequently, we integrate latent class modeling with location-scale modeling to devise a methodology capable of identifying meaningful subgroups based on both the mean and WS variability of longitudinal outcomes.

1.5 Summary

Overall, this dissertation is motivated by the need for standardized effect size measures, diagnostic tools, and extensions for location-scale models to form a more comprehensive methodology for intensive longitudinal data. The proposed methods are based on recent advances in MRM and MELS models [5, 13, 19, 33], as well as estimation software such as Stan [52].

The rest of this dissertation is organized as follows. Chapter 2 extends the existing method of defining R-squared measures for MRMs to MELS models and introduces an R package that derives and visualizes these measures. In Chapter 3, we develop a framework for researchers to detect influential subjects in intensive longitudinal data using MELS models. Chapters 2 and 3 have been published in *Statistics in Medicine* [63] and *BMC Medical Research Methodology* [64], respectively. Chapter 4 proposes a Bayesian model that identifies latent classes in both location trajectories and scale trajectories of intensive longitudinal data. A summary of the contributions of the three proposed methods and their potential extensions is included in Chapter 5.

CHAPTER 2

DEFINING R-SQUARED MEASURES FOR MIXED-EFFECTS LOCATION SCALE MODELS

2.1 Introduction

Modern data collection methods such as ecological momentary assessments (EMA) have allowed more detailed examination of subjects' heterogeneity both at the between-subject (BS) level (also known as level 2) and the within-subject (WS) level (also known as level 1) [47]. Hedeker et al. [15] extended the commonly used mixed-effects regression model (MRM) into the mixed-effects location scale (MELS) model that includes both random location effects and random scale effects. Random location effects refer to random subject effects on the mean of the response variable, and random scale effects refer to random subject effects on the WS variability of the response variable. While scale sometimes only refers to standard deviation, here it is on variance metric.

Increasingly, researchers are encouraged to report effect sizes in addition to p-values for their study results [53]. Standardized effect size measures are of particular interest as they allow direct comparison of different models. However, to the best of our knowledge, there are no existing standardized effect size measures specifically for MELS models.

Standardized effect size measures have been developed for MRMs. Earlier pseudo-R^2 measures evaluate a model's reduction in residual variance from the null model [6]. A problem with this approach is that it can result in negative values and thus become meaningless as shown by Snijders and Bosker [49]. Snijders and Bosker [49] then resolved this problem by constructing R^2 using model-implied variances. Recently, Rights and Sterba [40] introduced a comprehensive framework of R^2 measures for multilevel models using model-implied variances that measure both the total variance of the response variable explained and level-specific variances of the response variable explained. While this work primarily discusses

cross-sectional multilevel models, their later work accommodates specific features of longitudinal multilevel models [42]. Rights and Sterba also introduced supplementary visualization and R functions to help researchers implement their proposed framework.

Our study extends Rights and Sterba's framework to two-level MELS models. We develop frameworks of R^2 measures for two forms of MELS models, which differ in their characterization of the random location effects. In the first form, the model includes random subject intercepts and allows their variance to be influenced by both subject-level and observation-level covariates. Alternatively, the second form includes random subject intercepts and slopes of observation-level covariates. For both forms, R^2 measures are constructed for both the location model and the scale model. We also develop an R function, *R2MELS*, that allows calculation and visualization of R^2 measures specifically for MELS models.

2.1.1 Mixed-effects location-scale (MELS) models with random intercepts with covariate-influenced variance

To begin, we review the MELS model for a two-level continuous response variable y_{ij} ($i = 1, 2, \ldots, N$ subjects, $j = 1, 2, ..., n_i$ observations) proposed by Hedeker et al. [15]:

$$y_{ij} = \beta_0 + \boldsymbol{x}_{ij}^T \boldsymbol{\beta} + \nu_i + \epsilon_{ij}, \tag{2.1}$$

where β_0 is the fixed intercept of the location model, $\boldsymbol{x}_{ij}$ is the $p \times 1$ vector of fixed location effect covariates, and $\boldsymbol{\beta}$ is the corresponding $p \times 1$ vector of fixed location effects. In our context, all covariates are considered as random variables. BS heterogeneity is included via random intercepts ν_i, also recognized as the random location effects. ϵ_{ij} is the observation-level residuals and incorporates WS heterogeneity. ν_i and ϵ_{ij} are assumed to be normally distributed with mean 0 and variances $\sigma^2_{\nu_i}$ and $\sigma^2_{\epsilon_{ij}}$, respectively.

$\sigma^2_{\nu_i}$ is further modeled in log-linear form to account for different BS heterogeneity at

different values of covariates:

$$\sigma^2_{\nu_i} = \text{Var}(\nu_i|\boldsymbol{u}_{ij}) = \exp(\alpha_0 + \boldsymbol{u}_{ij}^T\boldsymbol{\alpha}), \tag{2.2}$$

where $\exp(\alpha_0)$ is the variance of ν_i when the covariates equal zero, or if the covariates have no influence on the BS variance. $\boldsymbol{u}_{ij}$ is the vector of covariates influencing $\sigma^2_{\nu_i}$, which can contain both subject-level and observation-level covariates. $\boldsymbol{\alpha}$ represents the vector of fixed effects associated with $\boldsymbol{u}_{ij}$ on $\sigma^2_{\nu_i}$.

For ϵ_{ij}, both heteroskedasticity at different covariate values and heteroskedasticity between subjects are allowed. The heteroskedasticity of ϵ_{ij} between subjects is included via a random scale effect ω_i, which is assumed to follow a normal distribution with a mean of 0 and a variance of σ^2_ω. $\sigma^2_{\epsilon_{ij}}$ is again modeled in log-linear form to ensure positive variance values:

$$\sigma^2_{\epsilon_{ij}} = \text{Var}(\epsilon_{ij}|\boldsymbol{w}_{ij}) = \exp(\tau_0 + \boldsymbol{w}_{ij}^T\boldsymbol{\tau} + \omega_i), \tag{2.3}$$

where $\exp(\tau_0)$ is the value of $\sigma^2_{\epsilon_{ij}}$ when the covariates and the random scale effect equal zero, or when there is neither any covariate influencing $\sigma^2_{\epsilon_{ij}}$ nor subject heteroskedasticity of ϵ_{ij}. $\boldsymbol{w}_{ij}$ is the vector of covariates influencing $\sigma^2_{\epsilon_{ij}}$. Similar to $\boldsymbol{u}_{ij}$, $\boldsymbol{w}_{ij}$ can contain both subject-level and observation-level covariates. $\boldsymbol{\tau}$ is the vector of fixed effects associated with $\boldsymbol{w}_{ij}$ on $\sigma^2_{\epsilon_{ij}}$.

For convenience, Equation 2.2 can be rewritten in terms of standardized random location effects, denoted as θ_{1i}:

$$y_{ij} = \beta_0 + \boldsymbol{x}_{ij}^T\boldsymbol{\beta} + \sigma_{\nu_i}\theta_{1i} + \epsilon_{ij}, \tag{2.4}$$

and Equation 2.3 can also be written in terms of standardized random scale effects, denoted as θ_{2i}:

$$\sigma^2_{\epsilon_{ij}} = \exp(\tau_0 + \boldsymbol{w}_{ij}^T\boldsymbol{\tau} + \sigma_\omega\theta_{2i}), \tag{2.5}$$

Since the random location effects ν_i and random scale effects ω_i are not necessarily uncorrelated, θ_{1i} and θ_{2i} are assumed to follow the following bivariate normal distribution:

$$\begin{pmatrix} \theta_{1i} \\ \theta_{2i} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho_{\nu\omega} \\ \rho_{\nu\omega} & 1 \end{pmatrix} \right), \tag{2.6}$$

where $\rho_{\nu\omega}$ is the correlation between the random location effect ν_i and the random scale effect ω_i.

2.1.2 MELS models with random slopes of observation-level covariates

Instead of having random intercepts and allowing covariates to influence their variance, random location effects can be modeled by random intercepts with constant variance and random slopes of observation-level covariates [34]:

$$y_{ij} = \beta_0 + \boldsymbol{x}_{ij}^T \boldsymbol{\beta} + \boldsymbol{z}_{ij}^T \boldsymbol{\nu}_i + \epsilon_{ij}, \tag{2.7}$$

where we use $\boldsymbol{z}_{ij}$ to represent an $(r+1) \times 1$ vector with the first element of 1 for the random intercept followed by r observation-level covariates for random slopes. $\boldsymbol{\nu}_i$, the corresponding $(r+1) \times 1$ vector of the random location effects, follows the following distribution:

$$\boldsymbol{\nu}_i \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_{\boldsymbol{\nu}}). \tag{2.8}$$

The first element of $\boldsymbol{\nu}_i$ is the random intercept, and the following elements are r random slopes associated with covariates in $\boldsymbol{z}_{ij}$. The variances of the random intercept and the random slopes, $\sigma^2_{\nu_0}$, $\sigma^2_{\nu_1}$, $\ldots$, $\sigma^2_{\nu_r}$, are scalars not influenced by any covariates.

In this form of MELS models, the fixed intercept, fixed location effects, and variance of observation-level residuals are modeled the same way as in Section 2.1.1.

Again, we can express the random effects as standardized random effects for convenience.

Namely,

$$\boldsymbol{\nu}_i = \begin{pmatrix} \sigma_{\nu_0}\theta_{0i} \\ \sigma_{\nu_1}\theta_{1i} \\ \vdots \\ \sigma_{\nu_r}\theta_{ri} \end{pmatrix}, \tag{2.9}$$

and

$$\omega_i = \sigma_\omega \theta_{\omega i}, \tag{2.10}$$

where $\theta_{0i}, \theta_{1i}, \ldots, \theta_{ri}$, and $\theta_{\omega i}$ follow a multivariate normal distribution with mean $\mathbf{0}$, and their variance-covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$ is a $(r+2) \times (r+2)$ matrix given as:

$$\boldsymbol{\Sigma}_{\boldsymbol{\theta}} = \begin{pmatrix} 1 & \rho_{\nu_{0i}\nu_{1i}} & \cdots & \cdots & \rho_{\nu_{0i}\omega_i} \\ \rho_{\nu_{0i}\nu_{1i}} & 1 & \cdots & \cdots & \rho_{\nu_{1i}\omega_i} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \cdots & \ddots & \rho_{\nu_{ri}\omega_i} \\ \rho_{\nu_{0i}\omega_i} & \rho_{\nu_{1i}\omega_i} & \cdots & \rho_{\nu_{ri}\omega_i} & 1 \end{pmatrix}, \tag{2.11}$$

where $\rho_{\nu_{pi}\nu_{qi}} (p \neq q; p, q = 0, 1, \ldots, r)$ is the correlation between ν_{pi} and ν_{qi}, and $\rho_{\nu_{pi}\omega_i}$ is the correlation between $\nu_{pi} (p = 0, 1, \ldots, r)$ and ω_i.

2.2 The proposed method

2.2.1 Decomposition of observation-level covariates

As mentioned, $\boldsymbol{x}_{ij}$, $\boldsymbol{u}_{ij}$, and $\boldsymbol{w}_{ij}$ can contain both subject-level covariates and observation-level covariates, and $\boldsymbol{z}_{ij}$ is solely composed of observation-level covariates. Since an observation-level covariate not centered at the subject level contains both BS variation and WS variation, and we assume that the covariates are multivariate normally distributed at each level, we first decompose each of $\boldsymbol{x}_{ij}$, $\boldsymbol{u}_{ij}$, $\boldsymbol{w}_{ij}$, and $\boldsymbol{z}_{ij}$ into a BS component and a WS component

before deriving their contribution to the variance of the response. For example,

$$\boldsymbol{x}_{ij} = \bar{\boldsymbol{x}}_i + (\boldsymbol{x}_{ij} - \bar{\boldsymbol{x}}_i), \tag{2.12}$$

where $\bar{\boldsymbol{x}}_i$ is the vector of subject-level fixed location effect covariates, which characterizes the BS components of observation-level fixed location effect covariates. Elements of $\bar{\boldsymbol{x}}_i$ are multivariate normally distributed with mean $\boldsymbol{\mu}_{\boldsymbol{x}}$ and variance $\boldsymbol{\Phi}_{\boldsymbol{x}}^{bs}$. Analogously, $(\boldsymbol{x}_{ij} - \bar{\boldsymbol{x}}_i)$ is the vector of WS components of observation-level fixed location effect covariates, and its elements are multivariate normally distributed with mean $\mathbf{0}$ and variance $\boldsymbol{\Phi}_{\boldsymbol{x}}^{ws}$.

Similarly, decomposition of covariates influencing the variance of the random intercepts (i.e., the BS variance) is given by $\boldsymbol{u}_{ij} = \bar{\boldsymbol{u}}_i + (\boldsymbol{u}_{ij} - \bar{\boldsymbol{u}}_i)$, where $\bar{\boldsymbol{u}}_i \sim \mathcal{N}(\boldsymbol{\mu}_{\boldsymbol{u}}, \boldsymbol{\Phi}_{\boldsymbol{u}}^{bs})$, and $(\boldsymbol{u}_{ij} - \bar{\boldsymbol{u}}_i) \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Phi}_{\boldsymbol{u}}^{ws})$. Decomposition of covariates influencing the WS variance is similarly: $\boldsymbol{w}_{ij} = \bar{\boldsymbol{w}}_i + (\boldsymbol{w}_{ij} - \bar{\boldsymbol{w}}_i)$, where $\bar{\boldsymbol{w}}_i \sim \mathcal{N}(\boldsymbol{\mu}_{\boldsymbol{w}}, \boldsymbol{\Phi}_{\boldsymbol{w}}^{bs})$, and $(\boldsymbol{w}_{ij} - \bar{\boldsymbol{w}}_i) \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Phi}_{\boldsymbol{w}}^{ws})$. Finally, decomposition of covariates for random slopes is given by $\boldsymbol{z}_{ij} = \bar{\boldsymbol{z}}_i + (\boldsymbol{z}_{ij} - \bar{\boldsymbol{z}}_i)$, where $\bar{\boldsymbol{z}}_i \sim \mathcal{N}(\boldsymbol{\mu}_{\boldsymbol{z}}, \boldsymbol{\Phi}_{\boldsymbol{z}}^{bs})$, and $(\boldsymbol{z}_{ij} - \bar{\boldsymbol{z}}_i) \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Phi}_{\boldsymbol{z}}^{ws})$. The recurring superscript bs in the variance matrices represents BS components, and ws represents WS components. Note that if one includes a random slope for an observation-level covariate in $\boldsymbol{x}_{ij}$, the same covariate will also occur in $\boldsymbol{z}_{ij}$, and the decomposition of this variable will be the same in $\boldsymbol{x}_{ij}$ and $\boldsymbol{z}_{ij}$.

2.2.2 Variance partitioning

Variance partitioning for MELS models with random intercepts with covariate-influenced variance

For the model described in Section 2.1.1,

$$\begin{aligned}\mathrm{Var}(y_{ij}) &= \mathrm{Var}(\beta_0 + \boldsymbol{x}_{ij}^T\boldsymbol{\beta} + \nu_i + \epsilon_{ij}) \\ &= \mathrm{Var}((\boldsymbol{x}_{ij} - \bar{\boldsymbol{x}}_i)^T\boldsymbol{\beta} + \bar{\boldsymbol{x}}_i^T\boldsymbol{\beta} + \nu_i + \epsilon_{ij}) \\ &= \mathrm{Var}((\boldsymbol{x}_{ij} - \bar{\boldsymbol{x}}_i)^T\boldsymbol{\beta}) + \mathrm{Var}(\bar{\boldsymbol{x}}_i^T\boldsymbol{\beta}) + \mathrm{Var}(\nu_i) + \mathrm{Var}(\epsilon_{ij})\end{aligned} \tag{2.13}$$

since $(\boldsymbol{x}_{ij} - \bar{\boldsymbol{x}}_i)$, $\bar{\boldsymbol{x}}_i$, ν_i, and ϵ_{ij} are independent.

The WS variance of the response variable explained by fixed location effects of WS components of observation-level covariates, denoted as *f1*, can be derived based on the property of the multivariate normal distribution [39]:

$$\mathrm{Var}((\boldsymbol{x}_{ij} - \bar{\boldsymbol{x}}_i)^T\boldsymbol{\beta}) = \boldsymbol{\beta}^T\boldsymbol{\Phi}_{\boldsymbol{x}}^{ws}\boldsymbol{\beta}. \tag{2.14}$$

Similarly, the BS variance of the response variable explained by fixed location effects of subject-level covariates and BS components of observation-level covariates, denoted as *f2*, can be expressed as:

$$\mathrm{Var}(\bar{\boldsymbol{x}}_i^T\boldsymbol{\beta}) = \boldsymbol{\beta}^T\boldsymbol{\Phi}_{\boldsymbol{x}}^{bs}\boldsymbol{\beta}. \tag{2.15}$$

In applying these formulas, we substitute $\boldsymbol{\Phi}_{\boldsymbol{x}}^{ws}$ and $\boldsymbol{\Phi}_{\boldsymbol{x}}^{bs}$ with their sample estimators.

By the law of total variance, the variance of the random intercepts is

$$\mathrm{Var}(\nu_i) = \exp(\alpha_0 + \boldsymbol{\mu}_{\boldsymbol{u}}^T\boldsymbol{\alpha} + \frac{\boldsymbol{\alpha}^T(\boldsymbol{\Phi}_{\boldsymbol{u}}^{ws} + \boldsymbol{\Phi}_{\boldsymbol{u}}^{bs})\boldsymbol{\alpha}}{2}), \tag{2.16}$$

denoted as v. Let $m = \exp(\alpha_0 + \boldsymbol{\mu}_{\boldsymbol{u}}^T\boldsymbol{\alpha})$ represent the variance of the random intercepts at

the mean (both on the BS level and the WS level) of all covariates influencing the variance of the random intercepts. Note that the mean of the WS component of a covariate is zero. Then, $v - m$ represents the variance of the random intercepts explained by covariates. To further decompose $v - m$, logarithmic transformation is taken:

$$\log(v) - \log(m) = \frac{\boldsymbol{\alpha}^T \boldsymbol{\Phi}_{\boldsymbol{u}}^{bs} \boldsymbol{\alpha}}{2} + \frac{\boldsymbol{\alpha}^T \boldsymbol{\Phi}_{\boldsymbol{u}}^{ws} \boldsymbol{\alpha}}{2}, \tag{2.17}$$

where $v1 = \frac{\boldsymbol{\alpha}^T \boldsymbol{\Phi}_{\boldsymbol{u}}^{ws} \boldsymbol{\alpha}}{2}$ represents the log-transformed variance of the random intercepts explained by fixed effects of WS components of observation-level covariates. $\frac{\boldsymbol{\alpha}^T \boldsymbol{\Phi}_{\boldsymbol{u}}^{bs} \boldsymbol{\alpha}}{2}$, denoted as $v2$, is the log-transformed variance of the random intercepts explained by fixed effects of subject-level covariates and BS components of observation-level covariates. We use sample estimators of $\boldsymbol{\mu}_{\boldsymbol{u}}$, $\boldsymbol{\Phi}_{\boldsymbol{u}}^{ws}$, and $\boldsymbol{\Phi}_{\boldsymbol{u}}^{bs}$ in place of these population parameters themselves in practice.

The variance of the observation-level residuals, also known as the scale of the response variable, is

$$\mathrm{Var}(\epsilon_{ij}) = \exp(\tau_0 + \boldsymbol{\mu}_{\boldsymbol{w}}^T \boldsymbol{\tau} + \frac{\boldsymbol{\tau}^T (\boldsymbol{\Phi}_{\boldsymbol{w}}^{ws} + \boldsymbol{\Phi}_{\boldsymbol{w}}^{bs}) \boldsymbol{\tau} + \sigma_\omega^2}{2}), \tag{2.18}$$

denoted as e. $e0 = \exp(\tau_0 + \boldsymbol{\mu}_{\boldsymbol{w}}^T \boldsymbol{\tau})$ is the variance of the observation-level residuals at the mean (both on the BS level and the WS level) of all covariates in the scale model, which is also the unexplained scale of the response variable, and $e - e0$ is the variance of the observation-level residuals explained by covariates and the random scale effects. Similar to the decomposition of $v - m$, $e - e0$ can be further decomposed on the logarithmic scale:

$$\log(e) - \log(e0) = \frac{\boldsymbol{\tau}^T \boldsymbol{\Phi}_{\boldsymbol{w}}^{bs} \boldsymbol{\tau}}{2} + \frac{\boldsymbol{\tau}^T \boldsymbol{\Phi}_{\boldsymbol{w}}^{ws} \boldsymbol{\tau}}{2} + \frac{\sigma_\omega^2}{2}, \tag{2.19}$$

where $e1 = \frac{\boldsymbol{\tau}^T \boldsymbol{\Phi}_{\boldsymbol{w}}^{ws} \boldsymbol{\tau}}{2}$ is the log-transformed scale of the response variable explained by fixed effects of WS components of observation-level covariates, $e2 = \frac{\boldsymbol{\tau}^T \boldsymbol{\Phi}_{\boldsymbol{w}}^{bs} \boldsymbol{\tau}}{2}$ is the log-transformed scale of the response variable explained by fixed effects of subject-level covariates and BS

components of observation-level covariates, and $d = \frac{\sigma_\omega^2}{2}$ is the log-transformed scale of the response variable explained by the random scale effects. When applied, the sample estimators of $\boldsymbol{\mu_w}$, $\boldsymbol{\Phi_w^{bs}}$, and $\boldsymbol{\Phi_w^{ws}}$ are used in Equations 2.18 and 2.19.

