Contributions to Mediation Analysis and First Principles Modeling for Mechanistic Statistical Analysis

by

Joseph R. Dickens

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Statistics)
in the University of Michigan
2020

Doctoral Committee:

Professor Kerby Shedden, Chair
Professor Johann Gagnon-Bartsch
Professor Chandra Sripada
Professor Gongjun Xu

Joseph R. Dickens

josephdi@umich.edu

ORCID iD: 0000-0002-3954-0771

ACKNOWLEDGEMENTS

I would like to begin by thanking Kerby, my advisor and committee chair. I learned a lot during our meetings and work over the past five years that I expect to use in my future career. I would additionally like to thank my committee members Johann, Chandra, and Gongjun for their contributions to my thesis work.

I would also like to thank the entirety of the Amidon research group, in particular Gordon, Gail, Marival, Bart, Kai and Paulo. My work with your team was my first collaborative experience as a statistician. I appreciated your patience and guidance and it was a pleasure to work with you.

Finally, I would like to thank the entirety of the UM Statistics department for its support during my studies. Liza and the department staff, in particular Judy, Gina, and Bebe, were all of tremendous help to me during my time in the department.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

GI gastrointestinal

FDA U.S. Food and Drug Administration

mg milligram

CDF cumulative distribution function

PGRS polygenic risk score

GLM generalized linear model

ABSTRACT

This thesis contains three projects that propose novel methods for studying mechanisms that explain statistical relationships. The ultimate goal of each of these methods is to help researchers describe how or why complex relationships between observed variables exist.

The first project proposes and studies a method for recovering mediation structure in high dimensions. We take a dimension reduction approach that generalizes the "product of coefficients" concept for univariate mediation analysis through the optimization of a loss function. We devise an efficient algorithm for optimizing the product-of-coefficients inspired loss function. Through extensive simulation studies, we show that the method is capable of consistently identifying mediation structure. Finally, two case studies are presented that demonstrate how the method can be used to conduct multivariate mediation analysis.

The second project uses tools from conditional inference to improve the calibration of tests of univariate mediation hypotheses. The key insight of the project is that the non-Euclidean geometry of the null parameter space causes the test statistic's sampling distribution to depend on a nuisance parameter. After identifying a statistic that is both sufficient for the nuisance parameter and approximately ancillary for the parameter of interest, we derive the test statistic's limiting conditional sampling distribution. We additionally develop a non-standard bootstrap procedure for calibration in finite samples. We demonstrate through simulation studies that improved evidence calibration leads to substantial power increases over existing methods. This project suggests that conditional inference might be a useful tool in evidence calibration for other non-standard or otherwise challenging problems.

In the last project, we present a methodological contribution to a pharmaceutical science study of *in vivo* ibuprofen pharmacokinetics. We demonstrate how model misspecification in a first-principles analysis can be addressed by augmenting the model to include a term corresponding to an omitted source of variation. In previously used first-principles models, gastric emptying, which is pulsatile and stochastic, is modeled as first-order diffusion for simplicity. However, analyses suggest that the actual gastric emptying process is expected to be a unimodal smooth function, with phase and amplitude varying by subject. Therefore, we adopt a flexible approach in which a highly idealized parametric version of gastric emptying is combined with a Gaussian process to capture deviations from the idealized

form. These functions are characterized by their distributions, which allows us to learn their common and unique features across subjects despite that these features are not directly observed. Through simulation studies, we show that the proposed approach is able to identify certain features of latent function distributions.

CHAPTER 1

Introduction

This thesis develops new methods for conducting mechanistic analyses. Most analyses carried out by statisticians and researchers assess whether a *statistical* relationship exists between two variables. A scientist may have several hypotheses that explain why a given statistical relationship exists. A mechanistic analysis allows the researcher to assess the evidence for or against their hypotheses. To do so, the researcher describes the system or phenomenon of interest through one or more processes or mechanisms and uses data to estimate the parameters of the mechanistic model. The researcher then assesses whether the parameter estimates are consistent with their mechanistic hypothesis. The power of mechanistic analyses are their ability to answer the question "why does y depend on x?"

We present two case studies in this thesis in order to highlight mechanistic analyses' strengths. The first example comes from the field of cognitive science. It is well-known that an individual's genes are predictive of their cognitive abilities or intelligence [1]. Of course, genes do not directly influence an individual's performance on cognitive tests, but rather influence how an individual's brain functions. A mechanistic analysis of cognitive ability identifies patterns of brain activity that are associated with both an individual's genes, and then show that these patterns of brain activity predict better performance on cognitive tests.

This mechanistic analysis is an example of mediation analysis. In this case study, one says that brain activity is a potential mediator of the genetic-intelligence pathway. In Chapter 2, we introduce a new methodology that efficiently identifies low-dimensional mediation

structure among higher-dimensional variables. We then apply the proposed method to data from a neuroimaging study to identify patterns of brain activity that mediate the association between an individual's genes and intelligence. In Chapter 3 we propose a test of univariate mediation hypotheses. A careful study of the testing problem leads to insights that allow us to design better calibrated tests. The improved calibration of the procedure will lead to the discovery of more mechanisms.

Our second example involves a case study from the pharmaceutical sciences. Before a generic drug product can be brought to market, it must be shown to behave equivalently to the reference drug product. The U.S. Food and Drug Administration (FDA) has created and implemented regulatory tests in order to establish bioequivalence of two drugs products. Unfortunately, many of these tests can be underpowered, meaning that they fail to show that two equivalent products meet the regulatory definition of bioequivalence.

Pharmaceutical scientists believe that the tests' low power is due to substantial between-trial variation in the *in vivo* environment. A research team at the University of Michigan conducted an intubation study to measure *in vivo* factors that affect drug concentrations in healthy human subjects [2]. In Chapter 4 we present a mechanistic case study which uses a compartmental model to describe *in vivo* drug concentrations. We introduce a new method capable of identifying mechanisms that cause variation *in vivo* drug concentrations. Results from our analysis suggest that a biologically plausible mechanism explains much of the observed heterogeneity in plasma drug concentrations.

These case studies underscore the importance of mechanistic analysis. Mechanistic analysis answers the important questions how and why one variable has an association with another. Two challenging statistical tasks in a mechanistic analysis are creating statistical models capable of describing mechanistic relationships and describing one's confidence that a mechanism is real. The methods proposed in Chapters 2, 3 and 4 address one or both of these tasks.

The remainder of Chapter 1 contains an introduction to mediation analysis. We include this section here because it provides background necessary for both Chapters 2 and 3.

1.1 Mediation Analysis Background

Chapters 2 and 3 introduce two new methods for identifying and testing for mediation relationship. In this section we provide an introduction to mediation analysis that will help the reader understand the analytic questions that pertain to both projects. A separate, project-specific introduction and literature review will begin each chapter.

Mediation analysis is a tool that helps one better understand statistical relationships that are established through exposure-outcome data analyses [3]. When the exposure is randomly assigned to observational units, one hopes to establish that the outcome causal depends on the exposure or treatment. Establishing that an association or causal relationship exists does not directly explain *how* the exposure is associated with the outcome. Mediation analysis is the the formal analysis of potential mediating variables, which represent mechanisms that explain statistical exposure-outcome associations.

To establish that a variable is a mediator, one assesses whether part of the exposure's association with the outcome "flows" through the mediating variable. This formally requires establishing that part of the exposure-outcome association is attributable to the mediator. Mediation analysis decomposes the "total effect" of the exposure on the outcome into two components. The first part is called the "indirect effect" and refers to changes in outcome that are associated with the exposure *through* the exposure's association with the mediator. The second part of the decomposition, the "direct" effect, is the remaining exposure-outcome association after accounting for the mediator's role. When the indirect effect is estimated to be statistically different than 0, the variable is declared to mediate the exposure-outcome association.

Figure 1.1 shows the typical mediation setting as a path diagram. The variables X, M,

and Y play the roles of an exposure, mediator, and outcome, respectively. The diagram makes it easier to visualize the direct and indirect effects. The indirect effect is the portion of the $X - Y$ association that flows through M (the $X \to M \to Y$ pathway in Figure 1.1). The direct effect reflects the portion of the $X \to Y$ relationship that is explained by mechanisms unrelated to M (the direct $X \to Y$ pathway in Figure 1.1). In many common settings, in order to establish that M is a mediator of the $X - Y$ association, one must show that both of the blue arcs in Figure 1.1 exist.

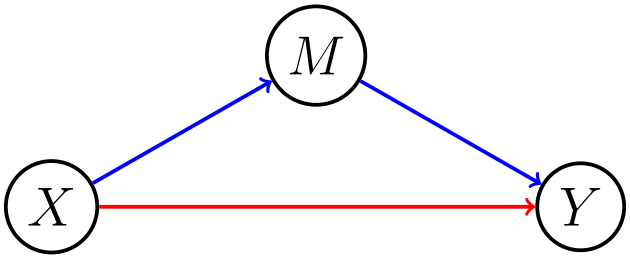


Figure 1.1: A typical mediation analysis setting in which the effect of a treatment X may be partially mediated through M.

1.2 Mediation literature review

Mediation analysis' origin lies in the work of the geneticist Sewell Wright, who introduced the technique of path analysis in the 1910s and 1920s [4, 5]. Wright developed path analysis in order to study heritability of traits in guinea pigs [4], although he was aware that the methodology is more broadly applicable. Path analysis represents functional relationships between variables through a diagram. Nodes represented variables and paths between nodes represented dependencies between variables. A modern reader would recognize Wright's diagrams as dependency graphs, although his work predated modern graph theory.

Wright realized that given a conceptual model of how variables co-relate and the necessary data to estimate partial correlations, one could estimate the relationship between any two variables in a path diagram [5]. Furthermore, the correlation between two variables

could be divided into unique contributions from each path connecting the two variables. Mediation analysis borrows this decomposition from path analysis when it divides the total effect of an exposure on an outcome into the indirect and direct effect.

Formal tests of mediation hypotheses were first introduced in the structural equation modeling (SEM) literature. The first such test was proposed by Michael Sobel, who derived a normal approximation to the sampling distribution of the indirect effect [6, 7]. He derived the approximation in the *linear structural equations model* (LSEM) setting [8, 9, 10, 11]. Let $X, M,$ and $Y \in \mathbb{R}$ play the roles of a treatment or exposure, a potential mediator, and outcome measure, respectively. Additionally, assume that the conditional mean of the hypothesized mediator M is linear in X, while the conditional mean of Y is linear in both X and M. The following system of equations describes these relationships and assumptions:

$$\begin{aligned} X &\sim F \\ M &= \gamma_1 + \alpha X + \epsilon_m \\ Y &= \gamma_2 + \gamma_3 X + \beta M + \epsilon_y, \end{aligned} \tag{1.1}$$

with $\mathbb{E}[\epsilon_m, \epsilon_y] = [0, 0]$, $\text{Cov}(\epsilon_m, \epsilon_y)$ diagonal, and F is a valid probability distribution.

For the LSEM, the indirect effect of X on Y is given by $\alpha\beta$. The Sobel test uses the delta method to approximate the sampling distribution of $\alpha\beta$ as a normal random variable. Under H_0, $\alpha\beta = 0$, and so $\hat{\alpha}\hat{\beta} \sim \mathcal{N}(0, \sigma^2_{\alpha\beta})$, where $\sigma^2_{\alpha\beta} = \hat{\alpha}^2\sigma^2_{\beta} + \beta^2\sigma^2_{\alpha}$ [6]. The covariance between $\hat{\alpha}\hat{\beta}$ is assumed to be small and is ignored when calculating the standard error of $\hat{\alpha}\hat{\beta}$. An $\tilde{\alpha}$-level test of the null hypotheses compares $Z = |\hat{\alpha}\hat{\beta}|/\sigma_{\alpha\beta}$ to the $1 - \tilde{\alpha}/2$ quantile of the standard normal distribution, rejecting H_0 whenever Z is larger than the critical value.

Another procedure for assessing whether M is a mediator in an LSEM was proposed in a paper that sought to clarify the difference between moderation and mediation [12]. The method was named after the paper's authors, and called the "Baron and Kenny" approach. This approach has become one of the most common ways of testing mediating hypotheses

after its introduction [8, 13]. They consider the following linear model in addition to those in 1.1:

$$Y = \phi_1 + \phi_2 X + \epsilon_3.$$

To decide if M is a mediator, according to the Baron and Kenny approach, one must test:

1. $H_0 : \phi_2 = 0$ to assess whether X and Y are significantly associated.
2. $H_0 : \alpha = 0$ to assess whether X and M are associated.
3. $H_0 : \beta = 0$ to assess whether Y and M are associated, controlling for X.

In order to declare that M is a mediator at the $\tilde{\alpha}$ level, The "Baron and Kenny" approach require that each null hypothesis is rejected at the $\tilde{\alpha}$ level.

A number of authors have pointed out the shortcomings of both the "Baron and Kenny" approach and the Sobel test [8, 13]. First, the inferential properties of the "Baron and Kenny" method are not well understood. In particular, the significance level of the "Baron and Kenny" approach is not obvious since it tests multiple hypotheses sequentially. Additionally, M can in fact be a mediating variable even if $\phi_2 = 0$, which occurs when $\gamma_3 = \alpha\beta$. Thus, regardless of the procedure's level, it can make incorrect inferences due to a fundamental flaw.

The performance of the Sobel test relies on the quality of the normal approximation to the sampling distribution of $\hat{\alpha}\hat{\beta}$. In two common scenarios, when the sample size n is small or either α or β is close to 0, the normal approximation can be quite bad. The Sobel test, like the "Baron and Kenny" approach, has been observed to be conservative under either of these scenarios.

Subsequently, many authors have proposed new methods for testing mediation hypotheses, which are reviewed in [14]. These methods are primarily resampling-based, which allows for better approximation of the sampling distribution of $\hat{\alpha}\hat{\beta}$ in finite samples. The simplest of these approaches uses a non-parametric bootstrap to estimate the sampling distribution of $\hat{\alpha}\hat{\beta}$ [13]. One version of this approach takes non-parametric bootstrap samples

of the original dataset, and uses each bootstrap sample to estimate the indirect effect $\alpha\beta$. After taking n_b bootstrap samples, one uses $\{(\hat{\alpha}\hat{\beta})_j\}_{j=1}^{n_b}$ to create a $(1-\tilde{\alpha})\%$ confidence interval for $\alpha\beta$. [14] suggests creating a percentile-based bootstrap confidence interval. If 0 lies outside of the interval, then M is declared to be a mediating variable of the $X-Y$ association. Throughout this paper, we will refer to this testing approach as the confidence interval-based approach.

The approach outlined in [13] became popular because its authors released SPSS and SAS macros that implemented its testing procedure. This made the method accessible to many scientists who otherwise would not have been able to implement the bootstrap-based estimation and testing routines.

Another re-sampling based method uses the counterfactual framework of causal inference to define mediation effects in terms of potential outcomes [15]. Using this framework permits the conditional distributions of M given X and Y given X and M to be non-linear, greatly increasing the applicability of the method. For a full description of the method, see [15]. Appendix A provides a review of studying mediating relationships in the counterfactual outcomes framework.

The method introduced in [15] is a quasi-Bayesian approach. Given statistical models for M and Y, one estimates the parameters $\theta = (\theta_1, \theta_2)$ of two statistical models and their sampling variance-covariance matrix. Next, one takes draws $\hat{\theta}^j = (\hat{\theta}_1^j, \hat{\theta}_2^j)$ from the sampling distribution of $\hat{\theta} = (\hat{\theta}_1, \hat{\theta}_2)$. Using the first statistical model and $\hat{\theta}_1^j$, the analyst takes parametric bootstrap samples of the mediator M with X fixed. This is followed by taking additional parametric bootstrap samples of the outcome Y using its statistical model and $\hat{\theta}_2^j$ with X fixed and M from the first parametric bootstrap. After repeating this process many times, one can estimate and create a $(1-\tilde{\alpha})\%$ confidence interval for the indirect effect by contrasting the correct potential outcomes (see [15] for details),

In Chapters 2 and 3 we make two methodological contributions to the mediation anal-

ysis literature. Chapter 2 introduces a new dimension reduction approach to studying mediating structure with vector-valued variables. Our approach is conceptually related to the classical method of canonical correlation analysis and is a principled, efficient method for uncovering low-dimensional mediation structure. Chapter 3 introduces a novel method for quantifying the certainty that a mediation effect is real, improving upon the approaches described in Section 1.2. Our approach uses conditional inference to better approximate the test statistic's sampling distribution. By doing so, our test of the indirect effect is better calibrated and therefore capable of discovering more true mediators.

CHAPTER 2

A Projection Pursuit Approach to Identify Low-Dimensional Mediation Structure from Higher Dimensional Data

2.1 Problem Definition

Recently, multivariate mediation methods have become a popular focus of mediation analysis research. These methods are necessary because researchers often have multiple mechanisms they would like to assess. In these settings, the researcher would like to consider all potential mediators simultaneously, and screen each variable's role in the $X \rightarrow Y$ association conditional on the other mediators. Existing multivariate mediation methods are model-based approaches, and are applicable when M is vector-valued and X and Y are scalars.

Existing multivariate mediation methods do not yet address the general multivariate mediation setting in which (X, M, Y) are each vector-valued. In this chapter, we present a novel approach for identifying low-dimensional mediation structure when (X, Y, M) are all potentially vector-valued variables. Our methodology reduces data based on an optimization goal. This is similar to the classical method canonical correlation analysis (CCA), and like CCA our results can be viewed in the context of various population models.

The optimization problem is based on maximizing an objective function motivated by

the classical "product-of-coefficients" estimate of the indirect effect. We will refer to the problem's objective function as the "mediation directions objective function." Like other projection pursuit methods, our choice of objective function is not derived from a statistical model. However, we believe that its choice is principled and intuitive, and will show that it is capable of identifying meaningful mediation structure.

The mediation directions objective function defines a non-convex optimization problem. Unlike CCA, our approach does not lead to a classically tractable optimization problem. We provide algorithms to find local optima of the objective function, and through simulation show that a local greedy descent algorithm using multiple initializations obtains good solutions. We will show that for several population models, the method quickly and consistently identifies the low-dimensional mediation structure.

Unlike existing methods, our approach is moment-based rather than likelihood-based. The mediation directions objective function was not derived from a particular statistical model's likelihood function, and the optimization variables do not necessarily correspond to parameters of a statistical model. Optimizing the mediation directions objective function is a one-step procedure. A few existing methods [16, 17] jointly estimate both the mediation parameters and the low-dimensional structure jointly, while others [18, 19] use two-step estimation schemes. Similar to other projection pursuit techniques, one might consider analyzing the projected exposures, mediators, and outcomes using univariate mediation methods. From this prospective, our approach becomes a two-step procedure as well.

The greatest strength of our methodology arises when two or more of $X, M,$ and Y are vector-valued. In such a setting, it is easy to imagine that the dominant associations between variables might not contain any mediation structure. We will later show through simulation that our method is still capable of identifying the present mediation structure. It is not immediately clear to us how a likelihood-based method would approach this problem. It would likely require specifying a complex, multilayer, factor-style model. *A priori* specification of the mediation structure in a multi-factor model would be difficult. In-

stead, such an approach would necessitate penalized estimation or post hoc analysis of the fitted model. For multivariate (X, M, Y) with non-dominant mediation structure, we believe our optimization-based methodology is a more attractive approach to identifying low-dimensional mediation structure.

2.1.1 A review of existing multivariate mediation analysis methods

Existing multivariate mediation methods are regression-based and address situations where X and Y are scalar-valued and M is vector-valued. Several authors have recently proposed projection- or factor-based methods for identifying low-dimensional mediation structure when both X and Y are scalar-valued. We present both of these approaches next, and conclude the section by contrasting our method with existing methods.

Regression-based approaches generalize the univariate mediation model by introducing additional mediating variables. Let M_i for $i = 1, \ldots, p$ denote p potential mediating variables, and let $X, Y \in \mathbb{R}$ denote a scalar exposure and outcome, respectively. Each M_i is independently modeled as a linear function of X and other covariates, omitted here for clarity of presentation. Y is modeled as linear function of X, each M_i, and other control covariates, again omitted. This gives us the following system of linear models:

$$\begin{aligned} M_1 &= \theta_1 + \alpha_1 X + \epsilon_1 \\ &\vdots \\ M_p &= \theta_p + \alpha_p X + \epsilon_p \\ Y &= \beta_0 + \gamma X + \sum_{i=1}^{p} \beta_i M_i + \tilde{\epsilon}. \end{aligned} \tag{2.1}$$

The standard triangle diagram used to describe mediation relationship is modified to add the additional $m - 1$ candidate mediators (Figure 2.1). In this type of analysis, conditional relationships between the mediators are not estimated. An indirect effect for each candidate mediator M_i is given by the product $\tau_i = \alpha_i \beta_i$ for $i = 1, \ldots, p$. The indirect effects in

this setting can be thought of as conditional indirect effects, as each β_i is the conditional association of a one-unit change in M_i controlling for X, all other candidate mediators, and any other covariates.

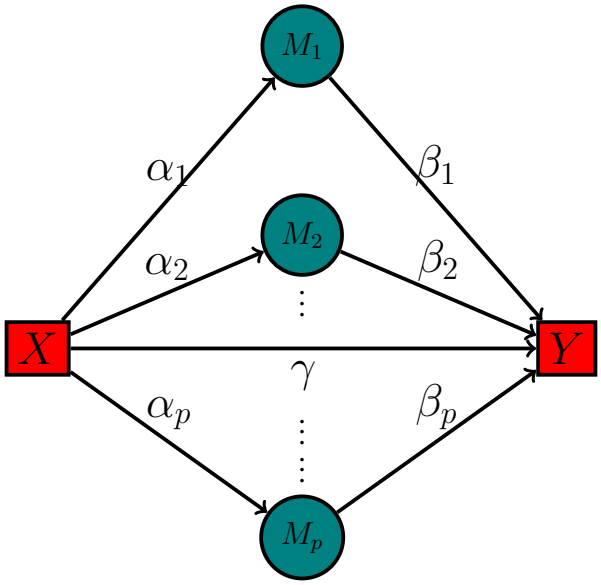


Figure 2.1: A path diagram for a typical multiple mediator model with scalar-valued X and Y.

The vector of candidate mediators can be high-dimensional, which is often the case in medical imaging or genetic applications. In such a setting, penalized regression is used to estimate a sparse set of variables that explain the $X \to Y$ association. One author proposed using the non-convex maximum concavity penalty (MCP) to identify a sparse set of mediating variables when M is high-dimensional [20]. The MCP estimator provides estimates of standard errors for each non-zero regression coefficient, which are used to calculated corrected p-values to test whether M_i is a mediating variable [20].

Other authors [19, 18, 16, 17] have proposed methods that identify a projection of the mediating variables. Let $w \in \mathbb{R}^p$ denote a vector, so the projected variate is $\tilde{M} = M^T w$. The vector w is estimated so that $\tilde{M}$ fits "optimally" into the linear structural equations model 1.1. Although the details vary, the papers [17, 16] approach the problem of estimating w via maximum likelihood estimation. A likelihood function $\ell(\cdot|X, M, Y)$ for the parameters of the LSEM Model 1.1 and w is defined. The likelihood function is then iter-

atively optimized by either an alternating least squares algorithm [16], or an EM algorithm [17]. These approaches seek to maximize ℓ jointly in terms of the LSEM parameters and w. The paper [17] includes a sparsity-inducing penalty on the components of w, which further distinguishes it from [16].

In contrast, [19] proposed using principal components analysis (PCA) to transform the p mediating variables into a smaller set of say k variates $\tilde{M}_j$ for $j = 1, \ldots, k$. The transformation via PCA takes place after residualizing each M_i against X. Then, given X, the variables $\tilde{M}_j$ and $\tilde{M}_{j'}$ are independent for $j \neq j'$. Since the $\tilde{M}_j$ are independent conditional on X, to estimate the indirect effect of X on Y through $\tilde{M}_j$, one fits k separate univariate mediation models using Model 1.1. One can also induce sparsity in the principal components, which means that only a few M_j have non-zero loadings on each component [18]. In addition to the joint estimation scheme described above, [17] proposed a two-stage estimator analogous to the approaches of [19, 18].

The remainder of this chapter is organized as follows. In Section 2.2 we introduce the mediation directions objective function and motivate its connection to univariate mediation analysis. In Section 2.3 we describe several computational algorithms for optimizing the mediation directions objective. In Sections 2.4 and 2.5 we introduce several multivariate statistical models with low-dimensional mediation structure and then show that we can consistently identify this structure through simulation. Finally, we conclude the chapter in Section 2.6.1 by using our methodology in two case studies.

2.2 The multivariate mediation objective function

Our objective function is inspired by the data reduction strategy of CCA and the product-of-coefficients estimate of traditional scalar-valued mediation analysis. Let $X \in \mathbb{R}^k$ denote a collection of treatment or exposure variables and $Y \in \mathbb{R}^m$ denote outcome measures. CCA finds the vectors $\beta \in \mathbb{R}^k$ of X and $\theta \in \mathbb{R}^m$ of Y such that $\text{Cor}(\beta^T X, \theta^T Y)$ is maximized.

Our focus is on a third variable $M \in \mathbb{R}^\ell$ that might mediate the $X - Y$ relationship. Let $\eta \in \mathbb{R}^\ell$ and suppose that we aim to maximize

$$\tau(\beta, \eta, \theta) = \text{Cor}(\beta^T X, \eta^T M) \times \text{Cor}(\eta^T M, \theta^T Y \mid X). \tag{2.2}$$

The criterion 2.2 is analogous to the traditional product of coefficients for standardized scalar X, M, and Y. Alternatively, one could formulate the optimization goal using the marginal correlation of $\eta^T M$ and $\theta^T Y$:

$$\tau_m(\beta, \eta, \theta) = \text{Cor}(\beta^T X, \eta^T M) \times \text{Cor}(\eta^T M, \theta^T Y), \tag{2.3}$$

where the subscript m denotes that the *marginal* correlation between $\eta^T M$ and $\theta^T Y$ is used rather than the conditional correlation. We recommend working with objective function 2.2 rather than 2.3 due to its closer connection to traditional univariate mediation analysis.

To operationalize 2.2, one can create the partial cross-correlation matrix of (M, Y) by residualizing M and Y with respect to X. Let P_X be the projection matrix onto the column space of X and form the matrices of residuals $\tilde{M} = M - P_X M$ and $\tilde{Y} = Y - P_X Y$. The partial cross-correlation matrix is then calculated using $\tilde{M}$ and $\tilde{Y}$.

Suppose that we have β^*, η^* and θ^* that maximize objective 2.2. These coefficients reduce the data in order to maximize the mediated effect described previously. Interpretations of the coefficients must be made in the context of the optimization goal that produced them, similar to the way one interprets CCA's canonical directions and PCA's loadings. In the multivariate mediation setting this means that the coefficients give the optimal linear combination of M that correlates with linear combinations of both X and Y given X.

We now introduce notation that allows us to precisely state our optimization goal. Let X, M, and Y all have mean 0. Define the following cross-covariance and covariance

matrices,

$$\tilde{\Sigma}_{XM} = \mathbb{E}[XM^T],\ \tilde{\Sigma}_{MY} = \mathbb{E}[MY^T|X],$$
$$\tilde{\Sigma}_X = \mathbb{E}[XX^T],\ \tilde{\Sigma}_M^1 = \mathbb{E}[MM^T],\ \tilde{\Sigma}_M^2 = \mathbb{E}[MM^T|X],\ \text{and}\ \tilde{\Sigma}_Y = \mathbb{E}[YY^T|X].$$

and assume that the second moments exist and are finite. Our optimization problem can be stated as follows:

$$\begin{aligned}\underset{\beta\in\mathbb{R}^k,\eta\in\mathbb{R}^l,\theta\in\mathbb{R}^m}{\text{maximize}}\ \tau(\beta,\eta,\theta) &= \text{Cor}(\beta^T X, \eta^T M) \times \text{Cor}(\eta^T M, \theta^T Y|X) \\ &= \beta^T \text{Cor}(X,M)\eta \times \eta^T \text{Cor}(M,Y|X)\theta \\ &= \frac{\beta^T \tilde{\Sigma}_{XM}\eta \cdot \eta^T \tilde{\Sigma}_{MY}\theta}{\sqrt{\beta^T \tilde{\Sigma}_X \beta \cdot \eta^T \tilde{\Sigma}_M^1 \eta \cdot \eta^T \tilde{\Sigma}_M^2 \eta \cdot \theta^T \tilde{\Sigma}_Y \theta}}.\end{aligned} \tag{2.4}$$

For $\mathcal{I} \in \{X, M, Y\}$, denote the square root $\tilde{\Sigma}_I^{1/2}$ of $\tilde{\Sigma}_\mathcal{I}$, where $\tilde{\Sigma}_\mathcal{I} = \tilde{\Sigma}_\mathcal{I}^{1/2}\tilde{\Sigma}_\mathcal{I}^{T/2}$. Define the transformations

$$a = \tilde{\Sigma}_X^{T/2}\beta,\ \ b = \left(\tilde{\Sigma}_M^1\right)^{T/2}\eta,\ \text{and}\ c = \tilde{\Sigma}_Y^{T/2}\theta. \tag{2.5}$$

These transformations allow us to perform a standard whitening of the variance-covariance matrices of X, M and Y. After transformation, we have an equivalent optimization problem in the optimization variables a, b and c:

$$\begin{aligned}\underset{a\in\mathbb{R}^k,b\in\mathbb{R}^l,c\in\mathbb{R}^m}{\text{maximize}}\ \tau(a,b,c) &= \frac{(\tilde{\Sigma}_X^{-T/2}a)^T \tilde{\Sigma}_{XM}(\tilde{\Sigma}_M^{-T/2}b) \cdot (\tilde{\Sigma}_M^{-T/2}b)^T \tilde{\Sigma}_{MY}(\tilde{\Sigma}_Y^{-T}c)}{\sqrt{a^Ta \cdot b^Tb \cdot b^T\left(\left(\tilde{\Sigma}_M^1\right)^{-1/2}\tilde{\Sigma}_M^2\left(\tilde{\Sigma}_M^1\right)^{-T/2}\right)b \cdot c^Tc}} \\ &= \frac{a^T\Sigma_{XM}b \cdot b^T\Sigma_{MY}c}{\sqrt{a^Ta \cdot b^Tb \cdot b^T\Sigma_M b \cdot c^Tc}} \\ &= \frac{a^T\Sigma_{XM}b}{\sqrt{a^Ta \cdot b^Tb}} \times \frac{b^T\Sigma_{MY}c}{\sqrt{b^T\Sigma_M b \cdot c^Tc}},\end{aligned} \tag{2.6}$$

where:

$$\Sigma_{XM} = \tilde{\Sigma}_X^{-1/2}\tilde{\Sigma}_{TM}\tilde{\Sigma}_M^{-T/2}, \;\; \Sigma_{MY} = \tilde{\Sigma}_M^{-1/2}\tilde{\Sigma}_{MY}\tilde{\Sigma}_Y^{-T/2}, \;\; \text{and}$$
$$\Sigma_M = \left(\tilde{\Sigma}_M^1\right)^{-1/2} \tilde{\Sigma}_M^2 \left(\tilde{\Sigma}_M^1\right)^{-T/2}.$$

This representation of the optimization objective allows us to describe populations for which the objective function is equal to zero for any β, η, and θ. This occurs when:

1. $\Sigma_{XM} = 0 \in \mathbb{R}^{k \times \ell}$, or

2. $\Sigma_{MY} = 0 \in \mathbb{R}^{\ell \times m}$.