For simplicity, here the coefficients $\beta's$, $\alpha's$, and $\tau's$ are assumed to be the same for the BS and WS components of covariates. However, in some cases, it may be desirable to allow the BS and the WS component of a covariate to have different effects. For this, one can substitute the corresponding coefficients of distinct BS and WS effects in the appropriate equations. Our R function described in Section 2.2.4 allows for this possibility. Users would need to specify the BS component and the WS component of a covariate as two distinct variables in their dataset and input their corresponding effect estimates into the function separately.

Variance partitioning for MELS models with random slopes of observation-level covariates

Here, we decompose the variance of the response variable based on the model described in Section 2.1.2.

$$\begin{aligned}
\text{Var}(y_{ij}) &= \text{Var}(\beta_0 + \boldsymbol{x}_{ij}^T\boldsymbol{\beta} + \boldsymbol{z}_{ij}^T\boldsymbol{\nu}_i + \epsilon_{ij}) \\
&= \text{Var}((\boldsymbol{x}_{ij} - \bar{\boldsymbol{x}}_i)^T\boldsymbol{\beta} + \bar{\boldsymbol{x}}_i^T\boldsymbol{\beta} + (\boldsymbol{z}_{ij} - \bar{\boldsymbol{z}}_i)^T\boldsymbol{\nu}_i + \bar{\boldsymbol{z}}_i^T\boldsymbol{\nu}_i + \epsilon_{ij}) \\
&= \text{Var}((\boldsymbol{x}_{ij} - \bar{\boldsymbol{x}}_i)^T\boldsymbol{\beta}) + \text{Var}(\bar{\boldsymbol{x}}_i^T\boldsymbol{\beta}) + \text{Var}((\boldsymbol{z}_{ij} - \bar{\boldsymbol{z}}_i)^T\boldsymbol{\nu}_i) + \text{Var}(\bar{\boldsymbol{z}}_i^T\boldsymbol{\nu}_i) + \text{Var}(\epsilon_{ij})
\end{aligned} \tag{2.20}$$

given the independence of $(\boldsymbol{x}_{ij} - \bar{\boldsymbol{x}}_i)$, $\bar{\boldsymbol{x}}_i$, $\boldsymbol{\nu}_i$, and ϵ_{ij} as well as the independence of $(\boldsymbol{z}_{ij} - \bar{\boldsymbol{z}}_i)$ and $\bar{\boldsymbol{z}}_i$.

The interpretations and derivations for $\text{Var}((\boldsymbol{x}_{ij} - \bar{\boldsymbol{x}}_i)^T\boldsymbol{\beta})$ (*f1*), $\text{Var}(\bar{\boldsymbol{x}}_i^T\boldsymbol{\beta})$ (*f2*), and $\text{Var}(\epsilon_{ij})$ (e) are the same as in Section 2.2.2, while the variance partitioning for the random

location effects are given by, with $\mathrm{Tr}(\cdot)$ representing the trace of a matrix:

$$\mathrm{Var}((\boldsymbol{z}_{ij} - \bar{\boldsymbol{z}}_i)^T \boldsymbol{\nu}_i) = \mathrm{Tr}(\boldsymbol{\Phi}_{\boldsymbol{z}}^{ws} \boldsymbol{\Sigma}_{\boldsymbol{\nu}}) \tag{2.21}$$

denoted as *v1*, which is the WS variance of the response variable explained by random slopes of WS components of observation-level covariates, and

$$\mathrm{Var}(\bar{\boldsymbol{z}}_i^T \boldsymbol{\nu}_i) = \mathrm{Tr}(\boldsymbol{\Phi}_{\boldsymbol{z}}^{bs} \boldsymbol{\Sigma}_{\boldsymbol{\nu}}) + \boldsymbol{\mu}_{\boldsymbol{z}}^T \boldsymbol{\Sigma}_{\boldsymbol{\nu}} \boldsymbol{\mu}_{\boldsymbol{z}}. \tag{2.22}$$

Here, $\mathrm{Tr}(\boldsymbol{\Phi}_{\boldsymbol{z}}^{bs} \boldsymbol{\Sigma}_{\boldsymbol{\nu}})$ is denoted as *v2*, which corresponds to the BS variance of the response variable explained by random slopes of BS components of observation-level covariates. Also, $\boldsymbol{\mu}_{\boldsymbol{z}}^T \boldsymbol{\Sigma}_{\boldsymbol{\nu}} \boldsymbol{\mu}_{\boldsymbol{z}}$ is denoted as m, which represents the BS variance of the response variable explained by the random intercepts at the mean of BS components of all covariates for random location effects. The derivations of Equation 2.21 and Equation 2.22 can be found in Rights and Sterba's work [42].

2.2.3 Defining R^2 measures

We develop R^2 measures for the total variance of the response variable, the level-specific variances of the response variable, and the scale of the response variable.

Table 2.1 details R^2 measures for the location part of the model described in Section 2.1.1, and Table 2.2 describes R^2 measures for the location part of the model presented in Section 2.1.2. R^2 measures for the scale model are illustrated in Table 2.3. The superscripts in parentheses denote the source(s) of variation. Also, the subscripts represent the denominators of the R^2 measures, meaning which part of the variance of the response variable that one is trying to explain. Namely, subscript t indicates total variance of the response variable and is calculated as $f1 + f2 + v + e$ for MELS models with random intercepts (with covariate-influenced variance), and $f1 + f2 + v1 + v2 + m + e$ for MELS models with random

slopes of observation-level covariates. Subscript w and subscript b represent WS variance of the response variable and BS variance of the response variable, respectively. For the model described in Section 2.1.1, the model-implied WS variance of the response variable is $f1 + e$, and the model-implied BS variance of the response variable is $f2 + v$. For the Section 2.1.2 model, the WS variance of the response variable equals $f1 + v1 + e$ while the BS variance of the response variable is calculated as $f2 + v2 + m$. Lastly, the subscript s represents variance of the observation-level residuals, which is denoted as e in both specifications of MELS models.

The R^2s defined can measure the variance of the response variable explained by single sources of variation. Namely, $R_t^{2(f1)}$, $R_t^{2(f2)}$, $R_t^{2(v1)}$,$R_t^{2(v2)}$,$R_t^{2(m)}$, $R_w^{2(f1)}$, $R_b^{2(f2)}$, $R_b^{2(v1)}$, $R_b^{2(v2)}$, and $R_b^{2(m)}$ in Table 2.1, $R_t^{2(f1)}$, $R_t^{2(f2)}$, $R_t^{2(v1)}$, $R_t^{2(v2)}$, $R_t^{2(m)}$, $R_w^{2(f1)}$, $R_w^{2(v1)}$, $R_b^{2(f2)}$, $R_b^{2(v2)}$, and $R_b^{2(m)}$ in Table 2.2, as well as $R_s^{2(e1)}$, $R_s^{2(e2)}$, and $R_s^{2(d)}$ in Table 2.3 are single-source R^2 measures. We also define R^2s that measure joint effects of multiple parts of the models. Specifically, $f = f1 + f2$ represents the variance of the response variable explained by fixed location effects of both subject-level covariates and observation-level covariates, v in Table 2.1 represents the variance of the response variable explained by the random intercepts, and $v = v1 + v2$ in Table 2.2 represents the variance of the response variable explained by the random slopes of both the WS components and BS components of observation-level covariates.

Since the proportion of the total variance of the response variable that is BS can be of interest to researchers applying MELS models, we add $R_t^{2(f2v)}$ in Table 2.1 and $R_t^{2(f2v2m)}$ in Table 2.2. These two R^2s measure the proportion of total variance of the response variable explained by BS location effects.

Table 2.1: Definitions and interpretations of R^2 measures for the location part of an MELS model with random intercepts with covariate-influenced variance. The coefficients and covariates refer to elements of the model needed to calculate the source of variation in the specific R^2 measure, i.e. what is labeled in the parenthesized superscript

Definition	**Coefficients**	**Covariates**	**Interpretation**
R^2 Measures for Total Variance of the Response Variable			
$R_t^{2(f1)} = \frac{f1}{f1 + f2 + v + e}$	$\boldsymbol{\beta}$	$(\boldsymbol{x}_{ij} - \bar{\boldsymbol{x}}_i)$	Proportion of total variance of the response variable explained by fixed location effects of WS components of observation-level covariates
$R_t^{2(f2)} = \frac{f2}{f1 + f2 + v + e}$	$\boldsymbol{\beta}$	$\bar{\boldsymbol{x}}_i$	Proportion of total variance of the response variable explained by fixed location effects of subject-level covariates and BS components of observation-level covariates
$R_t^{2(f)} = \frac{f1 + f2}{f1 + f2 + v + e}$	$\boldsymbol{\beta}$	$(\boldsymbol{x}_{ij} - \bar{\boldsymbol{x}}_i)$, $\bar{\boldsymbol{x}}_i$	Proportion of total variance of the response variable explained by fixed location effects
$R_t^{2(v1)} = \frac{\frac{v1}{v1+v2}(v-m)}{f1 + f2 + v + e}$	$\boldsymbol{\alpha}$	$(\boldsymbol{u}_{ij} - \bar{\boldsymbol{u}}_i)$	Proportion of total variance of the response variable explained by fixed effects of WS components of observation-level covariates on the variance of the random intercepts
$R_t^{2(v2)} = \frac{\frac{v2}{v1+v2}(v-m)}{f1 + f2 + v + e}$	$\boldsymbol{\alpha}$	$\bar{\boldsymbol{u}}_i$	Proportion of total variance of the response variable explained by fixed effects of subject-level covariates and BS components of observation-level covariates on the variance of the random intercepts
$R_t^{2(m)} = \frac{m}{f1 + f2 + v + e}$	$\alpha_0, \boldsymbol{\alpha}$	$\bar{\boldsymbol{u}}_i$	Proportion of total variance of the response variable explained by random intercepts at the mean of all covariates influencing the variance of the random intercepts
$R_t^{2(v)} = \frac{v}{f1 + f2 + v + e}$	$\alpha_0, \boldsymbol{\alpha}$	$(\boldsymbol{u}_{ij} - \bar{\boldsymbol{u}}_i)$, $\bar{\boldsymbol{u}}_i$	Proportion of total variance of the response variable explained by random location effects
$R_t^{2(fv)} = \frac{f1 + f2 + v}{f1 + f2 + v + e}$	$\boldsymbol{\beta}, \alpha_0, \boldsymbol{\alpha}$	$(\boldsymbol{x}_{ij} - \bar{\boldsymbol{x}}_i)$, $\bar{\boldsymbol{x}}_i$, $(\boldsymbol{u}_{ij} - \bar{\boldsymbol{u}}_i)$, $\bar{\boldsymbol{u}}_i$	Proportion of total variance of the response variable explained by both fixed location effects and random location effects
$R_t^{2(f2v)} = \frac{f2 + v}{f1 + f2 + v + e}$	$\boldsymbol{\beta}, \alpha_0, \boldsymbol{\alpha}$	$\bar{\boldsymbol{x}}_i$, $(\boldsymbol{u}_{ij} - \bar{\boldsymbol{u}}_i)$, $\bar{\boldsymbol{u}}_i$	Proportion of total variance of the response variable explained by BS location effects
R^2 Measures for BS Variance of the Response Variable			
$R_b^{2(f2)} = \frac{f2}{f2 + v}$	$\boldsymbol{\beta}$	$\bar{\boldsymbol{x}}_i$	Proportion of BS variance of the response variable explained by fixed location effects of subject-level covariates and BS components of observation-level covariates
$R_b^{2(v1)} = \frac{\frac{v1}{v1 + v2}(v-m)}{f2 + v}$	$\boldsymbol{\alpha}$	$(\boldsymbol{u}_{ij} - \bar{\boldsymbol{u}}_i)$	Proportion of BS variance of the response variable explained by fixed effects of WS components of observation-level covariates on the variance of the random intercepts
$R_b^{2(v2)} = \frac{\frac{v2}{v1 + v2}(v-m)}{f2 + v}$	$\boldsymbol{\alpha}$	$\bar{\boldsymbol{u}}_i$	Proportion of BS variance of the response variable explained by fixed effects of subject-level covariates and BS components of observation-level covariates on the variance of the random intercepts
$R_b^{2(m)} = \frac{m}{f2 + v}$	$\alpha_0, \boldsymbol{\alpha}$	$\bar{\boldsymbol{u}}_i$	Proportion of BS variance of the response variable explained by random intercepts at the mean of all covariates influencing the variance of the random intercepts
$R_b^{2(v)} = \frac{v}{f2 + v}$	$\alpha_0, \boldsymbol{\alpha}$	$(\boldsymbol{u}_{ij} - \bar{\boldsymbol{u}}_i)$, $\bar{\boldsymbol{u}}_i$	Proportion of BS variance of the response variable explained by random location effects
R^2 Measures for WS Variance of the Response Variable			
$R_w^{2(f1)} = \frac{f1}{f1 + e}$	$\boldsymbol{\beta}$	$(\boldsymbol{x}_{ij} - \bar{\boldsymbol{x}}_i)$	Proportion of WS variance of the response variable explained by fixed location effects of WS components of observation-level covariates

Table 2.2: Definitions and interpretations of R^2 measures for the location part of an MELS model with random slopes of observation-level covariates. The coefficients and covariates refer to elements of the model needed to calculate the source of variation in the specific R^2 measure, i.e. what is labeled in the parenthesized superscript

Definition	**Coefficients**	**Covariates**	**Interpretation**
R^2 Measures for Total Variance of the Response Variable			
$R_t^{2(f1)} = \frac{f1}{f1 + f2 + v1 + v2 + m + e}$	$\boldsymbol{\beta}$	$(\boldsymbol{x}_{ij} - \bar{\boldsymbol{x}}_i)$	Proportion of total variance of the response variable explained by fixed location effects of WS components of observation-level covariates
$R_t^{2(f2)} = \frac{f2}{f1 + f2 + v1 + v2 + m + e}$	$\boldsymbol{\beta}$	$\bar{\boldsymbol{x}}_i$	Proportion of total variance of the response variable explained by fixed location effects of subject-level covariates and BS components of observation-level covariates
$R_t^{2(f)} = \frac{f1 + f2}{f1 + f2 + v1 + v2 + m + e}$	$\boldsymbol{\beta}$	$(\boldsymbol{x}_{ij} - \bar{\boldsymbol{x}}_i)$, $\bar{\boldsymbol{x}}_i$	Proportion of total variance of the response variable explained by fixed location effects
$R_t^{2(v1)} = \frac{v1}{f1 + f2 + v1 + v2 + m + e}$	$\boldsymbol{\nu}_i$	$(\boldsymbol{z}_{ij} - \bar{\boldsymbol{z}}_i)$	Proportion of total variance of the response variable explained by random slopes of WS components of observation-level covariates
$R_t^{2(v2)} = \frac{v2}{f1 + f2 + v1 + v2 + m + e}$	$\boldsymbol{\nu}_i$	$\bar{\boldsymbol{z}}_i$	Proportion of total variance of the response variable explained by random slopes of BS components of observation-level covariates
$R_t^{2(m)} = \frac{m}{f1 + f2 + v1 + v2 + m + e}$	$\boldsymbol{\nu}_i$	$\bar{\boldsymbol{z}}_i$	Proportion of total variance of the response variable explained by random intercepts at the mean of BS components of all covariates for random location effects
$R_t^{2(vm)} = \frac{v1 + v2 + m}{f1 + f2 + v1 + v2 + m + e}$	$\boldsymbol{\nu}_i$	$(\boldsymbol{z}_{ij} - \bar{\boldsymbol{z}}_i), \bar{\boldsymbol{z}}_i$	Proportion of total variance of the response variable explained by random location effects
$R_t^{2(fvm)} = \frac{f1 + f2 + v1 + v2 + m}{f1 + f2 + v1 + v2 + m + e}$	$\boldsymbol{\beta}$, $\boldsymbol{\nu}_i$	$(\boldsymbol{x}_{ij} - \bar{\boldsymbol{x}}_i)$, $\bar{\boldsymbol{x}}_i$, $(\boldsymbol{z}_{ij} - \bar{\boldsymbol{z}}_i)$, $\bar{\boldsymbol{z}}_i$	Proportion of total variance of the response variable explained by both fixed location effects and random location effects
$R_t^{2(f2v2m)} = \frac{f2 + v2 + m}{f1 + f2 + v1 + v2 + m + e}$	$\boldsymbol{\beta}$, $\boldsymbol{\nu}_i$	$\bar{\boldsymbol{x}}_i$, $\bar{\boldsymbol{z}}_i$	Proportion of total variance of the response variable explained by BS location effects

2.2.4 Implementation in R

The commented code for an R function named *R2MELS* and descriptions of its arguments are provided in the supplementary files. The function is developed based on Rights and Sterba's *r2MLMlong* function [42]. Users input their parameter estimates of a MELS model, and the function will output two tables of variance partitioning results (one for the location part of the model, and the other for the scale part of the model), two tables of R^2 values (one for the location part of the model, and the other for the scale part of the model) as well as a stacked bar plot of the single-source R^2 values.

Table 2.2: Definitions and interpretations of R^2 measures for the location part of an MELS model with random slopes of observation-level covariates. The coefficients and covariates refer to elements of the model needed to calculate the source of variation in the specific R^2 measure, i.e. what is labeled in the parenthesized superscript (continued)

Definition	**Coefficients**	**Covariates**	**Interpretation**
	R^2 Measures for BS Variance of the Response Variable		
$R_b^{2(f2)} = \frac{f2}{f2 + v2 + m}$	$\boldsymbol{\beta}$	$\bar{\boldsymbol{x}}_i$	Proportion of BS variance of the response variable explained by fixed location effects of subject-level covariates and BS components of observation-level covariates
$R_b^{2(v2)} = \frac{v2}{f2 + v2 + m}$	$\boldsymbol{\nu}_i$	$\bar{\boldsymbol{z}}_i$	Proportion of BS variance of the response variable explained by random slopes of BS components of observation-level covariates
$R_b^{2(m)} = \frac{m}{f2 + v2 + m}$	$\boldsymbol{\nu}_i$	$\bar{\boldsymbol{z}}_i$	Proportion of BS variance of the response variable explained by random intercepts at the mean of BS components of all covariates for random location effects
$R_b^{2(v2m)} = \frac{v2 + m}{f2 + v2 + m}$	$\boldsymbol{\nu}_i$	$\bar{\boldsymbol{z}}_i$	Proportion of BS variance of the response variable explained by random location effects
	R^2 Measures for WS Variance of the Response Variable		
$R_w^{2(f1)} = \frac{f1}{f1 + v1 + e}$	$\boldsymbol{\beta}$	$(\boldsymbol{x}_{ij} - \bar{\boldsymbol{x}}_i)$	Proportion of WS variance of the response variable explained by fixed location effects of WS components of observation-level covariates
$R_w^{2(v1)} = \frac{v1}{f1 + v1 + e}$	$\boldsymbol{\nu}_i$	$(\boldsymbol{z}_{ij} - \bar{\boldsymbol{z}}_i)$	Proportion of WS variance of the response variable explained by random slopes of WS components of observation-level covariates
$R_w^{2(f1v1)} = \frac{f1 + v1}{f1 + v1 + e}$	$\boldsymbol{\beta}$, $\boldsymbol{\nu}_i$	$(\boldsymbol{x}_{ij} - \bar{\boldsymbol{x}}_i)$, $(\boldsymbol{z}_{ij} - \bar{\boldsymbol{z}}_i)$	Proportion of WS variance of the response variable explained by both fixed location effects and random slopes of WS components of observation-level covariates

Table 2.3: Definitions and interpretations of R^2 measures for the scale part of an MELS model. The coefficients and covariates refer to elements of the model needed to calculate the source of variation in the specific R^2 measure, i.e. what is labeled in the parenthesized superscript

Definition	**Coeficients**	**Covariates**	**Interpretation**
$R_s^{2(e1)} = \frac{\frac{e1}{e1 + e2 + d}(e - e0)}{e}$	$\boldsymbol{\tau}$	$(\boldsymbol{w}_{ij} - \bar{\boldsymbol{w}}_i)$	Proportion of variance of observation-level residuals explained by WS components of observation-level covariates
$R_s^{2(e2)} = \frac{\frac{e2}{e1 + e2 + d}(e - e0)}{e}$	$\boldsymbol{\tau}$	$\bar{\boldsymbol{w}}_i$	Proportion of variance of observation-level residuals explained by subject-level covariates and BS components of observation-level covariates
$R_s^{2(e1e2)} = \frac{\frac{e1 + e2}{e1 + e2 + d}(e - e0)}{e}$	$\boldsymbol{\tau}$	$(\boldsymbol{w}_{ij} - \bar{\boldsymbol{w}}_i), \bar{\boldsymbol{w}}_i$	Proportion of variance of observation-level residuals explained by covariates
$R_s^{2(d)} = \frac{\frac{d}{e1 + e2 + d}(e - e0)}{e}$	σ_ω	N/A	Proportion of variance of observation-level residuals explained by random scale effects

2.3 Simulation Study

To assess the validity of the proposed method, we conducted a small simulation study. Specifically, we fitted MELS models to 500 simulated datasets (200 subjects in each sample, 50 observations of each subject). For each simulated dataset, we allowed for two subject-level covariates, $\begin{pmatrix} x_{1i} \\ x_{2i} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 & 0.5 \\ 0.5 & 1.5 \end{pmatrix}\right)$, and three observation-level covariates that vary purely within-subjects, $\begin{pmatrix} x_{3ij} \\ x_{4ij} \\ x_{5ij} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 & 0.3 & 0.75 \\ 0.3 & 1.5 & 0.2 \\ 0.75 & 0.2 & 1 \end{pmatrix}\right)$.

The location model was specified as follows:

$$y_{ij} = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3ij} + \beta_4 x_{4ij} + \beta_5 x_{5ij} + \sigma_{\nu_i}\theta_{1i} + \epsilon_{ij}. \tag{2.23}$$

For the variance of the random intercept ν_i,

$$\sigma^2_{\nu_i} = \exp(\alpha_0 + \alpha x_{3ij}). \tag{2.24}$$

For the variance of the observation-level residuals,

$$\sigma^2_{\epsilon_{ij}} = \exp(\tau_0 + \tau_1 x_{3ij} + \tau_2 x_{4ij} + \tau_3 x_{5ij} + \sigma_\omega \theta_{2i}), \tag{2.25}$$

and

$$\begin{pmatrix} \theta_{1i} \\ \theta_{2i} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho_{\nu\omega} \\ \rho_{\nu\omega} & 1 \end{pmatrix}\right). \tag{2.26}$$

The generating parameters and average estimates from SAS PROC NLMIXED are summarized in Table 2.4, and the corresponding theoretical values of R^2 measures and average simulated values can be found in Table 2.5. As can be seen, all parameter estimates were

well recovered, and all average simulated R^2 values resemble their corresponding theoretical values. Namely, all differences between the average simulated values and their corresponding theoretical values are lower than 0.004.

Table 2.4: Generating parameters and mean parameter estimates from 500 simulations

Parameter	**True Value**	**Simulated Values Mean(SD)**
β_0	1	1.000(0.069)
β_1	-0.5	-0.497(0.020)
β_2	2	1.999(0.024)
β_3	1	1.000(0.013)
β_4	-2	-2.000(0.007)
β_5	3	3.000(0.010)
α_0	0.1	0.098(0.103)
α_1	0.4	0.401(0.012)
τ_0	0.2	0.199(0.061)
τ_1	0.3	0.300(0.013)
τ_2	-0.1	-0.101(0.013)
τ_3	0.5	0.502(0.018)
σ_ω^2	0.7	0.697(0.076)
$\rho_{\nu\omega}$	0.1	0.106 (0.075)

Table 2.5: Theoretical values of R^2 measures and average simulated values from 500 simulations

R^2 **Measure**	**Theoretical Value**	**Simulated Values Mean(SD)**
R^2 Measures for Total Variance of the Response Variable		
$R_t^{2(f1)} = \frac{f1}{f1 + f2 + v + e}$	0.661	0.663(0.015)
$R_t^{2(f2)} = \frac{f2}{f1 + f2 + v + e}$	0.203	0.201(0.017)
$R_t^{2(f)} = \frac{f1 + f2}{f1 + f2 + v + e}$	0.864	0.864(0.008)
$R_t^{2(v1)} = \frac{v\text{-}m}{f1 + f2 + v + e}$	0.007	0.007(0.001)
$R_t^{2(m)} = \frac{m}{f1 + f2 + v + e}$	0.041	0.041(0.004)
$R_t^{2(v)} = \frac{v}{f1 + f2 + v + e}$	0.048	0.048(0.005)
$R_t^{2(fv)} = \frac{f1 + f2 + v}{f1 + f2 + v + e}$	0.912	0.912(0.006)
$R_t^{2(f2v)} = \frac{f2 + v}{f1 + f2 + v + e}$	0.251	0.249(0.017)
R^2 Measures for BS Variance of the Response Variable		
$R_b^{2(f2)} = \frac{f2}{f2 + v}$	0.809	0.806(0.023)
$R_b^{2(v1)} = \frac{v-m}{f2 + v}$	0.028	0.029(0.004)
$R_b^{2(m)} = \frac{m}{f2 + v}$	0.163	0.165(0.020)
$R_b^{2(v)} = \frac{v}{f2 + v}$	0.191	0.194(0.023)
R^2 Measures for WS Variance of the Response Variable		
$R_b^{2(f1)} = \frac{f1}{f1 + e}$	0.883	0.883(0.008)
R^2 Measures for Scale of the Response Variable		
$R_s^{2(e1)} = \frac{\frac{e1}{e1+d}(e-e0)}{e}$	0.231	0.232(0.009)
$R_s^{2(d)} = \frac{\frac{d}{e1+d}(e-e0)}{e}$	0.256	0.254(0.024)

2.4 Examples

2.4.1 Example 1: Health Behaviors Data

Flueckiger et al. [11] collected intensive longitudinal data on 72 first-year psychology students from the University of Basel regarding their sleep quality, physical activity, positive and negative affect, learning goal achievement, and examination grades. We fit a MELS model on these data to examine how negative affect (NA) of the students and survey day influenced

their mean positive affect (PA), how survey day influenced the BS variance of PA, as well as how survey day influenced the WS variance of PA. PA and NA were measured on 7-point Likert scales in which 1 means "not at all" and 7 means "extremely". We used a grand-mean-centered and scaled version of survey day (Day_c), for which 1 unit indicates 1 week.

The location model is as follows:

$$PA_{ij} = \beta_0 + \beta_{NA} NA_{ij} + \beta_{Day\_c} Day\_c_{ij} + \sigma_{\nu_i}\theta_{1i} + \epsilon_{ij}. \tag{2.27}$$

The variance of the random subject intercepts and the variance of the observation-level residuals are modeled as

$$\sigma^2_{\nu_i} = \exp(\alpha_0 + \alpha_{Day\_c} Day\_c_{ij}) \tag{2.28}$$

and

$$\sigma^2_{\epsilon_{ij}} = \exp(\tau_0 + \tau_{Day\_c} Day\_c_{ij} + \sigma_\omega \theta_{2i}), \tag{2.29}$$

respectively. The random effects θ_{1i} and θ_{2i} follow the same bivariate normal distribution as specified in Equation 2.6. The parameter estimates from SAS PROC NLMIXED are included in Table 2.6, and a visualization of the proposed R^2 measures for this example is shown in Figure 2.1.

The first three bars on the left of Figure 2.1 correspond to the total variance, the BS variance, and the WS variance of PA, respectively. Within the bars, the red blocks represent proportion of the variance of PA explained by fixed location effects of WS components of observation-level covariates, $(NA_{ij} - \overline{NA}_i)$ and $(Day\_c_{ij} - \overline{Day\_c}_i)$, while the orange blocks can be interpreted as the proportion of the variance of PA, that is, the variance of observation-level residuals. Blue blocks correspond to the BS variance of PA. Specifically, the darkest blue blocks indicate the proportion of the variance of PA explained by the effect of $(Day\_c_{ij} - \overline{Day\_c}_i)$ on the variance of the random intercepts, and the lightest blue

Table 2.6: Parameter estimates of the MELS model on health behaviors data

Parameter	**Estimate**	**SE**	**T-value**	**P-value**
β_0	5.855	0.109	53.631	< 0.001
β_{NA}	-0.644	0.016	-39.900	< 0.001
β_{Day_c}	-0.077	0.015	-5.247	< 0.001
α_0	-0.376	0.176	-2.144	0.036
α_{Day_c}	0.139	0.031	4.531	< 0.001
τ_0	-0.603	0.093	-6.509	< 0.001
τ_{Day_c}	0.073	0.025	2.888	0.005
σ_ω	0.735	0.070	10.451	< 0.001
$\rho_{\nu\omega}$	-0.500	0.100	-5.009	< 0.001

represents the variance of the random intercepts at the means of both the BS component and the WS component of Day_c. The mid-blue blocks show the proportion of the variance of PA explained by the fixed location effects of the BS components of NA and Day_c. Grey blocks, which represent the proportion of Var(PA) explained by the variance of random intercepts explained by $\overline{Day\_c}_i$, are too small and thus almost invisible in the plot.

We can see that most (53.5%) of the variance of PA is within-subject, as represented by the red and orange blocks in the first column of Figure 2.1. For the WS variance of PA specifically, 42.3% is attributed to the fixed location effects of the WS components of NA and Day_c, which is visualized by the red block in the third column of Figure 2.1. In terms of the BS variance of PA, the random subject intercepts at the mean of both the BS component and the WS component of Day_c, which explain 63.1% of the BS variance of PA as indicated by the light blue portion of the middle column of Figure 2.1, are of particular importance. The R^2 measures for the scale model are summarized in the rightmost bar of

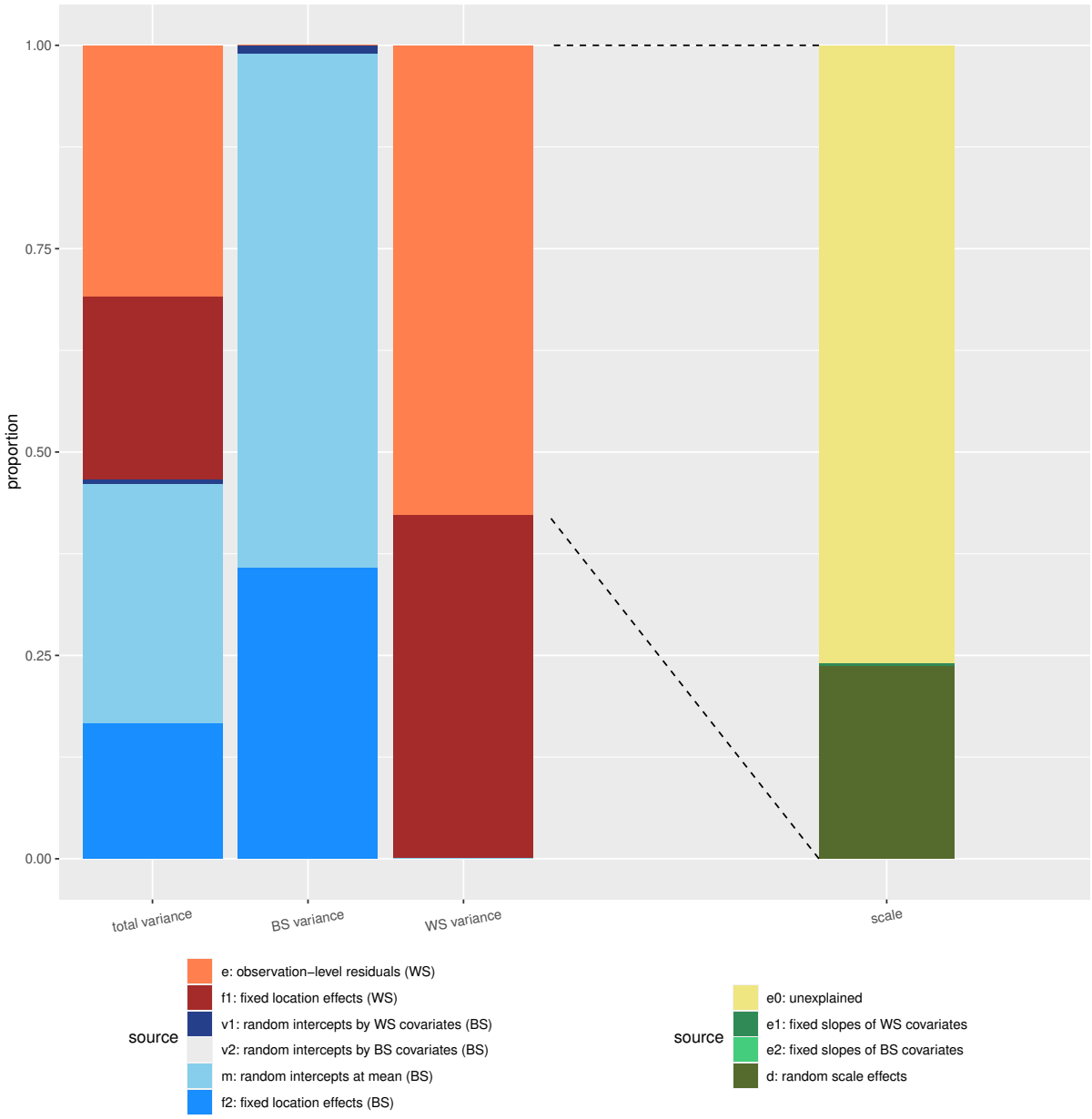


Figure 2.1: Variance partitioning for example 1: application to health behaviors data.

Figure 2.1. As shown by the proportion of the bottom dark olive green block in this bar, 23.6% of the scale of *PA* is explained by random scale effects. Less than 0.5% of the scale of *PA* is explained by covariates, which is made clear by the almost negligible proportion of the two green blocks in the middle of this bar.

2.4.2 Example 2: Depression Study Data

While the example in Section 2.4.1 presents a random intercepts model, in which its variance is modeled in terms of covariates, here we examine the application of our proposed R^2 framework to a MELS model with random slopes, as described in Section 2.1.2. The data are from Reisby et al.'s study [38] on the clinical responses of 66 depressed inpatients treated with anti-depressant medication. Here, we are interested in how patients' Hamilton depression score (*HamD*) changed following their weeks in the study (*week*). Additionally, *endog* is a subject-level dummy code that is coded as 1 if the patient had endogenous depression. The location model with response variable *HamD* and predictor *week*, controlled for *endog*, is specified as follows:

$$\begin{aligned} HamD_{ij} &= \beta_0 + \nu_{0i} + (\beta_{week} + \nu_{week,i})week_{ij} + \beta_{endog}endog_i + \epsilon_{ij} \\ &= \beta_0 + \sigma_{\nu_0}\theta_{0i} + (\beta_{week} + \sigma_{\nu_{week}}\theta_{1i})week_{ij} + \beta_{endog}endog_i + \epsilon_{ij}, \end{aligned} \tag{2.30}$$

where $\sigma_{\nu_0}\theta_{0i}$ is the individual deviation from the average intercept β_0, and $\sigma_{\nu_{week}}\theta_{1i}$ is the individual deviation from the average weekly change β_{week}.

The variance of the observation-level residuals is assumed to change over weeks and across subjects as well.

$$\sigma^2_{\epsilon_{ij}} = \exp(\tau_0 + \tau_{week}week_{ij} + \sigma_\omega\theta_{\omega i}). \tag{2.31}$$

The distribution of the standardized random effects is given by:

$$\begin{pmatrix} \theta_{0i} \\ \theta_{1i} \\ \theta_{\omega i} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho_{\nu_{0i}\nu_{week,i}} & \rho_{\nu_{0i}\omega_i} \\ \rho_{\nu_{0i}\nu_{week,i}} & 1 & \rho_{\nu_{week,i}\omega_i} \\ \rho_{\nu_{0i}\omega_i} & \rho_{\nu_{week,i}\omega_i} & 1 \end{pmatrix} \right), \tag{2.32}$$

where $\rho_{\nu_{0i}\nu_{week,i}}$ is the correlation between the random intercepts and the random slopes. Likewise, $\rho_{\nu_{0i}\omega_i}$ and $\rho_{\nu_{week,i}\omega_i}$ are the correlations between the random intercepts and the

random scale effects, and between the random slopes and the random scale effects, respectively.

Table 2.7 lists the parameter estimates for this model, and Figure 2.2 is the visualization of the proposed R^2 measures for this example. The meaning of each bar in Figure 2.2 can be interpreted as in Section 2.4.1, but for the variance of *HamD* instead of *PA*. The red blocks now represent the proportion of the variance of *HamD* explained by fixed location effects of the WS component of *week*. The orange blocks can be interpreted as the proportion of the variance of *HamD* that corresponds to observation-level residuals, and the mid-blue blocks indicate the proportion of the variance of *HamD* explained by fixed location effects of *endog* and the BS component of *week*. The newly added brown blocks correspond to the proportion of the variance of *HamD* explained by random slopes of WS components of *week*, and the dark blue blocks represent the proportion of the variance of *HamD* explained by random slopes of BS components of *week*. The light blue blocks still indicate the proportion of the variance of the response variable captured by the random intercepts but at the mean of $\overline{week}_i$.

As represented by the total proportion of the blue blocks in the first column of Figure 2.2, 35.8% of the variance of *HamD* is between-subjects. While random slopes of BS components of *week* explain very little BS variance of *HamD*, as $\overline{week}_i$ varies very little across subjects, the relative size of the light blue block in the second bar of Figure 2.2 shows that random subject intercepts at the mean of $\overline{week}_i$ explain 94.3% of BS variance of *HamD*. Fixed location effects of the WS component of *week* explain the most (47.5%, the relative space of the red block in the third bar of Figure 2.2) WS variance of *HamD* while another 16.3% is attributed to random slopes of the WS component of *week* as indicated by the brown block in that bar. 13.4% of the scale of *HamD* is explained by random scale effects while another 1.3% is explained by WS variation in *Week*. Less than 0.03% of the scale of *HamD* is explained by the BS component of *Week*.

Table 2.7: Parameter estimates of the MELS model on depression study data

Parameter	**Estimate**	**SE**	**T-value**	**P-value**
β_0	22.657	0.702	32.27	< 0.001
β_{week}	-2.357	0.201	-11.74	< 0.001
β_{endog}	1.533	0.924	1.66	0.102
τ_0	2.062	0.212	9.72	< 0.001
τ_{week}	0.100	0.067	1.51	0.137
σ_ω	0.538	0.116	4.65	< 0.001
σ_{ν_0}	3.154	0.481	6.56	< 0.001
$\sigma_{\nu_{week}}$	1.379	0.174	7.92	< 0.001
$\rho_{\nu_{0i}\nu_{week,i}}$	-0.207	0.180	-1.15	0.255
$\rho_{\nu_{0i}\omega}$	0.464	0.235	1.98	0.053
$\rho_{\nu_{week,i}\omega}$	-0.485	0.213	-2.28	0.026

2.5 Discussion

Our work extends Rights and Sterba's framework of defining R^2 measures for multilevel models [40, 42] to MELS models proposed by Hedeker et al. [15] and Nordgren et al. [34] The extended R^2 framework accommodates two special features of MELS modeling: (1) observation-level residual heteroskedasticity at different covariate values and across subjects; (2) inclusion of random location effects through either heteroskedastic random subject intercepts depending on covariates, or random subject intercepts and random subject slopes of observation-level covariates. We believe that this standardized effect size framework can facilitate the interpretation of MELS models and encourage wider use of this type of model.