This form of the optimization problem also represents a fully reduced version of the problem. For any (X, M, Y) with full-rank variance-covariance matrices, the optimization problem 2.4 can be transformed into an equivalent problem 2.6. The reduced problem has three inputs Σ_{XM}, Σ_{MY}, and Σ_M rather than the six in objective 2.4. Having solved the optimization problem 2.6, one can use the inverse transformations defined by 2.5 to find the solution in the original coordinates.

The connection between our objective function and CCA is clear in Equation 2.6. The objective function appears to be comprised of two CCA objective functions. However, the vector b appears in each term, which connects the two CCA subproblems. The coupling however is critical from a conceptual perspective. In order to qualify as a mediating projection of M, this projection must be simultaneously correlated with projections of X and Y. For the remainder of this proposal, we will use the terms "coefficients", "projections", and "mediation directions" interchangeably.

2.3 Computational aspects of estimating mediation directions

2.3.1 Characteristics of the multivariate mediation objective

The square of our objective function τ is a rational function, as both the numerator and denominator can be expressed in polynomial form. The two bilinear forms in the numerator of τ can be expressed as a polynomial in the variables $a \in \mathbb{R}^k$ and $b \in \mathbb{R}^\ell$ and $c \in \mathbb{R}^m$:

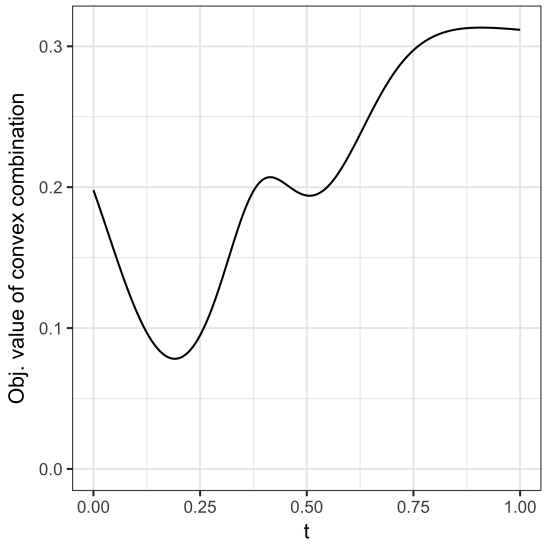
$$a^T \Sigma_{XM} b = \sum_{i=1}^{k} \sum_{j=1}^{\ell} a_i b_j \Sigma_{XM}(i,j) \text{ and } b^T \Sigma_{MY} c = \sum_{j=1}^{\ell} \sum_{h=1}^{m} b_j c_h \Sigma_{MY}(j,h). \tag{2.7}$$

The squared objective τ^2 is:

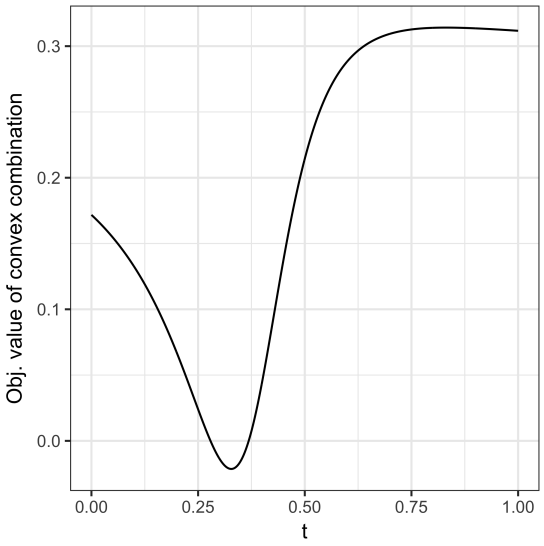
$$\tau(a,b,c)^2 = \frac{\left(\sum_{i=1}^{k} \sum_{j=1}^{\ell} a_i b_j \Sigma_{XM,ij}\right)^2 \left(\sum_{j=1}^{\ell} \sum_{h=1}^{m} b_j c_h \Sigma_{MY,jh}\right)^2}{\sum_{i=1}^{k} a_i^2 \cdot \sum_{j=1}^{l} b_j^2 \cdot \sum_{i=1}^{\ell} \sum_{j=1}^{\ell} b_i b_j \Sigma_M(i,j) \cdot \sum_{h=1}^{m} c_h^2}. \tag{2.8}$$

From 2.8, it is clear that $\tau(a,b,c)^2$ is a rational function. The polynomials in both numerator and denominator are homogeneous of degree 8. If we consider τ^2 restricted to each argument independently, then the polynomials are homogeneous of degree 2, 4, and 2 for a, b, and c, respectively. This is important, as it implies that the scaling of our optimization variables is irrelevant, thus without loss of generality they can be regarded as unit vectors.

Finally, it is easy to check that our function τ^2 is not jointly convex in a, b, and c. Additionally, restrictions of τ^2 to any of its arguments are not convex functions since the Hessians of the restrictions of τ^2 are indefinite matrices. Figures 2.2a and 2.2b show that τ is neither convex jointly in all of its optimization variables, nor when restricted to β, η, or θ.



(a) This figure plots the objective function value for convex combinations of two local optima. As one moves from $t = 0$ to $t = 1$, the function argument moves from one local optimum to the other. It is clear that the function is non-convex in this convex combination.



(b) The objective function τ is non-convex when restricted to η. Here we take convex combinations of the first and second mediation directions of M and again find that the objective function is non-convex.

Figure 2.2: The mediation directions objective function is a non-convex function.

2.3.2 Algorithms for estimating the mediation directions

Let n denote the number of independent, jointly observed samples of (X, M, Y). With a slight abuse of notation, these samples are organized into matrices $X \in \mathbb{R}^{n\times k}$, $M \in \mathbb{R}^{n\times \ell}$, and $Y \in \mathbb{R}^{n\times m}$, where the i^{th} observational unit's data is in the i^{th} row of each matrix. We replace the population moments in optimization problem 2.6 with their sample counterparts, denoted with a hat.

The scale-invariance property of the objective allows us to solve an equivalent constrained optimization problem 2.9 instead of the unconstrained problem 2.6. Through the course of development, we have found that algorithms designed to maximize the quotient perform better than algorithms that solve constrained versions of problem 2.6. In practice, we will always rescale our mediation directions to have unit length after optimization is complete. Below we give the sample version of the population optimization objective 2.6 and its constrained form.

$$\underset{a\in\mathbb{R}^k, b\in\mathbb{R}^l, c\in\mathbb{R}^m}{\text{maximize}} \frac{a^T\hat{\Sigma}_{XM}b}{\sqrt{a^Ta \cdot b^Tb}} \times \frac{b^T\hat{\Sigma}_{MY}c}{\sqrt{b^T\hat{\Sigma}_M b \cdot c^Tc}} \tag{2.9}$$

$$\begin{aligned} \underset{a\in\mathbb{R}^k, b\in\mathbb{R}^l, c\in\mathbb{R}^m}{\text{maximize}} \quad & a^T\hat{\Sigma}_{XM}b \times b^T\hat{\Sigma}_{MY}c \\ \text{subject to} \quad & \|a\|_2 \leq 1, \|b\| \leq 1, \|\Sigma_M^{T/2}b\| \leq 1, \|c\| \leq 1 \end{aligned} \tag{2.10}$$

Note that we assume that the transformations given in Equation 2.5 have been performed.

Not surprisingly due to the high order polynomial nature of the objective function, to date we have not found a simple and exact method for optimizing it. Figure 2.2 shows by example that the objective function is not convex. We propose several principled optimization methods, based on heuristics and relaxations, that are known to successfully optimize non-convex functions. Through simulations in Section 2.5, we show that these methods are capable of uncovering the correct structure under certain data-generating models. In the presentation below we begin with the more heuristic optimizers and then discuss several algorithms that rely on relaxations in order to approximately optimize 2.6 or 2.10.

Optimization Algorithm 1: Variation optimized mediators

As a first cut, it is reasonable to ask whether the directions of maximal variation in X, M, and Y also contain mediation structure. This heuristic optimizer uses the dominant loading vectors of the PCA transformations of X, M and Y to estimate β, η, and θ. They can also be used as starting values for other algorithms and provide a benchmark against which the optimal mediation directions are evaluated.

Optimization Algorithm 2: Correlation optimized mediators

This algorithm is based on the observation that the mediation objective function takes the form of two linked CCA problems. The algorithm decouples the two terms by introducing separate variables η_1 and η_2 to replace the shared variable η. The solutions to the separate CCA problems return the primary directions of covariation between X and M ($\hat{\beta}_{CCA}$ and $\hat{\eta}_1$) and between M and Y ($\hat{\eta}_2$ and $\hat{\theta}_{CCA}$) and we use these directions to estimate the leading

mediation directions. Specifically, our estimates of the mediation directions are:

$$\hat{\beta} = \hat{\beta}_{CCA}, \;\; \hat{\theta} = \hat{\theta}_{CCA}, \;\; \text{and} \;\; \hat{\eta} = \frac{t\hat{\eta}_1 + (1-t)\hat{\eta}_2}{\|t\hat{\eta}_1 + (1-t)\hat{\eta}_2\|_2}, \tag{2.11}$$

optimizing over all convex-combination of η_1 and η_2 to produce a common η that re-couples the two CCA problems.

Once again, we do not expect that these values will be optimal in a global setting. However, they may perform well statistically and may be used as starting values or as a basis of comparison for other estimates. In practice, this algorithm provides insight about the difficulty of the optimization problem. If $\hat{\beta}$, $\hat{\eta}$ and $\hat{\theta}$ perform well from an optimization perspective, it tells us that the two CCA problems share similar structure.

Optimization Algorithm 3: Greedy local descent

This algorithm, or class of algorithms, uses greedy local descent to optimize the additive inverse of the mediation directions objective. We expect to find that the performance of this class of algorithms largely depends on finding a quality starting value since the objective function is non-convex.

One advantage of this class of algorithms is that it is quite flexible, since they at most require finding an expression for the gradient of the objective function. Optimization strategies that can be utilized in this setting include:

- Optimize either the quotient or constrained versions of τ.
- Minimize $-\log(\tau(a,b,c))$ instead of $-\tau(a,b,c)$.
- Use log-barrier functions to enforce the norm constraints or relax the constraints and re-project to feasible region after each iteration.
- Use gradient-based or coordinate-wise, gradient-free methods.

In practice, almost any set of choices can be used together to produce slightly different algorithms. Although our function is not convex, it is relatively easy to differentiate with

respect to all of the optimization variables, regardless of specification.

We now describe one local descent algorithm that we have found to work well in practice. This algorithm will also be the method used in simulations in Section 2.5. The algorithm optimizes the log-transformed quotient formulation 2.9 using local gradient descent. Our optimization problem then takes the form:

$$\begin{aligned}
\underset{a\in\mathbb{R}^k, b\in\mathbb{R}^\ell, c\in\mathbb{R}^m}{\text{minimize}} \; L(a,b,c) &:= \log\left(\tau(a,b,c)\right) \\
&= \log\left(\frac{a^T\hat{\Sigma}_{XM}b}{\sqrt{a^Ta\cdot b^Tb}} \times \frac{b^T\hat{\Sigma}_{MY}c}{\sqrt{b^T\hat{\Sigma}_Mb\cdot c^Tc}}\right) \\
&= \log\left(a^T\hat{\Sigma}_{XM}b\right) + \log\left(b^T\hat{\Sigma}_{MY}c\right) - \\
&\quad \frac{1}{2}\left\{\log(a^Ta) + \log(b^Tb) + \log(b^T\hat{\Sigma}_Mb) + \log(c^Tc)\right\}
\end{aligned} \tag{2.12}$$

The gradients of $L(a,b,c)$ are:

$$\begin{aligned}
\nabla_a L &= -\frac{\hat{\Sigma}_{XM}b}{a^T\hat{\Sigma}_{XM}b} + \frac{a}{a^Ta} \\
\nabla_b L &= -\frac{\hat{\Sigma}_{XM}^Ta}{a^T\hat{\Sigma}_{XM}b} + \frac{\hat{\Sigma}_{MY}c}{b^T\hat{\Sigma}_{MY}c} + \frac{b}{b^Tb} + \frac{\hat{\Sigma}_Mb}{b^T\hat{\Sigma}_Mb} \\
\nabla_c L &= -\frac{\hat{\Sigma}_{MY}^Tb}{b^T\hat{\Sigma}_{MY}c} + \frac{c}{c^Tc}.
\end{aligned} \tag{2.13}$$

Rather than use a coordinate-wise descent algorithm, we concatenate the three search directions together and perform a backtracking line search over all three optimization variables at once. Our search direction is $\Delta = [\nabla_a L_t, \nabla_b L_t, \nabla_c L_t]^T \in \mathbb{R}^{k+\ell+m}$.

Given a feasible initialization, we have found that the algorithm quickly converges to a local optimum (See Figure 2.3). The algorithm continues until the ℓ_2-norm of the search direction Δ falls below a preset tolerance ϵ. When the problem dimension grows, we have found that it is more difficult to find solutions with very small gradient norms, and ϵ must be somewhat larger (on the order of $1e^{-5}$). Using accelerated methods also speeds convergence when the problem dimension grows.

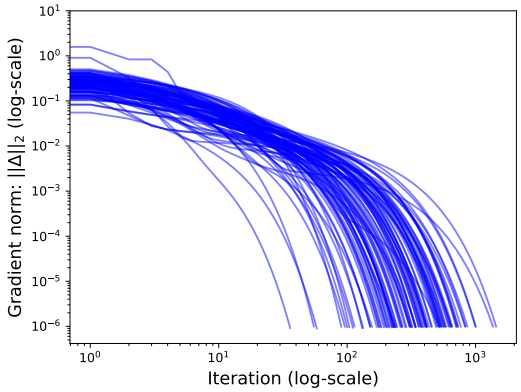


Figure 2.3: The ℓ_2-norm of the mediation directions objective function gradient ($\|\Delta\|_2$) plotted against iteration number for 100 unique starting values.

Optimization Algorithm 4: Relaxation via decoupling and parameter expansion

This algorithm decouples the two CCA problems by introducing a fourth optimization variable d in place of b in the second term of the quotient objective 2.6. However, we want to find a single projection of the mediators that is highly correlated with projections of both X and Y, so we penalize the ℓ_2 norm of the difference between b and d. The relaxed optimization problem is then given by:

$$\underset{a\in\mathbb{R}^k;b,d\in\mathbb{R}^\ell;c\in\mathbb{R}^m}{\text{maximize}} \; \tau(a,b,c,d) + \lambda\|b-d\|_2, \tag{2.14}$$

where $\lambda \geq 0$ controls the degree of relaxation. The objective function τ could take either quotient or constrained form.

We now apply this relaxation to the prototype algorithm described in Algorithm 3. The objective function from Algorithm 3 is augmented to include an additional optimization variable d and an ℓ_2 penalty on the difference between b and d.

$$\begin{aligned}\underset{a\in\mathbb{R}^k;b,d\in\mathbb{R}^\ell;c\in\mathbb{R}^m}{\text{minimize}} \; L_\lambda(a,b,c,d) = \; & -\log\left(a^T\hat{\Sigma}_{XM}b\cdot d^T\hat{\Sigma}_{MY}c\right) + \\ & \lambda\|b-d\|_2^2 - \frac{1}{2}\left(\log(a^Ta)+\log(b^Tb)+\log(b^T\hat{\Sigma}_Mb)+\log(c^Tc)\right). \end{aligned} \tag{2.15}$$

Once again, a local gradient descent method could be used to optimize 2.15. Simple changes to the gradient equations provided for Algorithm 3 produce gradient updates for 2.15 for fixed λ.

In practice, the primary difficulty of using this algorithm is determining how one should increase the penalty λ so that at termination b and d converge. Ideally, the solutions to each sub-problem are "smooth" as λ grows. One must decide how to increase λ so that the algorithm is both efficient and recovers the desired mediation structure.

Optimization Algorithm 5: Relaxation via matrix reparameterization

This algorithm uses the circulant property of the trace to reparameterize the objective function, and then strongly relaxes in the new parameterization. The relaxed problem has much higher dimension, but admits an exact solution using classical numerical methods. After finding the global solution to the relaxed problem, we project it back to the original search space. This algorithm borrows ideas from [21], who show that globally solving a heavily relaxed convex problem can outperform local optimization of the original non-convex problem.

Considering only the numerator of the quotient objective 2.6, we have

$$a^T\hat{\Sigma}_{XM}b \cdot b^T\hat{\Sigma}_{MY}c = \text{trace}\left(a^T\hat{\Sigma}_{XM}b \cdot b^T\hat{\Sigma}_{MY}c\right) = \text{trace}\left(\hat{\Sigma}_{XM}bb^T\hat{\Sigma}_{MY}ca^T\right). \quad (2.16)$$

Next we both relax and reparameterize the problem by introducing matrix-variate optimization variables $B \in \mathbb{R}^{\ell\times\ell}$ and $A \in \mathbb{R}^{m\times k}$ for bb^T and ca^T, respectively. The relaxation arises by allowing B and A to take on arbitrary values, rather than being constrained to rank one matrices. The relaxation quadratically increases the dimension of the parameters. A key question is whether the benefit of obtaining a relaxed problem with an exact solution is undone by the subsequent projection back to the feasible domain. Our objective function then becomes

$$\underset{B\in\mathbb{R}^{\ell\times\ell},\ A\in\mathbb{R}^{m\times k}}{\text{maximize}} \text{trace}\left(\hat{\Sigma}_{XM}B\hat{\Sigma}_{MY}A\right) \text{ s.t. } \|B\|_F \leq 1,\ \|A\|_F \leq 1. \tag{2.17}$$

Note that the overall dimension has increased from $\ell + k + m$ to $\ell^2 + km$, although this can be somewhat lessened by exploiting the symmetry constraint on B. We will see that the structure of 2.17 allows the norm constraints to be easily accommodated.

In order to show that the objective 2.17 does indeed have a globally optimal solution, we show that

$$\text{trace}\left(\hat{\Sigma}_{XM}B\hat{\Sigma}_{MY}A\right) = \text{vec}(B)\left(\hat{\Sigma}_{XM}^T \otimes \hat{\Sigma}_{MY}\right)\text{vec}(A). \tag{2.18}$$

For a cleaner presentation, let $\Sigma = \hat{\Sigma}_{XM}$ and $\Lambda = \hat{\Sigma}_{MY}$. For a general matrix P, denote the i^{th} row and j^{th} columns by $P_{i:}$ and $P_{:j}$.

Our first step is to find an expression for the diagonal elements of $\Sigma B\Lambda A \in \mathbb{R}^{k\times k}$. Using basic properties of matrix multiplication, we can write

$$[\Sigma B\Lambda A]_{ii} = \Sigma_{i:}^T B\Lambda A_{:i} = \Sigma_{i:}^T \begin{bmatrix} — & B_{1:}^T\Lambda & — \\ & \vdots & \\ — & B_{\ell:}^T\Lambda & — \end{bmatrix} A_{:i} = \sum_{j=1}^{\ell} \Sigma_{ij} b_j^T \Lambda A_{:i}. \tag{2.19}$$

The trace is the sum of all of the diagonal elements, which gives us:

$$\text{trace}(\Sigma B\Lambda A) = \sum_{i=1}^{k}\sum_{j=1}^{\ell} \Sigma_{ij} B_{j:}^T \Lambda A_{:i} = \sum_{j=1}^{\ell} B_{j:}^T \sum_{i=1}^{k} \Sigma_{ij}\Lambda A_{:i}. \tag{2.20}$$

We now introduce vectorized versions of A and B. Let $\text{vec}(A) = [a_1, a_2, \ldots, a_k]^T \in \mathbb{R}^{mk}$ and $\text{vec}(B) = [b_1^T, b_2^T, \ldots, b_\ell^T] \in \mathbb{R}^{\ell^2}$. We can then write

$$\sum_{i=1}^{k} \Sigma_{ij}\Lambda A_{:i} = \begin{bmatrix} \Sigma_{1j}\Lambda & \Sigma_{2j}\Lambda & \cdots & \Sigma_{kj}\Lambda \end{bmatrix} \text{vec}(A) = \tilde{\Lambda}_j \text{vec}(A). \tag{2.21}$$

Summing over all $j = 1 \ldots, \ell$ and we have:

$$\sum_{j=1}^{\ell} B_{j:}^T \tilde{\Lambda}_j \text{vec}(A) = \text{vec}(B) \begin{bmatrix} \tilde{\Lambda}_1 \\ \vdots \\ \tilde{\Lambda}_k \end{bmatrix} \text{vec}(A) = \text{vec}(B)\tilde{\Lambda}\text{vec}(A). \tag{2.22}$$

The ij^{th} block of $\tilde{\Lambda}$ is $[\tilde{\Lambda}]_{ij} = \sigma_{ji}\Lambda$, so $\tilde{\Lambda} = \Sigma^T \otimes \Lambda$.

Thus, the relaxed problem 2.17 is equivalent to maximizing a bilinear form, i.e.

$$\underset{B\in\mathbb{R}^{l\times l}, A\in\mathbb{R}^{m\times l}}{\text{maximize}} \ \text{trace}\left(\Sigma B \Lambda A\right) \ \text{s.t.} \ \|B\|_F \leq 1, \ \|A\|_F \leq 1. \equiv$$

$$\underset{B\in\mathbb{R}^{l\times l}, A\in\mathbb{R}^{m\times l}}{\text{maximize}} \ \text{vec}(B)\left(\Sigma^T \otimes \Lambda\right)\text{vec}(A) \ \text{s.t.} \ \|B\|_F \leq 1, \ \|A\|_F \leq 1, \tag{2.23}$$

which can be classically solved through the singular value decomposition (SVD) of $\Sigma^T \otimes \Lambda$. As a minor enhancement, we may exploit the underlying symmetry of B by introducing a matrix $\tilde{M}$ which sums the (i, j) and (j, i) elements of vec(B). Thus, our relaxed and reparameterized objective is:

$$\underset{B\in\mathbb{R}^{l\times l}, A\in\mathbb{R}^{m\times l}}{\text{maximize}} \quad \text{vec}(B)\tilde{M}(\Sigma^T \otimes \Lambda)\text{vec}(A) \ \text{s.t.} \ \|B\|_F \leq 1, \ \|A\|_F \leq 1.. \tag{2.24}$$

To obtain estimates of the mediation directions, we must project the solution of 2.24 so that the estimates of A and B are rank one matrices and B is symmetric. It then becomes possible to extract variable estimates for our original problem.

Our current algorithm first utilizes the leading left and right singular vectors u_1 and v_1 of $\tilde{M}(\Sigma^T \otimes \Lambda)$ and appropriately reshapes each to form the matrices B and A, respectively. Next, to re-project, we calculate SVD's of $B = U_B D_B V_B^T$ and $A = U_A D_A V_A^T$ and finally

estimate the mediation directions using the vectors:

$$b = u_{B,1}, \;\; a = v_{A,1}, \;\; \text{and} \;\; c = u_{A,1}. \tag{2.25}$$

In future work we will consider a related approach that augments the objective 2.24 by penalizing the nuclear norms of A and B. The nuclear norm penalties may lead to matrices A and B that better reflect the implicit rank one property of each optimization variable. This may improve the statistical performance of our estimators. However, the inclusion of the nuclear norm penalties means that 2.24 is no longer solvable using classical methods.

Discussion of the proposed algorithms

After considering many algorithms we found that the local gradient descent algorithm that minimizes the log-quotient problem 2.9 works well in practice. Although gradient descent converges after many iterations, the computational expense of function and gradient evaluations is low.

Since the objective function is scale invariant in β, η, and θ, the constrained form of the problem is less attractive than the quotient formulation. Each log-barrier optimization is roughly as expensive as solving the quotient formulation of the problem. Solving a sequence of these problems is clearly much more computationally expensive. Correctly tuning the log-barrier problems also proved to be challenging.

Finally, relaxation-based algorithms proved to be either computationally more expensive (Algorithm 4) or inconsistent (Algorithm 5). The matrix-reparameterization algorithm is an analytically neat reformulation, but we found that it was inconsistent in η. Perhaps by using nuclear norm penalties, discussed at the end of its description, one could produce an algorithm that consistently identifies the best projection of M as n grows. However, we have focused on problems of moderate dimension and the local gradient descent algorithm works well for these problems.

2.3.3 Finding additional mediation directions

The algorithms discussed above estimate the dominant mediation directions β_1, η_1 and θ_1. In practice, one may be interested in finding additional mediation directions. In this section, we suggest two sequential methods for finding additional mediation directions that are orthogonal to the existing mediation directions.

Residualization

The first method is the more straightforward of the two and similar to the preferred method of [16]. Suppose that β_1, η_1 and θ_1 are the first mediation directions. To find the secondary mediation directions, we project X, M, and Y onto the orthogonal complement of β_1, η_1, and θ_1, and then reapply our optimization algorithm to the residualized matrices. In general, if we want to find d mediating directions, we complete the following procedure.

Algorithm 2.1: Estimate multiple mediation directions via residualization

Data: Matrices X, M, and Y

Result: d triples of mediation vectors $\{\beta_1, \eta_1, \theta_1\}, \{\beta_2, \eta_2, \theta_2\}, \ldots, \{\beta_d, \eta_d, \theta_d\}$

1. Let $\beta_0 = 0$, $\eta_0 = 0$, and $\theta_0 = 0$.

for $j = 1, \ldots, d$ **do**

2.1 Form matrices $B_j = \begin{bmatrix}\beta_0 & \cdots & \beta_{j-1}\end{bmatrix}$, $H_j = \begin{bmatrix}\eta_0 & \cdots & \eta_{j-1}\end{bmatrix}$, and $\Theta_j = \begin{bmatrix}\theta_0 & \cdots & \theta_{j-1}\end{bmatrix}$

2.2 Create $\tilde{X}_j = (I - B_j(B_j^T B_j)^{-1} B_j^T)X$, $\tilde{M}_j = (I - H_j(H_j^T H_j)^{-1} H_j^T)M$, and $\tilde{Y}_j = (I - \Theta_j(\Theta_j^{-1}\Theta_j)^T \Theta_j^T)Y$.

2.3 Using $\tilde{X}_j$, $\tilde{M}_j$, and $\tilde{Y}_j$, estimate the j^{th} mediation directions β_j, η_j and θ_j.

Equality constrained optimization

Our second method forces subsequent mediation directions to be orthogonal by augmenting problems 2.12 and 2.15 with an additional equality constraint. Suppose that we have found

d unit-length, orthogonal mediation directions. Let the rows of matrices $B \in \mathbb{R}^{d\times k}$, $H \in \mathbb{R}^{d\times l}$, and $\Theta \in \mathbb{R}^{d\times m}$ contain the first d mediation directions of X, M, and Y respectively.

Then to find the next mediation directions β_{d+1}, η_{d+1}, and θ_{d+1}, we solve either problem 2.12 or 2.15 subject to:

$$B\beta_{d+1} = 0, \;\; H\eta_{d+1} = 0, \text{ and } \Theta\theta_{d+1} = 0. \tag{2.26}$$

These additional constraints are easily accommodated. We can parameterize the space of all solutions orthogonal to our existing directions using the eigenvectors of the projection matrices onto the complement spaces of those vectors. Specifically, we can equate the two sets

$$\{\beta \;:\; B\beta = 0\} = \{z \;:\; F_B z = 0\} \tag{2.27}$$

where the columns of F_B contain all of the eigenvectors of $I - B(B^TB)^{-1}B^T$ whose corresponding eigenvalues equal 1. Note that $I - B(B^TB)^{-1}B^T = I - BB^T$ since B is an orthonormal matrix. This approach to identifying a sequence of mediation directions is found in Algorithm 2.2.

2.3.4 Assessing the product-of-correlations estimate

The in-sample estimate of the product-of-correlations likely overestimates the population product-of-correlations. We use a bootstrap-based approach to assess both the degree of overfitting and whether the estimated mediation direction generalizes to independent data. The method uses the out-of-bootstrap sample (often called the "out-of-bag" data) as a test set with which to evaluate the generalizability of the estimated mediation directions.

In order to clearly present the analysis, let the subscripts b and t denote bootstrapped and out-of-bag quantities, respectively. For example, X_b and M_t denote a bootstrap sample of

Algorithm 2.2: Estimate multiple mediation directions via constrained optimization

Data: Matrices X, M, and Y

Result: d triples of mediation vectors $\{\beta_1, \eta_1, \theta_1\}, \{\beta_2, \eta_2, \theta_2\}, \ldots, \{\beta_d, \eta_d, \theta_d\}$

1. Let $\beta_0 = 0$, $\eta_0 = 0$ and $\theta_0 = 0$.

for $j = 1, \ldots, d$ **do**

2.1 Form the matrices

$$B = \begin{bmatrix} — & \beta_0^T & — \\ & \vdots & \\ — & \beta_{j-1}^T & — \end{bmatrix}, \quad H = \begin{bmatrix} — & \eta_0^T & — \\ & \vdots & \\ — & \eta_{j-1}^T & — \end{bmatrix}, \text{ and } \Theta = \begin{bmatrix} — & \theta_0^T & — \\ & \vdots & \\ — & \theta_{j-1}^T & — \end{bmatrix}.$$

2.2 Create the matrices F_B, F_H, F_Θ which contain a basis for the subspaces spanned by the projection matrices $I - B_j B_j^T$, $I - H_j H_j^T$, $I - \Theta_j \Theta_j^T$ respectively.

2.3 To find the j^{th} mediation directions, solve the problem:

$$\underset{z_a \in \mathbb{R}^{k-j-1}, z_b \in \mathbb{R}^{l-j-1}, z_c \in \mathbb{R}^{m-j-1}}{\text{minimize}} \tau(F_a z_a, F_b z_b, F_c z_c) \tag{2.28}$$

using the desired algorithm.

2.4 Set $a_j = F_a z_a$, $b_j = F_b z_b$, and $c = F_c z_c$.

the exposures and M_t represents an out-of-bag sample of the mediators. Bootstrap datasets will be used for estimation of the mediation directions, while the out-of-bag data will be used to estimate the out-of-sample performance of $\hat{\beta}$, $\hat{\eta}$, and $\hat{\theta}$.

Algorithm 2.3: Produce an out-of-sample product-of-correlations estimate

Data: Matrices X, M, and Y. The number of bootstrap samples to draw n_b.

Result: A out-of-sample estimate $\hat{\tau}_{os}$ of the product-of-correlations.

1. Calculate the in-sample product-of-correlations $\tau_s = \underset{\beta \in \mathbb{R}^k, \eta \in \mathbb{R}^l, \theta \in \mathbb{R}^m}{\arg\max} \tau(\beta, \eta, \theta)$ using matrices X, M, and Y.

2. **for** $j = 1, \ldots, n_b$ **do**

   2.1 Take a sample with replacement of size n from $\{1, \ldots, n\}$ and denote it b_j. Define $t_j = \{1, \ldots, n\} \setminus b_j$.

   2.2 Estimate $\hat{\beta}_j$, $\hat{\eta}_j$ and $\hat{\theta}_j$ by optimizing 2.33 using data $Z_{b_j} = (X_{b_j}, M_{b_j}, Y_{b_j})$.

   2.3 Calculate $\hat{\tau}_{os}^j = \widehat{\text{Cor}}(\hat{\beta}_j^T X_{t_j}, \hat{\eta}_j^T M_{t_j}) \times \widehat{\text{Cor}}(\hat{\eta}^T M_{t_j}, \hat{\theta}_j^T Y_{t_j} | \hat{\beta}_j^T X_{t_j})$.

end

3. Use $\hat{\tau}_{os} = \{\hat{\tau}_{os}^j, \ldots, \hat{\tau}_{os}^{n_b}\}$ to estimate out-of-sample product-of-correlations.