In this study, our defined R^2 measures assume a two-level MELS model. Future work

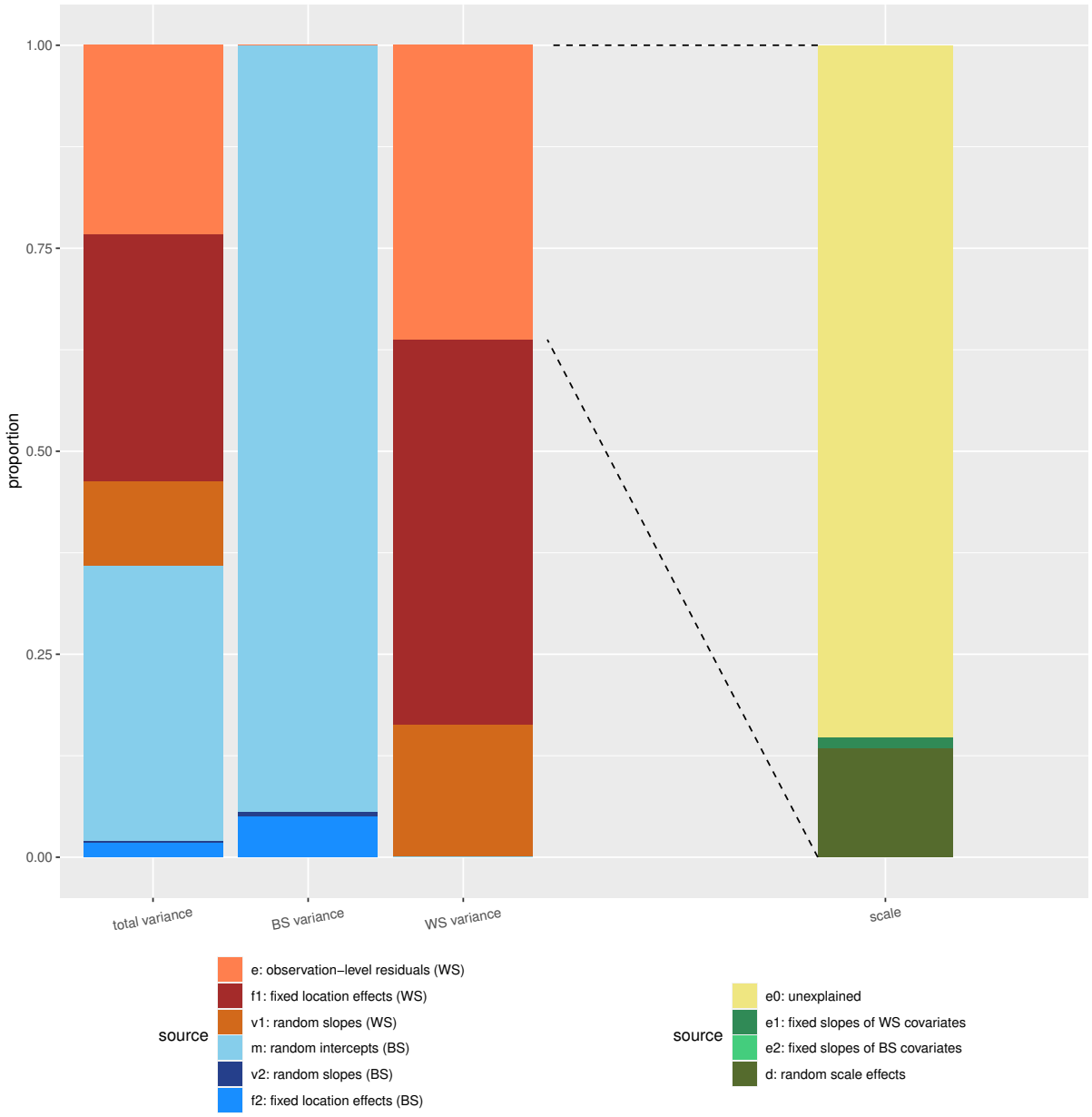


Figure 2.2: Variance partitioning for example 2: application to depression study data.

can extend these measures to three-level models, as developed by Lin et al. [22] in which WS variation of the response variable is further divided into WS variation between waves and WS variation within waves. Also, the proposed R^2 measures can be expanded to other kinds of outcomes, for example, count outcomes and ordinal outcomes. An application of a MELS model for ordinal data was discussed by Hedeker et al. [16]. Furthermore, future research

might take into consideration the autocorrelation of observation-level residuals. A recent development by Nestler [32] extends the MELS models to include AR(1) autocorrelation influenced by subject-level covariates and a random subject effect for the autocorrelation.

Currently, our framework focuses on point estimates of R^2 measures. While reporting the point estimates is conventional for R^2 measures, researchers interested in coverage and confidence intervals of these measures can apply bootstrapping methods to calculate these quantities. There is no existing literature on bootstrapping in MELS models to our knowledge, but researchers can refer to Goldstein's discussion on bootstrapping in multilevel models [14].

Lastly, this work focuses on defining R^2's for a single model, and comparisons of R^2's between different models are beyond the scope of this study. Researchers can refer to Rights and Sterba's recommendations on the use of R^2 differences in multilevel model comparisons for interpretation of differences between R^2 measures [41].

CHAPTER 3

DETECTING INFLUENTIAL SUBJECTS IN INTENSIVE LONGITUDINAL DATA USING MIXED-EFFECTS LOCATION SCALE MODELS

3.1 Introduction

In the past decades, intra-individual variability, also called within-subject (WS) heterogeneity or level-1 heterogeneity, of health behaviors and conditions has received increased attention [25, 54, 62]. Accordingly, new developments in statistical modeling, including mixed-effects location scale (MELS) models proposed by Hedeker et al. [15] and Nordgren et al. [34], allow researchers to model the mean, or location, and variability, or (square of the) scale of responses simultaneously.

Although these models can accommodate between-subject (BS) heterogeneity, also called level-2 heterogeneity, through random subject effects, there often exist subjects who are so different from the others that they need additional analyses and/or may affect model estimation. For example, in ecological momentary assessments (EMA) studies, a typical type of careless responses is that subjects give many consecutive items the same answer [27]. In this case, even if the data from these subjects do not deviate from any pattern in mean responses that subjects correctly responding to questions may show, they exhibit exceptionally low variability. Such abnormal behavior can distort the parameter estimates, especially the estimated variance of random scale effects. On the other hand, subjects who display exceptional consistency during behavioral studies, such as a study that aims at increasing subjects' physical activities, might be of interest as such a pattern may indicate good adherence to the study protocol. Being able to identify such subjects is important in terms of informing personalized healthcare. Especially, there have been discussions about using mixed-effects regression models (MRMs) to find individuals who behave differently

from others regarding their health conditions and outcomes [45, 46]. Since MELS models are an extension of standard MRMs, they can also perform similar functions and offer additional information about WS variability.

While there has been extensive literature discussing influential subjects and observations in MRMs [7, 19, 46] and software developed to implement these methods [33], to the best of our knowledge, there lack similar methods for the scale model of the response. As we will show in later sections, not considering WS variability in influence analysis can leave subject(s) with abnormal WS variability undetected. Hence, in the topic of influence analysis, WS variability deserves equal attention as the mean of the outcomes.

Therefore, we propose a framework for detecting influential subjects using MELS models. We examine each subject's influence on model fit as well as point estimates of parameters and their efficiency. The proposed method enables researchers to identify subjects who exhibit dubious patterns in their WS variability and/or mean, thus facilitating research on intra-individual variability. Unlike the traditional case-deletion approach of influential data detection, we adopt the method proposed by Langford and Lewis [19]. If a subject is being examined for its influence, it is removed from the estimation of level-2 effects and given subject-specific fixed effects. We will demonstrate how the estimation of the leave-one-out models can be carried out in SAS. Also, a health behavior study will be used as an example to illustrate the proposed methodology. Finally, we will assess the performance of our method and demonstrate its benefits over influence analysis using MRMs via simulated examples.

3.2 Methods

3.2.1 Leave-One-Out MELS Model

In this section, we first briefly review MELS models developed by Hedeker et al. [15]. For simplicity, only random intercepts with scalar variances are included in both the location

and the scale models described below. Nevertheless, more complicated models that include random slopes of time-varying covariates and/or have covariates influence the variances of the random effects are possible.

Suppose that subject i ($i = 1, 2, \ldots, N$) is measured at visit j ($j = 1, 2, \ldots, n_i$). The model for the response y_{ij} can be expressed as:

$$y_{ij} = \beta_0 + \nu_i + \boldsymbol{x}_{ij}^T \boldsymbol{\beta'} + \epsilon_{ij}, \tag{3.1}$$

where $\boldsymbol{x}_{ij}$ is a $p \times 1$ vector of time-varying covariates influencing the mean of y_{ij}, and ν_i is subject i's random intercept, indicating subject i's deviation from the fixed part of the model. β_0 is the fixed intercept, i.e. average response when all covariates equal 0, and $\boldsymbol{\beta'}$ is a $p \times 1$ vector of coefficients corresponding to $\boldsymbol{x}_{ij}$. ϵ_{ij} is the level-1 residual and follows a normal distribution with a mean of 0 and a variance of $\sigma^2_{\epsilon_{ij}}$. Later on, the notation $\boldsymbol{\beta}$ in influence measures represents a vector of which the first entry is β_0 and the remaining entries come from $\boldsymbol{\beta'}$.

In the MELS model, a log-linear sub-model is applied to the WS variance of the response to ensure a positive value:

$$\sigma^2_{\epsilon_{ij}} = \exp(\tau_0 + \omega_i + \boldsymbol{w}_{ij}^T \boldsymbol{\tau'}), \tag{3.2}$$

where $\boldsymbol{w}_{ij}$ is an $r \times 1$ vector of time-varying covariates influencing the scale of y_{ij}, and ω_i is the ith subject's deviation from the average log WS variance of y_{ij}. τ_0 is the average log WS variance when every covariate in $\boldsymbol{w}_{ij}$ is 0, and $\boldsymbol{\tau'}$ is an $r \times 1$ vector of coefficients of $\boldsymbol{w}_{ij}$. Likewise, $\boldsymbol{\tau}$ in influence measures refers to a vector of which the first entry is τ_0, and the remaining entries come from $\boldsymbol{\tau'}$.

The random location and scale effects can correlate with each other, and they are assumed

to follow a bivariate normal distribution:

$$\begin{pmatrix} \nu_i \\ \omega_i \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_\nu^2 & \sigma_{\nu\omega} \\ \sigma_{\nu\omega} & \sigma_\omega^2 \end{pmatrix} \right). \tag{3.3}$$

Given our focus on influence analysis at the subject level, also known as level 2 in longitudinal models, it's important to pause here and clarify that level-2 effects in MELS models comprise fixed effects of time-invariant covariates (including the fixed intercepts) and all random effects. This distinction lays the foundation for our forthcoming exploration of influence analysis using MELS models. The following paragraphs present how the leave-one-out MELS models are formed, i.e., how each subject is separated from the random effects and subject-level fixed effects and given subject-specific fixed effects, for subsequent influence analyses.

As suggested by Langford and Lewis [19], in order to separate the subject under evaluation, denoted as i^* ($i^* \in 1, 2, \ldots, N$), from level 2 of the location model, we exclude subject i^* from the estimation of the random effects and level-2 fixed effects. In this case, the only level-2 fixed effect is β_0. Then, subject i^* is given a subject-specific fixed effect denoted as c_{i^*}. Namely, the location model becomes

$$y_{ij} = \mathbb{1}(i \neq i^*) \times (\beta_{0(-i^*)} + \nu_i) + \boldsymbol{x}_{ij}^T \boldsymbol{\beta'}_{(-i^*)} + \mathbb{1}(i = i^*) \times c_{i^*} + \epsilon_{ij}, \tag{3.4}$$

where $\mathbb{1}(i = i^*)$ equals 1 for subject i^* and 0 for all other subjects, and $\mathbb{1}(i \neq i^*)$ equals 0 for subject i^* and 1 for all other subjects.

Note that for simplicity in illustrating leave-one-out MELS models, we assume that every covariate in $\boldsymbol{x}_{ij}$ and $\boldsymbol{w}_{ij}$ is time-varying. When some time-invariant covariates are also present, subject i^* should be separated from the estimation of their associated coefficients as well.

To also separate subject i^* from level 2 of the scale model described in Equation 3.2, the

leave-one-out scale model is as follows:

$$\sigma^2_{\epsilon_{ij}} = \exp(\mathbb{1}(i \neq i^*) \times (\tau_{0(-i^*)} + \omega_i) + \boldsymbol{w}^T_{ij}\boldsymbol{\tau'}_{(-i^*)} + \mathbb{1}(i = i^*) \times d_{i^*}). \tag{3.5}$$

Here, d_{i^*} is the i^{*th} subject's subject-specific fixed scale effect, and the variances of random location effects and random scale effects become $\sigma^2_{\nu(-i^*)}$ and $\sigma^2_{\omega(-i^*)}$ while their covariance is $\sigma_{\nu\omega(-i^*)}$ without subject i^*. Note that the subscript $(-i^*)$ here as well as in Equations 3.4 and 3.5 does not mean that subject i^* is entirely removed from modeling but instead only removed from the estimation of level-2 effects.

Overall, the algorithm to estimate a leave-one-out MELS model includes the following steps: (1) create an indicator variable that equals 1 for subject i^* and 0 otherwise; (2) use the opposite of the indicator variable in (1) as the covariate for the random intercepts; (3) following step (2), subject i^* will not be included in the random effect estimation, and the fixed location effect for the indicator variable would be c_{i^*} while the fixed scale effect is d_{i^*}. Our programming specifics and pertinent material will be explained in Section 3.3.

A separate leave-one-out model is estimated for every subject and compared with the model in which all subjects are treated the same. Besides allowing the detection of highly influential subjects, this leave-one-out structure can be viewed as a way to keep all subjects in the model despite influential subject(s). Using all available data to estimate the level-1 fixed effects has the benefit of increasing statistical power.

3.2.2 Influence Analysis

Influence on Model Fit

Given that the model described in Equations 3.1 and 3.2 are nested in the leave-one-out model described in Equations 3.4 and 3.5, likelihood-ratio test was used to examine subject i^*'s influence on the model fit. The test statistic, difference in deviance, denoted as LR_{i^*} for

subject i^*, can be calculated as follows:

$$LR_{i^*} = 2\ln\left(\frac{\mathcal{L}_{(-i^*)}}{\mathcal{L}}\right), \tag{3.6}$$

where $\mathcal{L}_{(-i^*)}$ represents the likelihood of the MELS model in which subject i^* is separated, and $\mathcal{L}$ represents the likelihood of the naive MELS model. Since N tests are conducted simultaneously, the false discovery rate (FDR) procedure [4] is applied to adjust for multiple comparisons. This specific multiple testing correction method is chosen because of its ability to control for both Type I and Type II errors [58].

Influence on Parameter Estimates

Cook's distance [9] is a well-known measure for the influence of data on the point estimates of a group of parameter estimates. Based on the structure of MELS models, we calculate the Cook's distances for three groups of parameters, namely, fixed location effects ($\boldsymbol{\beta}$), fixed scale effects ($\boldsymbol{\tau}$), and variances and covariance of random effects, denoted as $\boldsymbol{\eta}$ ($\boldsymbol{\eta} = [\sigma^2_\nu, \sigma^2_\omega, \sigma_{\nu\omega}]$). Following the formula of Cook's distance for multilevel models described by Snijder and Bosker [50], Cook's distance can be calculated as

$$C^{\boldsymbol{\gamma}}_{i^*} = \frac{1}{r_{\boldsymbol{\gamma}}}(\hat{\boldsymbol{\gamma}} - \hat{\boldsymbol{\gamma}}_{(-i^*)})^T \hat{\boldsymbol{\Sigma}}^{-1}_{\hat{\boldsymbol{\gamma}}(-i^*)}(\hat{\boldsymbol{\gamma}} - \hat{\boldsymbol{\gamma}}_{(-i^*)}), \tag{3.7}$$

where $\boldsymbol{\gamma}$ can be $\boldsymbol{\beta}$, $\boldsymbol{\tau}$, or $\boldsymbol{\eta}$. Here, $r_{\boldsymbol{\gamma}}$ is the number of parameters being examined, and $\boldsymbol{\Sigma}_{\hat{\boldsymbol{\gamma}}(-i^*)}$ is the variance-covariance matrix of $\hat{\boldsymbol{\gamma}}$ after subject i^* is separated from the random effect estimation. A large Cook's distance indicates a heavy influence on a specific group of parameter estimates.

Once a subject is determined to be influential on a particular group of parameter estimates, it is often of interest to investigate this subject's influence on each specific parameter estimate within this group. In this case, DFBETAS can be used. DFBETAS is the difference

between the estimate of a parameter obtained when all subjects are kept in the random effect and level-2 effect estimation and the parameter estimate when a subject is separated, divided by the standard error of the parameter estimate [12]. Its mathematical representation is as follows:

$$\text{DFBETAS}_{i*}^{\theta} = \frac{\hat{\theta} - \hat{\theta}_{(-i^*)}}{SE(\hat{\theta}_{(-i^*)})}, \tag{3.8}$$

where θ can be any single parameter in the model. DFBETAS can have both positive and negative values. A highly positive value indicates that the inclusion of the influential subject creates a positive bias for $\hat{\theta}$, and vice versa.

Influence on the Precision of Parameter Estimates

Influential subjects can also affect the variances of parameter estimates besides their point estimates. COVTRACE proposed by Christensen et al. [7] and COVRATIO described by Belsley et al. [3] are often used to measure such changes.

$\text{COVTRACE}_{i*}^{\boldsymbol{\gamma}}$ is calculated as the absolute value of the difference between the trace of the inverse ratio of variance-covariance matrix estimates with and without subject i^* in level-2 model and the number of parameters under investigation:

$$\text{COVTRACE}_{i*}^{\boldsymbol{\gamma}} = |\operatorname{Tr}(\hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\gamma}}}^{-1}\hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\gamma}}(-i^*)}) - r_{\boldsymbol{\gamma}}|. \tag{3.9}$$

The larger the COVTRACE value, the greater the influence of subject i^* on the precision of $\hat{\boldsymbol{\gamma}}$.

Meanwhile, $\text{COVRATIO}_{i*}^{\boldsymbol{\gamma}}$ is the inverse ratio of the determinant of the estimated variance-covariance matrix of $\boldsymbol{\gamma}$ with and without subject i^* in level-2 model:

$$\text{COVRATIO}_{i*}^{\boldsymbol{\gamma}} = \frac{\det(\hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\gamma}}(-i^*)})}{\det(\hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\gamma}}})}. \tag{3.10}$$

Again, $\boldsymbol{\gamma}$ is one of $\boldsymbol{\beta}$, $\boldsymbol{\tau}$, and $\boldsymbol{\eta}$. The determinant of a variance-covariance matrix is also known as the generalized variance, a measure of multidimensional scatter [60]. The examples in the next two sections will focus on COVRATIO because it provides information on both the magnitude and the direction of the influence. A value of COVRATIO above 1 indicates a loss in precision when separating subject i^*, and on the contrary, a value below 1 indicates a gain in precision.

All the procedures of influence analysis described in Section 3.2.2 are summarized in Table 3.1.

3.3 Results

3.3.1 Application to Health Behavior Study Example

Data collected by Flueckiger et al. [11] for a study on the association between health behaviors and learning goal achievements are used as an example to illustrate the proposed method.