The mean and standard deviation or a $1 - \alpha\%$ confidence interval can be computed from $\hat{\tau}_{os}$ and compared to τ_s to assess whether the method is finding generalizable mediation directions.

2.4 Multivariate mediation population models

This section introduce two models which will: a) allow us to more clearly define our conceptualization of multivariate mediation, b) clearly contrast our approach to studying mediation with traditional mediation analysis methods, and c) provide generative models for simulations studies.

2.4.1 A single-layer multivariate mediation model

We begin by describing a simple model where X, M, and Y each predominantly vary in a one-dimensional subspace of their domain and are contaminated with white noise.

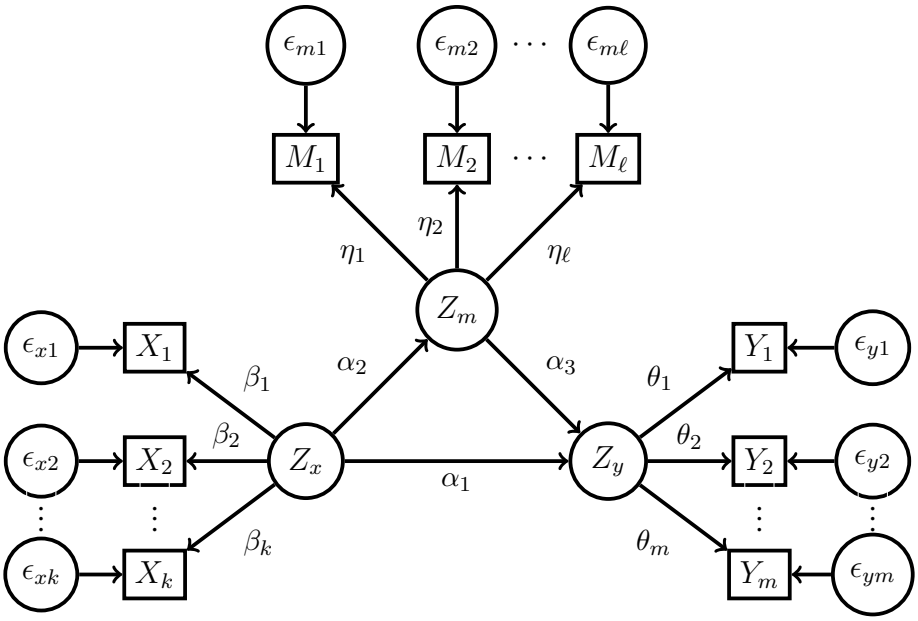


Figure 2.4: A graphical representation of the single-layer mediation model.

This implies that the mediation structure is contained in the projections of X, M, and Y onto their dominant axes. In this setting, we expect that we should be able to recover the underlying latent mediation directions (β, η, and θ) since the cross-covariance matrices will converge to rank one matrices as the number of observed samples n grows.

Figure 2.4 shows the proposed model in graphical form. Latent and observed random variables are represented with circular and rectangular nodes, respectively. In the center of the figure, the three latent variables Z_x, Z_m, and Z_y give rise the observations $X \in \mathbb{R}^k$, $M \in \mathbb{R}^\ell$, and $Y \in \mathbb{R}^m$. X, M, and Y are often called "indicators" of Z_x, Z_m, and Z_y respectively. We assume that a linear relationship exists between the latent variables and their indicators. For example, $X_j = \beta_j Z_x + \epsilon_{xj}$ for $j \in 1 \ldots, k$.

The parameters $\alpha = (\alpha_1, \alpha_2, \alpha_3)^T$ govern the associations between $Z = (Z_x, Z_m, Z_y)^T$ and therefore control whether mediation structure exists. We will describe the necessary conditions for the indirect effect to be equal to 0 at the population level after describing how one can generate data for this model. Model 1 gives a generative description of the graphical model pictured in Figure 2.4.

Model 1. *Let $\beta \in \mathbb{R}^k$, $\eta \in \mathbb{R}^l$ and $\theta \in \mathbb{R}^m$ be the fixed mediation directions. Let the random vector $Z = [Z_x, Z_m, Z_y]^T \in \mathbb{R}^3$ have mean 0 and covariance $\mathbb{E}(ZZ^T) = \Sigma$. We observe X, M, and Y where*

$$X = Z_x\beta + \epsilon_1,\ M = Z_m\eta + \epsilon_2,\ Y = Z_y\theta + \epsilon_3,$$

where ϵ_1, ϵ_2 and ϵ_3 also have mean 0 and covariance $\mathbb{E}\left(\epsilon_\nu \epsilon_\nu^T\right) = \sigma_\nu^2 I$ for $\nu = \{x, m, y\}$.

For this model, one might expect that the covariances between elements of Z control whether the population product-of-correlations is non-zero. However, due to the model's factor structure, this is not quite the case. For the remainder of this sections, we will assume that the additive errors ϵ_x, ϵ_m, and ϵ_y are all independent. Let

$$\text{Cov}(Z) = \begin{bmatrix} \tau_x^2 & \tau_{xm} & \tau_{xy} \\ \tau_{xm} & \tau_m^2 & \tau_{my} \\ \tau_{xy} & \tau_{my} & \tau_y^2 \end{bmatrix}. \tag{2.29}$$

Under the independence assumption, we have the following expressions for the cross-covariance terms:

$$\begin{aligned} \text{Cov}(X, M) &= \Sigma_{XM} = \tau_{xm}\beta\eta^T \\ \text{Cov}(M, Y|X) &= \Sigma_{MY} = \eta\left(\tau_{my} - \frac{\tau_{xm}\tau_{xy}}{\sigma_x^2}\left(1 - \frac{\tau_x^2}{\sigma_x^2 + \tau_x^2}\right)\right)\theta^T. \end{aligned} \tag{2.30}$$

The second equality is derived using standard conditioning formulas of the multivariate Gaussian distribution.

In Section 2.2, we discussed conditions under which the indirect mediation effect is 0. Condition 1, $\text{Cov}(X, M) = 0$, holds when $\tau_{xm} = 0$, meaning that Z_x and Z_m are independent of one another. Lastly, Condition 2, $\text{Cov}(M, Y|X) = 0$, holds when

$$\tau_{my} = \frac{\tau_{xm}\tau_{xy}}{\sigma_x^2}\left(1 - \frac{\tau_x^2}{\sigma_x^2 + \tau_x^2}\right). \tag{2.31}$$

The condition "M is independent of Y given X" is not sufficient for Equation 2.31. However, given $\tau_{xm}, \tau_{xy}, \sigma_x^2$, and τ_x^2, it does allow one to choose a τ_{my} such that $\text{Cov}(M, Y|X) = 0$. These conditions will play an important role in the simulation studies presented in Section 2.5.

2.4.2 A general multivariate mediation model with common cross-covariance bases

Although the single-layer mediation model is a useful tool for understanding our conception of multivariate mediation, data are likely to exhibit more complex structure, necessitating a richer class of models. We relax the single-layer structure of Model 1 by allowing the number of latent directions generating X, M, and Y to increase. However, under this model, the right singular vectors of $\mathbb{E}\left[XM^T\right]$ and the left singular vectors $\mathbb{E}\left[MY^T\right]$ are the same, up to a reordering of the vectors. A fully general model will allow the bases describing covariation of M with X and Y to be independent.

One can think of each "layer" of the general model as being represented by an independent single-layer model described in Model 1. Combining across layers produces the observed data (X, M, Y). A generative procedure for this model is described next.

Model 2. *Let $\beta^{(j)} \in \mathbb{R}^k$ for $j = 1, \ldots, p_1$, $\eta^{(j)} \in \mathbb{R}^l$ for $j = 1, \ldots, p_2$, and $\theta^{(j)} \in \mathbb{R}^m$ for $j = 1 \ldots, p_3$ be the fixed, mediation directions.*
We then generate the random vectors $Z^{(j)} = [Z_x^{(j)}, Z_m^{(j)}, Z_y^{(j)}]^T \in \mathbb{R}^3$ with $\mathbb{E}(Z) = 0$ and

$$\textit{Cov}(Z^{(j)}) = \Sigma_j$$

for $j = 1, \ldots, \max\{p_1, p_2, p_3\}$. We use the convention that when $j > p_k$ ($k = 1, 2, 3$), then

$Z_k^{(j)} = 0$ *(with a slight abuse of notation).*

Then the treatment, mediating and response variables X*,* M*, and* Y *are defined:*

$$X = \sum_{j=1}^{p_1} Z_x^{(j)} \beta^{(j)} + \epsilon_1^{(j)}, \;\; M = \sum_{j=1}^{p_2} Z_m^{(j)} \eta^{(j)} + \epsilon_2^{(j)}, \text{ and } Y = \sum_{j=1}^{p_3} Z_y^{(j)} \theta^{(j)} + \epsilon_3^{(j)},$$

where $\mathbb{E}(\epsilon_i^{(j)}) = 0$ *and Cov*$(\epsilon_i^{(j)}) = \Psi_i^{(j)}$.

Without loss of generality, assume for all $j < i$,

$$\text{diag}(\Sigma_i) \geq \text{diag}(\Sigma_j) \tag{2.32}$$

component-wise. This allows us to define dominant mediating directions, which will be useful when two or more mediating directions exist. If observed data are generated according to Model 2, then we would hope to not only estimate the leading mediation directions, but also the additional directions $(\beta_2, \eta_2, \theta_2), \ldots, (\beta_d, \eta_d, \theta_d)$. We will assess whether this is possible using Algorithm 2.2.

2.5 Simulation studies

2.5.1 Single layer consistency simulation studies

In this simulation study, we establish that the proposed method is capable of consistently identifying mediation structure when the data-generating model belongs to the class described in Section 2.4.2. We show that the method is capable of identifying mediation structure at lower levels or layers when the primary directions of variation in X, M, and Y do not contain mediation structure.

To this end, we consider six different data-generating populations with three layers ($d = 3$). Graphical models for the latent variables are displayed in Figure 2.5 for each of

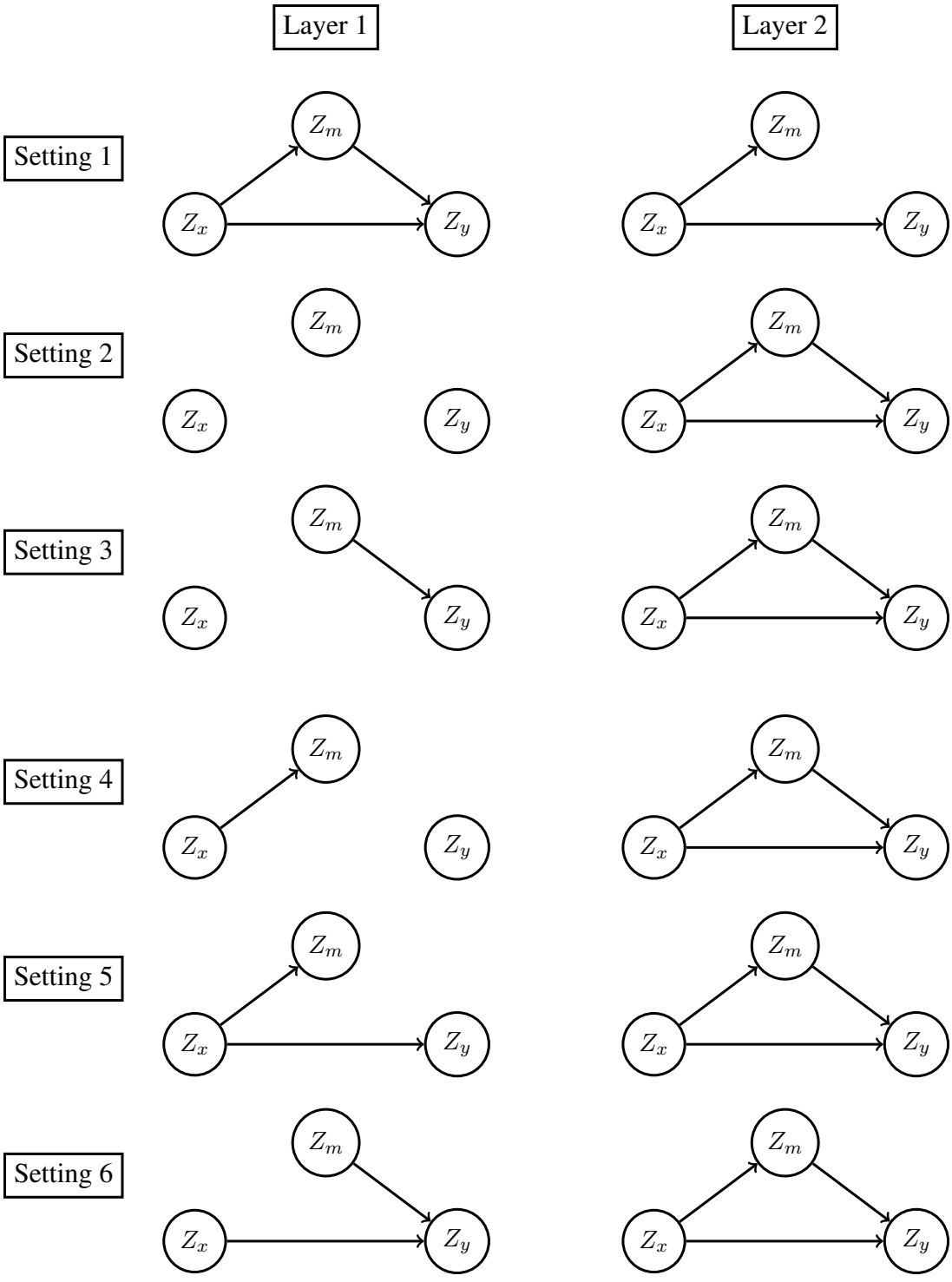


Figure 2.5: This figure shows graphical models of the relationship between latent variables at the first two levels of a general mediation model. A directed edge is present when the downstream variable depends on the upstream variable. To have a non-zero indirect effect, an edge must exist between both Z_m and Z_x and Z_m and Z_y within a single layer. In Setting 1, the mediation structure exists at Layer 1. For all other settings, the dominant mediation structure exists at Layer 2.

the six settings. We omit the graphical model for the third layer, as it is constant between settings and does not contain mediation structure. As described in Section 2.4.2, the layers are ordered so that the marginal variance of $Z^j \in \mathbb{R}^3$ are greater than the marginal variance of $Z^{j'} \in \mathbb{R}^3$ whenever $j < j'$. We set $\text{diag}(\text{Var}(Z^1)) = 25$, $\text{diag}(\text{Var}(Z^2)) = 9$, and $\text{diag}(\text{Var}(Z^3)) = 1$ for each simulation study.

In the first data-generating setting, the mediation structure exists in the first layer. For the remaining five data-generating populations, the mediation structure will exist at layer 2. The covariance structure of the layer 1 latent variables will change between these 5 data-generating populations, but will never contain mediation structure. Mediation structure exists at a given layer when an edge exists between Z_x and Z_m and between Z_m and Z_y.

The latent factors $\beta := \{\beta^1, \beta^2, \beta^3\}$, $\eta := \{\eta^1, \eta^2, \eta^3\}$, and $\theta := \{\theta^1, \theta^2, \theta^3\}$ are generated so that each has unit length and is orthogonal to all other vectors in its set. For example $(\eta^j)^T \eta^{j'} = 0$ whenever $j \neq j'$ and $(\eta^j)^T \eta^{j'} = 1$ when $j = j'$.

We choose to generate vectors in this manner as the orthogonality between layers causes the principle mediation directions to be vectors in the sets β, η and θ. When orthogonality does not hold, one can still find the population mediation directions by calculating the population variance-covariance matrix Σ of (X, M, Y). Using this matrix, one can use one of the algorithms described in Section 2.3 to find the population mediation directions. After identifying the population mediation directions, one again finds that the method consistently identifies these directions as n grows.

Finally, in order to assess the consistency of the method, we consider five problem dimensions and six different sample sizes. In this simulation, $|X| = |M| = |Y| = p$ with $p \in \{3, 5, 7, 10, 15\}$. The samples sizes considered are

$$n \in \{100, 250, 500, 5000, 10000, 100000\}.$$

For each data-generating population, problem dimension p, and sample size n, we con-

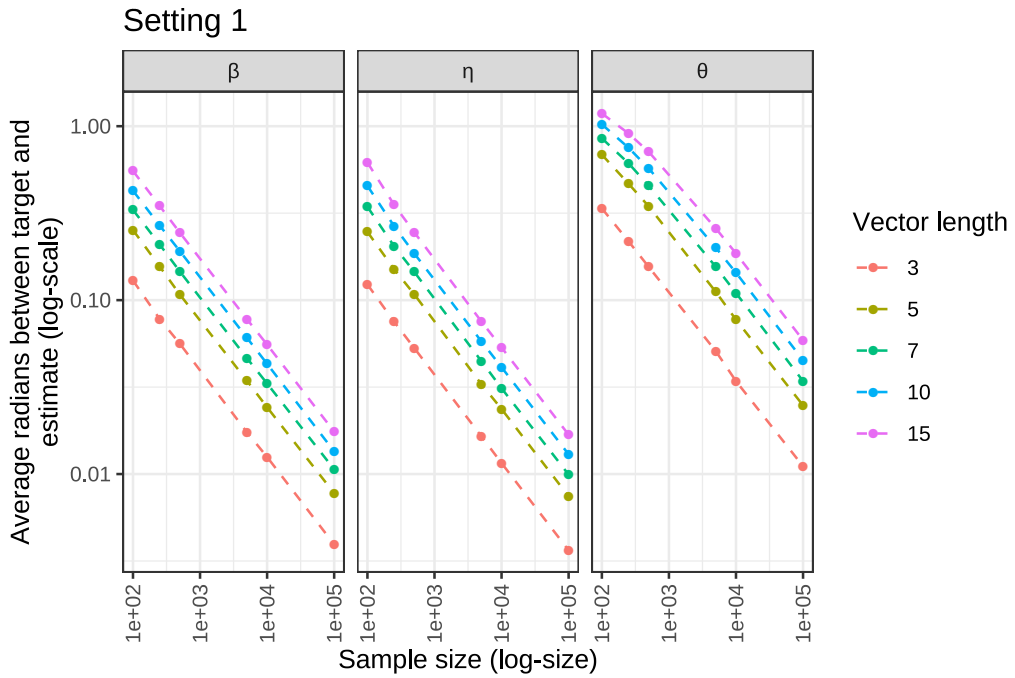


Figure 2.6: Single-layer simulation results for Setting 1

duct 1000 Monte Carlo trials. For each trial, we record the angle between the population mediation directions and the estimates.

Figure 2.6 (which consists of 6 separate plots) show results for all settings. The x-axes and y-axes represent the sample size n and average angle between our estimates and the population targets. Both are displayed on the log-scale. Panels within each figure display results for the vectors β, η, and θ. Finally, we use different colors to distinguish between different vector lengths.

The results of the simulation study are overwhelmingly positive. For all problem dimensions, our estimates converge to their targets as n grows. For each setting-vector-sample size combination, regressing the log error on the log sample size gives a slope estimate of roughly $-\frac{1}{2}$. This suggests that the algorithm has the expected root-n convergence rate. The convergence rates' constant terms (the intercept terms in the log-log regressions) depend on the data-generating population, vector β, η, and θ, and vector length.

This simulation study demonstrates that our proposed method is capable of identifying

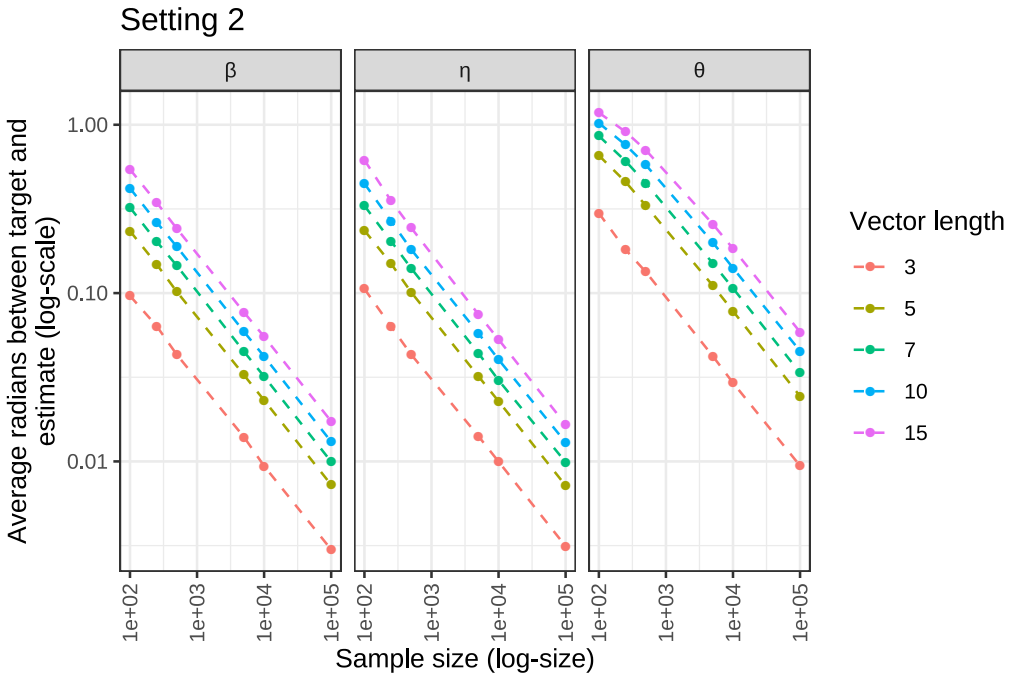


Figure 2.6: Single-layer simulation results for Setting 2

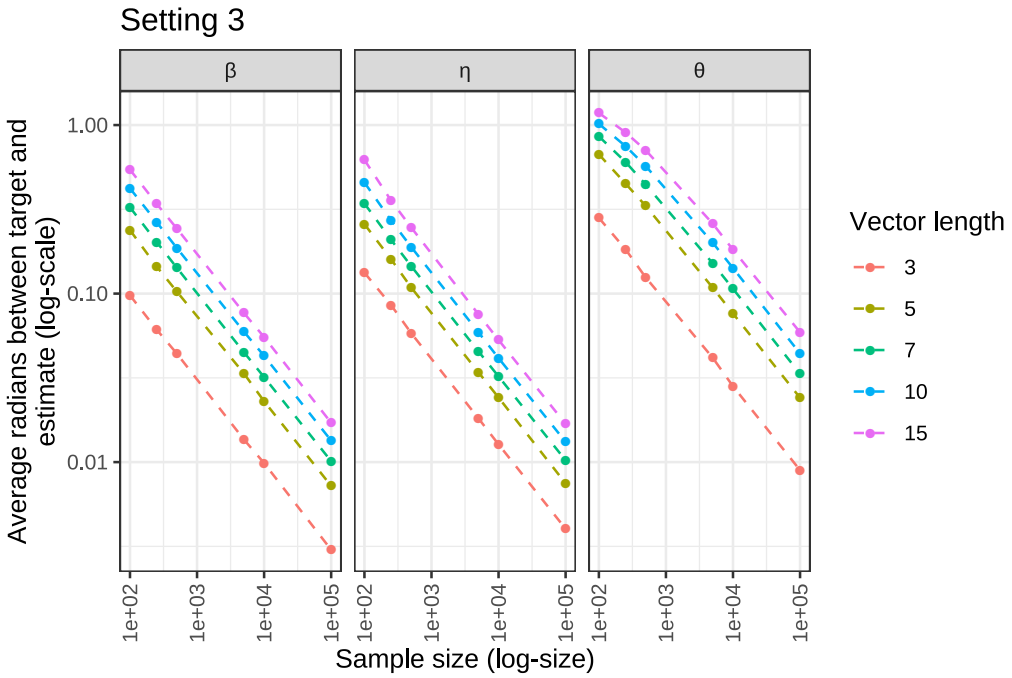


Figure 2.6: Single-layer simulation results for Setting 3

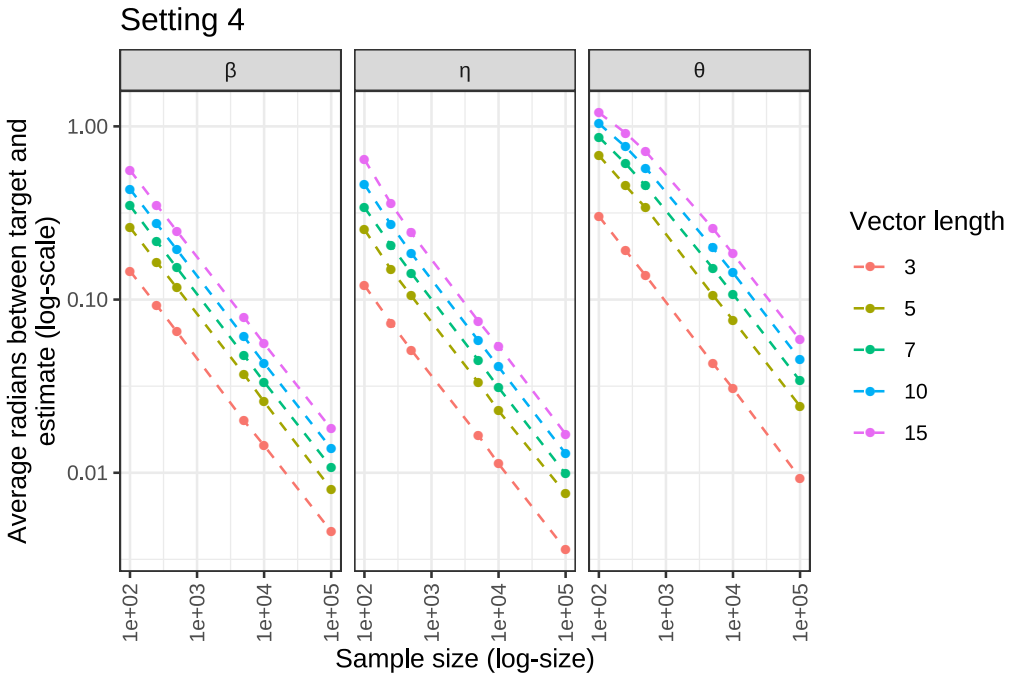


Figure 2.6: Single-layer simulation results for Setting 4

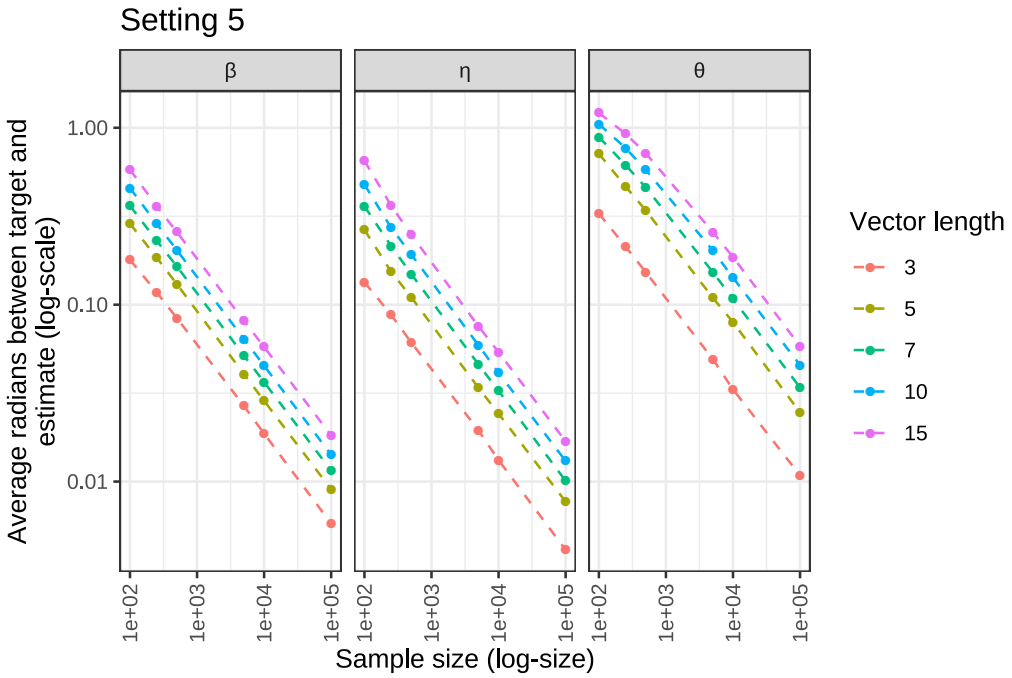


Figure 2.6: Single-layer simulation results for Setting 5

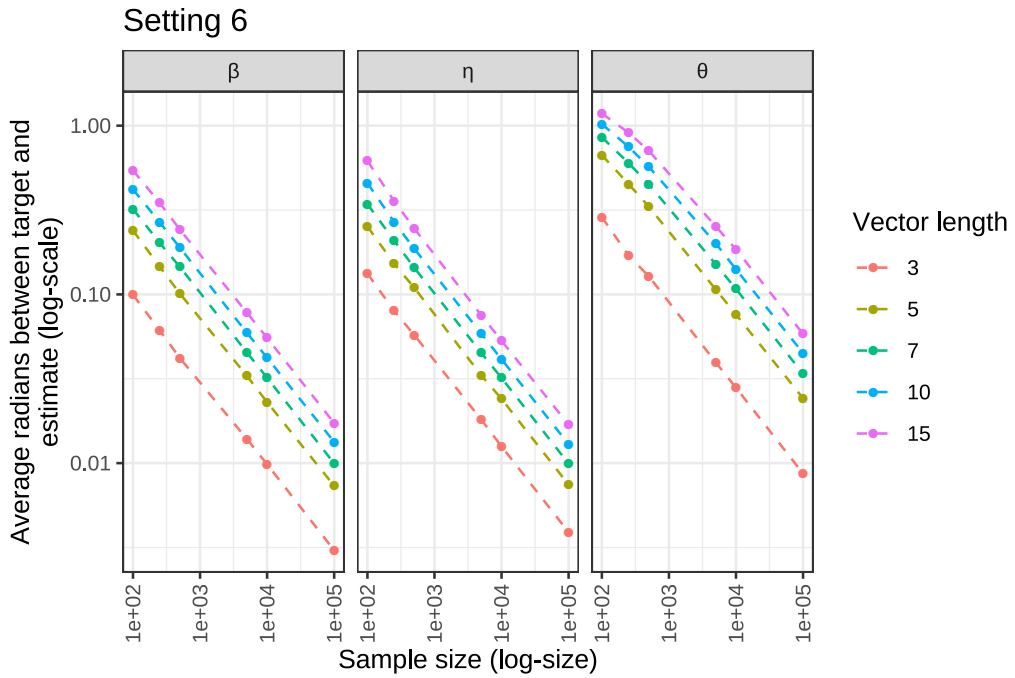


Figure 2.6: Single-layer simulation results for Setting 6

meaningful multivariate mediation structure, even when the mediation structure is obscured by higher variance, non-mediation structure. Our method has the expected statistical convergence rate and appears to have similar convergence rates between data-generating populations. Furthermore, the average error is not estimated to drastically differ between Setting 1 and Settings 2 through 5 after controlling for vector and problem size.

2.5.2 Multiple layer consistency simulation studies

Having established the proposed method's capacity to identify the population mediation structure, we now explore whether it can identify multiple mediation layers using the Algorithm 2.2. This simulation study uses a design similar to the first simulation study, and considers four different data-generating populations.

Each data-generating population contains four layers of factor structure for each variable X, M, and Y. Two of the four layers contain mediation structure among projections of X, M, and Y. At the non-mediation levels, we allow other association structures between

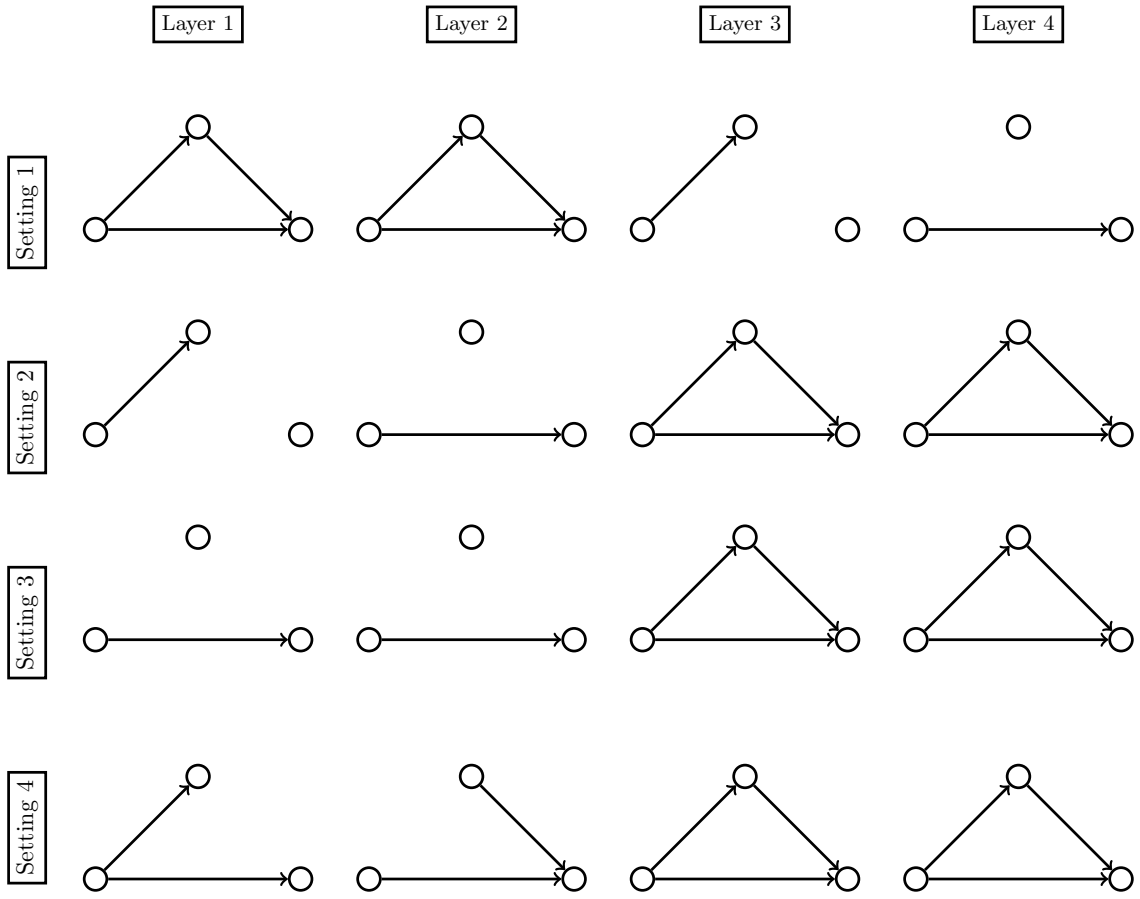


Figure 2.7: This figure shows graphical model of the latent variable associations for the second simulation study. Each data-generating population (Settings 1 through 4) has four layers of structure. Associations between variables are represented by arcs. Note that we omit the node labels in this figure in order to make the presentation cleaner. The orientation of the nodes is identical to past figures including Figures 2.4 and 2.5.

projections of X, M, and Y. In Setting 1, the mediation structure is contained in layers 1 and 2. For Settings 2 through 4, the mediation structure occurs at levels 3 and 4. The larger product-of-correlations always occurs at the higher-variance layer. For example, for Setting 1, the layer 1 product-of-correlations is larger than the layer 2 product-of-correlations. Graphical models of the latent variables for each data-generating population and layer are shown in Figure 2.7.

Factor generation for this simulation study is identical to the simulation study in Section 2.5.1. The marginal standard deviations of the latent variables are 3.0, 2.5, 2.0, and 1.0 for

layers 1, 2, 3, and 4. Again, the "first" layer is the one with the greatest marginal variance. For this simulation study, we consider three vector lengths: $p \in \{5, 10, 15\}$ and

$$n \in \{250, 500, 1000, 2500, 5000, 10000, 100000\}.$$

For each data-generating population, vector length and sample size, we generate 1000 synthetic data sets. For each data set, we estimate the first two mediation directions and record the angle between the estimates and their population targets. We expect that as n grows, the angle between an estimate and its target should decrease for both the first and second mediation layers.

Figure 2.8 (split into 4 separate plots) shows the convergence results for each of the four settings. We plot the angle between an estimate and its target on the y-axis and the sample size n on the x-axis (both on the log-scale). Each plot is split into three panels, which give results for vectors β, η and θ, respectively. Line color denotes the value of p, and line type denotes which the mediation layer (solid and dashed for the dominant and non-dominant layers, respectively).

Overall, the results of the simulation study are promising and suggest that the proposed method is able to correctly identify multiply layers of mediation structure. The first plot in Figure 2.5.2 gives representative results. Similar to results from Section 2.5.1, we have root-n convergence of estimates to their targets. The convergence rate constants depend on p, the vector, data-generating population and layer.

Convergence occurs more quickly when β, η, and θ when p is small. Estimates of θ appear to converge more slowly than estimates of β and η. Interestingly, estimates of the first-layer mediation directions (β_1, η_1, θ_1) converge more quickly than estimates of the second-layer mediation directions (β_2, η_2, θ_2) for each data-generating population. The second-layer mediation directions still have root-n convergence rates to their targets in each setting.

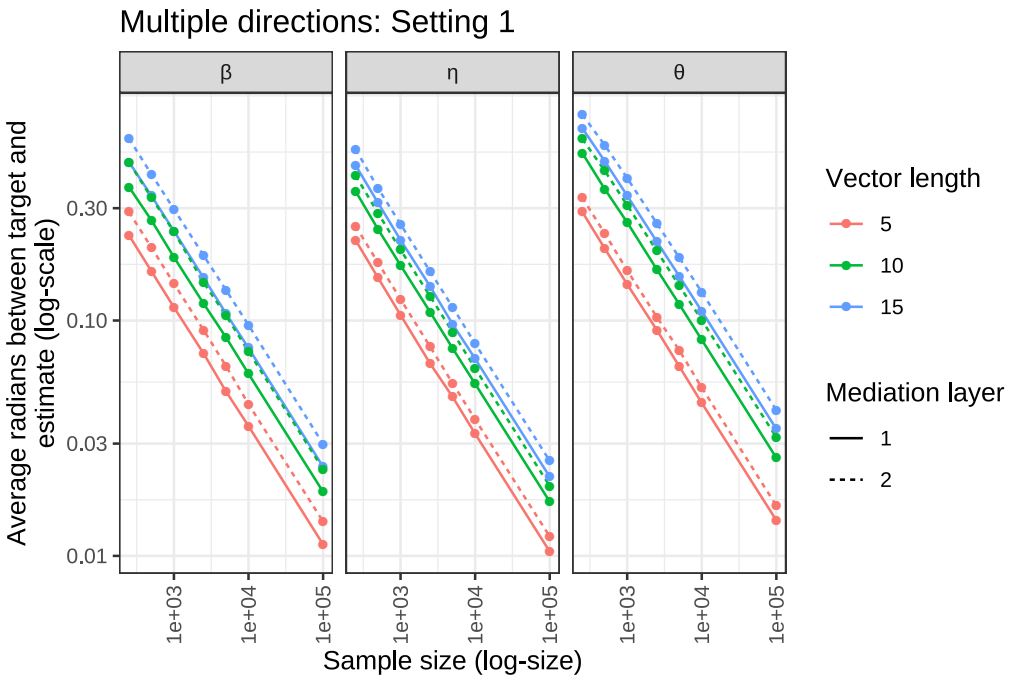


Figure 2.8: Multi-layer simulation results for Setting 1

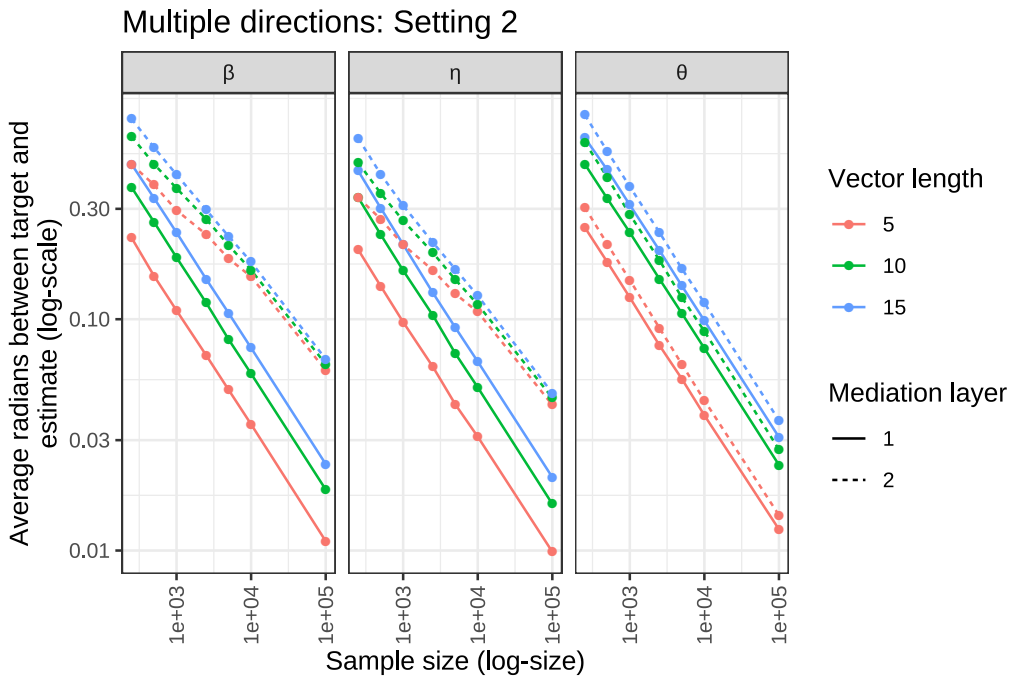


Figure 2.8: Multi-layer simulation results for Setting 2

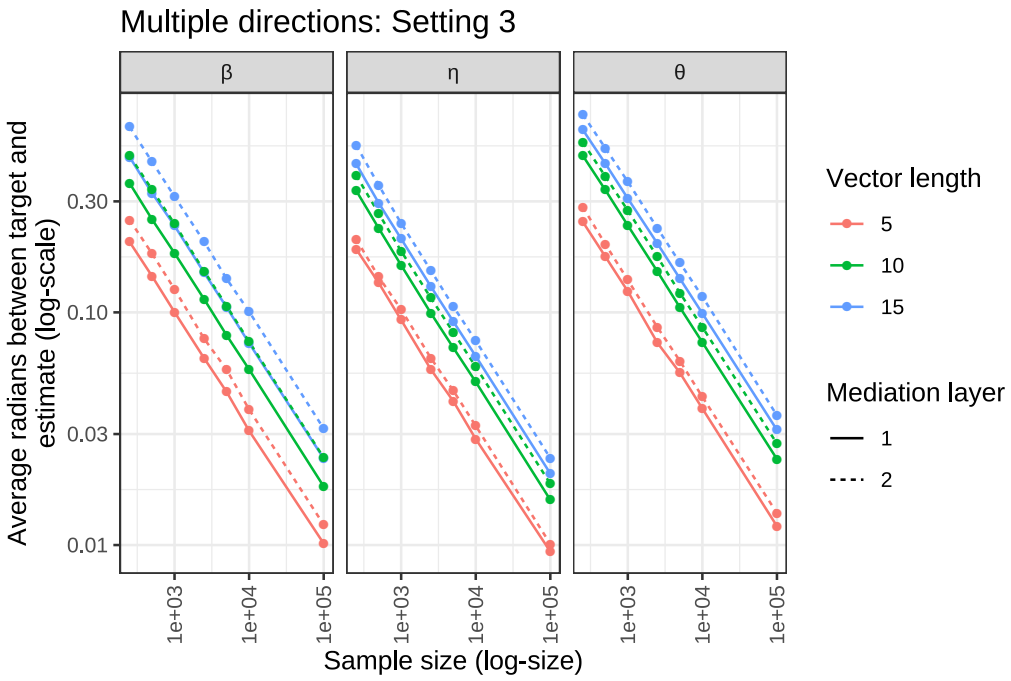


Figure 2.8: Multi-layer simulation results for Setting 3

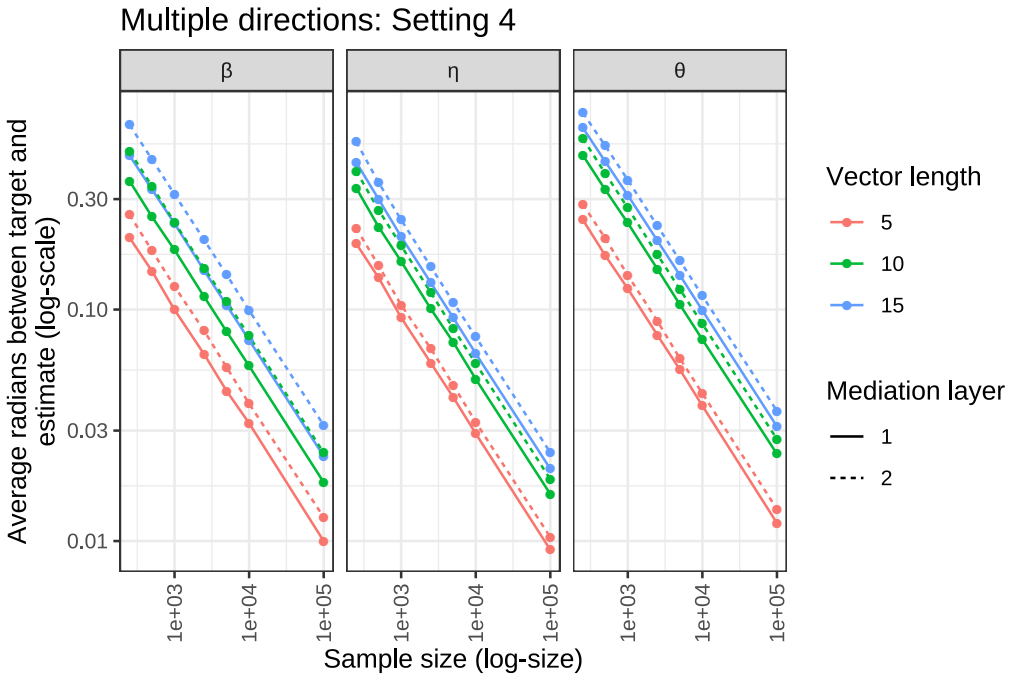


Figure 2.8: Multi-layer simulation results for Setting 4

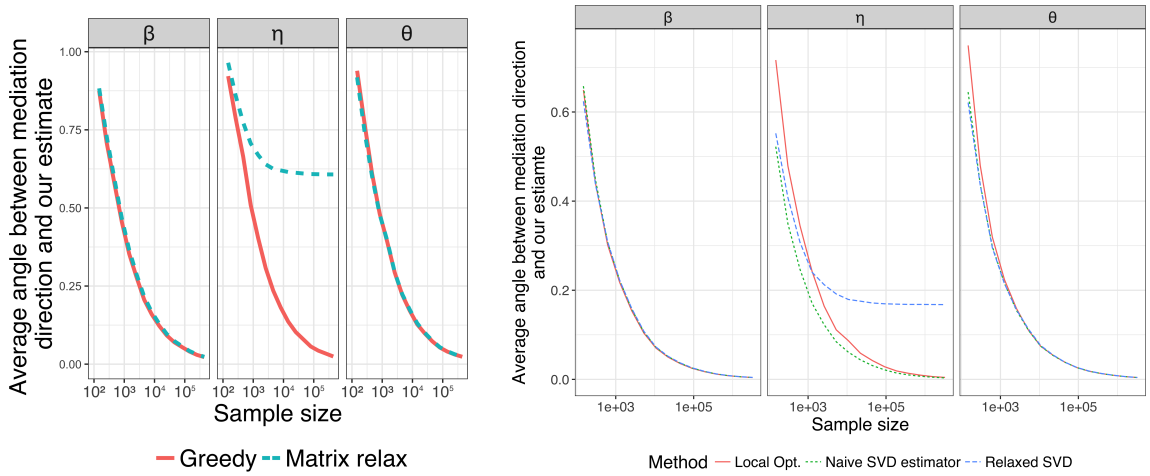
Simulation results for Setting 2 are different than results for the other data-generating populations. First layer estimates converge to their targets as expected, but estimates of β_2 and η_2 do not appear to have the expected root-n scaling. It is unclear why this occurs. One possible explanation is that in this setting we infrequently converged to a sub-optimal local optimum. To test this explanation, we re-ran the simulation with multiple initializations. However, using multiple initializations did not produces estimates with the expected root-n convergence. Strangely, the vector θ still has the expected root-n convergence.

At the time of writing, we have not identified a satisfactory explanation of the observed non-root-n convergence and we plan to continue studying this example. However, we do not believe that the result undermines the primary findings of the simulation study. First, the method is capable of identifying multiple layers of mediation structure, even when that structure is obscured by higher-variance factor structures without mediation information. Second, the precision of additional mediation direction estimates is lower for a fixed sample size n. Inexact estimation of the primary mediation directions leads subsequent estimates to converge more slowly.

2.5.3 A comparison of algorithms

We briefly present simulation study results that led us to use a greedy local method rather than Algorithm 5 (relaxation via matrix reparameterization algorithm) from Section 2.3. We will refer to this algorithm as the "matrix relaxation algorithm" for the remainder of this section. The data-generating populations were identical to Settings 1 and 2 from the single-layer simulations in Section 2.5.1. The mediation structure is the dominant factor structure in Setting 1. In Setting 2, the mediation structure occurs in the second layer and the factor structure in the first layer does not induce associations between X, M, and Y. Each simulation study uses 1000 Monte Carlo repetitions for each sample size. Results are found in Figure 2.9.

Results for Settings 1 and 2 are nearly identical. In both cases, we see that the conver-



(a) Simulation results for Setting 1

(b) Simulation results for Setting 2

Figure 2.9: A comparison of algorithms designed to solve the optimization problem 2.6. Both figures above plot the angle between the an estimated vector and its target against sample size (log-scale). Results are averaged over 1000 Monte Carlo trials. The left and right hand panels show results for Setting 1 and 2 respectively. Color and line type are used to differentiate between algorithms.

gence rates for the vectors β and η are nearly identical between the local greedy and matrix relaxation algorithms. However, the matrix relaxation algorithm is inconsistent for η for both data-generating populations. Its estimate converges to a vector that is not equal to the population mediation direction η.

We expect that by modifying Algorithm 5 to include nuclear norm penalties, one could create a matrix relaxation algorithm that is consistent for η in addition to β and θ. This variety of algorithm could prove to be very useful as the dimensions of X, M, and Y increase. However, the local greedy algorithm has proven to be sufficient for the moderate problem sizes that we have considered.

2.5.4 Estimating the population mediated effect

The purpose of this section and its simulation study is to better understand the how estimated mediation directions generalize to other samples from the same data-generating population. In moderate dimensional problems, we expect that our method likely over-

fits the observed data and overestimates the product-of-correlations. Most previous work on multivariate mediation analysis has attempted to perform inference on the product-of-coefficients using bootstrap methods.

More rigorous approaches to assessing the multivariate product-of-correlations assess either the out-of-sample generalizability of the estimated directions or the stability of the estimated vectors. The latter reduces to creating a confidence envelope on the p-dimensional unit sphere. We are uncertain as to how one should create such an envelope, so we instead offer one measure that assesses the stability of the estimated directions. Our approach appears at the end of Section 2.3.

At a high level, the approach uses a nonparametric bootstrapped data set to estimate the mediation directions $\hat{\beta}$, $\hat{\eta}$, and $\hat{\theta}$. The so-called "out-of-bag" observations, those observations which do not appear in the bootstrap dataset, are used to produce an out-of-sample estimate of the product-of-correlations. Let τ be the population product-of-correlations, and let $\hat{\tau}_s$, and $\hat{\tau}_{os}$ denote an in-sample and out-of-sample estimates of τ, respectively.

For this simulation study, we considered six different data-generating populations which are described in Table 2.1. They provide a mixture of different mediation structures and problem sizes. For each population, we consider six sample sizes

$$n \in \{100, 250, 500, 1000, 2000, 5000\}.$$

For a fixed population and sample size n, we generate 100 synthetic datasets. For each

Table 2.1: Data-generating populations for simulation study assessing overfitting

Setting	Problem size (k, ℓ, m))	Description of mediation structure
1	(10, 10, 10)	Dominant mediation structure in lower level
2	(10, 10, 10)	Dominant mediation structure in first level
3	(1, 10, 1)	Only mediation structure in first level
4	(1, 75, 1)	Only mediation structure in first level
5	(1, 10, 1)	No mediation structure
6	(1, 75, 1)	No mediation structure

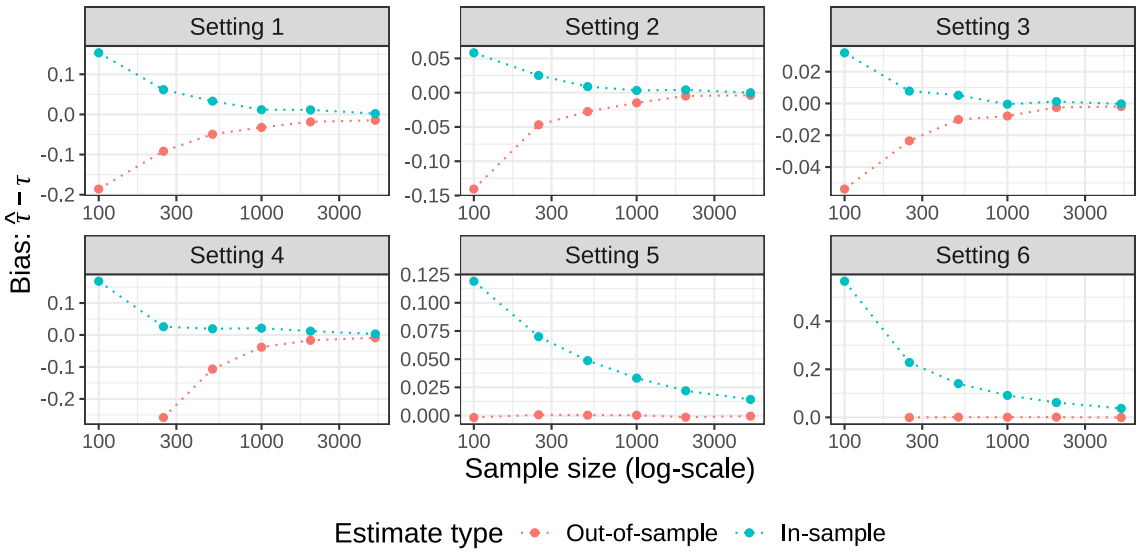


Figure 2.10: In-sample and out-of-sample bias of the product-of-correlations estimates plotted against sample size for the 6 simulation settings described in Table 2.1. Each panel plots the estimated bias against sample size using line color to distinguish between in-sample and out-of-sample estimates of the product-of-correlations.

synthetic data set, we estimate τ using both the in-sample estimate τ_s and the mean of $n_b = 250$ out-of-sample estimates: $\tau_{os} = \frac{1}{n_b}\sum_{i=1}^{n_b} \tau_{os}^i$. Using the 100 estimates of the in-sample and out-of-sample product-of-correlations, we calculated in the in-sample and out-of-sample bias

$$\hat{\tau}_s - \tau = \frac{1}{100}\sum_{j=1}^{100}\left(\hat{\tau}_s^j - \tau\right) \quad \text{and} \quad \hat{\tau}_{os} - \tau = \frac{1}{100}\sum_{j=1}^{100}\left(\hat{\tau}_{os}^j - \tau\right),$$

respectively. Figure 2.10 shows the estimated in-sample and out-of-sample bias for all simulation settings.

Across all simulation settings, as n increases both the in- and out-of-sample estimates converge to the true population value of τ. For all settings, when n is small, the out-of-sample estimates are upward biased, meaning that the method overestimates the true population product-of-correlations. In Settings 1 through 4, drastic overestimates of the product-of-correlations leads to mediation direction estimates that generalize poorly out-of-sample. In these regimes, the out-of-sample estimates $\hat{\tau}_{os}$ underestimate the population product-

of-correlations. The exception to this rule occurs when the true product-of-correlations is equal to 0. In Settings 5 and 6 where this is the case, $\hat{\tau}_{os} \approx 0$ for all n.

We draw two conclusions from this simulation study. First, when sample sizes are small, the method substantially overestimates the true product-of-correlations. This can occur even when the population product-of-correlations is equal to 0 (see Setting 6, sample size = 100 in Figure 2.10). Second, in regimes where the method drastically overestimates the product-of-correlations, the out-of-sample estimate $\hat{\tau}_{os}$ is a downward biased estimator of τ. This gives us a useful diagnostic procedure to assess whether we are dramatically overfitting the observed data. If the difference between in- and out-of-sample estimates of τ is large, then we're likely overfitting. Future work might address how to regularize the objective function so that estimators are more generalizable.

2.6 Case studies

2.6.1 Illustration via media perception study

We use data from a study called *Financial Crisis: A Longitudinal Study of Public Response* [22], which was funded by the National Science Foundation. The goal of the study was to understand how public perception of risk changed during and after the 2008 financial crisis. A panel of 800 respondents was given surveys at eight time points (September, 29 2008, October 8, 2008, November 5, 2008, December 6, 2008, March 21, 2009, June 30, 2009, October 6, 2009, and August 9, 2011). At least 600 panelists participated in each survey administration and 325 individuals completed all of the eight surveys. To illustrate our methodology, we consider only two of these time points (September and December 2008) and therefore do not make use of the longitudinal nature of the data.

The panelists were asked about a variety of topics, but for this illustration we will focus on the connections between media consumption, attitudes about the economy, and perception of the future. In the media studies literature, media consumption is thought to

influence perceived risk in addition to attitudes about the topics directly receiving coverage. In the fall of 2008, the media covered the economy and financial crisis closely, so it is reasonable to expect that people who consumed more media might have more negative attitudes about the economy and a perception of greater future economic risk. It is also possible that people who have more negative attitudes about the economy, regardless of the cause, will also have a greater perception of future economic risk.

Six media variables measured the amount of time a panelist spent consuming a certain type of media (e.g. T.V., radio, and internet). The financial attitudes variables each measure the degree to which the panelist felt an emotion toward the economy. Since the country was in the midst of the largest economic downturn since the Great Depression, each of the six attitude questions asked about a negative emotion. Finally, the risk perception variables tried to capture how a panelist felt about his or her economic future. For example, panelists were asked whether the crisis would limit their future or cause them to postpone making a large purchase. Again, we used six variables to capture their risk perception. The study surveys used Likert-type scales to collect responses about the variables described in the proceeding paragraph. Self explanatory variable names appear in Tables 2.2 and 2.3.

If one applied traditional scalar-valued mediation analysis techniques to these data, one would be forced to either (a) perform several scalar-valued mediation analyses on the variables of greatest interest, (b) somehow summarize the data into single measures of media consumption, economic attitudes and risk perceptions, or (c) use PCA or factor analysis to produce new variables. In any case, the data must be first processed or *ad hoc* choices must be made in order to perform mediation analysis. Applying our multivariate mediation method allows one to forgo making these choices and attempts to find the strongest possible mediating relationship among the variables.

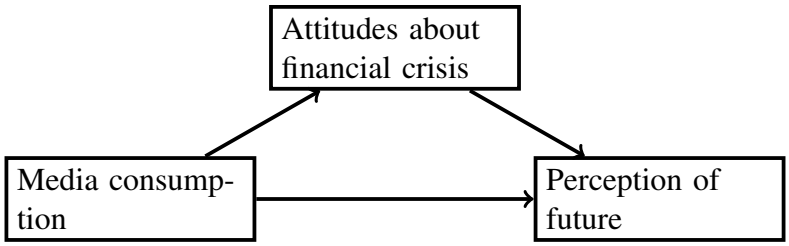
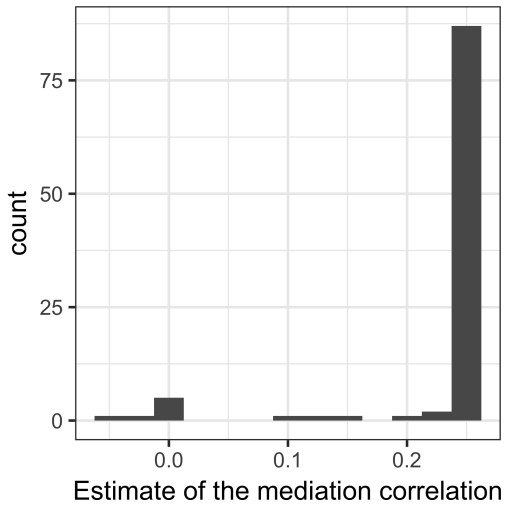
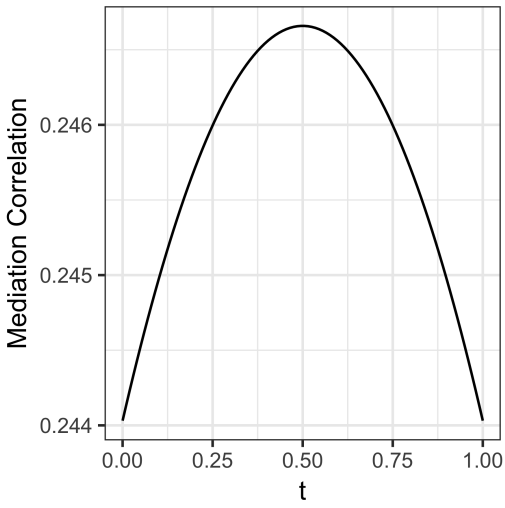


Figure 2.11: A path diagram representing a potential mediating relationship in a risk perception study.



(a) The local greedy optimizer converged to different local optima depending on starting values.



(b) Algorithm 2 achieves a similar mediation correlation. This suggests that the two CCA problems are closely linked.

Results

We use Algorithm 5 and Algorithm 3 to estimate the mediation directions, and obtained similar results from both estimators. We observe that the local optimizer occasionally converged to suboptimal local maximums for different initializations (see Figure 2.12a). Otherwise, both algorithms find similar optima and their estimates of the mediation correlation τ were approximately identical. In Tables 2.2 and 2.3 we report the estimates given by the matrix-variate relaxation (Algorithm 5).