During the 32-day span of the study, 72 students answered questions about their sleep quality (SQ), physical activity, positive and negative affect, learning goal achievement (LGA), and examination grades. The following analysis focuses on complete observations with values for all variables, so our analysis includes 1788 observations from 62 subjects. One of the major findings from the original study is that better SQ is positively associated with LGA. Hence, we chose SQ as the covariate for both the mean and the scale models. Namely, the location model for the example can be expressed as

$$\mathrm{LGA}_{ij} = \beta_0 + \nu_i + \beta_{\mathrm{SQ}}\mathrm{SQ}_{ij} + \epsilon_{ij}, \tag{3.11}$$

Table 3.1: Influence analysis framework for MELS models

Influence measure	**Influence sub-category**	**Formula**
Influence on model fit		
Difference in deviance		$LR_{i^*} = 2\ln\left(\frac{\mathcal{L}_{(-i^*)}}{\hat{\mathcal{L}}}\right)$
Influence on point estimates of a group of parameters		
Cook's distance	Fixed location effect estimates	$C_{i^*}^{\boldsymbol{\beta}} = \frac{1}{r_{\boldsymbol{\beta}}}(\hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}_{(-i^*)})^T \hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\beta}}(-i^*)}^{-1}(\hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}_{(-i^*)})$
	Fixed scale effect estimates	$C_{i^*}^{\boldsymbol{\tau}} = \frac{1}{r_{\boldsymbol{\tau}}}(\hat{\boldsymbol{\tau}} - \hat{\boldsymbol{\tau}}_{(-i^*)})^T \hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\tau}}(-i^*)}^{-1}(\hat{\boldsymbol{\tau}} - \hat{\boldsymbol{\tau}}_{(-i^*)})$
	Variances and covariances of random effects	$C_{i^*}^{\boldsymbol{\eta}} = \frac{1}{r_{\boldsymbol{\eta}}}(\hat{\boldsymbol{\eta}} - \hat{\boldsymbol{\eta}}_{(-i^*)})^T \hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\eta}}(-i^*)}^{-1}(\hat{\boldsymbol{\eta}} - \hat{\boldsymbol{\eta}}_{(-i^*)})$
Influence on point estimate of a single parameter		
DFBETAS		$\text{DFBETAS}_{i^*}^{\theta} = \frac{\hat{\theta} - \hat{\theta}_{(-i^*)}}{SE(\hat{\theta}_{(-i^*)})}$
Influence on variances and covariances of a group of parameters		
COVTRACE	Fixed location effect estimates	$\text{COVTRACE}_{i^*}^{\boldsymbol{\beta}} = \lvert \text{Tr}(\hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\beta}}}^{-1}\hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\beta}}(-i^*)}) - r_{\boldsymbol{\beta}} \rvert$
	Fixed scale effect estimates	$\text{COVTRACE}_{i^*}^{\boldsymbol{\tau}} = \lvert \text{Tr}(\hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\tau}}}^{-1}\hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\tau}}(-i^*)}) - r_{\boldsymbol{\tau}} \rvert$
	Variances and covariances of random effects	$\text{COVTRACE}_{i^*}^{\boldsymbol{\eta}} = \lvert \text{Tr}(\hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\eta}}}^{-1}\hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\eta}}(-i^*)}) - r_{\boldsymbol{\eta}} \rvert$
COVRATIO	Fixed location effect estimates	$\text{COVRATIO}_{i^*}^{\boldsymbol{\beta}} = \frac{\det(\hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\beta}}(-i^*)})}{\det(\hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\beta}}})}$
	Fixed scale effect estimates	$\text{COVRATIO}_{i^*}^{\boldsymbol{\tau}} = \frac{\det(\hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\tau}}(-i^*)})}{\det(\hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\tau}}})}$
	Variances and covariances of random effects	$\text{COVRATIO}_{i^*}^{\boldsymbol{\eta}} = \frac{\det(\hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\eta}}(-i^*)})}{\det(\hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\eta}}})}$

and the scale model, in which the variance of ϵ_{ij} is being modeled, is

$$\sigma^2_{\epsilon_{ij}} = \exp(\tau_0 + \omega_i + \tau_{\text{SQ}}\text{SQ}_{ij}), \tag{3.12}$$

where SQ was measured on a 4-point Likert scale in which 1 means very bad, and 4 means very good. LGA was measured on a 5-point Likert scale in which 0 represents having not achieved the goals at all and 4 represents having achieved the goals completely. We treated LGA and SQ as continuous, consistent with the approach taken in the original study. The approximate normal distribution of LGA has been validated through our exploratory analysis. Nevertheless, it is important for the readers to exercise caution when generalizing the results beyond the original LGA range.

The Cook's distances and COVRATIOs of all subjects are visualized in Figure 3.1. After applying the FDR procedure, subjects 7, 12, 49, and 69 are determined to have a statistically significant influence on the model fit based on likelihood-ratio tests. Influence analysis results of these four subjects are summarized in Table 3.2. According to the Cook's distances, we can see that subjects 7 and 12 have high influence on the fixed scale effects, subject 69 has high influence on the fixed location effects, while subject 49 has high influence on both. In particular, subject 7 has a large DFBETAS for τ_{SQ}, subject 49 has the largest DFBETAS for β_0 and the smallest DFBETAS for τ_0, and subject 69 has the smallest DFBETAS for β_0 and the 2nd largest DFBETAS for β_{SQ}. All four subjects have high influence on the point estimates of the variances and covariance of the random effects. Specifically, separation of subjects 7, 12, and 49 shrinks $\hat{\sigma}^2_\omega$; separation of subjects 49 and 69 shrinks $\hat{\sigma}^2_\nu$; separation of subject 7 and 69 decreases $\hat{\sigma}_{\nu\omega}$ while separation of subject 12 and 49 increases $\hat{\sigma}_{\nu\omega}$. Moreover, excluding each of these subjects from the random effect estimation also shrinks the corresponding generalized variances of the groups of parameters that they have influence on the point estimates, suggesting that these four subjects have caused a loss in model precision in the original model.

We visualize the SQ and LGA of these four subjects in Figure 3.2 to see how the data correspond with the influence analysis findings. While all except two of subject 7's self-reported SQ ratings have a value of 4, the LGA of this subject was highly variable when

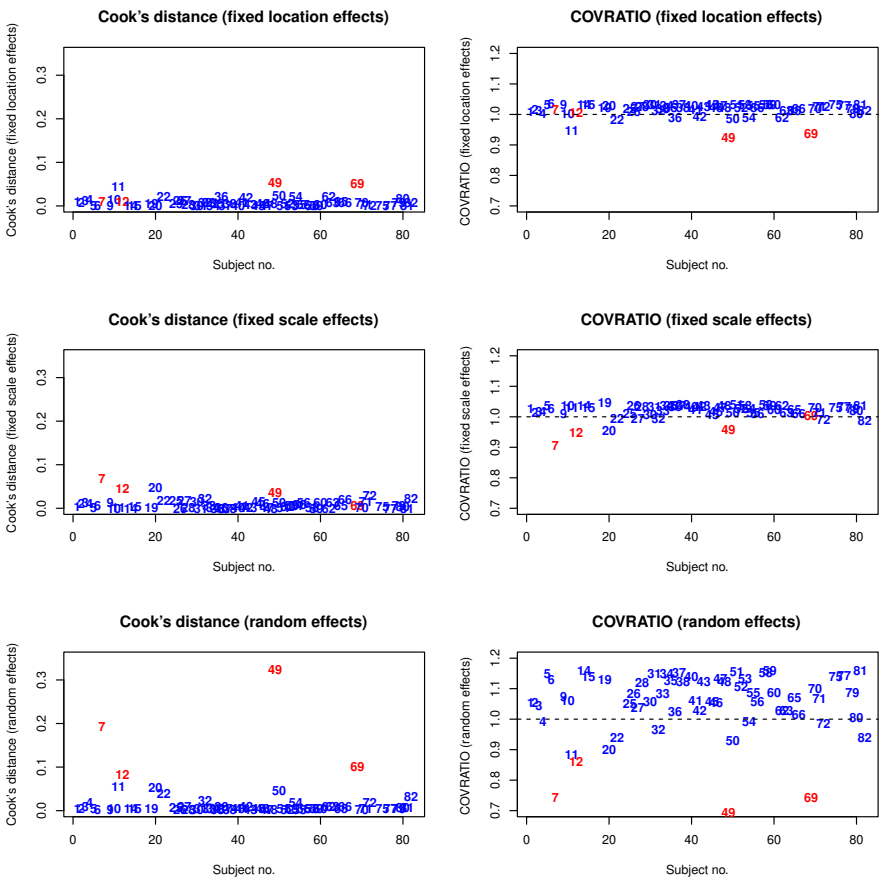


Figure 3.1: Cook's distances and COVRATIOs for the health behavior data. Points in red are the four subjects influential on the fit of the MELS model of health behavior data.

SQ equaled 4. This is consistent with the influence analysis result that giving subject 7 subject-specific fixed effects shrinks $\hat{\tau}_{\text{SQ}}$ and $\hat{\sigma}^2_\omega$. While subject 49 had consistently high LGA, subject 69 had all low LGA values at 0 and 1. Such patterns in the data correspond to subject 49's large DFBETAS_{β_0} and subject 69's small DFBETAS_{β_0}. Given subjects 12, 49, and 69 have such great influence on different components of the model and the fact that they have the lowest variability in LGA across all subjects, it could be worthwhile to investigate whether the similar answers for LGA are a result of careless responses or simply outstanding consistency in actual LGA.

Sensitivity analyses were conducted by excluding the scale model, i.e., using standard

Table 3.2: Influence analysis results for health behavior data

Subject	**Influence measure**	**Results**
7	Cook's distance	- Largest $C^{\boldsymbol{\tau}}$ (0.069) - 2nd largest $C^{\boldsymbol{\eta}}$ (0.194)
	DFBETAS	- 3rd largest DFBETAS$^{\beta_{SQ}}$ (0.055) - Largest DFBETAS$^{\tau_{SQ}}$ (0.163) - Largest DFBETAS$^{\sigma^2_\omega}$ (0.604) - Largest DFBETAS$^{\sigma_{\nu\omega}}$ (0.325)
	COVRATIO	- Smallest COVRATIO$^{\boldsymbol{\tau}}$ (0.905) - 3rd smallest COVRATIO$^{\boldsymbol{\eta}}$ (0.743)
12	Cook's distance	- 3rd largest $C^{\boldsymbol{\tau}}$ (0.045) - 4th largest $C^{\boldsymbol{\eta}}$ (0.084)
	DFBETAS	- 4th largest DFBETAS$^{\beta_0}$ (0.096) - 2nd largest DFBETAS$^{\sigma^2_\omega}$ (0.392) - 2nd smallest DFBETAS$^{\sigma_{\nu\omega}}$ (-0.343)
	COVRATIO	- 2nd smallest COVRATIO$^{\boldsymbol{\tau}}$ (0.948) - 4th smallest COVRATIO$^{\boldsymbol{\eta}}$ (0.862)
49	Cook's distance	- Largest $C^{\boldsymbol{\beta}}$ (0.054) - 4th largest $C^{\boldsymbol{\tau}}$ (0.037) - Largest $C^{\boldsymbol{\eta}}$ (0.323)
	DFBETAS	- Largest DFBETAS$^{\beta_0}$ (0.194) - Smallest DFBETAS$^{\tau_0}$ (-0.140) - Largest DFBETAS$^{\sigma^2_\nu}$ (0.553) - 4th largest DFBETAS$^{\sigma^2_\omega}$ (0.325) - Smallest DFBETAS$^{\sigma_{\nu\omega}}$ (-0.737)
	COVRATIO	- Smallest COVRATIO$^{\boldsymbol{\beta}}$ (0.925) - 4th smallest COVRATIO$^{\boldsymbol{\tau}}$ (0.959) - Smallest COVRATIO$^{\boldsymbol{\eta}}$ (0.694)
69	Cook's distance	- 2nd largest $C^{\boldsymbol{\beta}}$ (0.052) - 3rd largest $C^{\boldsymbol{\eta}}$ (0.102)
	DFBETAS	- Smallest DFBETAS$^{\beta_0}$ (-0.281) - 2nd largest DFBETAS$^{\beta_{SQ}}$ (0.121) - 2nd largest DFBETAS$^{\sigma^2_\nu}$ (0.422) - 3rd largest DFBETAS$^{\sigma_{\nu\omega}}$ (0.269)
	COVRATIO	- 2nd smallest COVRATIO$^{\boldsymbol{\beta}}$ (0.939) - 2nd smallest COVRATIO$^{\boldsymbol{\eta}}$ (0.741)

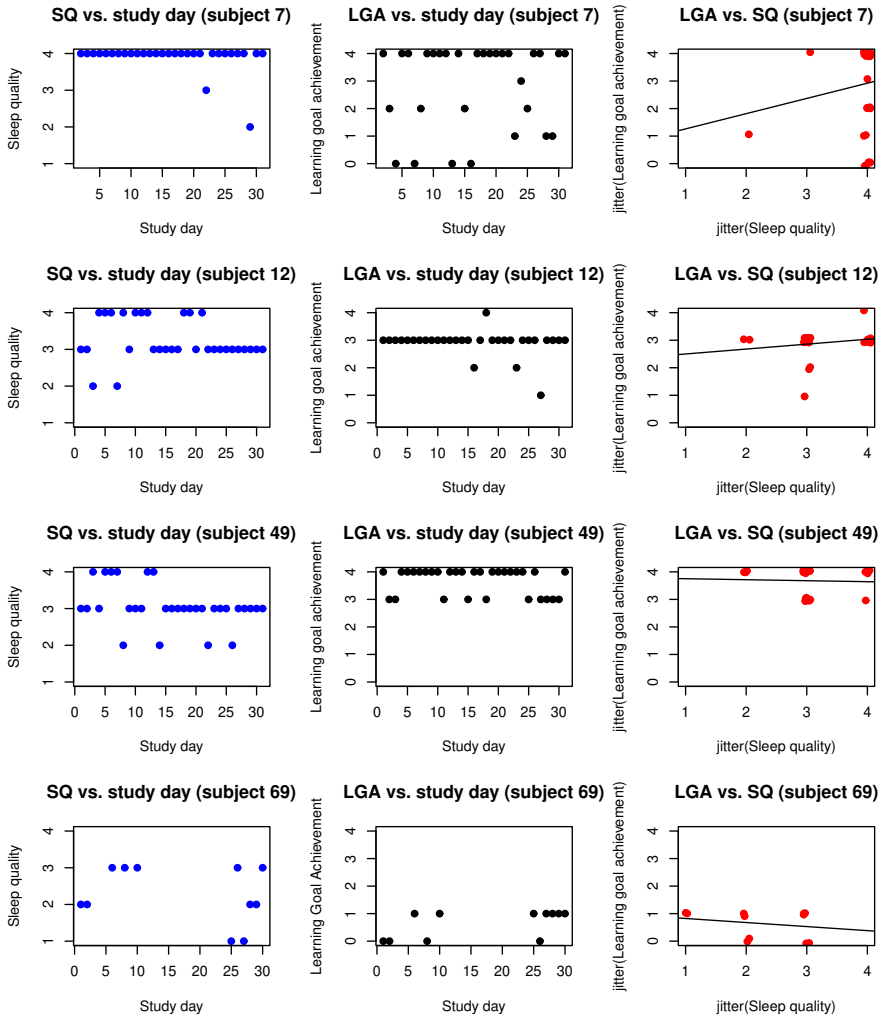


Figure 3.2: Data from subjects influential on the fit of the MELS model of health behavior data.

MRMs. The sensitivity analysis results still acknowledge the strong influence of subjects 49 and 69 on the location model but do not reveal significant influence from subjects 7 and 12. These findings are consistent with results in Section 3.3.2, that is, subjects with influential data in terms of scale can often be neglected if influence analyses are conducted via MRMs only.

PROC NLMIXED in SAS OnDemand for Academics (SAS Institute Inc.) was used to estimate both the original and the leave-one-out MELS models. All the influence analyses using the results from SAS were carried out in R version 4.2.0 (R Core Team). The SAS

codes used to estimate MELS models of the health behavior study dataset are provided in the supplementary files. The first part of the codes estimate MELS models that treat all subjects equally, and the second part of the codes were used to estimate leave-one-out MELS models.

3.3.2 Simulation Study

To further illustrate the advantages of conducting influence analyses using MELS models compared to analyses using MRMs with no scale components, we generated simulated examples in two different scenarios, with 500 datasets created for each scenario. Every dataset contains 50 subjects that follow the structure described in Equations 3.1 and 3.2 before designating some subject(s) to be influential. Both scenarios' responses resemble daily physical activity time in minutes, so negative values were removed. Examples of one regular subject and one artificially influential subject simulated in each scenario are illustrated in Figure 3.3.

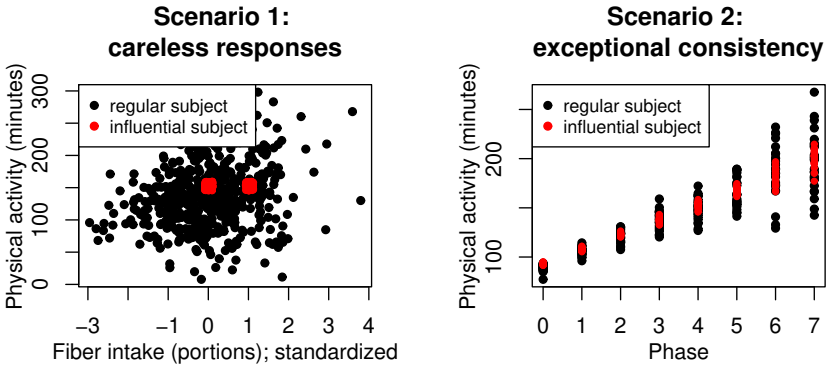


Figure 3.3: Simulated data examples. Points in black represent data from one regular subject, and points in red represent data from one influential subject.

The proposed method requires fitting a MELS model, for which estimation can be time-consuming, for every subject in the dataset. Hence, all the MELS models involved in the simulation study were estimated using a fast estimation algorithm for MELS models, FastRegLS, developed by Gill and Hedeker [13]. One or more of the models failed to converge

properly for 1 to 18 datasets (mean = 7.67) in the six simulation studies, and these datasets are not included in the result summary. One drawback of the FastRegLS algorithm is that it doesn't offer log-likelihoods. Hence, the simulation studies primarily concentrate on point estimates, along with the variances and covariances of parameter estimates.

Simulation Scenario 1

In the first scenario, the time-varying covariate, portions of fiber intake, is continuous and was simulated based on a $\mathcal{N}(3,1)$ distribution, after which absolute values were taken to ensure non-negativity. The values of the parameters used to generate the simulated datasets are as follows: $\beta = [100, 10]$; $\tau = [7, 0.5]$; $\eta = [20.09, 1, 0.45]$. In other words, the model used to generate data except for the influential one(s) is as follows:

$$\begin{aligned} y_{ij} &= 100 + \nu_i + 10x_{1ij} + \epsilon_{ij}, \\ \sigma^2_{\epsilon_{ij}} &= \exp(7 + \omega_i + 0.5x_{1ij}), \\ \begin{pmatrix} \nu_i \\ \omega_i \end{pmatrix} &\sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 20.09 & 0.45 \\ 0.45 & 1 \end{pmatrix} \right). \end{aligned} \tag{3.13}$$

There are 500 observations from each subject, and the covariate is standardized for influence analysis to facilitate model estimation. This scenario focuses on the case of careless responses, so we first designated subject 1 to have all responses at 150 and 155 and all covariates at 3 and 4 before standardization. The percentage of simulations in which each influence measure detects subject 1 to be the most influential are summarized in the first column of Table 3.3. In summary, subject 1 always has the smallest COVRATIO$^{\boldsymbol{\tau}}$ and the largest DFBETAS$^{\sigma^2_\omega}$. It also almost always has the largest DFBETAS$^{\tau_0}$. The second column in Table 3.3 shows the percentage of detection when we specify the first three subjects to have such careless responses, and the three subjects still almost always have the smallest

COVRATIO$^{\boldsymbol{\tau}}$'s and the largest DFBETAS$^{\sigma^2_\omega}$'s. Note that only all three influential subjects among the top three are counted as one detection.

Simulation Scenario 2

For scenario 2, the covariate values were simulated to represent discrete time periods with values that range from 0 to 7, and each subject had 25 observations from each time period before removing negative values. The simulated datasets were generated according to the following structure:

$$\begin{aligned} y_{ij} &= 100 + \nu_i + 15x_{1ij} + \epsilon_{ij}, \\ \sigma^2_{\epsilon_{ij}} &= \exp(3 + \omega_i + 0.6x_{1ij}), \\ \begin{pmatrix} \nu_i \\ \omega_i \end{pmatrix} &\sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 20.09 & 0.22 \\ 0.22 & 0.25 \end{pmatrix} \right). \end{aligned} \tag{3.14}$$

The influential subject(s) are subject(s) with exceptional consistency throughout the behavioral study and are simulated with an intercept of 0 in the scale model. In the simulation study with the first subject like this, the subject almost always has the largest $C^{\boldsymbol{\tau}}$, the smallest COVRATIO$^{\boldsymbol{\tau}}$, the largest DFBETAS$^{\tau_0}$, and the largest DFBETAS$^{\sigma^2_\omega}$. The same results remain in terms of COVRATIO$^{\boldsymbol{\tau}}$ and DFBETAS$^{\sigma^2_\omega}$ when the number of artificial influential subjects increases to three.

As shown in the third column of Table 3.3, using standard MRMs, none of the influence measures recognize the influences of the designated influential subjects by more than half of the time in either scenario. Therefore, in order to successfully identify an influential subject in these scenarios, it is essential to take the scale model into consideration.

Table 3.3: Percentage of detection using MELS models and MRMs and different influence measures. For scenarios with the first subject as the influential subject, "large(%)" represents the percentage of simulations in which this subject has the largest value of the respective influence measure among all subjects, and "small(%)" represents the percentage of simulations in which this subject has the smallest value of the respective influence measure among all subjects. For scenarios with the first three subjects as the influential subjects, "large(%)" represents the percentage of simulations in which all three of these subjects have the largest values of the respective influence measure among all subjects, and "small(%)" represents the percentage of simulations in which all these three subjects have the smallest values of the respective influence measure among all subjects. Results in the column named "MRMs" were obtained through analyses using MRMs with no scale components. The simulated datasets used in the MELS model and MRM analyses are the same in each scenario.