Estimates appear to be fairly similar between time points. The large coefficients in each mediation direction have the same sign, suggesting a coherent effect. The variables with large factor loadings are listed below:

Table 2.2: September 2008 coefficients (N=739)

Media Consumption		Financial Attitudes		Perception of future	
Variable	**Coefficients**	**Variable**	**Coefficients**	**Variable**	**Coefficients**
Hrs Internet	0.17	Angry	-0.06	Adjust to Economy	0.27
Hrs Newspaper	0.08	Anxious	0.01	Change Investments	0.21
Hrs Radio	0.14	Fearful	-0.59	Limit Future	-0.87
Hrs Talk	-0.33	Sad	-0.39	Take no Action	0.15
Hrs TV	-0.80	Stressed	-0.62	Postpone Purchases	-0.31
Numb. of people talked to abt. crisis	-0.45	Worried	-0.34	Control Future	-0.01

Table 2.3: December 2008 Coefficientss (N=610)

Media Consumption		Financial Attitudes		Perception of future	
Variable	**Coefficients**	**Variable**	**Coefficients**	**Variable**	**Coefficients**
Hrs Internet	0.39	Angry	-0.13	Adjust to Economy	0.34
Hrs Newspaper	-0.17	Anxious	-0.13	Change Investments	0.34
Hrs Radio	0.17	Fearful	-0.03	Limit Future	-0.78
Hrs Talk	-0.15	Sad	-0.21	Take no Action	0.12
Hrs TV	-0.66	Stressed	-0.36	Postpone purchases	-0.38
Numb. of people talked to abt. crisis	-0.58	Worried	-0.89	Control future	-0.09

- Media: the number of hours spent discussing the crisis and watching TV, and number of people one talked to about the crisis,

- Attitudes: fearful, sad, stressed and worried, and

- Lifestyle: limit future opportunities, postpone large purchases.

This suggests that people who watched more TV and talked more to others, were slightly more likely to have negative attitudes about the economy and believed that the crisis would impact their future. Interestingly, this component appeared to be negatively associated with likelihood of changing investments and beliefs about ability to adjust to the crisis. The amount of time spent on the internet and reading the newspaper did not appear to contribute substantially to the mediation pathway. For September, our estimate of the mediation correlation τ was 0.25.

In the December analysis, the vector loadings were largely unchanged. The variable "internet" received a larger positive weight. This suggests that in 2008, people who used the internet might not have been doing so to monitor financial news. The number of hours spent talking to people appeared to have grown less important by December as well. Coefficients among the attitude variables have shifted to place a greater emphasis on "worry" and a lower weighting on "stress." The coefficients of variables measuring the respondents' attitudes about their future did not change to a great degree. In December, the estimate of the mediation correlation decreased to 0.17.

Estimate diagnostics

Even though it appears that we have found interpretable directions of mediation that describe how the variables covary, we additionally want to see whether these directions describe variation within subsets of variables. We apply a common diagnostic tool used in conjunction with CCA to verify that our mediation directions do describe substantial variation in X, M, and Y. When we compare the mediation directions β, η, and θ returned by

Algorithm 5, we find that they are 35%, 63%, and 74% as variable as the optimal projections of X, M, and Y found through PCA of X, M, and Y. In these cases, it's clear that the mediation directions that we uncovered are not the primary directions of variation. They do however describe non-trivial amounts of variation in the data. In Figure 2.12b we plot estimates of the mediation correlation for different convex combinations $\hat{\eta}_1$ and $\hat{\eta}_2$ (see Algorithm 2). The correlation-optimized mediators achieve a similar mediation correlation as the matrix-variate and greedy descent algorithms. The performance of Algorithm 2 was not due to the fact that the canonical directions were close to each other (1.2 radians apart), but because both canonical directions performed well alone (mediation correlation estimates of 0.24 for each vector). The principal directions of variation do not appear to describe mediation structure (Algorithm 1) with mediation correlation $\tau = 0.13$. For these data, the mediation structure appears to be very closely tied to the joint correlation structure, but not the within-variable covariance structure.

2.6.2 Mediation of genetic factors and cognitive ability via brain activity measured by neuroimaging

Our second real data example assesses whether brain activity patterns mediate the association between an individual's genes and their performance on a battery of cognitive tests. At a high level, one expects that the impact of an individual's genes on their cognitive ability is due to how one's genes' affect functional brain patterns. Here we use novel data and our proposed method to study the gene-brain activity-intelligence pathway. We aim to find clear patterns of brain activity that are both associated with an individual's genetics and their cognitive capabilities.

The data for this analysis were collected as part of the *Adolescent Brain Cognitive Development* (ABCD) study [23]. The ABCD Study is a long-term, longitudinal observational study which collects and makes available the data necessary to better understand the complex process of cognitive development during adolescence. Study participants were

recruited at ages 9 and 10. Recruited individuals will be followed into early adulthood with biannual follow-ups, which will produce a rich, complex dataset with which to study cognitive development. In addition to assessing an individual's cognitive ability, the study collects family and environmental variables, genetic data and medical measures, including functional magnetic resonance imaging (fMRI). This analysis makes uses of the genetic, fMRI and aptitude test data.

Data description

At baseline, the ABCD study protocol requires the collection of participant DNA, functional brain connectivity, and neurocognition data, in addition to many other measures.

Each participant's blood sample was used to perform full genome sequencing. Using the full genome sequence, the primary study team collapsed the genetic data into a univariate polygenic risk score (PGRS). The PGRS was trained to be predictive of cognitive ability [1]. The PGRS weights were estimated on independent data.

A brain imaging study assessed each ABCD study participant's structural and functional brain connectivity patterns. The task-based assessment used three tasks which assess a wide range of cognitive abilities, including impulse control, emotional regulation, and reward processing. The primary research team performed a factor analysis to reduce voxel-level task-based fMRI data to 75 brain features. These 75 brain features will be used in our mediation analysis.

Finally, participants took a battery of 11 cognitive tests meant to measure different aspects of neurocognition at baseline [24]. Again, the primary study team performed a factor analysis on the results of the 11 tests in order to produce a univariate summary of cognitive performance for each individual. This measure will be referred to as G, which stands for "general intelligence."

Analysis plan

Our analysis use the PGRS, 75 brain features, and cognition score G to assess whether patterns of brain connectivity during tasked-based tests mediate the PGRS - intelligence pathway. Our analysis assumes the following structure between variables:

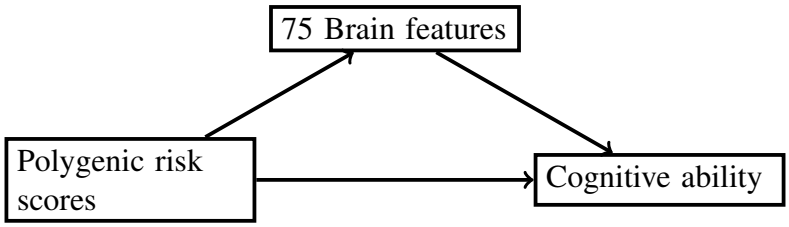


Figure 2.12: A path diagram representing a potential mediating relationship for the ABCD data analysis.

The analysis is limited to subjects with European ancestry, since the PGRS does not generalize beyond this subset. Two analytic datasets were considered. The first includes all 2530 subjects, while the latter excludes 61 subjects who had outlier brain-feature data. The choice of dataset does not change the key results of the analysis. We will use the terms "full" and "restricted" to refer to the datasets with 2530 and 2469 subjects, respectively, and present results for each dataset.

Both an unadjusted and adjusted analysis were performed. The adjusted analysis controlled for a subject's household income, parental education, age, gender and parental marital status. Once again, our findings did not meaningfully vary between the adjusted and unadjusted analysis. Here we present only the unadjusted analysis results.

Results

Let $X_i \in \mathbb{R}$, $M_i \in \mathbb{R}^{75}$, and $G_i \in \mathbb{R}$ denote subject i's the PGRS, 75 brain features, and cognition score for $i = 1, \ldots, n$, with $n \in \{2530, 2469\}$. We will denote a single subject's data by $Z_i = (X_i, M_i, Y_i)^T \in \mathbb{R}^{77}$. Because X and G are both scalar valued, we aim to identify a single linear projection $\eta^T M$ that captures patterns of brain activity that mediate the genetic-cognition pathway. The mediation directions objective function is maximized in order to estimate η:

Table 2.4: Estimated correlation coefficients and correlation coefficient products resulting from the optimization of the mediation directions objective.

Dataset	$\hat{\rho}_1$	$\hat{\rho}_2$	$\hat{\tau}$
Full	0.176	0.356	0.063
Restricted	0.178	0.353	0.063

$$\underset{\eta \in \mathbb{R}^{75}}{\text{maximize}} \ \tau(\eta) = \rho_1 \times \rho_2 := \frac{\Sigma_{xm}\eta}{\sqrt{\sigma_x^2 \times \eta^T \Sigma_m \eta}} \times \frac{\eta^T \Sigma_{mg|x}}{\sqrt{\eta^T \Sigma_{m|x} \eta \times \sigma_{g|x}^2}}, \tag{2.33}$$

where

- $\Sigma_{xm} \in \mathbb{R}^{1\times 75}$ denotes the cross-covariance matrix between the PGRS and the brain features.
- $\Sigma_{mg|x} \in \mathbb{R}^{75\times 1}$ denotes the conditional cross-covariance matrix between the brain features and cognition measures.
- σ_x^2 and $\sigma_{g|x}^2 \in \mathbb{R}$ are the variances of the PGRS and the cognition score given the PGRS.
- $\Sigma_m, \Sigma_{m|x} \in \mathbb{R}^{75\times 75}$ are the covariance and conditional covariance matrices of brain features, the latter given the PGRS.

The mediated effect

Solving the optimization problem 2.33 produces the following estimates of τ, ρ_1 and ρ_2 given in Table 2.4. The estimated correlation between the PGRS and the cognition scores is $\rho_{tot} = \widehat{\text{Cor}}(X, Y) = 0.171$. The in-sample estimate of the proportion of the total effect of the PGRS on the cognition score mediated by $\hat{\eta}^T M$ is

$$\rho/\rho_{tot} = 0.063/0.171 \times 100 \approx 37\%.$$

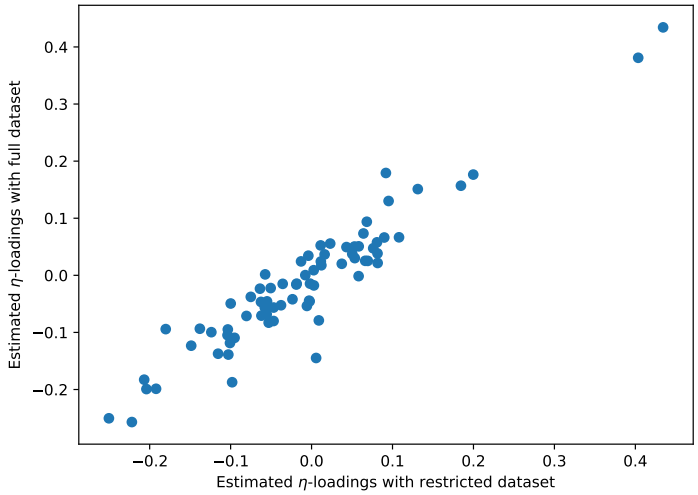


Figure 2.13: A comparison of η-loadings between the full and reduce analytic datasets. The figure shows that the estimated loadings do not appear to strongly depend on the inclusion or exclusion of high-influence subjects.

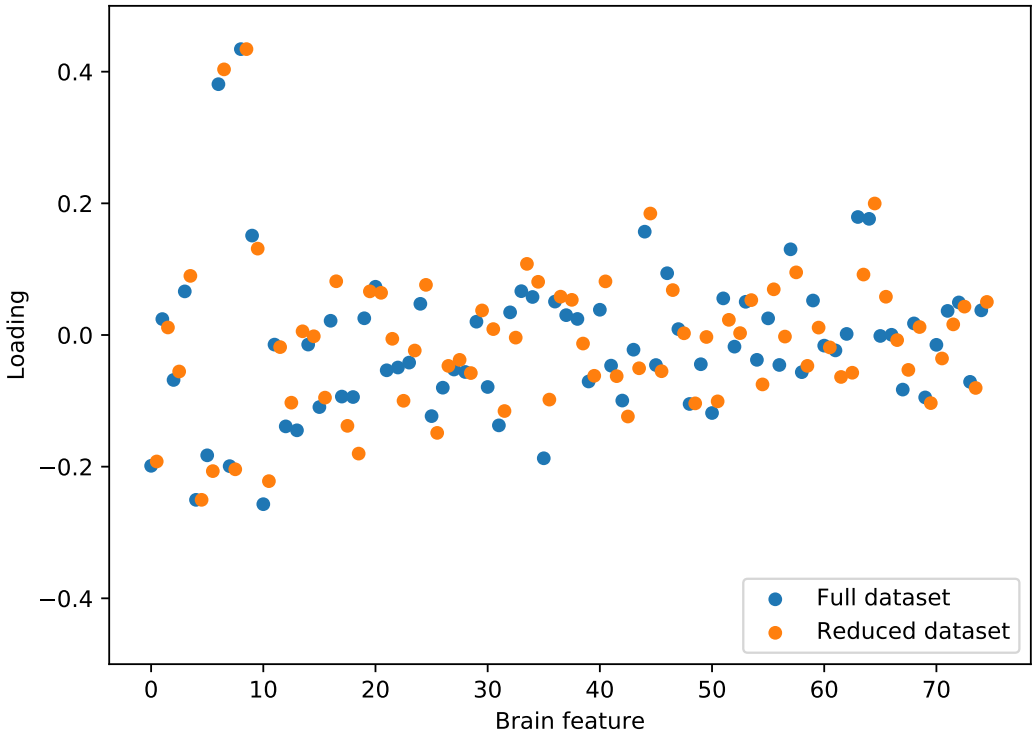


Figure 2.14: A comparison of η-loadings between the full and reduce analytic datasets ordered by brain feature. The estimated loadings ares similar between the full and reduced datasets, with only a few values in either $\hat{\eta}$ greater than 0.2 in magnitude.

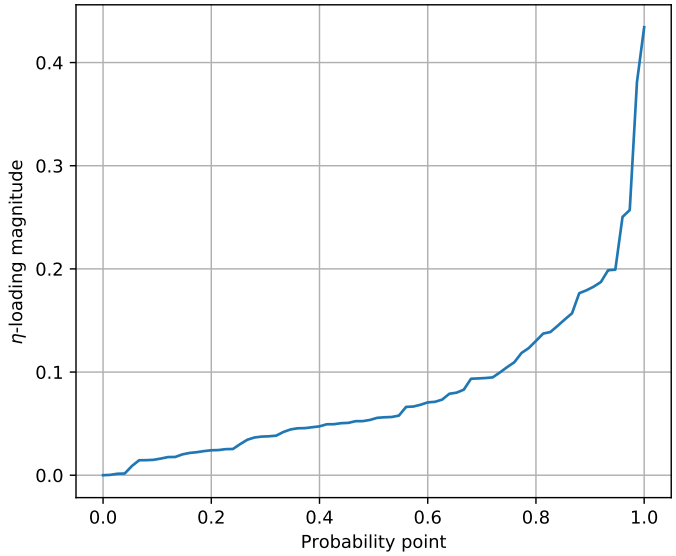


Figure 2.15: The empirical distribution of the η-loadings. The quantiles of the magnitude of η-loadings against probability points. The observed distribution of η-loadings is right skewed, with approximately 25% of loadings having magnitude greater than 0.1, and 5% having magnitude greater than 0.20.

Figures 2.13 and 2.14 show the estimated η-loadings using the full and reduced datasets. The magnitude of the loading estimates between analyses are similar. This suggests that the individuals with outlying brain feature scores do not influence the analysis substantially.

Figure 2.15 shows the estimated inverse-CDF function of the distribution of η-loadings (the η-loading quantiles are plotted against their corresponding probability points). The figure shows that the observed distribution of loadings is right skewed, which means that the mediated direction is primarily driven by brain activity differences in a small number of brain features. Because the brain features are identified only up to a change in sign, we are unable to interpret the signs of the loadings without the help of the primary research team's visualization tools.

The estimated mediation direction has found a direction in the data with moderate variation after projection. Figure 2.16 shows the estimated inverse-CDF function of the eigenvalues of Cov(M). The red, dashed line plots the observed variance of $\hat{\eta}^T M$, showing

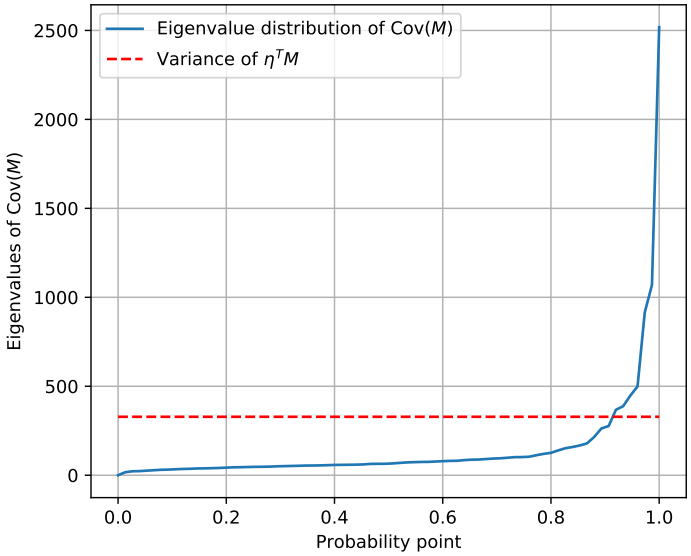


Figure 2.16: The empirical distribution of the eigenvalues of Cov(M). The empirical quantiles of the eigenvalues of Cov(M) are plotted against probability points. The variance of the brain feature projection $\hat{\eta}^T M$ is plotted as the horizontal red line as a point of reference.

that the variance of the estimated mediation variable falls at roughly the 90th percentile of the eigenvalue distribution. The ratio of Var($\eta^T M$) to the largest eigenvalue of Cov(M) is approximately 1/10. Importantly, the identified mediating direction describes non-trivial variation in the data. If the projection found a direction of M with little observed variance, the identified direction would be less believable.

Assessing overfitting

Table 2.5 compares using in- and out-of-sample estimates of the correlation coefficients and their product. The interval after the mean estimate is a 95% confidence interval based on 10,000 bootstrapped datasets. Most importantly, the 95% confidence interval for τ created using out-of-sample data does not include 0, suggesting that the identified brain activity patterns do indeed mediate the genetic-intelligence pathway. The out-of-sample estimate of the proportion of the total effect mediated by the variable $\hat{\eta}^T M$ is $0.023/0.171 = 0.13$.

The results of our simulation studies in Section 2.5.4 suggest that the out-of-sample estimate of the product-of-correlations is likely too small. We saw that in regimes where the in-sample estimates overfit the observed data, the out-of-sample estimate of τ was downward biased. Here we have a large relative difference between the in- and out-of-sample estimates of τ, which suggests that our estimate is downward biased. Additionally, the marginal distributions of the brain features are heavy tailed relative to a Gaussian distribution. This remained true after removing individuals who were clearly outliers. Because the method uses the data's second moments, the observations that fall in the tails have "high-leverage." They have greater influence on the estimated η direction during the in-sample portion of Algorithm 2.3. This causes the estimated η to generalize poorly out-of-sample.

Scientific findings

The results of this analysis suggest that the PGRS - general intelligence pathway is mediated by contrasting activation patterns in large-scale brain networks. Figure 2.18 shows a heat map of the estimated mediation direction loadings mapped onto the brain cortex. Increased activity in regions colored orange was positively associated with larger PGRS and greater general intelligence. Conversely, PGRS and general intelligence scores were inversely associated with activation in regions colored blue.

Interestingly, these activation patterns have interpretations in terms of large-scale brain activation networks. Orange and blue regions predominantly belong to the fronto-parietal network (FPN) and default mode network (DMN), respectively. It has been well documented in the literature that increased activity in the FPN and decreased activity in the

Table 2.5: In-sample and out-of-sample estimates and confidence intervals of the correlation coefficients.

	ρ_1	ρ_2	τ
In-sample	0.22 (0.18, 0.25)	0.36 (0.33, 0.39)	0.079 (0.06, 0.09)
Out-of-sample	0.08 (0.03, 0.13)	0.28 (0.22, 0.34)	0.023 (0.009, 0.037)

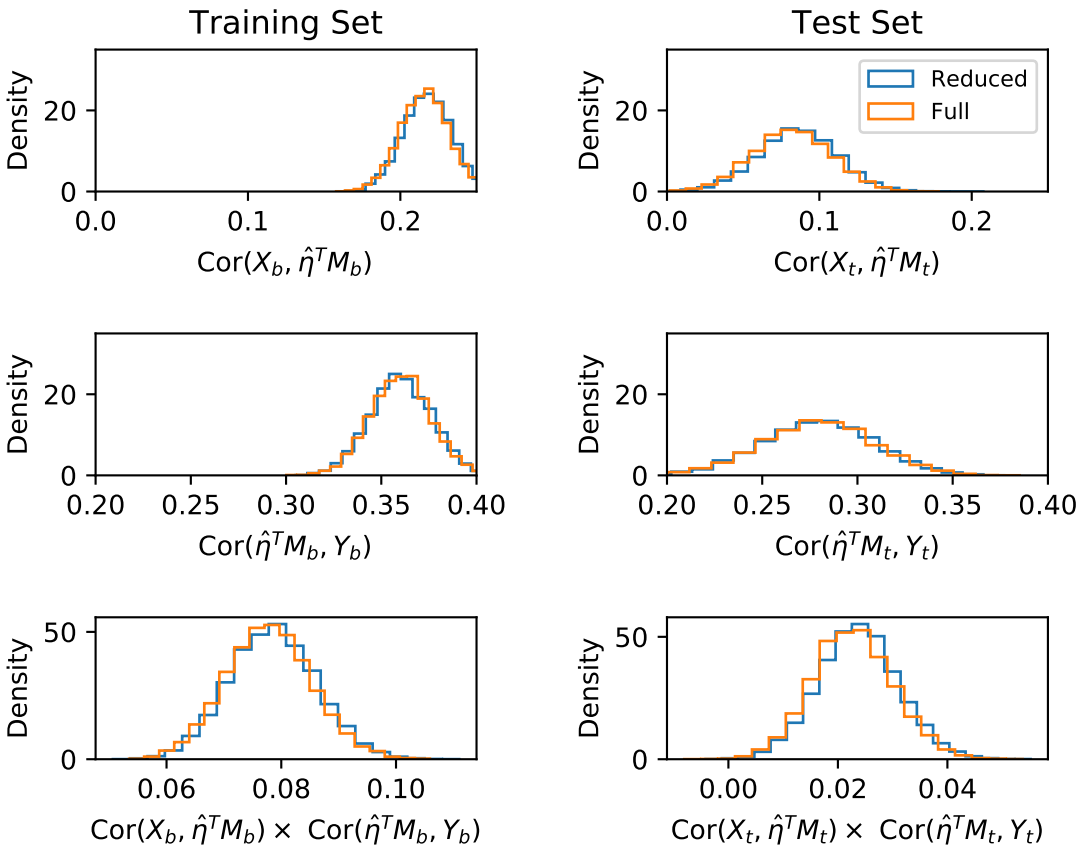


Figure 2.17: Histograms of the bootstrapped estimates of ρ_1, ρ_2 and ρ for the in-sample (training) and out-of-sample (testing) datasets.

DMN during cognitively challenging tasks are associated with greater general intelligence. Our analysis connects this well-studied association to genetic factors that are predictive of intelligence. Taken together, it seems that the contrasting activation patterns of the FPN and DMN partially explain the PGRS - general intelligence association.

However, these findings should not be overstated. In particular, both the in- and out-of-sample correlations between the PGRS and the one-dimensional brain feature are small. This should not surprising, since the PGRS was created to predict general intelligence not variation in large-scale brain network activation. Our findings suggest both that (a) another mechanism explains the remainder of the genetic - general intelligence association, and (b) other genetic or environmental factors explain variation in FPN and DMN activation that

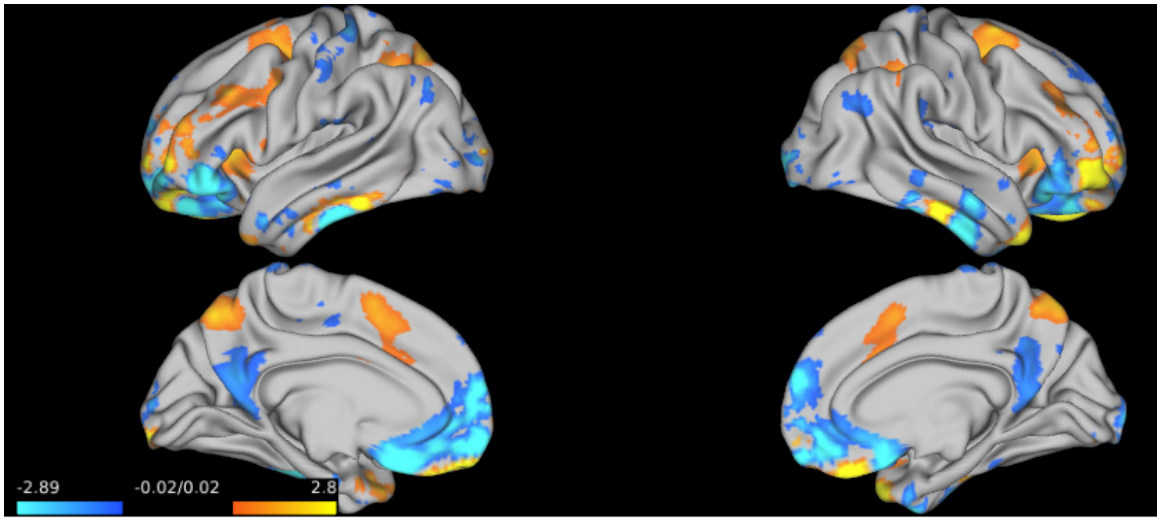


Figure 2.18: A brain heat map of activity patterns that appear to mediate the PGRS-intelligence association.

leads to greater general intelligence.

2.7 Discussion

In this chapter we proposed and studied the first fully optimization-based approach to identifying multivariate mediation structure. Existing multivariate mediation methods are model-based and must estimate parameters unrelated to the mediation structure. For problems of moderate dimension, specifying a full joint model is difficult to do. A major strength of our approach is that it is capable of identifying the most prominent mediation structure without specifying a full model fir the data. However, compared to model-based approaches, the proposed optimization-based approach estimates the mediation directions less efficiently.

The proposed method's limitations are in its lack of generalizability. One such extension would be modifying the objective function so that it regularizes the mediation directions. The fMRI case study data exhibited two scenarios in which regularization would likely improve estimates: when the data are either (a) high-dimensional or (b) have heavy tails. One would hope that regularized optimization would improve the out-of-sample gen-

eralizability of the estimated directions. We currently are unsure of how one might regularize this objective function due to its scale-invariability.

Regularization would also open up "$p > n$" applications where our objective function is currently undefined. In these settings, the variance-covariance matrices in the denominator of the objective function are not positive definite. Depending on the structure and what is known about the data-generating mechanism, one might consider using model-based regularization, or use penalized covariance estimators to estimate these inputs. These approaches could be used in combination with an approach that also regularizes the mediation directions.

Another useful extension of this methodology would modify the objective function to work with non-second-moment-based estimates of dependencies between X, M, and Y. Mediation analyses are frequently used by social scientists who collect Likert-scaled data via surveys. Analysts generally treat these data as continuous (as we did in the economic data case study) but alternative measures of dependence between variables might improve estimates of mediation structure. One potential solution would use copulas to estimate the dependence between discrete or ordinal variables.

We believe that the proposed method should be viewed as a part of a complete analysis. Because the mediation directions objective function links X to Y through M, the method can fail to identify projections of X and Y that are highly correlated. This could potentially be an issue if the variables in M are not the mechanisms that explain the exposure-outcome associations. Without additional multivariate data analyses, such as canonical correlation analysis and factor analysis, one would have a less-than-complete picture of the associations between the variables. Additional analyses also act as a check to make sure that the estimated mediation directions are directions in the data with reasonable variance.

CHAPTER 3

Conditional Methods for Non-Regular Inference Problems with Applications to Testing Mediation Hypotheses

3.1 Introduction

The primary aim of most univariate mediation analyses is to characterize the role of the hypothesized mediator in the exposure-outcome association. The first step in this process is assessing whether the indirect effect is non-zero. For the linear structural equation model introduced in Section 1.1, this reduces to testing $H_0 : \alpha\beta = 0$ versus $H_A : \alpha\beta \neq 0$. Tests of the null hypothesis have been shown to be conservative over a subspace of the null-parameter space [8, 13]. To the best of our knowledge, no proposed procedure for testing the indirect mediation effect is properly calibrated over the entire full parameter space.

This chapter proposes a procedure with improved calibration for assessing evidence against the null hypothesis $\alpha\beta = 0$. In order to achieve better test calibration, we first identify the fundamental challenge with testing in this setting. We combine the insight gained from this perspective with ideas from conditional inference to produce a new test. The proposed procedure is broadly applicable, as it uses likelihood ratios to summarize the evidence against the null hypothesis, and can be used when the mediator and outcome models belong to the class of generalized linear models. We show through simulation that

the procedure improves calibration relative to existing methods. Against certain alternative hypotheses where $\alpha\beta \neq 0$, the proposed procedure also has greater power than existing methods.

3.1.1 Motivation for a conditional test of the indirect effect

In a review of Ronald A. Fisher's work [25], David Hinkley stated that the point of conditional inference is to "condition the sampling distribution of a test statistic to only the relevant subsets of the general experimental subspace." If one adopts this point of view, then the calibration of a sampling distribution should account for the population that generated the observed sample. To motivate our use of conditioning for a test of the indirect effect, we show that data from different regions of the experimental subspace have different log-likelihood ratio sampling distributions.

Although most tests of the indirect effect create a $(1 - \tilde{\alpha})\%$ confidence interval for the parameter $\alpha\beta$, we choose to use likelihood ratios to assess the evidence that M is a mediator. Additionally, the likelihood ratio approach provides the perspective necessary to understand why tests of the indirect effect are conservative. We will include details of the log-likelihood ratio test in Section 3.3. For now, we generate likelihood ratio test statistics for three different populations in which the indirect effect is 0.

Figure 3.1 is a quantile-quantile plot (QQ-plot) that compares the quantiles of the log-likelihood ratio test statistic to the χ^2_1 asymptotic reference distribution for 3 data generating populations. Only Population 3's test statistic quantiles (blue line) are well approximated by the χ^2_1 reference distribution. Population 1 and 2's test statistic quantiles (red and green lines) are smaller than the reference distribution's. Figure 3.1 shows that inference using the χ^2_1 reference distribution will not be well calibrated for all null data-generating populations.

The summary of conditional inference found in [25] and Fisher's first description of conditional inference [26] suggest that we attempt to tailor the reference distribution to the log-likelihood ratio's sampling distribution. To do so, Fisher suggested that we should use

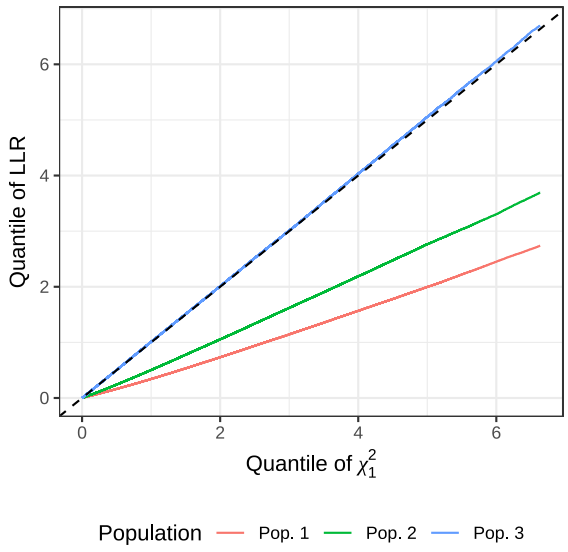


Figure 3.1: QQ-plots comparing three data-generating populations' log-likelihood ratios to the χ^2_1-distribution. Line color distinguishing between populations.

additional "ancillary" information in the data to locate our data in the full experimental design space. The following three insights allowed us to make our inferences about the indirect mediation effect more relevant to the observed data:

1. A nuisance parameter causes the log-likelihood ratio's sampling distribution to vary over the null parameter space.

2. There exists a sufficient statistic for the nuisance parameter which is approximately ancillary for the parameter that determines if H_0 is true.

3. By conditioning on the approximately ancillary statistic, we are able to tailor our inference to the relevant region of the experimental space.

We again defer giving a proper definition of the approximately ancillary statistic A to Section 3.3. To demonstrate A's impact on the sampling distribution, for each simulated dataset used to create the QQ-plot in Figure 3.1, we calculate both the log-likelihood ratio and its corresponding ancillary statistic. We then sort (λ, A) pairs into four groups, using the quartiles of A within each population as the cut points for the groups. Figure 3.2

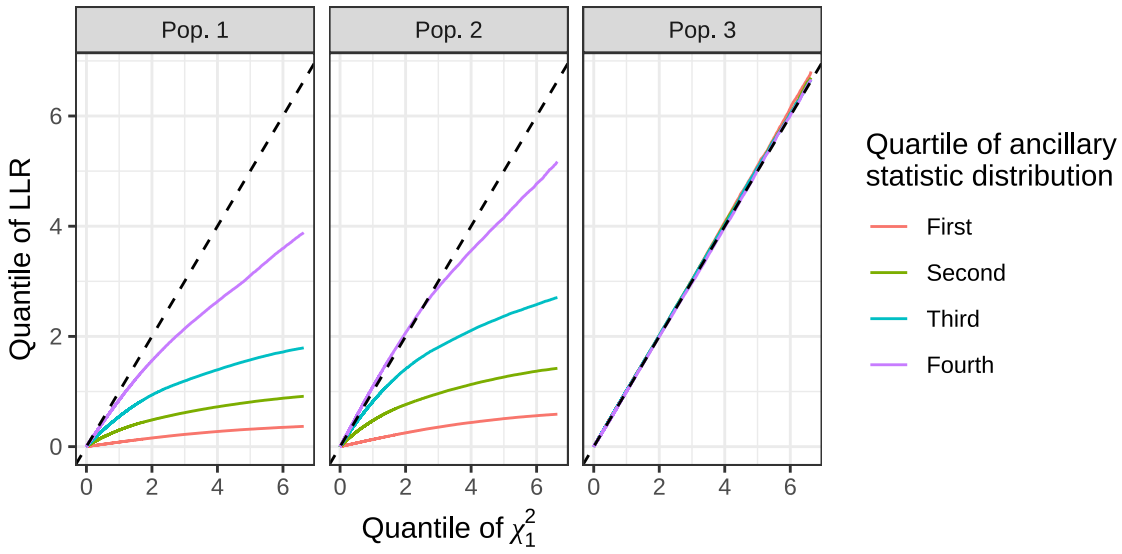


Figure 3.2: QQ-plots of the log-likelihood ratios quantiles against χ^2_1 quantiles stratified by ancillary statistic A. Results are split into panels for each data-generating population (by panel) and by quartile of A by color.

compares the quantiles of the log-likelihood ratios to the χ^2_1 reference distribution after stratifying by quartile of A for each population. Figure 3.2 shows that the distribution of the log-likelihood ratio depends strongly on A for Populations 1 and 2. The QQ-plots for the log-likelihood ratios vary strongly between quartiles of A. Thus, if A indeed contains little information about the truth of H_0, we will be able to make our inferences more relevant to the observed data by conditioning on A.

Finally, we show that by conditioning on A, the sampling distribution of the log-likelihood ratio is more similar between populations. Figure 3.3 shows 2 QQ-plots comparing quantiles of Population 1 and 2 log-likelihood ratios. After approximately conditioning on A, the sampling distribution of λ is more similar between Populations 1 and 2. This finding will be useful when we design our procedure for learning the conditional distribution of λ given A.

The remainder of this chapter is organized as follows. In Section 3.2 we review existing approaches to testing mediation hypotheses, conditional inference, and likelihood ratio

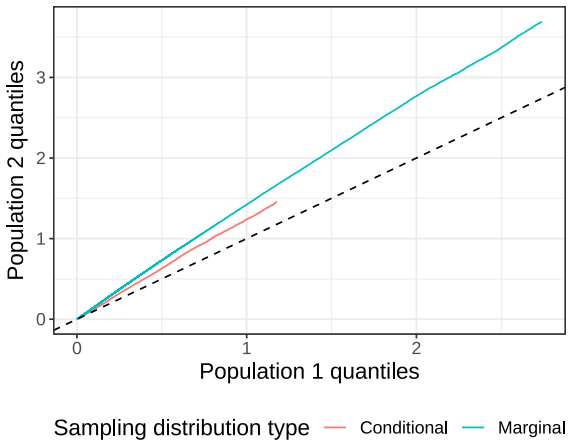


Figure 3.3: Marginal and conditional QQ-plots of the log-likelihood ratios for Populations 1 and 2. The green line shows the relationship between log-likelihood ratios marginal over A, while the red line restricts the QQ-plot to all (λ, A) pairs such that $A \approx 1.25$. Since the red line falls closer to the black, dashed 1-1 line than the green, the conditional sampling distributions are more similar than the marginal sampling distributions.

testing at singularities. We then introduce our conditional procedure in Section 3.3. In Section 3.4 we derive and characterized the limiting distribution of the log-likelihood ratio's conditional sampling distribution. We then present extensive simulation studies in Section 3.5, which demonstrate that the proposed method does in fact improve test calibration. We close with a discussion of the method in Section 3.6.

3.2 Relevant background

3.2.1 Conditional inference

Conditional inference is an inferential approach that attempts to tailor inferences to the characteristics of the observed data. Conditional inference is not a strictly frequentist approach to inference. David Hinkley writes in [25] that "inasmuch as inference proceeds by relating a given experimental outcome to a series of hypothetical repetitions which generates frequency distributions for statistics, that series of repetitions should be as relevant as possible to the data at hand." A simple example makes it clear what one means by "relevant

to the data at hand."

Suppose one is interested in estimating the probability p that a machine produces a defective product. The experimenter will either observe 25 or 100 products from the machine with equal probability. On the day of the test, it is determined that the experimenter observes 100 products. The experimenter is now left to decide how to calibrate the precision of his or her estimate.

A strict frequentist calculation of the standard error would marginalize over the random variable that determines whether 25 or 100 products are assessed for defects. Most individuals would choose instead to condition on the fact that they observed 100 products and ignore the sampling behavior of $\hat{p}$ when 25 products are observed. This choice is logical because it makes our inference more relevant for the data that were eventually produced by the study.

R.A. Fisher first formally described conditional inference in [27, 26]. For Fisher, conditional inference is closely related to identifying an ancillary statistic for the parameter of interest. At a high level, Fisher described ancillary statistics as quantities that contain little or no information about the parameters of interest.

"Ancillary" is a difficult concept to define precisely, in part because it has two meanings in the conditional inference context. In [28] Kalbfleisch clearly defined and delineated the differences between *experimental* and *mathematical* ancillaries.

The example given previously in which one of two studies was randomly selected and conducted is an example of an experimental ancillary. An experimental ancillary is any quantity whose distribution is independent of the phenomenon under study. In our example, whether the experimenter observed $n = 25$ or $n = 100$ products is independent of the defect rate p. Often by the time an analysis begins, one will not think of an experimental ancillary as random. This means that there is less uncertainty as to how inference should proceed when dealing with experimental ancillaries.

Fisher focused his development of conditional inference around mathematical ancillar-

ies. They are a result of the statistical model chosen for the data. Unlike experimental ancillaries, a mathematical ancillary may be formed from quantities that contain information about the quantity of interest. For example, consider $X_1, X_2 \overset{iid}{\sim} \mathcal{N}(\mu, 1)$. The quantity $A = X_1 - X_2$ is ancillary for μ since the distribution of A is free of μ. However, both X_1 and X_2 clearly contain information relevant to estimating μ.

Fisher described mathematical ancillaries as the "configurations" of the sample [26]. He believed that by accounting for the configuration of a particular sample, one could produce an inference more relevant to the observed data. Suppose that $X \sim F_\theta$ where $\theta = (\mu, \psi)$ and let (T, A) be a sufficient statistic for θ. If the likelihood of θ given X factors over T and A:

$$f(T, A; \theta) = f(T|A; \mu) f(A; \psi),$$

then we say that A is ancillary for μ, since the likelihood for μ does not depend on A. If one wishes to perform inference on μ, then ψ is often called a *nuisance parameter* since ψ is not the focus of the inference but must be accounted for when performing inference on μ. Fisher argued that inference for μ should be carried out using the conditional distribution $T|A$ rather than the marginal distribution of T.

Conditional inference examples

Fisher's exact test is the most widely known example of conditional inference and was one of the examples he chose to motivate the conditional approach to inference [26]. Fisher's exact test is most commonly used to analyze 2x2 contingency table, and assesses dependence between the two categorical variables. The dependence is often estimated using an odds ratio. A typical 2x2 contingency table is given in Table 3.1 of two variables X and Y taking values X_1 and X_2 and Y_1 and Y_2, respectively.

The test is described as "exact" because when the row and column totals of a 2x2 table $(a+c, b+d, a+b, c+d)$ with independent rows and columns are fixed, then one can model the cell counts using a hypergeometric distribution. Fisher argued that when the row and

	X_1	X_2	Row total
Y_1	a	b	$a+b$
Y_2	c	d	$c+d$
Column total	$a+c$	$b+d$	$a+b+c+d$

Table 3.1: An example of a 2x2 contingency table.

column totals are random, these quantities contain little information for the parameter of interest, or are nearly ancillary [26]. With random margins, the cell counts can be modeled as a multinomial random variable and the conditional distribution of the cell counts given the observed table margins remains hypergeometric. Fisher argued that ease of inference with a closed form sampling distribution outweighed the fact that the table margins are not exactly ancillary for the odds ratio.

Since its introduction, statisticians have disagreed about whether using Fisher's exact test is appropriate. The prevailing consensus is that conditioning is reasonable even when the table margins are not fixed. This is in large part due to two findings. First, a number of examples have shown that approximate, unconditional p-values can often exhibit undesired properties [29]. Second, it has been shown that the margins contain very little information about the odds ratio. In fact, [30] showed that the margins are nearly ancillary, and therefore conditioning has little negative impact on the inferential performance of Fisher's exact test.

A second example uses conditioning to estimate logistic regression models for stratified or matched data. Conditional logistic regression [31] was developed to study disease risk in case-control study designs. Let $Y_j \in \{0, 1\}$ record whether individual j has a disease, often called a case. In a case-control study design, each case is matched with d controls who do not have the disease. Together, a case and its assigned controls make up a single stratum. Conditional logistic regression is capable of accommodating both multiple cases and controls per stratum.

Cases are generally matched to the controls who have similar levels of variables believed to be relevant to the risk of developing the disease (e.g. age or sex). The primary goal of the analysis is to study how risk varies with other covariates $x \in \mathbb{R}^p$. Suppose we

have k cases, each matched with d controls. Within group or stratum i, the first individual 0 is the case and individuals $1, \ldots, d$ are the controls. Within group i, the probability of having the disease is modeled by

$$P(Y_{ij} = 1 | X = x_{ij}) = \left(1 + \exp(\alpha_i + \beta^T x_{ij})\right)^{-1},$$

where α_i is captures the baseline risk of having the disease in stratum i. The baseline risk α_i is assumed to depend on the variables used to stratify and create case-control groupings, and therefore will vary across strata.

In this application, the parameters α_i act as nuisance parameters, since the investigators are not directly interested in estimating each stratum's baseline risk, but wish to establish how the covariates x relate to risk of disease. When there are many strata, estimates of stratum-level effects are in fact inconsistent [32]. Because the number of cases in each stratum was determined before the analysis began, the number of cases within a stratum can be treated as ancillary. Suppose that in stratum i a single case occurred. After conditioning on the fact that only one case exists in the stratum, stratum i's the conditional likelihood is free of α_i:

$$\begin{aligned} P\left(Y_{i,0} = 1, Y_{i,1} = \cdots = Y_{i,d} = 0 | X_{i,0}, \ldots, X_{i,d}, \sum_{j=1}^{d} Y_{i,j} = 1\right) \\ = \frac{\left(1 + \exp(\alpha_i + \beta^T X_{i,0})\right)^{-1}}{\sum_{j=0}^{d} \left(1 + \exp(\alpha_i + \beta^T X_{i,j})\right)^{-1} \prod_{j' \neq j} \left(1 + \exp(\alpha_i + \beta^T X_{i,j'})\right)^{-1}} \\ = \frac{1}{1 + \sum_{j=1}^{d} \left(1 + \exp\left(\beta^T (X_{i,j} - X_{i,0})\right)\right)^{-1}}. \end{aligned} \quad (3.1)$$

The last line of 3.1 gives an expression for stratum i's contribution to the likelihood function of β. Summing across all strata gives the full conditional likelihood of β. The conditional likelihood is free of α_i for all i, which allows for maximum conditional likelihood estimation of the β-vector without estimating the baseline risk of developing the

disease within each stratum.

These applications highlight two strengths of conditional procedures. First, they may be reasonably robust if a statistic is nearly ancillary and, second, conditioning can allow inference to proceed without estimation of nuisance parameters. Our use of conditional inference is different from these two examples in one important way. In these examples, after conditioning on an approximately ancillary statistic for each example, either an closed form for the sampling distribution exists or nuisance parameters do not need to be estimated. In other words, inference is made simpler by conditioning.

Inference for the conditional indirect effect is not simplified by conditioning, but rather requires numeric methods to estimate the conditional distribution of the log-likelihood ratio test statistic. Unconditionally, one can use an asymptotic χ^2_1 approximation (although it might not be a good approximation) to the sampling distribution. Rather than gaining simplicity by conditioning, a conditional test of the indirect mediation effect is more relevant for the observed data. The unconditional χ^2_1 approximation may be appropriate for some samples but not others, whereas the conditional approach will be better calibrated for all samples. In fact, we believe that one of this chapter's contributions is showing that conditional inference can be a useful tool even when conditioning does not lead to simplified inference.

3.2.2 Likelihood ratio tests at singularities

Tests of the indirect effect are poorly calibrated in finite samples due to a singularity in the null parameter space. It suffices for our purposes to define a singularity to be a point in a subspace with locally non-Euclidean geometry. In other words, the subspace is not smooth at the singularity in the sense that a tangent surface to the subspace does not exist at the singularity.

For our application, a singularity exists at the origin. Let

$$\Theta_0 = \{(\alpha, \beta) \in \mathbb{R}^2 \mid \alpha\beta = 0\}$$

denote the null-parameter space of the hypothesis H_0 restricted to $\theta = (\alpha, \beta)$. The set Θ_0 is equal to the coordinate axes of the real plane. At the origin, which we will denote $\tilde{\theta}_0$ of the real plane, a tangent surface to Θ_0 does not exist since the null parameter space intersects itself at $\tilde{\theta}_0$.

Before we derive the limiting distribution of a likelihood ratio test at $\tilde{\theta}_0$, let us consider the general situation of conducting a likelihood ratio test at a singularity. This problem has been studied by [33] and more recently [34]. Suppose that we are interested in testing a hypothesis about a parameter $\theta \in \Theta$. Suppose that $H_0 : \theta \in \Theta_0$, with $\Theta_0 \subset \Theta$ and at the true parameter value $\theta_0 \in \Theta_0$, the space Θ_0 is non-Euclidean. As a result, a tangent plane to Θ_0 at θ_0 does not exist and so the standard χ^2 asymptotics of the log-likelihood ratio test need not hold. Chernoff showed that in such settings, the distribution of the log-likelihood ratio test statistic converges to the distribution of the squared distance of the projection of a normal random variable onto the tangent cone to Θ_0 at θ_0 [33]. For a formal definition of the tangent cone, see [34]. This distribution is often not a χ^2 distribution. In such a setting, using the standard χ^2 reference distribution will lead to incorrectly calibrated inference.

Returning to the test of the indirect mediation effect, the only non-Euclidean location in Θ_0 is $\tilde{\theta}_0$. The tangent cone to $\tilde{\theta}_0$ is equal to the full null-parameter space Θ_0. The log-likelihood ratio test statistic at $\tilde{\theta}_0$ is equal in distribution to the squared distance of the projection of a bivariate normal random variable Z onto Θ_0. The projection is accomplished by setting the smaller component of Z to 0. Thus, $\lambda \sim \min(Z_1^2, Z_2^2)$. Since $Z_1, Z_2 \overset{iid}{\sim} \mathcal{N}(0, 1)$, $Z_1^2, Z_2^2 \overset{iid}{\sim} \chi_1^2$. Let G be the distribution function of the minimum of two independent χ_1^2 random variables. When the data-generating parameter is $\tilde{\theta}_0$, then the asymptotic distribution of the log-likelihood ratio test statistic will be equal to G. At all

other $\theta_0 \in \Theta_0$, the asymptotic sampling distribution will be the standard χ^2_1 distribution. Thus, the asymptotic sampling distribution of the likelihood ratio test statistic depends on $\theta_0 \in \Theta_0$.

This insight is difficult to operationalize into an improved inferential procedure. This is due to the discontinuity in θ_0 between asymptotic sampling distributions. For all sequences $\{\theta_0^j\}_{j=1}^{\infty} \in \Theta_0 \setminus \{\tilde{\theta}_0\}$ such that $\theta_0 \to \tilde{\theta}_0$, the likelihood ratio converges to its expected χ^2_1 sampling distribution. However, at the limit of the sequence, the asymptotic sampling distribution is equal to the minimum of two independent χ^2_1 random variables. For a fixed sample size n, the distribution of the likelihood ratio varies smoothly with $\theta_0 \in \Theta_0$. A procedure that improves test calibration must be capable of approximately learning the sampling distribution. We achieve this by utilizing ideas from conditional inference.

3.3 A conditional test of the indirect mediation effect

3.3.1 The indirect mediation effect

Suppose that $X, M, Y \in \mathbb{R}$ and that the dependence structure between (X, M, Y) is given by the graphical model in Figure 1.1. Let $M \sim F_{M|X}$ and $Y \sim G_{Y|X,M}$, where F and G are distribution functions that depend on X and (X, M), respectively. Additionally assume that $X \sim H$ for some distribution function H. These can be quite general models, but in order to define the indirect effect $\mathbb{E}[Y|X, M]$ must exist. For this joint model for (X, M, Y), we use the definition of the indirect mediation effect found in [35, 15].

Definition 3.3.1. The indirect mediation effect of changing the exposure X from x to x' on Y is:

$$\tau(x', x) = \int \mathbb{E}[Y|X = x, M = m] \left\{ dF_{M|X=x'}(m) - dF_{M|X=x}(m) \right\},$$

where $F_{M|X=x}$ is the conditional distribution function of M given $X = x$.

In order to evaluate the hypothesis test $H_0 : \alpha\beta = 0$ using a likelihood ratio, we must be able to perform constrained maximum likelihood estimation such that the indirect effect is equal to zero. Using Definition 3.3.1, we can describe sufficient conditions under which the indirect effect is equal to zero and M is not a mediator.

First, if Y and M are conditionally independent given X, then $\mathbb{E}[Y|X = x, M = m]$ does not depend on m. In this case, the indirect effect is equal to:

$$\begin{aligned}
\tau(x', x) &= \int \mathbb{E}[Y|X = x, M = m] \left\{ dF_{M|X=x'}(m) - dF_{M|X=x}(m) \right\} \\
&= \mathbb{E}[Y|X = x] \int \left\{ dF_{M|X=x'}(m) - dF_{M|X=x}(m) \right\} \\
&= \mathbb{E}[Y|X = x](1 - 1) \\
&= 0.
\end{aligned}$$

Second, consider the case where X and M are independent, which implies that the conditional distribution of M given X is equal to the marginal distribution of M: $F_{M|X} = F_M$. Under this assumption, it is easy to see that the indirect effect is again equal to zero:

$$\begin{aligned}
\tau(x', x) &= \int \mathbb{E}[Y|X = x, M = m] \left\{ dF_{M|X=x'}(m) - dF_{M|X=x}(m) \right\} \\
&= \int \mathbb{E}[Y|X = x, M = m] dF_{M|X=x'}(m) - \int \mathbb{E}[Y|X = x, M = m] dF_{M|X=x}(m) \\
&= \int \mathbb{E}[Y|X = x, M = m] dF_M(m) - \int \mathbb{E}[Y|X = x, M = m] dF_M(m) \\
&= 0.
\end{aligned}$$

Thus, the indirect effect is equal to zero if either (a) M and X are marginally independent or (b) M and Y are conditionally independent given X. Optimization over Θ_0 will require that either condition (a) or (b) is met. Note also that these conditions are sufficient

for the indirect effect to be equal to zero for all x and x'.

Because these conditions are not specific to any one statistical model, they are broadly applicable and easily operationalized. Without the conditions, for each joint statistical model defined in terms of the distribution functions $\{H, F, G\}$, one would use Definition 3.3.1 to calculate the indirect effect either in closed form or evaluated numerically. Then, one would maximize the joint-likelihood subject to the constraint that the indirect effect is equal to zero. Assessing marginal and conditional independence is often much easier than performing constrained optimization.

3.3.2 Generalized linear model-based mediation models

We now describe the class of mediation models that we study in this chapter. Both the conditional models for M and Y will belong to the class of generalized linear model (GLM). The method should apply to other statistical models that are equipped with both a likelihood function and describe independence relationships through the nullity of a single parameter. GLMs meet both of these requirements and are frequently used in practice, making their selection natural. We will refer to these models as "GLM mediation models" for the remainder of the chapter.

Definition 3.3.2. Let X, M and Y play the roles of a univariate exposure, potential mediator, and outcome variables, respectively. Let $Z \in \mathbb{R}^p$ be a vector of additional covariates.

Suppose that $M \sim F_m$ and $Y \sim F_y$, where F_m and F_y are both distribution functions of an exponential family. Define $\eta_m = \gamma_1 + \alpha X + Z^T\lambda_1$ and $\eta_y = \gamma_2 + \gamma_3 X + \beta M + Z^T\lambda_2$, and let

$$\mathbb{E}[M|X] = g_m^{-1}(\eta_m) \text{ and } \mathbb{E}[Y|X, M] = g_y^{-1}(\eta_y),$$

and assume that the conditional variance of M and Y are each a function of their respective conditional mean and an additional scale parameter.

Let $\phi = (\gamma_1, \alpha, \lambda_m, \sigma_m)$ and $\psi = (\gamma_2, \gamma_3, \beta, \lambda_y, \sigma_y)$, where σ_m and σ_y are the GLM scale parameters. Let $\theta = (\phi, \psi)$ be the full parameter vector. The joint probability density or mass function over M and Y given X factorizes as $f(m, y|x) = f_m(y|x, m)f_y(m|x)$, where f_m and f_y are the probability distribution or mass functions of F_m and F_y, respectively. The log-likelihood function of the GLM mediation model is:

$$\ell(\theta; M, Y|X) = \ell_m(\phi; M|X) + \ell_y(\psi; Y|M, X).$$

Under the GLM mediation model, X and M are marginally independent if $\alpha = 0$ and M and Y are conditionally independent given X if $\beta = 0$. We saw above that if either of these two independence conditions are met, then the indirect effect defined in Definition 3.3.1 is equal to 0. As a result, testing whether M is a mediator in the GLM mediation model requires testing $H_0 : \alpha\beta = 0$.

3.3.3 GLM mediation model parameter estimation

Our likelihood-based approach to testing the indirect effect depends on likelihood ratios, which requires parameter estimation over both the null and alternative parameter spaces. Calculation of the log-likelihood of the alternative model is trivial, as no constraint is placed on the parameters. Define $\theta = (\phi, \psi)$, where ϕ and ψ are the parameters of the M and Y GLMs, respectively. We have shown that the likelihood for θ separates over ϕ and ψ. Because the likelihood separates over ϕ and ψ and either $\alpha = 0$ or $\beta = 0$ over the null parameter space, the constrained maximum likelihood estimator must select between two possible null models.

First, we formally define $\Theta_0 = \{\theta \in \Theta \mid \alpha\beta = 0\}$. Then $\hat{\theta} \in \Theta_0$ if and only if $\hat{\alpha} = 0$ or $\hat{\beta} = 0$. This gives rise to two null models: one in which $\hat{\alpha} = 0$ and the other in which $\hat{\beta} = 0$. Because the likelihood factorizes over ϕ and ψ, whenever one parameter is equals 0, its counterpart is equal to its unconstrained maximum likelihood estimator.

To maximize the GLM mediation model log-likelihood, we estimate four models: two

each for the M and Y GLMs. For each outcome, an unconstrained and a constrained model will be fit. We will denote the unconstrained and constrained maximum likelihood estimators by $\hat{\phi}_1$ and $\hat{\psi}_1$, and $\hat{\phi}_0$ and $\hat{\psi}_0$, respectively. To reiterate, the subscripts "1" and "0" denote unconstrained and constrained models. The four sub-models combine to make two joint-null models: $\hat{\theta}_0^m = (\hat{\phi}_0, \hat{\psi}_1)$ and $\hat{\theta}_0^y = (\hat{\phi}_1, \hat{\psi}_0)$. For each null model, $\hat{\alpha}\hat{\beta} = 0$.

The two null sub-models have log-likelihood values equal to:

$$\ell_m^0 := \ell((\hat{\theta}_m^0; M, Y|X) = \ell_m(\hat{\phi}_0; M|X) + \ell_y(\hat{\psi}_1; Y|M, X)$$

and

$$\ell_y^0 := \ell(\hat{\theta}_y^0; M, Y|X) = \ell_m(\hat{\phi}_1; M|X) + \ell_y(\hat{\psi}_0; Y|M, X).$$

The null-model likelihood is then $\ell_0 = \max(\ell_m^0, \ell_y^0)$, and $\hat{\theta}_0 = \underset{\theta_0 \in \{\hat{\theta}_m^0, \hat{\theta}_y^0\}}{\arg\max}\ \ell(\theta_0)$. Conversely, the alternative model likelihood is $\ell_A := \ell_m(\hat{\phi}_1; M|X) + \ell_y(\hat{\psi}_1; Y|M, X)$.

The log-likelihood ratio test statistic of $H_0 : \alpha\beta = 0$ can take one of two forms depending on whether $\hat{\alpha}$ or $\hat{\beta}$ is equal to 0 at the constrained maximum likelihood estimate:

$$\lambda = 2(\ell_A - \ell_0) = 2(\ell_A - \max\{\ell_m^0, \ell_y^0\}) = 2 \times \begin{cases} \ell_m(\hat{\phi}_1; M|X) - \ell_m(\hat{\phi}_0; M|X), & \hat{\theta}_0 = \hat{\theta}_0^m \\ \ell_y(\hat{\psi}_1; Y|M, X) - \ell_y(\hat{\psi}_0; Y|M, X), & \hat{\theta}_0 = \hat{\theta}_0^y \end{cases}. \tag{3.2}$$

Depending on which parameter is null, the test statistic λ is a log-likelihood ratio testing either $H_0 : \alpha = 0$ or $H_0 : \beta = 0$. If we define

$$\lambda_m = 2\left(\ell_m(\hat{\phi}_1; M|X) - \ell_m(\hat{\phi}_0; M|X)\right)$$

and

$$\lambda_y = 2\left(\ell_y(\hat{\psi}_1; Y|M, X) - \ell_y(\hat{\psi}_0; Y|M, X)\right),$$

then the log-likelihood ratio test statistic of the indirect mediation effect is $\lambda = \min\{\lambda_m, \lambda_y\}$. The test statistic of the indirect mediation effect selects the null model which incurs the smallest penalty to the joint likelihood.

3.3.4 The impact of a nuisance parameter on the likelihood ratio sampling distribution

Suppose that for a given $\theta \in \Theta_0$, $\alpha\beta = 0$, but either $\alpha \neq 0$ or $\beta \neq 0$. The sampling distribution of the test statistic $\lambda = \min\{\lambda_m, \lambda_y\}$ changes as a nuisance parameter (the non-zero element of (α, β)) varies. The dependence of the sampling distribution on the nuisance parameter can be understood both in terms of the "either-or" characteristic of the test statistic and the geometry of the null parameter space. Both are described, as each helps build intuition for the proposed conditional procedure.

First, we consider the effect of taking the minimum of two likelihood ratio test statistics by considering the common setting in which both GLM families are Gaussian. It can be shown that λ_m and λ_y converge in distribution to non-central χ^2_1 random variables with non-centrality parameters

$$\mu_m = \left(\frac{\sqrt{n}\alpha\sigma_x}{\sigma_m}\right)^2 \text{ and } \mu_y = \left(\frac{\sqrt{n}\beta\sigma_m}{\sigma_y}\right)^2$$

as n grows. These approximations will be derived in Section 3.4.1.

Without loss of generality, suppose that $\alpha = 0$, so $\mu_m = 0$ and $\mu_y > 0$. Our aim is to relate the tail probabilities $\mathbb{P}(\lambda > x)$ to the tail probabilities of χ^2_1-distribution so that we can assess the effect of β on the tail probabilities. In order for the χ^2_1 tail probabilities to be valid p-values, we need $\mathbb{P}(\lambda > x) = \mathbb{P}(\chi^2_1 > x)$. We derive the tail probability under the assumption that λ_m and λ_y are asymptotically independent. Although this has not been

proven, numerous simulation studies have failed to disprove the assumption. Let p_x denote the upper tail probability, $\mathbb{P}(\chi^2_1 > x)$, of the χ^2_1 distribution.

Since $\alpha = 0$, $\mu_m = 0$. Thus, $\lambda_m \sim \chi^2_1$ so $\mathbb{P}(\lambda_m > x) = p_x$. The upper tail probability is:

$$\begin{aligned} \mathbb{P}(\lambda > x) &= \mathbb{P}(\lambda_m > x, \lambda_y > x) \\ &= \mathbb{P}(\lambda_m > x)\mathbb{P}(\lambda_y > x). \\ &= p_x \times \mathbb{P}(\lambda_y > x). \end{aligned} \tag{3.3}$$

This makes it clear that $\mathbb{P}(\lambda > x) = p_x$ only if $\mathbb{P}(\lambda_y > x) = 1$. As $\mu_y \to \infty$, $\mathbb{P}(\lambda_y > x) \to 1$ for all $x > 0$. Therefore, χ^2_1 tail probabilities will not be valid p-values when μ_y is small. Using the standard reference distribution will result in a likelihood ratio test with level less than its target level whenever $\mathbb{P}(\lambda_y > x) < 1$.

For fixed n, σ_m and σ_y, μ_y depends only on β. Thus, in this setting β acts as a nuisance parameter. The sampling distribution of λ depends on the value of β even though the value of β does not affect the truth of H_0 since $\alpha = 0$ in the population.

A geometric interpretation of the "either-or" nature of the test statistic also helps clarify marginal tests' conservatism. With a slight abuse of notation, we restrict the null parameter space to the (α, β) plane and define $\Theta_0 = \{(\alpha, \beta) \in \mathbb{R}^2 \mid \alpha\beta = 0\}$. Next, the asymptotic behavior of λ is described, which varies bases on the data-generating parameter $\theta_0 \in \Theta_0$.