	MELS models		**MRMs**
	Single	Multiple (3)	Single
Scenario 1			
$C^{\boldsymbol{\tau}}$ [large(%)]	75.9	16.7	-
COVRATIO$^{\boldsymbol{\tau}}$[large(%)/small(%)]	0/100	0/99.6	-
DFBETAS$^{\tau_0}$ [large(%)/small(%)]	99.4/0	25.4/0.4	-
DFBETAS$^{\tau_1}$ [large(%)/small(%)]	5.4/34.6	1.4/15.7	-
DFBETAS$^{\sigma^2_\omega}$ [large(%)/small(%)]	100/0	99.6/0	-
C$^{\boldsymbol{\beta}}$ [large(%)]	3.5	5.8	0
COVRATIO$^{\boldsymbol{\beta}}$[large(%)/small(%)]	15.1/0.8	2.0/0.8	41.6/0
DFBETAS$^{\beta_0}$ [large(%)/small(%)]	0.8/1.7	0.6/0.6	0/0
DFBETAS$^{\beta_1}$ [large(%)/small(%)]	21.0/44.4	3.6/22.7	6.8/13.9
DFBETAS$^{\sigma^2_\nu}$ [large(%)/small(%)]	0.6/0	1.0/0	0/20.1
Scenario 2			
$C^{\boldsymbol{\tau}}$ [large(%)]	99.8	62.3	-
COVRATIO$^{\boldsymbol{\tau}}$[large(%)/small(%)]	0/100	0/99.8	-
DFBETAS$^{\tau_0}$ [large(%)/small(%)]	99.8/0	69.9/0	-
DFBETAS$^{\tau_1}$ [large(%)/small(%)]	11.3/16.0	0.4/0.8	-
DFBETAS$^{\sigma^2_\omega}$ [large(%)/small(%)]	100/0	100/0	-
C$^{\boldsymbol{\beta}}$ [large(%)]	21.9	0	1.2
COVRATIO$^{\boldsymbol{\beta}}$ [large(%)/small(%)]	0/2.6	0/0	2.0/1.2
DFBETAS$^{\beta_0}$ [large(%)/small(%)]	1.4/2.0	0/0	1.4/1.6
DFBETAS$^{\beta_1}$ [large(%)/small(%)]	38.9/44.9	0/0.8	1.4/0.6
DFBETAS$^{\sigma^2_\nu}$ [large(%)/small(%)]	1.4/6.3	0/0	1.2/2.0

3.4 Discussion

In this study, we have discussed procedures to detect subject(s) influential in model fit and parameter estimates in MELS models. This approach allows researchers to identify subjects influential on the scale structure of the outcome being modeled in addition to the location structure. We hope that our method is able to help researchers, especially researchers interested in studying intra-individual variability, better identify interesting or troublesome subjects that they want to study further in an EMA study or a large-scale longitudinal clinical trial. After determining which subjects are considered influential, researchers can keep these subjects in the analysis using our leave-one-out MELS model described in Section 3.2.2, and a summary of other common ways of dealing with influential data can be found in a recent work by Aguinis et al. [1]. The proposed method can also benefit analyses not carried out in MELS models by providing researchers with a better understanding of both the location and the scale structures of their data during exploratory analyses.

Our study has focused on subject-level influence analyses, so one possible extension is the detection of influential observations in MELS model. Given that there could be influential observation(s) within a non-influential subject, estimating a separate model for each observation might be required. Because of the enormous number of observations in intensive longitudinal data, such methods can be extremely computationally intensive. This article has also focused on maximum likelihood estimates of 2-level MELS models for normally distributed continuous outcomes. Further development can extend the proposed framework to accommodate models with 3 or more levels [21, 22], models on outcomes with more complicated structures [5], and Bayesian estimation approaches [36]. Future work will also extend to ordinal MELS models [16].

To plan for further data analysis based on the influence analysis results, we recommend readers go through all the results from our framework and use their domain knowledge to decide whether specific subject(s) are considered influential and need further analysis.

However, we understand that one might want cutoffs to guide their judgment. Rule-of-thumb cut-off values are $\frac{4}{N}$ for Cook's distances, $1 \pm 3(\frac{r_\gamma}{N})$ for COVRATIO, and $\frac{2}{\sqrt{N}}$ for DFBETAS [3]. Also, a possible future direction of this work is to improve existing cut-off values to be more suitable for MELS models.

The simulation studies have showcased that at least some of the influence measures are able to capture all influential subjects in the case that multiple of them coexist. Nevertheless, not all measures perform the same, which might be attributed to the masking effect [2]. Therefore, we again recommend readers carefully examine all influence measures mentioned in the framework, and a future step of this study will be improving individual influence measures to overcome any possible masking effect.

CHAPTER 4

A LATENT CLASS LOCATION-SCALE REGRESSION MODEL WITH AN APPLICATION TO CALORIE INTAKE DATA

4.1 Attributions

This study was conducted in collaboration with Dr. Juned Siddique and Dr. Bonnie Spring from Northwestern University. Dr. Siddique provided critical feedback on statistical analysis involved in this study. Both Dr. Siddique and Dr. Spring provided critical feedback on drafts of the manuscript and made revisions.

4.2 Introduction

With richer information collected by longitudinal studies, attention has been drawn to distinguishing homogeneous subgroups within the study population. Earlier efforts primarily focused on introducing latent classes (LCs) to the mean, also called location, structure of longitudinal data. Verbeke and Lesaffre [57], Muthén and Shedden [29], and Xu [61] developed mixed-effects models in which the random effects follow a mixture distribution. For longitudinal data, these are often called growth mixture models (GMMs). Alternatively, Nagin and Land [30] introduced latent class growth analysis (LCGA), which can be considered a special case of GMM, that assumes the growth trajectories within each latent subgroup are homogeneous, and all the heterogeneity between subjects is captured by their LC memberships. Computer software, including Mplus [28] and PROC TRAJ in SAS [18] have been developed to help researchers apply GMMs and LCGA.

Building on this, Elliott [10] suggested that the idea of LCs can be extended to within-subject (WS) variability, also called intraindividual variability or scale, of a longitudinal outcome. The approach described by Elliot models the location using a cubic spline re-

gression with distinct residual variances for different LCs. Jiang et al. [17] integrated LC modeling for both location and scale. They used a joint modeling approach consisting of two components. First, they model the longitudinal outcome, incorporating LCs in both the location and scale models. Second, they utilize the results from the LC model to inform the modeling of the binary primary outcome. However, both papers described above focus on the case of constant WS variability over time and assume that no covariates, in particular time, influence the WS variability.

In longitudinal behavioral interventions, researchers may find the time trajectory of WS variability, in addition to the mean, particularly intriguing. This heightened interest arises from the anticipation of more consistent behavior over time due to an intervention or, conversely, greater variability. Mixed-effects location scale (MELS) models have been used for modeling such trajectories while accounting for between-subject (BS) heterogeneity in intraindividual variability via random effects [15]. For example, a recent study by Sandhu and Leckie [44] examined subjects' trajectory of fluctuation in pain intensity after orthodontic separator placement. Another study on adolescent smokers found that WS variability in the mood of those adolescent smokers decreased significantly over time [35]. Instead of assuming continuous variation among subjects like in the MELS models, subgrouping can serve as a preliminary step for investigating factors influencing temporal variability in subjects. According to Weller et al. [59], LC analysis helps identify subgroups that may benefit from interventions due to their common attributes.

Therefore, this study proposes a Bayesian location-scale regression with LCs in the trajectories of both mean and intraindividual variability. The model assumes that each subject belongs to one of several subgroups, each characterized by distinct mean (or location) trajectories, and to another set of subgroups, each with unique WS variability (or scale) trajectories. In Section 4.5, we illustrate how the proposed model can be applied to calorie intake data collected by a weight loss management study. This example demonstrates

how LCs in intraindividual variability trajectory contribute to precision care by revealing additional insights into subjects' consistency in calorie intake. In Section 4.6, the proposed method is validated through simulation studies based on the real-life application. We also demonstrate how the parameters of this model can be estimated using Stan [52]. The accessibility of Stan, with its user-friendly interface, enables researchers with limited experience in Bayesian coding to adopt the proposed method readily. A code template is provided in the supplementary files to facilitate implementation.

4.3 Motivating example

Our methods are motivated by the SMART trial, a longitudinal weight loss management study that utilized mobile health (mHealth) tools [51]. Four hundred participants were enrolled in the SMART study and randomized to be in either the mHealth-only treatment group or mHealth plus weekly coaching treatment group. Those classified as non-responders during the 2, 4, and 8 weeks follow-ups were re-randomized to add one more mHealth component or both an mHealth component and a traditional weight loss treatment component. Participants in this study were assigned calorie goals based on their baseline weight as part of the weight loss initiatives. Treatment was discontinued at the end of week 12. Since all subjects received some form of treatment, which also varied over time, and the primary focus of the proposed method in this study is on time trajectories, treatment will not be included in the analysis in Section 4.5. Instead, we will briefly explore the association between the original treatment assignment and the estimated trajectories in calorie intake, based on the latent class classifications of subjects derived from the model that focuses on time trajectories.

We used data on daily reported total calorie intake from the first 12 weeks of the study. We used week as our timing variable and treated it as linear, i.e. week = 0 for week 1 of the study, week = 1 for week 2, etc. Because the focus of this study is on WS variability

trajectory, only participants with at least two weeks of observations were kept in our models. A specific week of a participant was included if this participant had at least three observations that week. The final dataset contains 25784 observations from 379 participants, with an average of 68.03 observations per participant. The range of observations per participant in the final dataset is from 9 to 84.

On average, the participants consumed 1299 calories per day, with a standard deviation of 499 calories. The minimum reported value, median reported value, and maximum reported value are 1.8, 1248.6, and 9061.1, respectively. Investigating subgroups based on both mean and intraindividual variability trajectories can shed light on the diverse calorie intake journey of our participants. Furthermore, exploring potential associations between these trajectory subgroups and various participant characteristics and behaviors may offer valuable insights into underlying factors contributing to these individuals' differences in calorie intake.

4.4 Method

4.4.1 Location-scale regression with latent classes in mean and intraindividual variability trajectories

The model proposed in this study is a location-scale regression model that utilizes latent classes to quantify BS heterogeneity in both the location and the scale of the outcome. In the proposed model, trajectories within each LC are assumed homogeneous, as described in the LCGA framework.

The longitudinal outcome Y_{ij} ($i = 1, 2, \ldots, N$ subjects; $j = 1, 2, \ldots, n_i$ occasions), when assuming a linear trajectory in its mean, is modeled, conditional on the latent location class of subject i, as follows:

$$[Y_{ij}|L_i = \ell] = \beta_{0\ell} + \beta_{1\ell} t_{ij} + \epsilon_{ij}, \tag{4.1}$$

where $L_i = 1, \cdots, L$ follows a Multinomial($\pi_1^{location}, \cdots, \pi_L^{location}$) distribution and indi-

cates the location LC that subject i belongs to. Accordingly, $\beta_{0\ell}$ represents the average response for the lth location LC when $t_{ij} = 0$, and $\beta_{1\ell}$ represents the average linear response slope over time for the lth location LC.

The observation-level residuals, ϵ_{ij}, are assumed to be normally distributed with a mean of zero. Their variance, also known as intraindividual variability or WS variability, is fitted using a log-linear model as follows:

$$[\sigma^2_{\epsilon_{ij}}|S_i = s] = \exp(\tau_{0s} + \tau_{1s}t_{ij}), \tag{4.2}$$

where $S_i = 1, \cdots, S$ also follows a Multinomial$(\pi_1^{scale}, \cdots, \pi_S^{scale})$ distribution and indicates the scale class that subject i belongs to. τ_{0s} is the average log WS variance for the sth scale LC at baseline, and τ_{1s} represents the average trend in intraindividual variability on the log scale for the sth scale LC .

Denote observations from subject i as $\boldsymbol{Y}_i$. Assuming a priori independence between location LCs and scale LCs, we integrate over the likelihood of $\boldsymbol{Y}_i$ conditional on subject i's location LC and scale LC memberships to obtain their marginal likelihood:

$$L(\boldsymbol{Y}_i) = \Sigma_{\ell=1}^{L}\Sigma_{s=1}^{S}\pi_\ell^{location}\pi_s^{scale}L(\boldsymbol{Y}_i|L_i = \ell, S_i = s), \tag{4.3}$$

where the conditional likelihood is based on that $Y_{ij}|L_i = \ell, S_i = s$ follows the normal distribution with a mean of $\beta_{0\ell} + \beta_{1\ell}t_{ij}$ and a variance of $\exp(\tau_{0s} + \tau_{1s}\mathrm{t}_{ij})$. The overall marginal log-likelihood for the study population is obtained by summing the marginal log-likelihoods of all subjects together. Posterior LC membership probabilities for subject i can

be derived following Bayes' theorem.

$$\eta_{i,\ell}^{location} = P(L_i = \ell|\boldsymbol{Y}_i) = \frac{\pi_\ell^{location}\Sigma_{s=1}^{S}\pi_s^{scale}L(\boldsymbol{Y}_i|L_i = \ell, S_i = s)}{L(\boldsymbol{Y}_i)}; \tag{4.4}$$

$$\eta_{i,s}^{scale} = P(S_i = s|\boldsymbol{Y}_i) = \frac{\pi_s^{scale}\Sigma_{\ell=1}^{L}\pi_\ell^{location}L(\boldsymbol{Y}_i|L_i = \ell, S_i = s)}{L(\boldsymbol{Y}_i)}. \tag{4.5}$$

Classification of each subject is then determined based on the class with the highest posterior mean probability. This can be expressed as $\tilde{L}_i = \text{argmax}_\ell \hat{\eta}_{i,\ell}^{location}$, and $\tilde{S}_i = \text{argmax}_s \hat{\eta}_{i,s}^{scale}$, where $\hat{\eta}_{i,\ell}^{location}$ and $\hat{\eta}_{i,s}^{scale}$ are the posterior means of $\eta_{i,\ell}^{location}$ and $\eta_{i,s}^{scale}$ described in Equations 4.4 and 4.5.

4.4.2 Bayesian estimation with Stan

Models described in this study were estimated using a Hamiltonian Monte Carlo (HMC) algorithm implemented by the R interface to Stan [52]. The Stan programs program in the supplementary files is portable to Stan interfaces in multiple programming languages. We chose Bayesian estimation over maximum likelihood estimation to avoid heavy calculation of the first- and second-order partial derivatives [26]. Also, according to Lee and Song [20], Bayesian estimation is more suitable for smaller sample sizes than maximum likelihood estimation and hence can be applied to studies of a wider range of sizes.

Default flat priors were used for all parameters. To temper the potential problem of label switching [37], we adopted the suggestion from the user manual of Stan [52], that is, first to run one chain of each model and use the posterior means of parameters from this step as starting values for the model with four chains. The starting values for the first chain were derived from separate regressions for the mean and WS variability of the outcome. Subsequently, the four MCMC chains were run for 2000 iterations, with the first 500 as warm-up iterations. We determined model convergence through trace plots and R-hat convergence diagnostics [56].

4.4.3 Choice of numbers of latent classes

Leave-one-out cross-validation (LOO) is a highly recommended method for assessing Bayesian models. In contrast to the Akaike information criterion (AIC) and the deviance information criterion (DIC), LOO offers several advantages: it fully embraces Bayesian principles by utilizing the entire posterior distribution, remains unaffected by parameterization choices, and is applicable even in cases involving singular models. LOO is also favored over the widely applicable information criterion (WAIC) due to its robustness in scenarios with weak priors and/or influential observations and its capability to offer approximate standard errors [55].

In this study, the *loo* package in R [55] was applied to calculate LOO expected log pointwise predictive density (elpd) values based on the fitted Stan models. The package uses Pareto-smoothed importance sampling (PSIS) to compute elpd in an efficient, accurate, and stable manner. A higher $elpd_{loo}$ value is better while accounting for model complexity.

4.5 Application to weight loss management study example

We now revisit the calorie intake data described in Section 4.3 to demonstrate how the proposed method can uncover subgroups with different trajectories in their mean and WS variability. The outcome, total calorie intake, was log-transformed to reduce skewness and improve estimation performance. With sample size and interoperability in mind, we analyzed the dataset with the proposed model with up to three location LCs and three scale LCs. We have also observed that a larger number of classes leads to increased data sparsity in the classification results. Inspection of the trace plots showed reasonable convergence for all nine models inspected. The $elpd_{loo}$ values of all nine models are listed in Table 4.1, and the model with three location LCs and three scale LCs was chosen. In other words, the location part of the final model can be expressed as

$$[\log(\text{total\_calories}_{ij})|L_i = \ell] = \beta_{0\ell} + \beta_{1\ell}\text{week}_{ij} + \epsilon_{ij},\ \ell = 1, 2, 3. \tag{4.6}$$

And conditional on participant i's latent variability trajectory class S_i, the scale model for participant i is as follows:

$$[\sigma^2_{\epsilon_{ij}}|S_i = s] = \exp(\tau_{0s} + \tau_{1s}\text{week}_{ij}),\ s = 1, 2, 3. \tag{4.7}$$

The estimated parameters and their 95% credible intervals for the final model described above can be found in Table 4.2. Visual representations of the location and scale trajectories are presented in Figure 4.1. Note that the numbers are small for the slopes due to log-normal modeling; that is, these numbers represent multiplicative trends.

Table 4.1: $elpd_{loo}$ for location-scale regressions with different numbers of LCs.

Number of LCs in the scale model	Number of LCs in the location model		
	1	2	3
1	-21338.0	-20226.1	-19847.0
2	-19933.4	-18860.6	-18591.5
3	-19594.9	-18534.2	-18248.4

Participants in all three location classes exhibit decreases in calorie intake over time. Participants in the first LC tend to start with lower calorie intake on average ($\beta_0|L_i = 1 = 6.916$) and decrease more compared to the other two LCs. The slopes of mean calorie intake are approximately parallel for location LCs 2 and 3, while the expected log(total calories) at baseline is 7.15 for participants in the 2nd location LC and 7.38 for participants in the 3rd location LC. These findings imply that it might be easier to reduce calorie intake when it initially starts at a lower level, highlighting the potential significance of cultivating healthy eating habits from the outset. For all three scale LCs, calorie intake for each participant also tends to become more variable over time, with the slope of WS variability being more positive for the 2nd scale LC. Thus, for participants in scale LC 2, interventions can potentially

Table 4.2: Parameter estimates and credible intervals for the application to SMART weight loss management study.

Location part of the model				
Location LC	Parameter	Estimate (95% credible interval)	Parameter	Estimate (95% credible interval)
1	$\beta_0 \vert L_i = 1$	6.916 (6.891, 6.940)	$\beta_1 \vert L_i = 1$	-0.023 (-0.028, -0.019)
2	$\beta_0 \vert L_i = 2$	7.151 (7.132, 7.171)	$\beta_1 \vert L_i = 2$	-0.011 (-0.013, -0.009)
3	$\beta_0 \vert L_i = 3$	7.375 (7.347, 7.413)	$\beta_1 \vert L_i = 3$	-0.009 (-0.012, -0.005)
Scale part of the model				
Scale LC	Parameter	Estimate (95% credible interval)	Parameter	Estimate (95% credible interval)
1	$\tau_0 \vert S_i = 1$	-2.411 (-2.492, -2.326)	$\tau_1 \vert S_i = 1$	0.024 (0.013, 0.036)
2	$\tau_0 \vert S_i = 2$	-1.738 (-1.793, -1.684)	$\tau_1 \vert S_i = 1$	0.058 (0.047, 0.067)
3	$\tau_0 \vert S_i = 3$	-0.775 (-0.879, -0.681)	$\tau_1 \vert S_i = 3$	0.022 (0.008, 0.037)

prioritize enhancing the consistency in calorie intake. The average WS variability for scale LC 2 at baseline ($\exp(\tau_0|S_i = 2)$), which equals 0.176, is between scale LC 1 ($\exp(\tau_0|S_i = 1) = 0.090$) and scale LC 2 ($\exp(\tau_0|S_i = 3) = 0.461$).