In $\mathbb{R}^2$, Θ_0 is equal to the coordinate axes. At the origin ($\alpha = \beta = 0$), the null parameter space is locally non-Euclidean (a tangent plane to the null parameter space does not exist at the origin). This geometric feature, often called a singularity, gives rise to non-standard asymptotics. Using results from [34, 33], one can show that if $\alpha = \beta = 0$, then $\lambda_n \xrightarrow{d} \min\{Z_1, Z_2\}$ as $n \to \infty$, where $Z_1, Z_2 \overset{i.i.d.}{\sim} \chi^2_1$. For details of this derivation, see Section 3.2.2. Standard results show that when either $\alpha \neq 0$ or $\beta \neq 0$ but $\alpha\beta = 0$, then $\lambda_n \xrightarrow{d} \chi^2_1$ as $n \to \infty$. However, in finite samples, the convergence in distribution can be slow (see Figure

3.4). The distance between the data-generating parameter $\theta \in \Theta_0$ and the origin impacts the sampling distribution of λ in finite samples. Again, this leads us to the conclusion that the non-zero regression coefficient acts as a nuisance parameter.

3.3.5 A conditional test of the indirect effect

Our aim is to create a conditional procedure that mitigates the effect of the nuisance parameter on the likelihood ratio test's level. Using Fisher's language, we would like to account for a particular dataset's "configuration" when performing inference. Above, we showed that the non-zero regression coefficient affects the sampling distribution of our test statistic. A conditional approach to inference will attempt to account for this configuration.

To account for the problem configuration, we propose conditioning on $A = \max\{\lambda_m, \lambda_y\}$, which will be called the "complementary likelihood ratio test statistic." The complementary likelihood ratio test statistic is a natural choice because it measures the distance in log-likelihood units from the unconstrained MLE to the region of the null parameter space that the unconstrained MLE was not projected to.

The conditional distribution of λ given A will vary less as the non-zero regression coefficient changes. Additionally, since $A > \lambda$ by definition, conditioning on A results in a truncated distribution supported on the interval $(0, A)$. Although the conditional distribution of $\lambda|A$ will still depend on the data-generating parameter $\theta \in \Theta_0$, given A, $\lambda|A$ has the same support for all $\theta \in \Theta_0$.

To the best of our knowledge, the distribution of $\lambda|A$ does not exist in a known analytical form in finite samples. In Section 3.4, we will derive an asymptotic approximation to the conditional distribution of $\lambda|A$. The approximation will depend on two assumptions that will not necessarily hold in finite samples. First, it requires that λ_m and λ_y are independent. Second, the χ^2 approximation to the log of the likelihood ratio test statistics must be reasonable. We will also propose a non-standard bootstrap approach to learning the distribution of λ given A which does not rely on either of these assumptions.

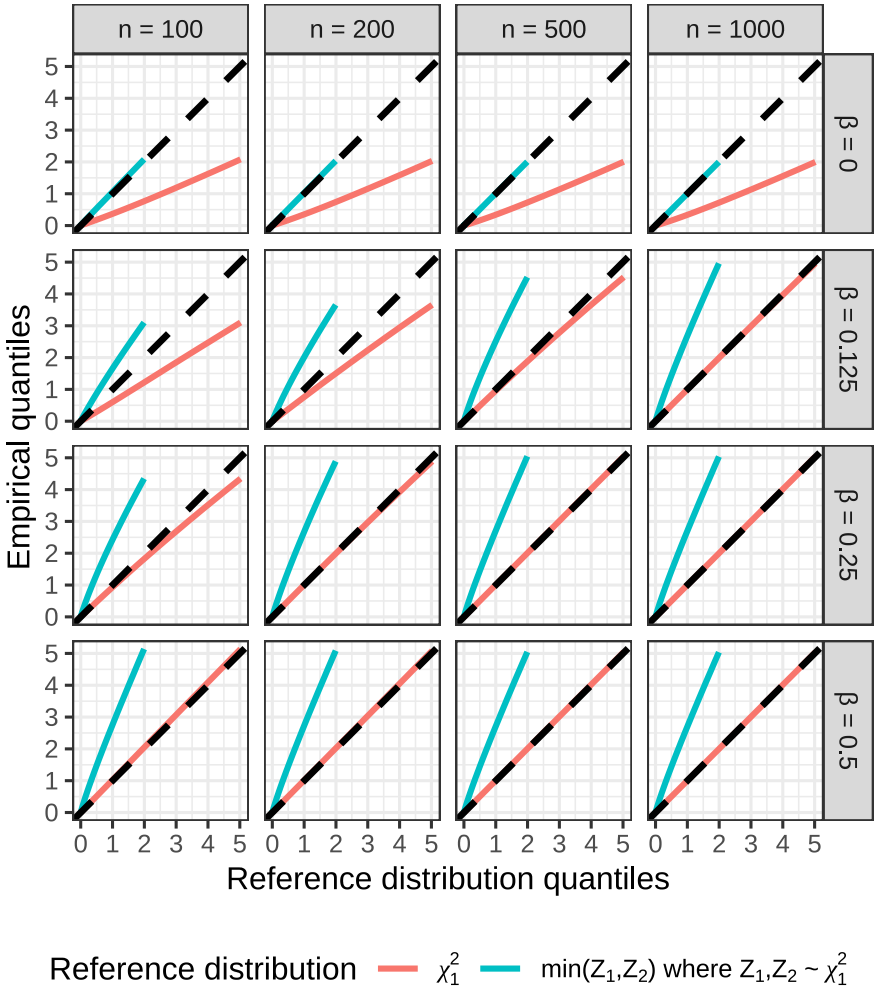


Figure 3.4: QQ-plots of empirical likelihood ratios' quantiles plotted against two theoretical reference sampling distributions. Results are stratified by data-generating population (rows) and sample size (columns). For each setting, $\alpha = 0$, so the likelihood ratio test statistics (LRT) are sampled from a population in which the indirect effect is zero. When both $\alpha = \beta = 0$, the LRTs behave like the minimum of two independent χ^2_1 random variables for all sample sizes, as theory predicts. As β increases, there's a transition from this reference distribution to the expected χ^2_1 reference distribution. The transition occurs more quickly for larger sample sizes.

A bootstrap approach to learning the distribution of λ given A

To begin we offer a high-level overview of the proposed bootstrap approach that explains the intuition underlying each part of the procedure. The method is made up of two rounds of bootstrapping. The first round takes a non-parametric bootstrap sample of the observed data. This captures uncertainty in the "true" null-space ($\alpha = 0$ or $\beta = 0$). When the data-generating parameter $\theta \in \Theta_0$ is close to the (α, β)-origin, over repeated sampling, the constrained MLE will be projected to the $\alpha = 0$ and $\beta = 0$ axes.

The second round of bootstrap sampling produces samples of (λ, A) from populations in which $\alpha\beta = 0$ is true. After taking a non-parametric bootstrap sample of the observed data, the procedure fits the null GLM-mediation model to the bootstrap data. Parametric bootstrap samples are then drawn from the fitted null model. The procedure uses the parametric bootstrap samples to produce (λ, A) pairs for which the no-mediation hypothesis is true in the population.

Finally, the test of the indirect effect is evaluated by estimating the conditional distribution $\hat{F}(\lambda|A)$ using the bootstrap samples of (λ, A). Then, using the observed values of the likelihood ratios (λ_{obs}, A_{obs}), an approximate p-value can be calculated using $\hat{F}$:

$$p = 1 - \hat{F}^{-1}(\lambda_{obs}|A = A_{obs}).$$

Before describing several approaches to estimating the conditional distribution $F(\lambda|A)$, the two-level bootstrap procedure is given in Algorithm 3.1. Suppose that the observed data is denoted $Z = \{(X_i, M_i, Y_i)\}_{i=1}^n$. We will use $\mathbb{P}$ to denote a GLM-mediation model.

Algorithm 3.1: Learning the conditional distribution of λ given A.

Result: Test decision for $H_0 : \alpha\beta = 0$ at level α

Data: $Z = \{(X_i, M_i, Y_i)\}_{i=1}^n$

1. Calculate λ_{obs} and A_{obs} from Z.

for $j = 1, \ldots, n_b$ **do**

2.1 Take a non-parametric bootstrap sample Z_j from Z.

2.2 Let $\hat{\theta}_0^j$ be the estimated null model using Z_j

for $k = 1, \ldots, n_p$ **do**

2.2.1 Take a parametric bootstrap sample $\tilde{Z}_{jk}$ from $\mathbb{P}(\hat{\theta}_0^j)$

2.2.2 Using $\tilde{Z}_{jk}$, estimate λ_{jk} and A_{jk}

end

end

3. Estimate the conditional cumulative distribution function (CDF) $F(\lambda^*|A = a)$ using $\{(\lambda_{jk}, A_{jk}) : j = 1, \ldots, n_b, k = 1, \ldots, n_p\}$.

4. If $\lambda_{obs} > F^{-1}(1 - \alpha|A = A_{obs})$, reject H_0 at level α.

To conclude this section, we describe several methods for estimating the conditional distribution of λ given A. To the best of our knowledge, the conditional distribution does not have an known analytical form. Thus, we will rely on statistical models to estimate conditional quantiles or p-values. These models will localize our estimate of the conditional distribution to the region where $A \approx A_{obs}$ using the samples of (λ, A) produced by the bootstrap procedure. In the next section, we will show that the conditional quantiles of λ given A are smooth but neither linear nor monotone, making estimation nontrivial.

Proposed methods for estimating the distribution of λ given A

1. **The unconditional bootstrap p-value**. If we choose not to condition on the value of the ancillary statistic, we can use the ordinary bootstrap p-value:

$$p_{boot} = \frac{\#\{j : \ell_j > \lambda_{obs}\}}{\tilde{n}}.$$

We will show through simulation that these p-values lead to conservative tests when H_0 is true and θ_0 is close to the (α, β)-origin.

2. **Plug-in p-value**. It was shown earlier that $\mathbb{P}(\lambda > x) = \mathbb{P}(\lambda_m > x)\mathbb{P}(\lambda_y > x)$. Without loss of generality, assume that $\hat{\alpha}_0 = 0$, $\hat{\beta}_0 \neq 0$, so that asymptotically $\lambda_m \sim \chi^2_1$ and $\lambda_y \sim \chi^2_1(\mu_y)$. Using $\hat{\theta}_0$, determine the plug-in estimator $\hat{\mu}_y$ of μ_y, and use it to calculate $p = \mathbb{P}(\chi^2_1 > \lambda_{obs})\mathbb{P}(\chi^2_1(\hat{\mu}_y) > \lambda_{obs})$.

3. **Kernel regression**. The conditional p-value $\hat{\mathbb{P}}(\lambda^* > \lambda | A = a)$ given bootstrap samples $\{\ell_j, a_j\}_{j=1}^{\tilde{n}}$ is defined as follows:

$$p_{cb} = \frac{\#\{j : \ell_j > \lambda_{obs}, a_j = A_{obs}\}}{\tilde{n}}.$$

Of course, in practice, $a_j \neq A_{obs}$ for any j. In order to estimate, p_{cb} one could introduce Gaussian kernel weights w_j for $j = 1, \ldots, \tilde{n}$ where

$$w_j = \frac{e^{-h(a-a_j)^2}}{\sum_{k=1}^{\tilde{n}} e^{-h(a-a_k)^2}}.$$

These weights will have the effect of localizing the estimator. A bootstrap sample (ℓ_j, a_j) will receive greater weight when its ancillary statistic falls near A_{obs}.

The kernel regression p-value is defined as follows:

$$p_{cb} = \sum_{j=1}^{\tilde{n}} w_j \times \mathbb{I}(\ell_j > \lambda_{obs}),$$

where $\mathbb{I}$ returns 1 when $\ell_j > \lambda_{obs}$ and is 0 otherwise.

4. **Quantile regression**. Finally, one could estimate the conditional quantile function of λ^* given that $A = a$. Suppose that you wish to evaluate the null hypotheses using an α level test. Let $Q_\alpha(a) = F_\lambda^{-1}(\alpha \mid A = a)$ be the conditional quantile function ($\mathbb{P}(\lambda < Q_\alpha(a)|A = a) = \alpha$). Using the $\tilde{n}$ bootstrap samples, we estimate $Q_{1-\alpha}$ and the critical value $\hat{\lambda}_{1-\alpha,a_{obs}} = \hat{Q}_{1-\alpha}(a_{obs})$. We then will choose to reject the null hypothesis when $\lambda_{obs} > \hat{\lambda}_{1-\alpha,a_{obs}}$.

   We will see that the conditional quantile function of λ given A is non-monotone and non-linear. As a result, we propose using quantile smoothing splines in order to estimate the conditional quantile function. Our hope is that a regression approach will use information efficiently from the bootstrap samples, requiring fewer bootstrap iterations to create a properly calibrated test.

3.4 An asymptotic approximation of the conditional sampling distribution

Although the sampling distribution of λ given A is unknown in finite samples, it can be approximated by its limiting distribution. In certain settings, it may be appropriate or one might want to calculate approximate p-values using the limiting distribution instead of the bootstrap procedure described in the previous section.

3.4.1 The asymptotic marginal distributions of λ_m and λ_y

The asymptotic approximation of the tests' power function requires knowing the asymptotic distribution of λ_m and λ_y, which we will now derive. This amounts to determining the asymptotic distribution of a likelihood ratio test of the nullity of a single parameter for a generalized linear model. The test statistic's distribution will be derived in the setting of the conditional model for Y given X and M, to make the presentation clearer.

Let $Y \sim F$, where F is an exponential family and suppose that $\mathbb{E}[Y|X, M] = g^{-1}(\beta_0 + \beta_1 X + \beta_2 M) = \mu$ and $\text{Var}[Y|X, M] = V(\mu, \sigma^2)$. Let $\theta = (\beta_0, \beta_1, \beta_2, \sigma^2)$, and define $\ell(\theta|Y)$ to be the likelihood function of F evaluated at θ. Our interest is in deriving the distribution of the likelihood ratio test statistic testing $H_0 : \beta_2 = 0$. When H_0 is true, standard results show that the likelihood ratio converges to a χ^2_1 random variable. When $\beta_2 \neq 0$, the distribution of the likelihood ratio will converge to a non-central χ^2_1 random variable with non-centrality parameter μ. The rest of this section derives an expression for this non-centrality parameter. The derivation uses results from Chapter 17 Section 2 of [36].

The null parameter space $\Theta_0 = \{(a_1, a_2, 0, b) \mid a_1, a_2 \in \mathbb{R}, b \in \mathbb{R}^+\}$ is linear, meaning that the tangent space is identical for all $\theta_0 \in \Theta_0$. The projection operator onto the tangent plane V for all $\theta_0 \in \Theta_0$ is given by

$$P_0 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{3.4}$$

The projector onto $V^\perp$ will be denoted $Q_0 = I_4 - P_0$.

Define $\Delta = (0, 0, \sqrt{n}\beta_2, 0)^T$. The non-centrality parameter μ has the following expression:

$$\mu = \Delta^T Q_0 \left(P_0 + Q_0 I(\theta)^{-1} Q_0\right)^{-1} Q_0 \Delta, \tag{3.5}$$

where $I(\theta)$ is the Fisher information of θ. All elements of the matrix product $Q_0 I(\theta)^{-1} Q_0$ are equal to zero except the $(3, 3)$ entry, which is equal to (after root n scaling) the asymptotic sampling variance of $\hat{\beta}_2$, say ν. Then, $(P_0 + Q_0 I(\theta)^{-1} Q_0)^{-1} = \text{diag}(1, 1, 1/\nu, 1)$. Pre- and post-multiplication by $\Delta^T Q_0$ gives

$$\mu = \frac{n\beta_2^2}{\nu}. \tag{3.6}$$

Unsurprisingly, the non-centrality parameter grows with both the sample size n and the signal-to-noise ratio β_2^2/ν. In the Gaussian linear model setting, if X and M are independent, then

$$\mu = \frac{n\beta_2^2\sigma_m^2}{\sigma_y^2}, \tag{3.7}$$

which is the form that is used in Section 3.4. In practice, one can estimate $I(\theta)$ using the observed Fisher information

$$\mathcal{I}(\theta) = -\nabla_\theta^2 \ell(\theta)\big|_{\theta=\hat{\theta}}, \tag{3.8}$$

and use this estimate to form a plug-in estimate of the non-centrality parameter μ.

We have shown that the asymptotic sampling distributions of λ_m and λ_y are non-central χ_1^2 distributions with non-centrality parameters μ_m and μ_y, taking the form given in 3.5. Our later analysis will make use of the distributions to develop some intuition for the power we expect to have for a fixed data-generating population and sample size.

3.4.2 The asymptotic independence of λ_m and λ_y

Our derivation of the asymptotic power function of the conditional test depends on the independence of the two log-likelihood ratios λ_m and λ_y. Because the proposed bootstrap procedure does not depend on the proposition, we do not attempt to prove Proposition 1, but do offer a small simulation study that suggests that the proposition is indeed true or nearly true.

Proposition 1:

The log-likelihood ratio test statistics λ_m and λ_y are independent as $n \to \infty$.

For a variety of parameter settings $(\alpha, \beta) \in \Theta_0$ and sample sizes n, we sampled $n_s = 10,000,000$ independent copies of (λ_m, λ_y). Using the set $S = \{(\lambda_m, \lambda_y)_i, i = 1, \ldots, n_s\}$, we estimated the probabilities

$$p_g := \mathbb{P}(\lambda_y < x | \lambda_m > a) \text{ and } p_\ell := \mathbb{P}(\lambda_y < x | \lambda_m < a). \tag{3.9}$$

for fixed a and $x \in \mathbb{R}$. Note that if the random variables λ_m and λ_y are independent, then the two probabilities should be equal at the population level for any value of a and x.

The data-generating parameters are varied in order to show that across all combinations, for large enough n, $p_\ell/p_g \to 1$. We set

$$(\alpha, \beta) \in \{(0, b) : b \in \{0.0, 0.125, 0.25, 0.5, 0.75, 1.0\}\},$$

and consider eight samples sizes n, equally spaced between 10^2 and 10^4 on the $\log_{10}$-scale. We show results for a single value of a, chosen to be the empirical median of the generated λ_m statistics. Additionally, we consider three values of x, denoted x_1, x_2 and x_3, which represent the quartiles of the empirical λ_m distribution. Figure 3.5 plots the ratio

$$p_\ell/p_g = \frac{\mathbb{P}(\lambda_y < x_j \mid \lambda_m < a_{0.5})}{\mathbb{P}(\lambda_y < x_j \mid \lambda_m > a_{0.5})}$$

against the sample size n on the log-scale for different quantiles x_j and data-generating populations. If λ_m and λ_y are asymptotically independent, then $p_\ell/p_g \to 1$.

The results of the simulation study (see Figure 3.5) suggest that Proposition 1 is true. For each value of x and data-generating population (indexed by β), as n increases, the ratio of probabilities $p_\ell/p_g \to 1$. This appears to hold for all values of a, although the results for a single a are included here. Convergence is affected by both the data-generating population and the quartile (x_1, x_2, or x_3).

When both $\alpha = \beta = 0$, the ratio $p_\ell/p_g \approx 1$ for all values of n, suggesting that the statis-

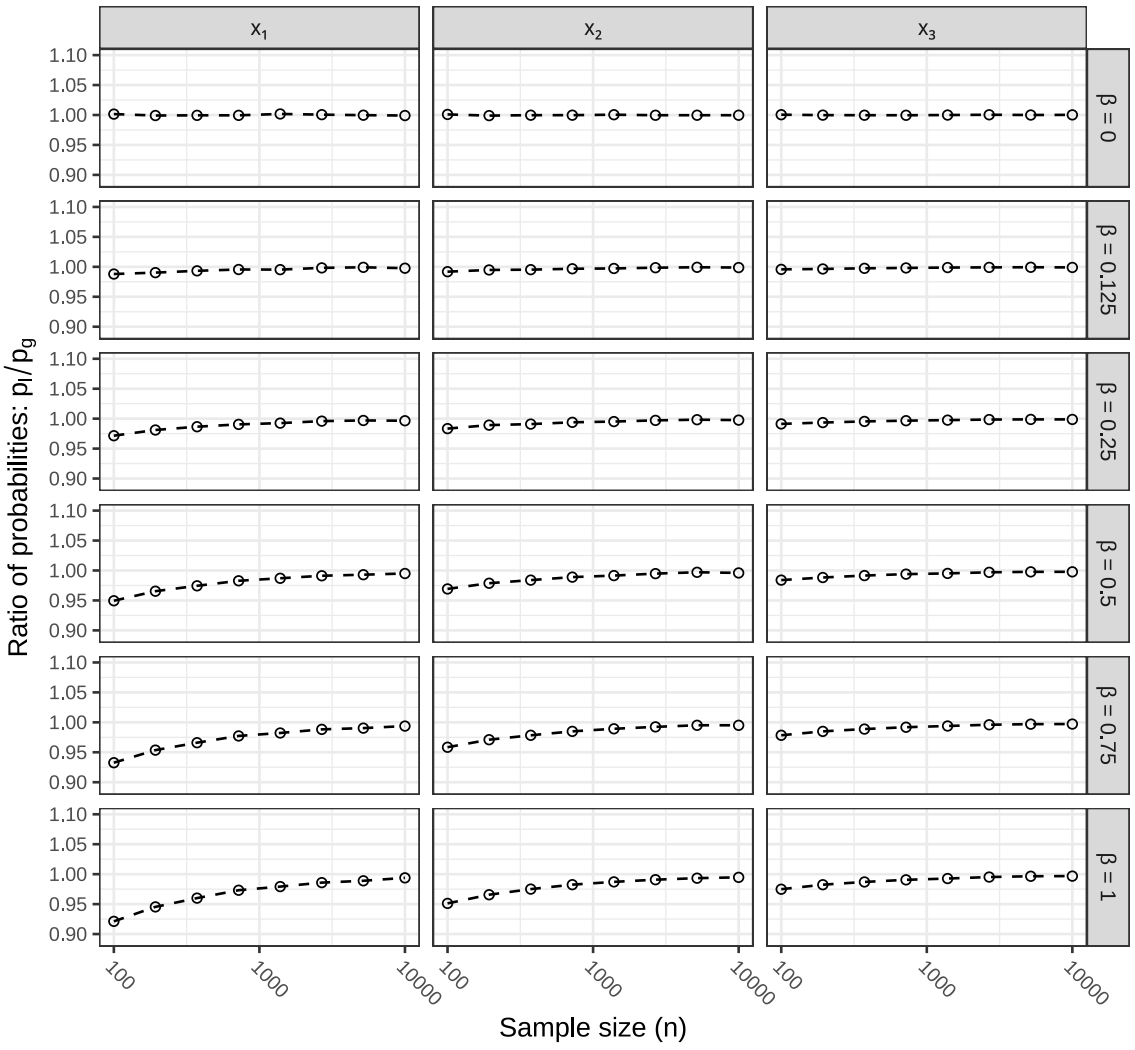


Figure 3.5: Empirical evidence that the log-likelihood ratios λ_m and λ_y are asymptotically independent. Each panel plots the ratio p_ℓ/p_g (defined in Section 3.4.2) against sample size n for different data-generating populations (rows) and different quantiles of the λ_m distribution (columns). For large enough n, the ratio of probabilities are all approximately 1, suggesting that λ_m and λ_y are asymptotically independent.

tics are independent in finite samples in this setting. This is expected as M is independent of X and M is independent of Y both marginally and conditional on X. Convergence appears to be slower as $\beta > 0$ grows across all values of x. This simulation study suggest that λ_m and λ_y are asymptotically independent, justifying our derivation of the power function of the conditional test.

3.4.3 Calculating approximate p-values via the asymptotic sampling distribution of λ given A

In order to derive the power function of the conditional test, we must be able to calculate tail probabilities of the distribution of λ given A. This amounts to determining the conditional distribution function of the minimum given the maximum of two independent random variables. We provide a general derivation of this distribution, and later will apply the derivation to the setting where both random variables are χ^2_1 random variables, potentially with non-zero non-centrality parameters.

Let $X \sim F$ and $Y \sim G$ be independent and continuous distributed random variables. Let

$$F(z) = \mathbb{P}(X < z) \text{ and } G(z) = \mathbb{P}(Y < z).$$

and let the respective density functions be denoted by f and g.

Define $\lambda = \min\{X, Y\}$ and $A = \max\{X, Y\}$. Our goal is to describe the conditional distribution of λ given A. We begin by finding the joint distribution and density functions of (λ, A). Our approach is to determine the joint distribution function $\mathbb{P}(\lambda < w, A < z)$ and then differentiate with respect to w and z in order to find the joint density. When $w < z$, the event $\{\lambda < w, A < z\}$ occurs if either $\{X < w, Y < z\}$ or $\{X < z, Y < w\}$. This then gives us:

$$\begin{aligned}
\mathbb{P}(\lambda < w, A < z) &= \mathbb{P}\big(\{X < w, Y < z\} \cup \{X < z, Y < w\}\big) \\
&= \mathbb{P}\big(\{X < w, Y < z\}\big) + \mathbb{P}\big(\{X < z, Y < w\}\big) - \\
&\qquad\qquad \mathbb{P}\big(\{X < w, Y < z\} \cap \{X < z, Y < w\}\big) \\
&= \mathbb{P}\big(\{X < w, Y < z\}\big) + \mathbb{P}\big(\{X < z, Y < w\}\big) - \mathbb{P}\big(\{X < w, Y < w\}\big) \\
&= F(w)G(z) + F(z)G(w) - F(w)G(w).
\end{aligned}$$

The density of (λ, A) is then equal to the first mixed partial derivative of $\mathbb{P}(\lambda < w, A < z)$:

$$\begin{aligned}
f_{(\lambda,A)}(w, z) &= \frac{\partial^2}{\partial w \partial z} \mathbb{P}(\lambda < w, A < z) \\
&= \frac{\partial^2}{\partial w \partial z} \left\{F(w)G(z) + F(z)G(w) - F(w)G(w)\right\} \\
&= f(w)g(z) + f(z)g(w).
\end{aligned}$$

Next, we turn to the marginal distribution of A. Here we have the familiar expression

$$\mathbb{P}(A < z) = \mathbb{P}(X < z, Y < z) = F(z)G(z),$$

the last inequality holding due to the independence of X and Y. We again differentiate with respect to z in order to determine the density function of A:

$$f_A(z) = \frac{\partial}{\partial z} \mathbb{P}(A < z) = \frac{\partial}{\partial z} F(z)G(z) = f(z)G(z) + F(z)g(z).$$

Thus, the joint density function of (λ, A) is:

$$f_{\lambda|A}(x|a) = \frac{f_{(\lambda,A)}(x,a)}{f_A(a)} = \frac{f(w)g(z) + f(z)g(w)}{f(a)G(a) + F(a)g(a)}.$$

To close this section, we calculate $\mathbb{P}(\lambda > w|A = a)$ when $w < a$. Using the joint density function, we have that

$$\begin{aligned}
\mathbb{P}(\lambda > w|A = a) &= \int_w^a f_{\lambda|A}(x|a)\partial x \\
&= \frac{1}{f(a)G(a) + F(a)g(a)} \int_w^a f(w)g(a) + f(a)g(w)\partial x \\
&= \frac{\big(F(a) - F(w)\big)g(a) + \big(G(a) - G(w)\big)f(a)}{f(a)G(a) + F(a)g(a)} \\
&= 1 - \frac{F(w)g(a) + G(w)f(a)}{G(a)f(a) + F(a)g(a)}.
\end{aligned} \tag{3.10}$$

Equation 3.3 gives a formula for a p-value for the conditional procedure when F and G are χ^2_1 distributions with non-centrality parameters μ_m and μ_y.

3.4.4 Characterizing the conditional distribution of λ given A

Using the asymptotic approximation to the sampling distribution of $\lambda|A$ derived in Sections 3.4.1 and 3.4.3, we estimate the conditional quantiles when the non-centrality parameter $\mu = 20.0$, which corresponds to the setting where $n = 500$, $\alpha = 0$, $\beta = 1/25$ and the asymptotic sampling variance of β is equal to 1. Let F_a to be the conditional distribution function of λ given $A = a$, so that $F_a(x) = \mathbb{P}(\lambda < x|A = a)$. Denote the inverse CDF of $\lambda|A = a$ by $F_a^{-1}(\alpha') = q$ so that $F_a(q) = \alpha'$. Figure 3.6 plots $F_a^{-1}(\alpha')$ for $\alpha' \in \{0.5, 0.9, 0.95, 0.99\}$.

The conditional distribution of λ given $A = a$ has non-monotone, non-convex quantiles in a. Since $A > \lambda$, the distribution of λ given $A = a$ is truncated at a. As the value of a increases, the conditional quantiles of λ increase as well. For large a, the conditional quantile function becomes nearly constant for all α', since once A is large, the effect of

truncation on the conditional distribution is negligible. In fact, in this region, the conditional distribution of λ given A behaves approximately like a χ^2_1 random variable. In fact, for moderate a the conditional quantile function decreases. This is due to the fact that the conditional quantiles are inflated for moderate a, as the conditional quantile function is a mixture of one central χ^2_1 and one non-central χ^2_1 random variable.

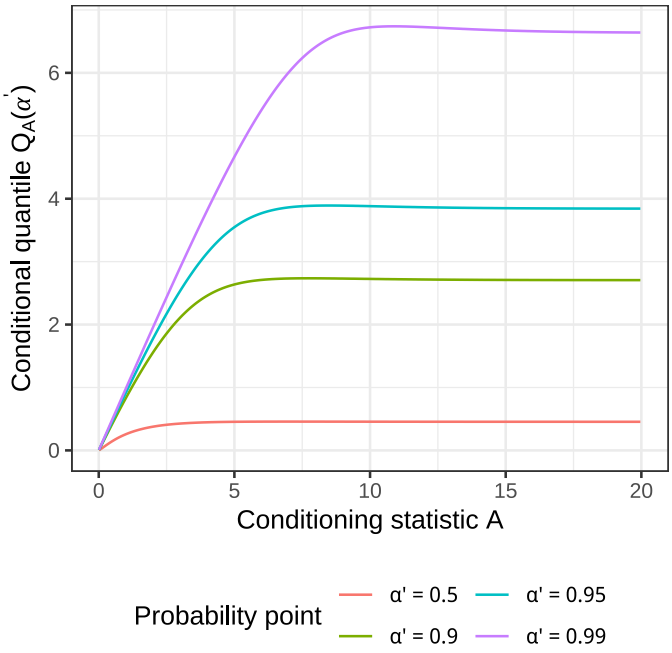


Figure 3.6: The conditional quantiles of λ plotted against A based on the asymptotic sampling distribution of λ given A. Conditional quantiles for four probability points $\tilde{\alpha}$ are shown. This conditional quantile function is specific to data-generating populations in which the non-zero non-centrality parameter equals 20. The conditional quantile function is non-monotone and non-convex in A. As A increases, the conditional quantile functions converge to the quantiles of the χ^2_1 distribution. This is because when A is large, λ given A behaves like a χ^2_1 random variable.

3.4.5 An asymptotic approximation of the power function

We begin this section by deriving an expression for the asymptotic power of the conditional test and then compare the theoretical power functions of the conditional and marginal likelihood ratio tests. Although competing procedures do not use the marginal likelihood ratio

test, asymptotic analysis of confidence interval-based tests would be difficult and is beyond the scope of this chapter. Comparisons between our conditional test and confidence interval-based approaches will be made using simulation studies in Section 3.5.