The numbers of participants classified into each of the nine combinations of location LCs and scale LCs are listed in Table 4.3. The procedure for classification is described in the last paragraph of Section 4.4.1. The most prevalent combination is location LC 2 and scale LC 2, which is the class that starts off with medium mean and variability and increases a lot in calorie intake variability over time. A limited number of participants are

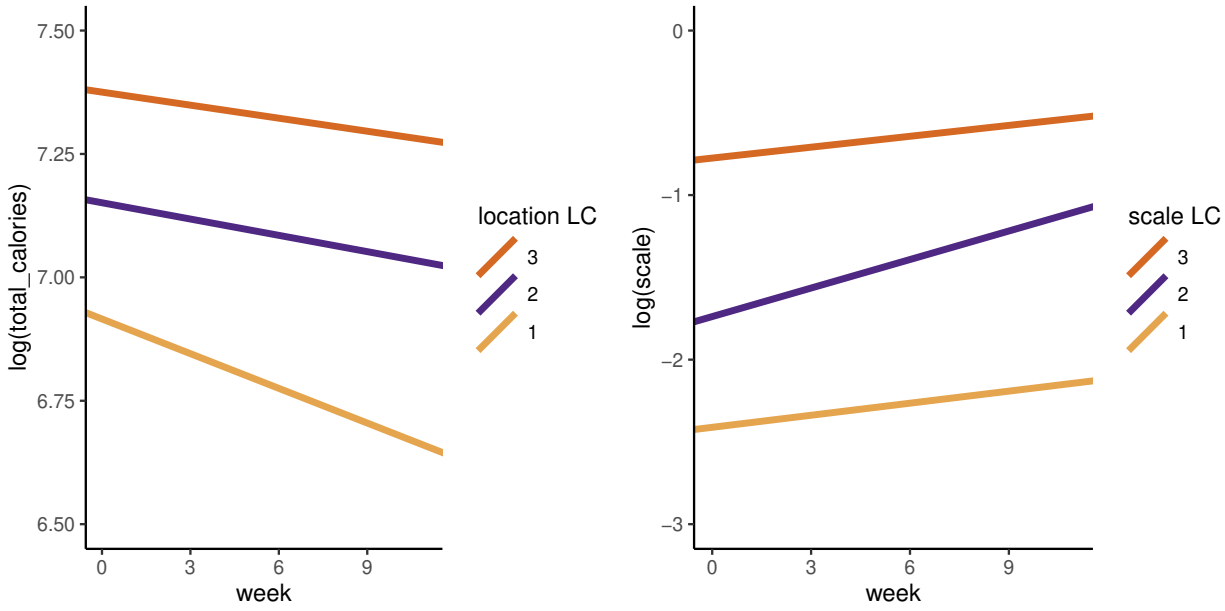


Figure 4.1: Estimated location and scale trajectories for each LC.

categorized within the groups denoted by location LC 1 and scale LC 1, as well as location LC 3 and scale LC 3. This observation suggests that within the study population, it is relatively infrequent for individuals to (1) initially exhibit low calorie intake with low WS variability while simultaneously experiencing a significant decrease in calorie intake over time; (2) commence with high calorie intake with high WS variability. These findings might be signs that high calorie intake is a relatively stable behavior and that calorie intake is resistant to change.

Figure 4.2 presents data from participants with high posterior probabilities of belonging to each LC combination. Consistent with the parameter estimates, participants 60, 243, and 58, whose posterior probabilities of belonging to location LC 1 are high, exhibited low mean calorie intake at baseline and a more negative trend over time in terms of mean calorie intake; on the contrary, participants 2, 101, and 241, with high posterior probabilities of belonging to location LC 3, started off with high mean calorie intake that did not decrease as much. From top to bottom of Figure 4.2, participants belonging to scale LCs 1, 2, and 3

Table 4.3: Distribution of subjects among LC combinations.

	Location LC			
Scale LC	1	2	3	**Total**
1	3 (0.8%)	51 (13.5%)	24 (6.3%)	**78 (20.6%)**
2	65 (17.2%)	97 (25.6%)	38 (10.0%)	**200 (52.8%)**
3	64 (16.9%)	31 (8.2%)	6 (1.6%)	**101 (26.6%)**
Total	**132 (34.8%)**	**179 (47.2%)**	**68 (17.9%)**	

displayed an upward pattern in terms of baseline WS variability in calorie intake. It is also obvious from the plots that the WS variability in calorie intake increased for participants 243, 234, and 101, who belong to scale LC 2.

Sociodemographic characteristics, baseline weight and height, as well as original treatment assignment of the participants, are summarized with regards to their LC classification in Table 4.4 and 4.5. For location, a participant's LC membership is significantly associated with their biological gender, race, weight at baseline, and height at baseline. There are more males and a lower percentage of Black or African American participants in location LC 3 compared to the other two location LCs. Both mean weight and mean height at baseline were the lowest for participants classified into location LC 1, higher for those in location LC 2, and highest for those in location LC 3. Scale LC membership is significantly associated with age at screening, weight at baseline, and highest level of education at baseline. Participants in scale LC 1 were on average older and lighter at baseline. Participants in scale LC 3 had lower education levels on average at baseline.

4.6 Simulation study

In this section, we use simulated data to evaluate the validity of our method and estimation procedure. Because of the computationally expensive nature of the model estimation process,

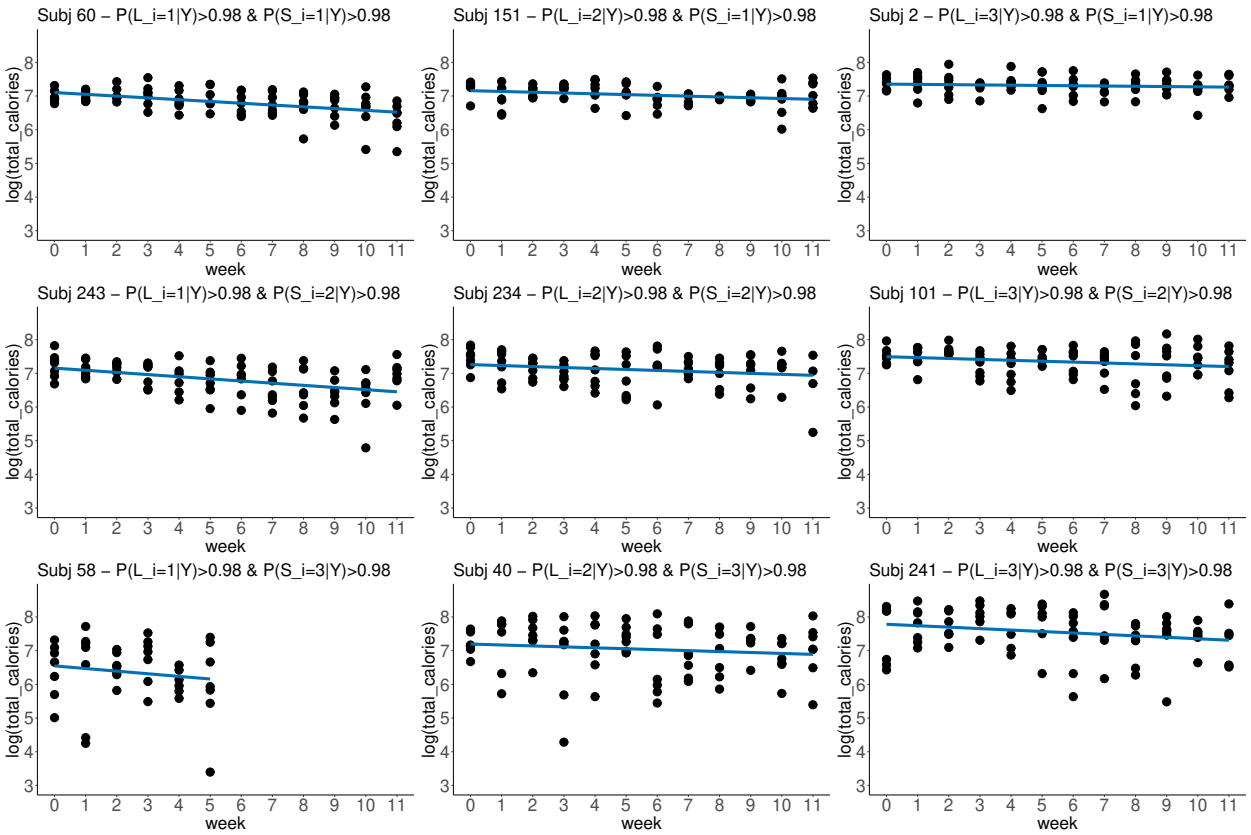


Figure 4.2: Subjects with high posterior probability of belonging to each LC combination. Blue lines represent trajectories of mean log(total calories) from subject-specific linear regressions.

the simulation study was limited to 100 datasets per scenario.

First, we generated simulated datasets based on parameter values similar to the estimates from the final model of the SMART study example. We used these datasets to examine the performance of our proposed model on data similar to the example. For every subject, we considered data across 12 time periods. Within each time period, we generated ten observations following the distribution assumed by the final model of the SMART study example. Each dataset includes 250 subjects, with each subject being randomly assigned to one of three latent location classes and one of three latent scale classes, each selection having an equal probability. The true values of the generating parameters are shown in Table 4.6.

Table 4.4: Characteristics and original treatment assignment of subjects in the weight loss management study by location LCs determined by posterior classifying probabilities. P-values were obtained from Chi-squared tests of independence for categorical variables and ANOVA for continuous variables. Observations with missing values for each respective variable are excluded.

Characteristic	Location LC 1 (N=132)	Location LC 2 (N=179)	Location LC 3 (N=68)	P-value
Biological gender				
Female	114 (86.4%)	134 (74.9%)	40 (58.8%)	<0.001
Male	18 (13.6%)	41 (22.9%)	28 (41.2%)	
Age at screening (years)				
Mean (SD)	40.6 (11.9)	40.0 (10.9)	39.1 (10.5)	0.672
Ethnicity				
Hispanic or Latino	21 (15.9%)	17 (9.5%)	7 (10.3%)	0.213
Not Hispanic or Latino	110 (83.3%)	160 (89.4%)	59 (86.8%)	
Race				
White	81 (61.4%)	129 (72.1%)	54 (79.4%)	0.027
Black or African American	34 (25.8%)	34 (19.0%)	5 (7.4%)	
Others	12 (9.1%)	12 (6.7%)	5 (7.4%)	
Weight at baseline (lbs)				
Mean (SD)	203.5 (35.5)	213.8 (33.6)	227.7 (32.8)	<0.001
Height at baseline (in)				
Mean (SD)	64.6 (6.5)	66.1 (3.2)	68.0 (3.3)	<0.001
Highest level of education at baseline				
High school graduate/Some college/Technical school	23 (17.4%)	19 (10.6%)	5 (7.4%)	0.075
College graduate	109 (82.6%)	160 (89.4%)	63 (92.6%)	
Employment status at baseline				
Employed	113 (85.6%)	165 (92.2%)	58 (85.3%)	0.123
Unemployed/Homemaker/Student/Retired	19 (14.4%)	14 (7.8%)	10 (14.7%)	
Relationship status at baseline				
Married/living with partner	63 (47.7%)	98 (54.7%)	44 (64.7%)	0.072
Divorced/Widowed/Separated/Never married	69 (52.3%)	81 (45.3%)	24 (35.3%)	
Original treatment assignment				
mHealth only	69 (52.3%)	80 (44.7%)	35 (51.5%)	0.362
mHealth + coaching	63 (47.7%)	99 (55.3%)	33 (48.5%)	

These datasets were analyzed using (1) location-scale regressions with three location LCs and three scale LCs, (2) location-scale regressions with three location LCs and one scale LC, and (3) regressions incorporating three location LCs without incorporating a scale

Table 4.5: Characteristics and original treatment assignment of subjects in the weight loss management study by scale LCs determined by posterior classifying probabilities. P-values were obtained from Chi-squared tests of independence for categorical variables and ANOVA for continuous variables. Observations with missing values for each respective variable are excluded.

Characteristic	Scale LC 1 (N=78)	Scale LC 2 (N=200)	Scale LC 3 (N=101)	P-value
Biological gender				
Female	56 (71.8%)	159 (79.5%)	73 (72.3%)	0.315
Male	21 (26.9%)	40 (20.0%)	26 (25.7%)	
Age at screening (years)				
Mean (SD)	43.1 (10.8)	39.3 (10.8)	39.0 (11.8)	0.024
Ethnicity				
Hispanic or Latino	4 (5.1%)	26 (13.0%)	15 (14.9%)	0.106
Not Hispanic or Latino	73 (93.6%)	171 (85.5%)	85 (84.2%)	
Race				
White	61 (78.2%)	138 (69.0%)	65 (64.4%)	0.083
Black or African American	9 (11.5%)	38 (19.0%)	26 (25.7%)	
Others	7 (9.0%)	18 (9.0%)	4 (4.0%)	
Weight at baseline (lbs)				
Mean (SD)	204.1 (32.9)	211.7 (35.1)	221.9 (34.9)	0.003
Height at baseline (in)				
Mean (SD)	65.9 (3.7)	65.7 (5.7)	66.3 (3.4)	0.571
Highest level of education at baseline				
High school graduate/Some college/Technical school	7 (9.0%)	19 (9.5%)	21 (20.8%)	0.011
College graduate	71 (91.0%)	181 (90.5%)	80 (79.2%)	
Employment status at baseline				
Employed	69 (88.5%)	183 (91.5%)	84 (83.2%)	0.099
Unemployed/Homemaker/Student/Retired	9 (11.5%)	17 (8.5%)	17 (16.8%)	
Relationship status at baseline				
Married/living with partner	51 (65.4%)	103 (51.5%)	51 (50.5%)	0.079
Divorced/Widowed/Separated/Never married	27 (34.6%)	97 (48.5%)	50 (49.5%)	
Original treatment assignment				
mHealth only	31 (39.7%)	97 (48.5%)	56 (55.4%)	0.114
mHealth + coaching	47 (60.3%)	103 (51.5%)	45 (44.6%)	

component. For easy comparison, Table 4.6 presents the averages of three key metrics for these models: the bias relative to true values, the average width of their corresponding 95% credible intervals, and the coverage rate of the true values. No label switching was detected

Table 4.6: Comparison of results from 100 simulations of location-scale regressions with LCs. AIW: average 95% credible interval width.

Parameter	True value	(1) 3 location LCs + 3 scale LCs			(2) 3 location LCs + 1 scale LC			(3) 3 location LCs + no scale component		
		Bias	AIW	Coverage (%)	Bias	AIW	Coverage (%)	Bias	AIW	Coverage (%)
$\beta_0 \vert L_i = 1$	6.92	4.88e-04	0.031	95.7	1.04e-04	0.038	97	5.94e-04	0.040	96.9
$\beta_0 \vert L_i = 2$	7.15	-4.20e-04	0.031	93.5	-5.11e-04	0.038	93	4.31e-05	0.040	90.6
$\beta_0 \vert L_i = 3$	7.38	7.46e-04	0.031	97.8	1.54e-03	0.038	99	8.59e-04	0.041	99.0
$\beta_1 \vert L_i = 1$	-0.02	-4.38e-05	0.005	93.5	-2.08e-05	0.006	97	-7.01e-05	0.006	96.9
$\beta_1 \vert L_i = 2$	-0.01	7.52e-05	0.005	94.6	-4.56e-05	0.006	96	-1.21e-04	0.006	97.9
$\beta_1 \vert L_i = 3$	-0.01	-4.56e-05	0.005	98.9	4.878e-05	0.006	97	8.05e-05	0.006	95.8
$\tau_0 \vert S_i = 1$	-2.41	1.05e-03	0.103	94.6	0.98	0.060	0	-	-	-
$\tau_0 \vert S_i = 2$	-1.74	2.18e-03	0.106	93.5	0.31	-	0	-	-	-
$\tau_0 \vert S_i = 3$	-0.78	5.23e-03	0.107	94.6	-0.65	-	0	-	-	-
$\tau_1 \vert S_i = 1$	0.02	1.44e-04	0.016	97.8	0.01	0.009	0	-	-	-
$\tau_1 \vert S_i = 2$	0.06	-7.34e-05	0.016	98.9	-0.03	-	0	-	-	-
$\tau_1 \vert S_i = 3$	0.02	4.93e-05	0.016	95.6	0.01	-	0	-	-	-

for models (1) and (2), while model (1) failed to converge for eight datasets. Label switching happened for four out of 100 datasets for model (3). Therefore, results presented in Table 4.6 are based on 92 datasets for model (1) and 96 datasets for model (3).

The first thing to notice in the simulation results are the small biases of the parameter estimates compared to their true values for Model (1), together with coverage rates around 95%, which suggest that the proposed method is suitable for data like those from the weight loss management study example. The AIWs are wider for all parameters in both Model (2) and Model (3) compared to Model (1).

Simulation results for Model (2) indicate that excluding latent classes (LCs) when modeling intraindividual variability can lead to unreliable scale parameter inferences. Specifically, we compared posterior means and corresponding 95% credible intervals for parameters τ_0 and τ_1 in Model (2) with true values of each of the three scale LCs. These 95% credible intervals fail to encompass the true values for any of the three LCs. The average estimated value for τ_0 is -1.43, while for τ_1, it is 0.03, both approximating the mean across the three LCs. This suggests that the model neglects the heterogeneity in initial values and the subsequent variability trends by not incorporating LCs in its scale component.

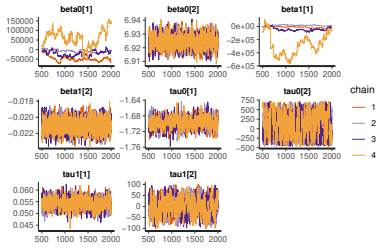
The average correct classification rates for the location component are notably high across all three models, at approximately 98.6%. Specifically, in Model (1), an average of 96.0% of subjects attain posterior classification probabilities exceeding 90% for a particular location LC. Also for Model (1), a similar high performance is observed in its scale part: 99.8% of subjects are correctly classified on average across all datasets, with an average proportion of 99.6% of subjects achieving posterior classification probabilities of at least 90% for a specific scale LC. These results suggest that the proposed model is highly effective in accurately classifying nearly all subjects with respect to both their location and scale trajectories.

Another objective of our simulation studies is to show that the proposed method does not over-identify LCs. Hence, we also simulated (1) datasets that assume both location and

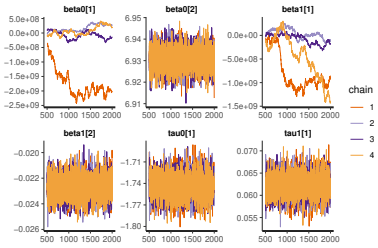
scale are homogeneous for all subjects and (2) datasets with two location LCs and two scale LCs. We then analyzed these datasets with our location-scale regression model with one more LC in either the location or scale part of the model or both. We found that these models would not converge, as shown by the trace plots in Figure 4.3. Specifically, when the model specifies one more location LC than in the dataset, the chains do not stabilize around a certain value but instead show a visible pattern for one set of β's. The chains also diverge from each other over iterations instead of converging to a common value. When there is one more scale LC in the model than in the dataset, the fluctuation is substantially wide for a pair of τ_0 and τ_1. Both findings occur when there are both an additional location LC and an additional scale LC in the model. These results provide compelling evidence that the LCs identified in Section 4.5 are a manifestation of actual heterogeneity in the data, rather than artifacts resulting from modeling limitations.

4.7 Discussion

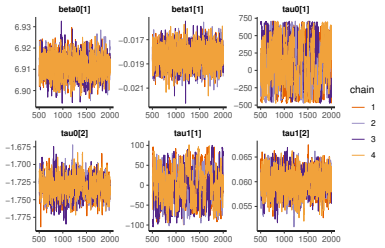
In our investigation, we explore the application of location-scale regressions incorporating LCs to model both mean and intraindividual variability trajectories. Through simulation studies, we substantiate the necessity of employing LC-based models when analyzing data exhibiting LC structures. This approach yields more precise parameter estimates and enhances the extraction of valuable insights from complex datasets characterized by LCs. Consequently, our research contributes a practical, user-friendly tool to researchers interested in subject subgrouping based on location-scale trajectory patterns. Instead of time trajectory, the proposed model can be used to conduct subgrouping in terms of the fixed effects of various time-varying covariates on mean and WS variability. In order to do so, researchers can simply replace the time covariate in the model with the specific time-varying covariate of relevance. In randomized controlled trials, baseline equivalence might be established, and analysis will focus on changes during the course of the trials. Our proposed model can be



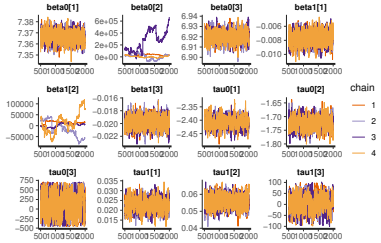
(a) Model with 2 location LCs and 2 scale LCs on data with no LC.



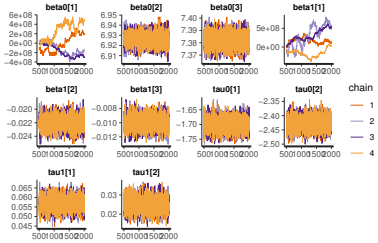
(b) Model with 2 location LCs and 1 scale LC on data with no LC.



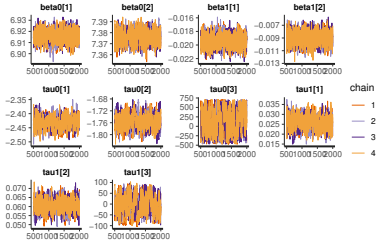
(c) Model with 1 location LC and 2 scale LCs on data with no LC.



(d) Model with 3 location LCs and 3 scale LCs on data with 2 location LCs and 2 scale LCs.



(e) Model with 3 location LCs and 2 scale LCs on data with 2 location LCs and 2 scale LCs.



(f) Model with 2 location LCs and 3 scale LCs on data with 2 location LCs and 2 scale LCs.

Figure 4.3: Markov chain trace plots for models with more LCs than in the simulated data.

easily adjusted to assume homogeneous starting points, or intercepts, and have the LCs determined by slopes only.

In the weight loss management study example, our proposed model divided the participants into three subgroups with different trajectories in location and another three subgroups with different trajectories in scale. Notably, prior research has substantiated the positive impact of dietary consistency on weight loss outcomes [43]. Consequently, there exists an opportunity for in-depth investigation within the subgroup displaying a more noticeable increasing trend in dietary variability. Such an inquiry aims to elucidate the underlying factors driving this transition, with the potential to inform strategies that can be employed to promote dietary consistency within this subgroup. However, it is important to acknowledge that calorie intake in this study relies on self-reporting. For instance, we see some really low daily

reported total calorie intake values in Figure 4.2. Under-reporting dietary intake in weight loss management initiatives is a known issue [8, 24]. Further research may be necessary to determine whether the observed variability in calorie intake reflects actual fluctuations or inconsistencies in reporting.

While this study focuses on linear trajectories in both location and scale, more complicated modeling of either mean and/or WS variability of outcome is possible. Yet, it is essential for readers to consider that these extensions may entail increased computational demands and a higher likelihood of convergence issues. Also, more observations from each subject might be required to achieve meaningful parameter estimates. Also, it is worth noting an alternative modeling approach, namely mixed location scale hidden Markov models [23], that emphasizes the sequential relationship of scale. In such sequential scale models, the focus is on time serial dependence of intraindividual variability over time, offering a distinct perspective on within-subject variability. Another assumption made in this study is the a priori independence between location LCs and scale LCs. A potential extension of this study could explore the possibility of incorporating correlated LCs and investigate the implications of such correlations.

The next steps of this study also include expanding the proposed model to include covariate(s) for the LC membership probabilities and building a joint model in which the primary outcome of a study, for example, weight loss in our example study, can be further modeled as a function of the LCs of subject behaviors. By integrating these components, the model can function as an analytical pipeline, examining the association between factors like treatment assignment and subjects' LC membership in behavior, followed by investigating the connection between the LC and the study's outcomes.

CHAPTER 5

CONCLUSIONS AND FUTURE DIRECTIONS

5.1 Summary and Conclusions

In this dissertation, three novel statistical methodologies were introduced to enhance existing location-scale models for analyzing intensive longitudinal data, aiming to achieve improved interpretability and statistical robustness. In Chapters 2 and 3, R^2 measures and influence analysis framework were designed for MELS models. And in Chapter 4, we introduced LCs into both the location and the scale components of the models. The validity and applicability of all three methods were assessed through simulation studies. Additionally, accompanying computer programs were provided to facilitate their implementation and usage.

In the field of public health sciences, the proposed methods have great potential for applications in psychology, behavioral studies, and clinical studies that collect wearable device data and EMA data. By leveraging standardized effect sizes, influence analysis, and LCs, researchers can gain deeper insights into BS heterogeneity within these datasets. Specifically, this dissertation has demonstrated the applications of the proposed methods in depression studies, health behavior studies, and lifestyle intervention studies.

5.2 Future Directions

While the methodologies described in this dissertation primarily target normally distributed univariate outcomes, the increasing availability of intensive longitudinal data across diverse fields offers opportunities to explore outcomes with different distributional assumptions. I am interested in the potential extensions of this dissertation to accommodate outcomes with different distributional assumptions, thereby enhancing flexibility in data analysis.

5.2.1 Multivariate Outcomes

Several correlated outcomes often need to be considered simultaneously to answer a complex research question. For instance, in quantitative psychology and behavioral studies, multivariate analysis techniques are particularly useful in analyzing mental traits, the development of cognitive skills and health behaviors, etc. In the clinical field, symptom development and treatment outcomes can also be represented by multivariate random variables.

All three methods elucidated in this dissertation have the potential for extension from univariate to multivariate analyses, accommodating correlations among multiple outcomes. However, such extensions may necessitate advancements in model estimation techniques to mitigate the challenges posed by increased model complexity, sample size requirements, and computational burden associated with multivariate regressions.

5.2.2 Non-Normal Outcome

One type of non-normal outcome to consider for future developments is zero-inflated non-negative outcomes, which often occur in studies about physical activity, substance abuse, health care utilization, etc. Recent advancement has enabled MELS modeling of these kinds of data [5, 48], so it is straightforward to extend the R^2 measures and the influence analysis framework described in Chapters 2 and 3 to consider zero-inflation as well.

We can also add a logistic part with LCs to the location-scale regression described in Chapter 4 to accommodate excess zeros. Let ζ_{ij} denote the logit of subject i's probability of conducting having non-zero outcome at time j, and μ_{ij} denote the mean of subject i's

non-zero outcome at time j, we have

$$[\zeta_{ij}|Z_i = z] = \gamma_{0z} + \gamma_{1z}t_{ij}, \tag{5.1}$$

$$[\mu_{ij}|L_i = \ell] = \beta_{0\ell} + \beta_{1\ell}t_{ij}, \tag{5.2}$$

$$\text{and } [\sigma^2_{\epsilon_{ij}}|S_i = s] = \exp(\tau_{0s} + \tau_{1s}t_{ij}), \tag{5.3}$$

where Z_i stands for subject i's LC membership in terms of their proportion of non-zero outcome, γ_{0z} represents the average proportion of non-zero outcome for the zth logistic LC at baseline, and γ_{1z} represents the average linear trend of how the proportion of non-zero outcome change over time for the zth logistic LC.

Adding this additional component not only allows for more accurate parameter estimates but can also be informative in many applications. For example, in the context of modeling PA, LCs derived from the logistic model can provide insights into subgroups based on exercise frequency.

As mentioned in the discussion sections of Chapters 2 and 3, there has been the development of MELS models for ordinal outcomes. Ordinal outcomes are particularly useful in surveys, including EMA surveys, as many responses are recorded using Likert-type scales. Therefore, the extension of methods described in this dissertation can become useful analytical tools for survey data analysis.

5.2.3 Joint Modeling

The model for zero-inflated non-negative outcomes described in Section 5.2.2 can be considered as an example of joint models for longitudinal data that concurrently captures the frequency, magnitude, and variability of health behaviors. In addition to modeling a longitudinal outcome itself using joint models, joint modeling of longitudinal data and a primary health outcome, usually time-to-event or binary, is also popular. The latter describes the dy-

namic relationship between longitudinal processes and the occurrence of events or outcomes over time.

To combine location-scale regressions with a primary outcome sub-model, we can employ a MELS model for the longitudinal data. The random effects derived from the MELS model can then serve as predictors in the primary outcome model. Or if an LC model described in Chapter 4 is applied for the longitudinal data instead, LCs can be used as predictors in the primary outcome sub-model. Utilizing standardized effect sizes such as R^2 measures and conducting influence analysis can enhance the interpretability and robustness of these more complex models, so extending the methods outlined in this dissertation for R^2 measures and influence analysis would represent a valuable contribution to advancing the field of joint modeling.

REFERENCES

[1] Herman Aguinis, Ryan K Gottfredson, and Harry Joo. Best-practice recommendations for defining, identifying, and handling outliers. *Organizational Research Methods*, 16 (2):270–301, 2013.

[2] AC Atkinson. Masking unmasked. *Biometrika*, 73(3):533–541, 1986.

[3] David A Belsley, Edwin Kuh, and Roy E Welsch. *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*. John Wiley & Sons, Hoboken, NJ, 2005.

[4] Yoav Benjamini and Yosef Hochberg. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57(1):289–300, 1995.

[5] Shelley A Blozis, Melissa McTernan, Jeffrey R Harring, and Qiwen Zheng. Two-part mixed-effects location scale models. *Behavior Research Methods*, 52(5):1836–1847, 2020.

[6] Anthony S Bryk and Stephen W Raudenbush. *Hierarchical Linear Models: Applications and Data Analysis Methods*. Sage, Thousand Oaks, CA, 1992.

[7] Ronald Christensen, Larry M Pearson, and Wesley Johnson. Case-deletion diagnostics for mixed models. *Technometrics*, 34(1):38–45, 1992.

[8] Susan Connor. Underreporting of dietary intake: Key issues for weight management clinicians. *Current Cardiovascular Risk Reports*, 14:1–10, 2020.

[9] R Dennis Cook. Detection of influential observation in linear regression. *Technometrics*, 19(1):15–18, 1977.

[10] Michael R Elliott. Identifying latent clusters of variability in longitudinal data. *Biostatistics*, 8(4):756–771, 2007.

[11] Lavinia Flueckiger, Roselind Lieb, Andrea H Meyer, and Jutta Mata. How health behaviors relate to academic performance via affect: An intensive longitudinal study. *PLoS One*, 9(10):e111080, 2014.

[12] John Fox and Georges Monette. *An R and S-Plus Companion to Applied Regression*. Sage, Thousand Oaks, CA, 2002.

[13] Nathan Gill and Donald Hedeker. Fast estimation of mixed-effects location-scale regression models. *Statistics in Medicine*, 42(9):1430–1444, 2023.

[14] Harvey Goldstein. Bootstrapping in multilevel models. In *Handbook of Advanced Multilevel Analysis*, pages 171–180. Routledge, New York, NY, 2011.

[15] Donald Hedeker, Robin J Mermelstein, and Hakan Demirtas. An application of a mixed-effects location scale model for analysis of ecological momentary assessment (EMA) data. *Biometrics*, 64(2):627–634, 2008.

[16] Donald Hedeker, Hakan Demirtas, and Robin J Mermelstein. A mixed ordinal location scale model for analysis of ecological momentary assessment (EMA) data. *Statistics and its Interface*, 2(4):391, 2009.

[17] Bei Jiang, Michael R Elliott, Mary D Sammel, and Naisyin Wang. Joint modeling of cross-sectional health outcomes and longitudinal predictors via mixtures of means and variances. *Biometrics*, 71(2):487–497, 2015.

[18] Bobby L Jones, Daniel S Nagin, and Kathryn Roeder. A SAS procedure based on mixture models for estimating developmental trajectories. *Sociological Methods & Research*, 29(3):374–393, 2001.

[19] Ian H Langford and Toby Lewis. Outliers in multilevel data. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 161(2):121–160, 1998.

[20] Sik-Yum Lee and Xin-Yuan Song. Evaluation of the bayesian and maximum likelihood approaches in analyzing structural equation models with small sample sizes. *Multivariate Behavioral Research*, 39(4):653–686, 2004.

[21] Xue Li and Donald Hedeker. A three-level mixed-effects location scale model with an application to ecological momentary assessment data. *Statistics in Medicine*, 31(26): 3192–3210, 2012.

[22] Xiaolei Lin, Robin J Mermelstein, and Donald Hedeker. A 3-level bayesian mixed effects location scale model with an application to ecological momentary assessment data. *Statistics in Medicine*, 37(13):2108–2119, 2018.

[23] Xiaolei Lin, Robin Mermelstein, and Donald Hedeker. Mixed location scale hidden markov model for the analysis of intensive longitudinal data. *Health Services and Outcomes Research Methodology*, 20:222–236, 2020.

[24] Jennie Macdiarmid and John Blundell. Assessing dietary intake: who, what and why of under-reporting. *Nutrition Research Reviews*, 11(2):231–253, 1998.

[25] Stuart WS MacDonald, Lars Nyberg, and Lars Bäckman. Intra-individual variability in behavior: links to brain structure, neurotransmission and neuronal activity. *Trends in Neurosciences*, 29(8):474–480, 2006.

[26] Charles E McCulloch. *Generalized Linear Mixed Models*. Beachwood, OH: Institute of Mathematical Statistics, 2003.

[27] Adam W Meade and S Bartholomew Craig. Identifying careless responses in survey data. *Psychological Methods*, 17(3):437, 2012.

[28] Bengt Muthén and Linda Muthén. Mplus. In Wim J van der Linden, editor, *Handbook of Item Response Theory*, pages 507–518. Chapman and Hall/CRC, New York, NY, 1 edition, 2017.

[29] Bengt Muthén and Kerby Shedden. Finite mixture modeling with mixture outcomes using the EM algorithm. *Biometrics*, 55(2):463–469, 1999.

[30] Daniel S Nagin and Kenneth C Land. Age, criminal careers, and population heterogeneity: Specification and estimation of a nonparametric, mixed Poisson model. *Criminology*, 31(3):327–362, 1993.

[31] John R Nesselroade. Interindividual differences in intraindividual change. In Linda M. Collins and John L. Horn, editors, *Best Methods for the Analysis of Change: Recent Advances, Unanswered Questions, Future Directions*, chapter 6, page 92–105. American Psychological Association, Washington, DC, 1991.

[32] Steffen Nestler. An extension of the mixed-effects growth model that considers between-person differences in the within-subject variance and the autocorrelation. *Statistics in Medicine*, 41(3):471–482, 2022.

[33] Rense Nieuwenhuis, HF Te Grotenhuis, and BJ Pelzer. Influence.ME: tools for detecting influential data in mixed effects models. *The R Journal*, 4(2), 2012.

[34] Rachel Nordgren, Donald Hedeker, Genevieve Dunton, and Chih-Hsiang Yang. Extending the mixed-effects model to consider within-subject variance for ecological momentary assessment data. *Statistics in Medicine*, 39(5):577–590, 2020.

[35] Thomas M Piasecki, Donald Hedeker, Lisa C Dierker, and Robin J Mermelstein. Progression of nicotine dependence, mood level, and mood variability in adolescent smokers. *Psychology of Addictive Behaviors*, 30(4):484, 2016.

[36] Philippe Rast, Scott M Hofer, and Catharine Sparks. Modeling individual differences in within-person variation of negative and positive affect in a mixed effects location scale model using BUGS/JAGS. *Multivariate Behavioral Research*, 47(2):177–200, 2012.

[37] Richard A Redner and Homer F Walker. Mixture densities, maximum likelihood and the EM algorithm. *SIAM Review*, 26(2):195–239, 1984.

[38] Niels Reisby, Lars F Gram, Per Bech, Adam Nagy, Gorm Odden Petersen, Jørgen Ortmann, Ilse Ibsen, Sven J Dencker, Ove Jacobsen, Ole Krautwald, et al. Imipramine: clinical effects and pharmacokinetic variability. *Psychopharmacology*, 54(3):263–272, 1977.

[39] Alvin C Rencher and G Bruce Schaalje. *Linear Models in Statistics*. John Wiley & Sons, Hoboken, NJ, 2008.

[40] Jason D Rights and Sonya K Sterba. Quantifying explained variance in multilevel models: An integrative framework for defining R-squared measures. *Psychological Methods*, 24(3):309, 2019.

[41] Jason D Rights and Sonya K Sterba. New recommendations on the use of R-squared differences in multilevel model comparisons. *Multivariate Behavioral Research*, 55(4): 568–599, 2020.

[42] Jason D Rights and Sonya K Sterba. Effect size measures for longitudinal growth analyses: Extending a framework of multilevel model R-squareds to accommodate heteroscedasticity, autocorrelation, nonlinearity, and alternative centering strategies. *New Directions for Child and Adolescent Development*, 2021(175):65–110, 2021.

[43] Diane L Rosenbaum, Leah M Schumacher, Katherine Schaumberg, Amani D Piers, Monika E Gaspar, Michael R Lowe, Evan M Forman, and Meghan L Butryn. Energy intake highs and lows: how much does consistency matter in weight control? *Clinical Obesity*, 6(3):193–201, 2016.

[44] Satpal S Sandhu and George Leckie. Orthodontic pain trajectories in adolescents: Between-subject and within-subject variability in pain perception. *American Journal of Orthodontics and Dentofacial Orthopedics*, 149(4):491–500, 2016.

[45] Mona Sharifi, Gareth Marshall, Roberta Goldman, Sheryl L Rifas-Shiman, Christine M Horan, Renata Koziol, Richard Marshall, Thomas D Sequist, and Elsie M Taveras. Exploring innovative approaches and patient-centered outcomes from positive outliers in childhood obesity. *Academic Pediatrics*, 14(6):646–655, 2014.

[46] Lei Shi and Gemai Chen. Detection of outliers in multilevel models. *Journal of Statistical Planning and Inference*, 138(10):3189–3199, 2008.

[47] Saul Shiffman, Arthur A Stone, and Michael R Hufford. Ecological momentary assessment. *Annual Review of Clinical Psychology*, 4:1–32, 2008.

[48] Juned Siddique, Michael J Daniels, Gül Inan, Samuel Battalio, Bonnie Spring, and Donald Hedeker. Joint modeling the frequency and duration of accelerometer-measured physical activity from a lifestyle intervention trial. *Statistics in Medicine*, 42(28):5100–5112, 2023.

[49] Tom AB Snijders and Roel J Bosker. Modeled variance in two-level models. *Sociological Methods & Research*, 22(3):342–363, 1994.

[50] Tom AB Snijders and Roel J Bosker. *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling*. Sage, Thousand Oaks, CA, 2011.

[51] Bonnie Spring, Angela Fidler Pfammatter, Laura Scanlan, Elyse Daly, Jean Reading, Sam Battalio, Harvey Gene McFadden, Donald Hedeker, Juned Siddique, and Inbal Nahum-Shani. An adaptive behavioral intervention for weight loss management: A noninferiority randomized clinical trial. *JAMA*, in press.

[52] Stan Development Team. Stan modeling language users guide and reference manual, version 2.31, 2022. URL `http://mc-stan.org/`.

[53] Gail M Sullivan and Richard Feinn. Using effect size—or why the p value is not enough. *Journal of Graduate Medical Education*, 4(3):279–282, 2012.

[54] Catrine Tudor-Locke, Sandra A Ham, Caroline A Macera, Barbara E Ainsworth, Karen A Kirtland, Jared P Reis, and C Dexter Kimsey Jr. Descriptive epidemiology of pedometer-determined physical activity. *Medicine and Science in Sports and Exercise*, 36(9):1567–1573, 2004.

[55] Aki Vehtari, Andrew Gelman, and Jonah Gabry. Practical bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, 27:1413–1432, 2017.

[56] Aki Vehtari, Andrew Gelman, Daniel Simpson, Bob Carpenter, and Paul-Christian Bürkner. Rank-normalization, folding, and localization: An improved $\hat{R}$ for assessing convergence of MCMC (with discussion). *Bayesian Analysis*, 16(2):667–718, 2021.

[57] Geert Verbeke and Emmanuel Lesaffre. A linear mixed-effects model with heterogeneity in the random-effects population. *Journal of the American Statistical Association*, 91 (433):217–221, 1996.

[58] Koen JF Verhoeven, Katy L Simonsen, and Lauren M McIntyre. Implementing false discovery rate control: increasing your power. *Oikos*, 108(3):643–647, 2005.

[59] Bridget E Weller, Natasha K Bowen, and Sarah J Faubert. Latent class analysis: a guide to best practice. *Journal of Black Psychology*, 46(4):287–311, 2020.

[60] Samuel S Wilks. Certain generalizations in the analysis of variance. *Biometrika*, 24 (3/4):471–494, 1932.

[61] Weichun Xu. *Mixtures in Random-Effects Regression Models*. PhD thesis, University of Illinois at Chicago, 1995.

[62] Lei Yu, Tianhao Wang, Robert S Wilson, Sue Leurgans, Julie A Schneider, David A Bennett, and Patricia A Boyle. Common age-related neuropathologies and yearly variability in cognition. *Annals of Clinical and Translational Neurology*, 6(11):2140–2149, 2019.

[63] Xingruo Zhang and Donald Hedeker. Defining R-squared measures for mixed-effects location scale models. *Statistics in Medicine*, 41(22):4467–4483, 2022.

[64] Xingruo Zhang and Donald Hedeker. Detecting influential subjects in intensive longitudinal data using mixed-effects location scale models. *BMC Medical Research Methodology*, 23(1):237, 2023.