Theoretical power of the conditional test

We wish to approximate the power of the conditional procedure testing $H_0 : \alpha\beta = 0$ at level $\tilde{\alpha}$. The derivation will rely on both Proposition 1 and the asymptotic χ^2 behavior of likelihood ratio test statistics. The following expression is equal to the power of the conditional test:

$$\mathbb{P}(\text{reject } H_0) = \mathbb{P}(\text{reject } H_0|\lambda = \lambda_m)\mathbb{P}(\lambda = \lambda_m)+\mathbb{P}(\text{reject } H_0|\lambda = \lambda_y)\mathbb{P}(\lambda = \lambda_y). \quad (3.11)$$

Assuming that Proposition 1 is true,

$$\begin{aligned} \mathbb{P}(\lambda = \lambda_m) &= \mathbb{P}(\lambda_m < \lambda_y) \\ &= \int_0^\infty \mathbb{P}(\lambda_m < x) f_{\mu_y}(x) dx, \end{aligned} \quad (3.12)$$

where f_{μ_y} is the density function of the $\chi^2_1(\mu_y)$ distribution and $\mathbb{P}(\lambda_m < x)$ is given by the CDF of the $\chi^2_1(\mu_m)$ distribution.

Next, we turn to finding an expression for $\mathbb{P}(\text{reject } H_0|\lambda = \lambda_m)$. Because the conditional procedure conditions on the value $A = \max\{\lambda_m, \lambda_y\}$, we first determine the conditional probability $\mathbb{P}(\text{reject } H_0|\lambda = \lambda_m, A = a)$ and then integrate over the distribution of A to get $\mathbb{P}(\text{reject } H_0|\lambda = \lambda_m)$.

First, the critical value of the test must be found for fixed $A = \lambda_y$. Assuming that H_0 is true and since $\lambda = \lambda_m$, the asymptotic approximations are $\lambda_m \sim \chi^2_1(0)$ and $\lambda_y \sim \chi^2_1(\mu_y)$. Using the results of Section 3.4.3, the distribution function of λ_m given $A = a$, which we denote G_0, can be expressed as the density and distribution function of the χ^2_1 and $\chi^2_1(\mu_y)$

distributions. The inverse CDF G_0^{-1} gives the $1-\tilde{\alpha}$ quantile, $q_{1-\tilde{\alpha},a}$ of the distribution λ given $A = a$, which can be determined numerically using standard root-finding algorithms.

The conditional power depends on the true distribution of λ_m, which may be a non-central χ_1^2 distribution. Again using results for Section 3.4.3, the true sampling distribution of λ_m given A can be determined, which we will denote G_1. Then $\mathbb{P}(\text{reject } H_0|\lambda = \lambda_m, A = a) = 1 - G_1(q_{1-\tilde{\alpha},a})$. Integrating the function $1 - G_1(q_{1-\tilde{\alpha},a})$ against the $\chi^2(\mu_y)$ gives

$$\mathbb{P}(\text{reject } H_0|\lambda = \lambda_m) = \int_0^\infty (1 - G_1(q_{1-\tilde{\alpha},a}))f_{\mu_y}(a)da.$$

This derivation is symmetric in λ_m and λ_y, and is also valid for the term $\mathbb{P}(\text{reject } H_0|\lambda = \lambda_y)$. Putting these parts together gives a calculable expression for $\mathbb{P}(\text{reject } H_0)$, which is a function of μ_m and μ_y, say $P_c(\mu_m, \mu_y)$.

As a point of comparison, we also derive the power function of the marginal likelihood ratio test. Let $\tilde{q}_{1-\tilde{\alpha}}$ be the $1-\tilde{\alpha}$ quantile of the χ_1^2 distribution. If one were to naively conduct a likelihood ratio test, the χ_1^2 would be the asymptotic reference distribution of the test statistic. Again, assuming that Proposition 1 is true, the power of the marginal test is

$$\begin{aligned} \mathbb{P}(\text{reject}H_0) &= \mathbb{P}(\lambda_m > \tilde{q}_{1-\tilde{\alpha}}, \lambda_y > \tilde{q}_{1-\tilde{\alpha}}) \\ &= \mathbb{P}(\lambda_m > \tilde{q}_{1-\tilde{\alpha}})\mathbb{P}(\lambda_y > \tilde{q}_{1-\tilde{\alpha}}). \end{aligned} \tag{3.13}$$

The tail probabilities can be calculated using the distribution functions of the $\chi_1^2(\mu_m)$ and $\chi_1^2(\mu_y)$ distributions. The marginal power function will be denoted $P_m(\mu_m, \mu_y)$.

Figure 3.7 compares the approximations to the asymptotic power functions of the conditional and marginal likelihood ratio tests. The power functions will be denoted $P_c(\mu_m, \mu_y)$ and $P_m(\mu_m, \mu_y)$, respectively. The power of each procedure is determined through numeric integration of the power functions given above. The conditional test is more powerful than the marginal test over all considered values of μ_m and μ_y. However, the conditional test is also anti-conservative (see the panels labeled "μ_2 = 0" in right hand column of Figure 3.7).

This analysis provides two important insights into the nature of this testing problem.

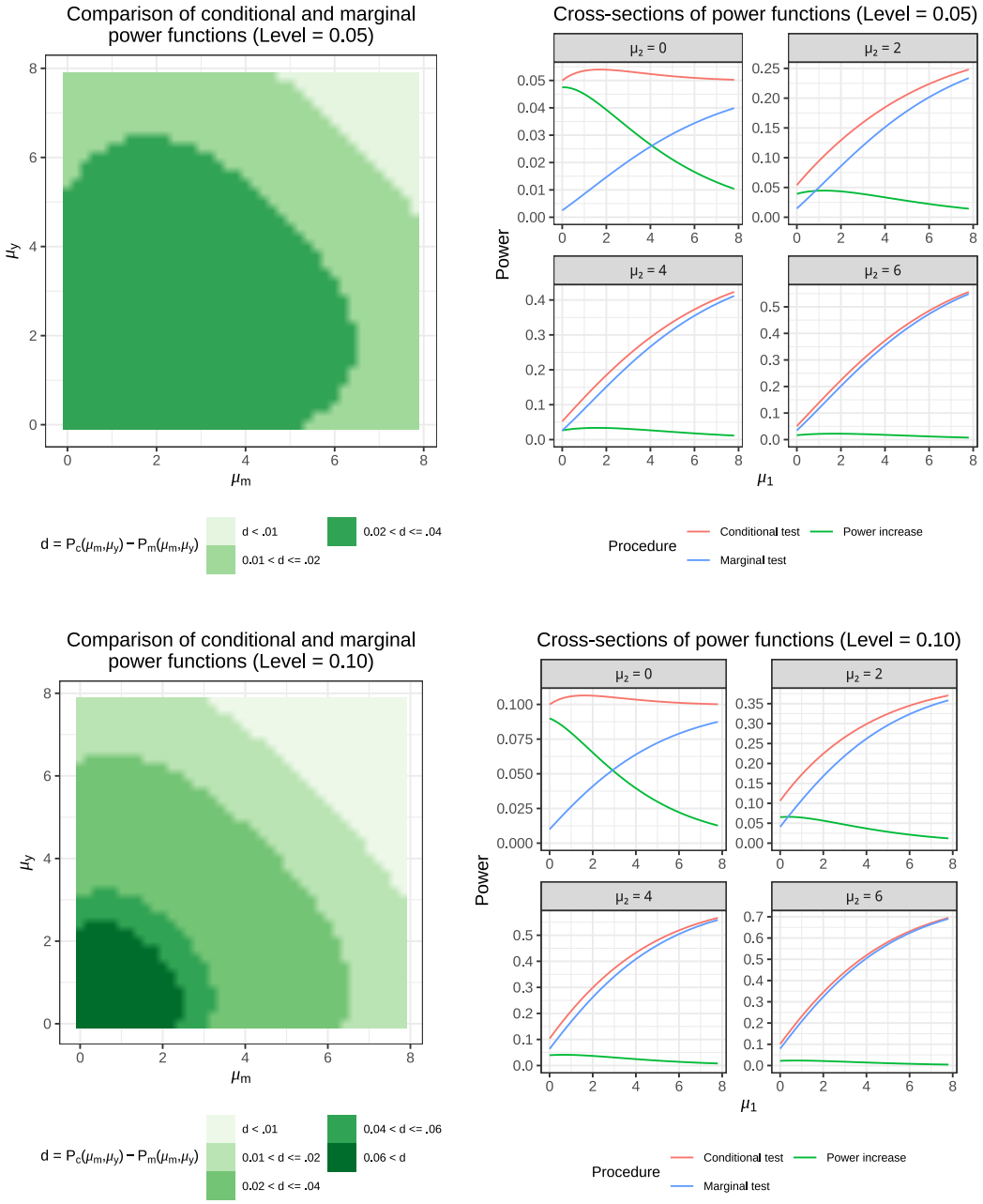


Figure 3.7: Comparisons of the conditional and marginal power functions over a grid of (μ_m, μ_y) pairs. Two test levels are considered; 0.05 and 0.10% for the top and bottom rows, respectively. The right hand column shows the power increase of the conditional test over the marginal test. Darker green colors denote larger increases in power. The right column plots power of the tests against the first non-centrality parameter μ_1, stratified over four values of $\mu_2 \in \{0.0, 2.0, 4.0, 6.0\}$. The conditional (red line), marginal (green line), and the difference in power (green line) are plotted against μ_1.

First, the conditional approach is more powerful than the marginal likelihood ratio test for certain null hypotheses. Simulation studies will offer additional evidence that this is the case for two other marginal tests. Second, the analysis shows that the theoretical power function exceeds its nominal level for a range of data-generating parameter settings in the null parameter space. The power function exceeds its nominal level by at most 8 and 6 percent for tests at the 0.05 and 0.10 levels, respectively. Simulation studies will show that the proposed bootstrap procedure partially accounts for the anti-conservatism.

The anti-conservatism is caused by conditioning on the incorrect likelihood ratio test statistic and using the non-null likelihood ratio as the test statistic. The theoretical power function converges to the test's nominal level, since the probability of conditioning on the incorrect likelihood ratio decreases as the non-zero non-centrality parameter grows. In these settings, the test statistic's sampling distribution is stochastically larger than the bootstrapped estimate of the sampling distribution, resulting in anti-conservative test. The bootstrap procedure attempts to calibrate the sampling distribution of the test statistic, but does not achieve perfect recalibration as demonstrated in the following section.

3.5 Simulation studies

This section presents simulation studies to demonstrate that the conditional test has desirable properties. First, we show that the method approximately achieves its target level and has better power than competitor methods. We then explore how the performance of the method depends on the choice of the bootstrap parameters.

3.5.1 Approximate level control

This simulation study demonstrates that the conditional inference approach approximately achieves its level when the null hypothesis $H_0 : \alpha\beta = 0$ is true. Data are generated from the Gaussian linear model:

$$
\begin{aligned}
M &= \gamma_1 + \alpha X + \epsilon_m \\
Y &= \gamma_2 + \gamma_3 X + \beta M + \epsilon_y,
\end{aligned}
\tag{3.14}
$$

where ϵ_m and ϵ_y are independent with variances σ_m^2 and σ_y^2, respectively.

The behavior of the conditional likelihood ratio test is a function of the two non-centrality parameters. For this statistical model, the non-centrality parameters are:

$$
\mu_m = \left(\frac{\sqrt{n}\alpha\sigma_x}{\sigma_m} \right)^2 \text{ and } \mu_y = \left(\frac{\sqrt{n}\beta\sigma_m}{\sigma_y} \right)^2. \tag{3.15}
$$

Therefore, without loss of generality, we assume that $\sigma_x = \sigma_m = \sigma_y = 1$, so that $\mu_m = n\alpha^2$ and $\mu_y = n\beta^2$.

Two sample sizes, $n = 250$ and $n = 1,000$ will be used in this simulation study. The regression coefficients α and β will be set so that the same non-centrality parameters are considered for each sample size. The non-centrality parameters will fall in the set $\mu = \{0.0, \ldots 25.0\}$. Since $\alpha\beta = 0$ for all data-generating populations, at most one non-centrality parameter will be non-zero for each data-generating population.

For each sample size and data-generating population, 20,000 Monte Carlo synthetic datasets were generated. The conditional inference procedure was run on each using 500 nonparametric and 20 parametric bootstrap samples for a total of 10,000 total bootstrap samples of the statistics (λ_b, A_b). Conditional p-values were calculated using the kernel weights for three effective sample sizes: 250, 500, and 1,000. These effective sample sizes correspond to using 2.5, 5, and 10% of the total bootstrap samples when conditioning. We compare the level of the conditional test in each setting to the level of the bootstrap confidence interval-based test, which is the most commonly used procedure for testing the indirect effect. This test declares that M is a mediating variable at the $\tilde{\alpha}$-level if 0 falls outside of a $(1-\tilde{\alpha})\%$ confidence interval for $\alpha\beta$. The conditional inference test rejects the null hypothesis $H_0 : \alpha\beta = 0$ at the $\tilde{\alpha}$ level if the conditional bootstrap p-value was less than $\tilde{\alpha}$.

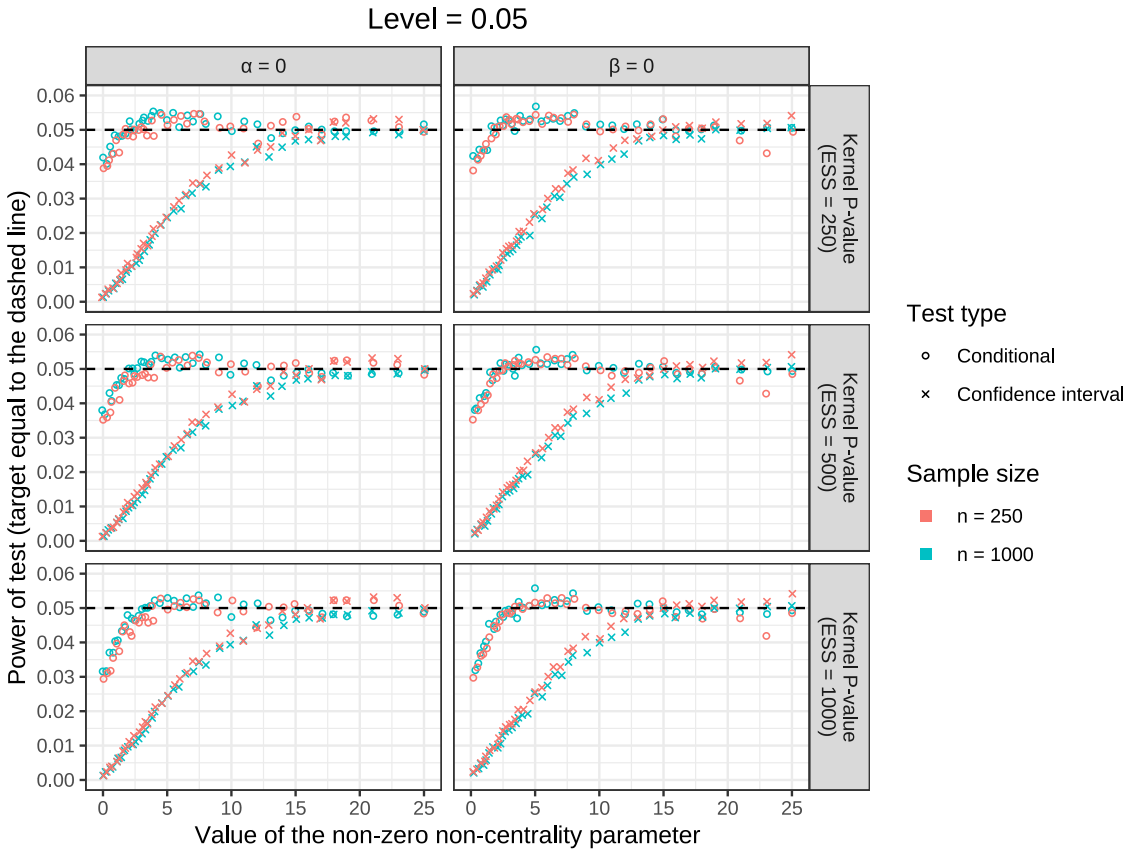


Figure 3.8: Results from a simulation study exploring the calibration of the conditional procedure at the $\tilde{\alpha} = 0.05$ level when $\alpha\beta = 0$ in the population. The estimated level of both the conditional and marginal tests of the indirect effect are plotted against the value of the non-zero non-centrality parameter. Results are stratified by data-generating population (columns) and the effective sample size used to calculate conditional p-values (rows). Results for two sample sizes are included (distinguished by color) and tests (distinguished by point shape). Overall, the conditional procedure is better calibrated than the marginal procedure.

Overall, the results shown in Figures 3.8 and 3.9 are positive. Each method is conservative for small values of the non-zero non-centrality parameter, which we will denote by μ for the remainder of this section. Consider the case when both $\alpha = \beta = 0$. In this setting, both λ_m and λ_y are marginally $\chi^2_1(0)$ random variables. However, the bootstrap procedure will sample (λ, A) from populations in which one non-centrality parameter is greater than 0. As a result, the conditional quantiles of the bootstrapped estimate of the sampling distribution will be larger than the true sampling distribution, leading to a conservative test.

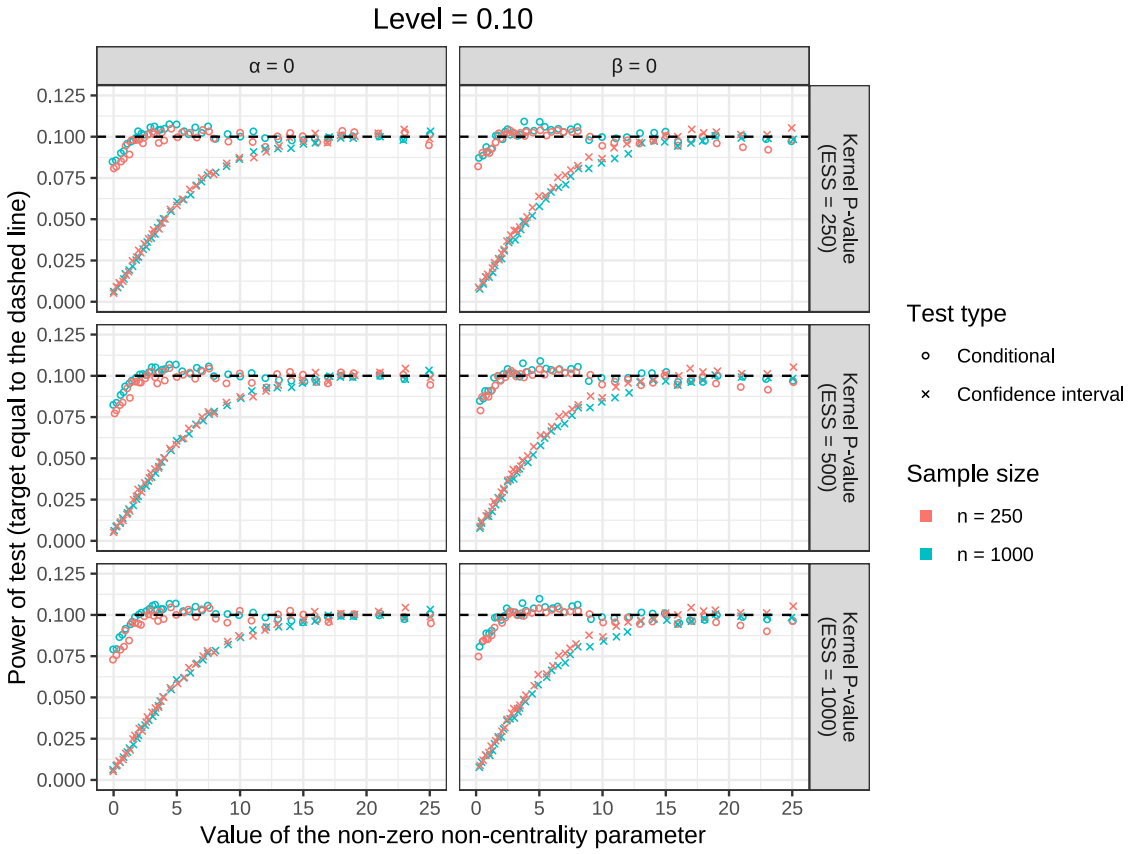


Figure 3.9: Results from a simulation study exploring the calibration of the conditional procedure at the $\tilde{\alpha} = 0.10$ level when $\alpha\beta = 0$ in the population. The estimated level of both the conditional and marginal tests of the indirect effect are plotted against the value of the non-zero non-centrality parameter. Results are separated by data-generating population (columns) and the effective sample size used to calculate conditional p-values (rows). Results for two sample sizes are included (distinguished by color) and tests (distinguished by point shape). Overall, the conditional procedure is better calibrated than the marginal procedure.

As μ grows, the bootstrapped distribution will both under- and overestimate the true non-zero non-centrality parameter. As a result, the estimates of the conditional quantile function will be approximately correct across many tests. However, Figures 3.8 and 3.9 show that at both the 0.05 and 0.10 levels, the procedure is anti-conservative by up to 0.5-1.0 % when $\mu \in (2, 8)$. In this range, there is a nontrivial probability that a $\chi^2_1(0)$ random variable will be greater than a $\chi^2_1(\mu)$ random variable. When this event occurs, the estimated sample distribution of λ given A, estimated using the bootstrap procedure, will

have smaller conditional quantiles, producing an anti-conservative test.

The conditional procedure is anti-conservative for all considered effective sample sizes. Alternative methods of calculating conditional p-values and increasing the number of bootstrap samples used to estimate the conditional p-value did not completely resolve the anti-conservatism, but decreased its magnitude. Overall, it is our assessment that the anti-conservative nature of the procedure for certain values of μ is real and not an artifact. As a point of reference, the Monte Carlo simulation standard error for each estimate of the tests' levels is approximately 0.2%. This means that greatest violations of the level are more than two standard errors away from their target level.

For $\mu > 8$, the estimated level of the test appears to converge back to its target level. In this region of the parameter space, the probability of conditioning on the incorrect likelihood ratio vanishes. In this region we are able to correctly learn the conditional sampling distribution of the test statistic λ given A.

The confidence interval-based test is much more conservative than the conditional procedure for small values of μ. In fact, it does not reach its target level until $\mu > 15$. The confidence interval-based test is valid for all values of μ, meaning that we do not have reason to believe that the procedure is anti-conservative for any value of μ. Selecting between the methods therefore requires deciding whether anti-conservatism of 0.5 and 1.0 % over a portion of the null-parameter space is better or worse than substantial conservatism over a much larger region of the null parameter space. Our next simulation will show that if one elects to use the conditional inference procedure, they gain a moderate amount of power in exchange for the tests' anti-conservatism.

3.5.2 Performance of tests based on the asymptotic sampling distribution of λ given A

In Section 3.4.3 we derived the asymptotic sampling distribution of λ given A under the assumption that λ and A are independent. In this section, we present the results of a simu-

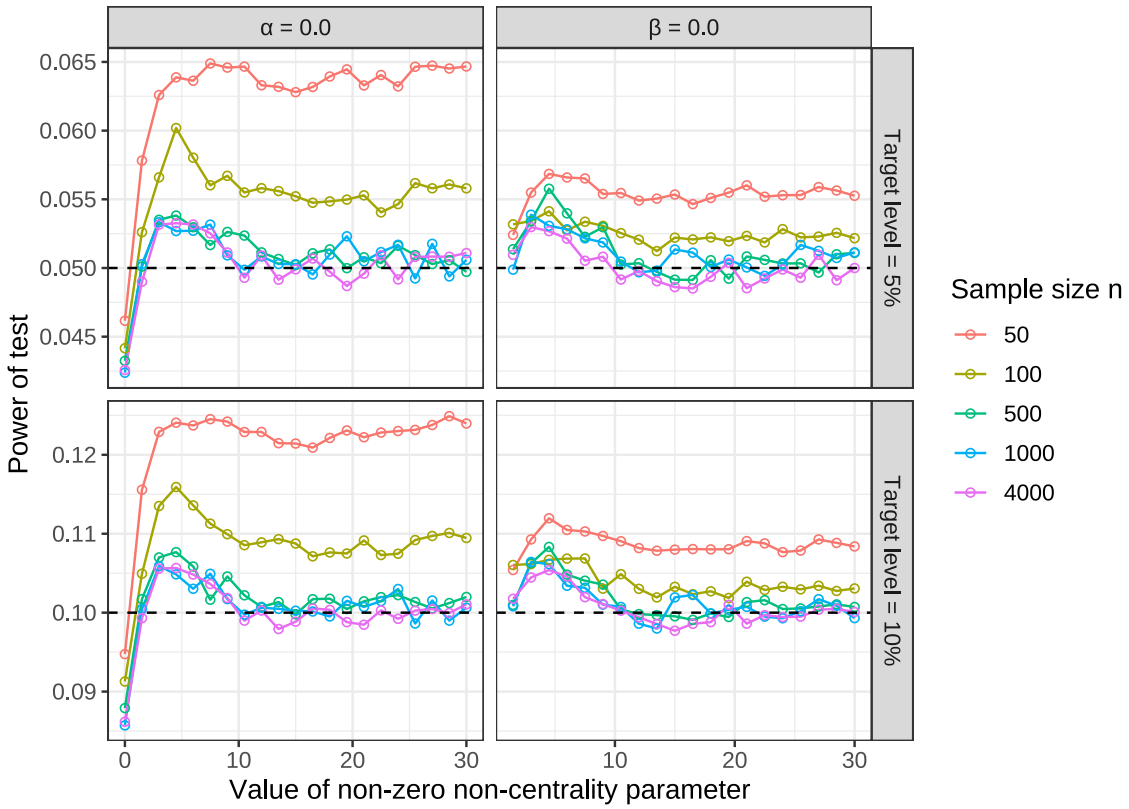


Figure 3.10: Estimated level of conditional test using asymptotic p-values. Each panel shows power of the test plotted against the value of the non-zero χ^2_1 non-centrality parameter. Results are stratified by null parameter (α and β) and by target level (5% and 10%). Line and point color are used to designate the sample size n. The dashed black lines indicate the target level of the test.

lation study that uses the derived sampling distribution to calculate approximate p-values.

Here we are only interested in exploring the level of the procedure over the null space. As a result for all considered settings, $\alpha\beta = 0$ in the population. We generate data with sample sizes equal to $n \in \{50, 100, 500, 1000, 4000\}$ in order to assess the effect of sample size on the test level. We generate data for settings in which both α and β are 0, and generate data so that the non-zero χ^2 non-centrality parameter falls at 20 regularly space points in the interval $[0, 30]$. Remember that the non-centrality parameter $\mu_\phi = (\sqrt{n}\phi/\nu_\phi)^2$, where n is the sample size, ϕ is the value of the regression coefficient, and ν_ϕ is the Fisher information of ϕ. For fixed n and ν_ϕ, one can solve ϕ. We control μ_ϕ rather than than ϕ so that simulations are comparable across sample sizes. For each simulation setting 100,000

data sets were created to estimate the test's level.

Figure 3.10 presents the simulation results. There are several consistent patterns across all four panels (combinations of null parameters and test level). First, the level of the test depends on sample size. When n is small, the level of the test exceeds its nominal level for almost every value of the non-zero non-centrality parameter. It appears that after $n \geq 500$, the approximation is reasonable and does not dramatically improve as n grows. The anti-conservatism with small n can primarily be attributed to the fact that the χ^2_1 approximation to the log-likelihood ratios is not appropriate. Secondly, when $\beta = 0$ and $\alpha \neq 0$, tests using the asymptotic p-values are less anti-conservative than when $\alpha = 0$ and $\beta \neq 0$. This difference is due to the dependence between the λ and A when $\beta \neq 0$, which violates the independence assumption underlying the approximate p-value calculation.

Overall, the results from the simulation studies suggest that the use of the asymptotic sampling distribution of λ given A is reasonable when n is moderately large. In such settings, there's only weak dependence between λ and A and the χ^2_1 approximation is good. We saw that the approximate p-values are anti-conservative by at mot 0.5% for non-centrality parameters less than 10 when n is large. Once the non-centrality parameter exceeds 10, the tests achieve their nominal level.

3.5.3 A comparison of power functions

We now consider the power of the conditional procedure against alternative hypotheses and compare the power of the conditional method to the power of a confidence interval-based test. This simulation study takes place in the same setting as Simulation 1. To simplify the presentation, we present results of the conditional procedure using an effective sample size of 500 to estimate the conditional distribution of λ given A. Additionally, we limit the simulation study to values of $\mu_\alpha, \mu_\beta \in \{0.0, \ldots, 8.0\}$, since the largest difference between the methods occurs in this region.

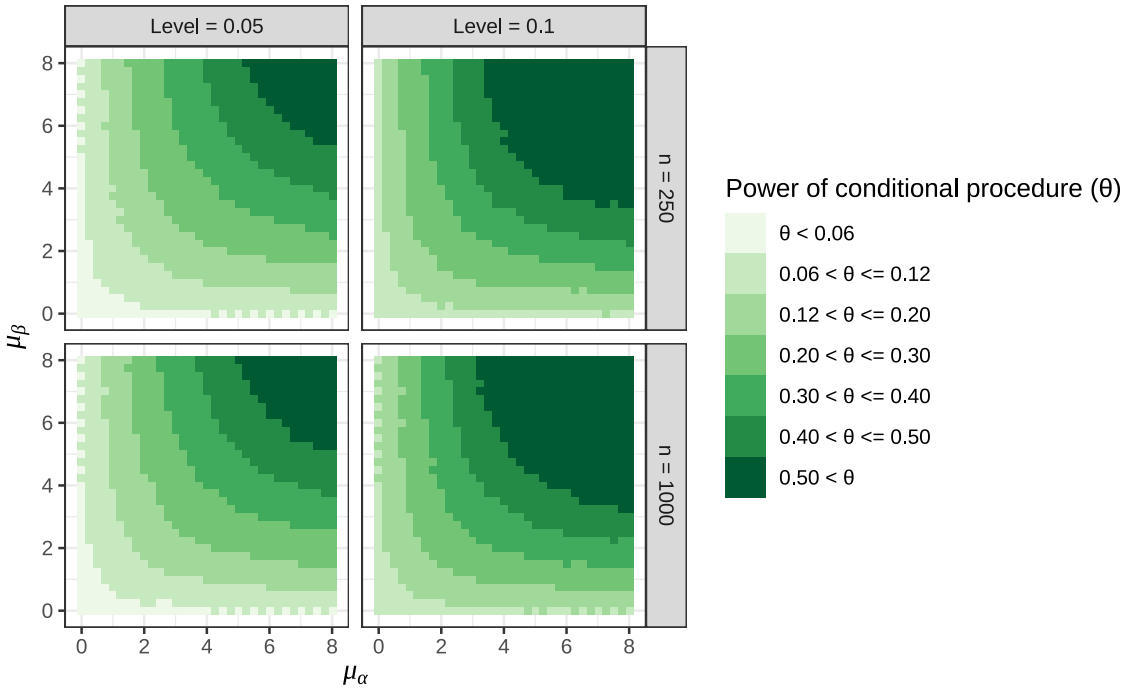


Figure 3.11: A simulation-based estimate of the power function of the conditional test over a grid of μ_α and μ_β values. Two test levels were considered (varying across columns) and two sample sizes were considered (varying by rows). Lighter green areas represent areas of lower power, while dark green regions indicate areas of greater power.

Figure 3.11 shows a simulation study-based estimate of the discretized power function of the conditional procedure. For any fixed value on either the μ_α or μ_β axes, the power function increases as one moves away from the coordinate axis. As either non-centrality parameter grows, the level sets of the power function appear to approach an asymptote which is parallel to either the μ_α or μ_β axis. The power of the conditional procedure does not appear to depend greatly on the sample size n after fixing μ_α and μ_β (each non-centrality parameter is a function of the sample size n).

Figure 3.12 plots the difference in the power of the conditional and confidence interval-based procedures, denoted δ in the figure legend. The power gain δ is discretized to aid interoperability.

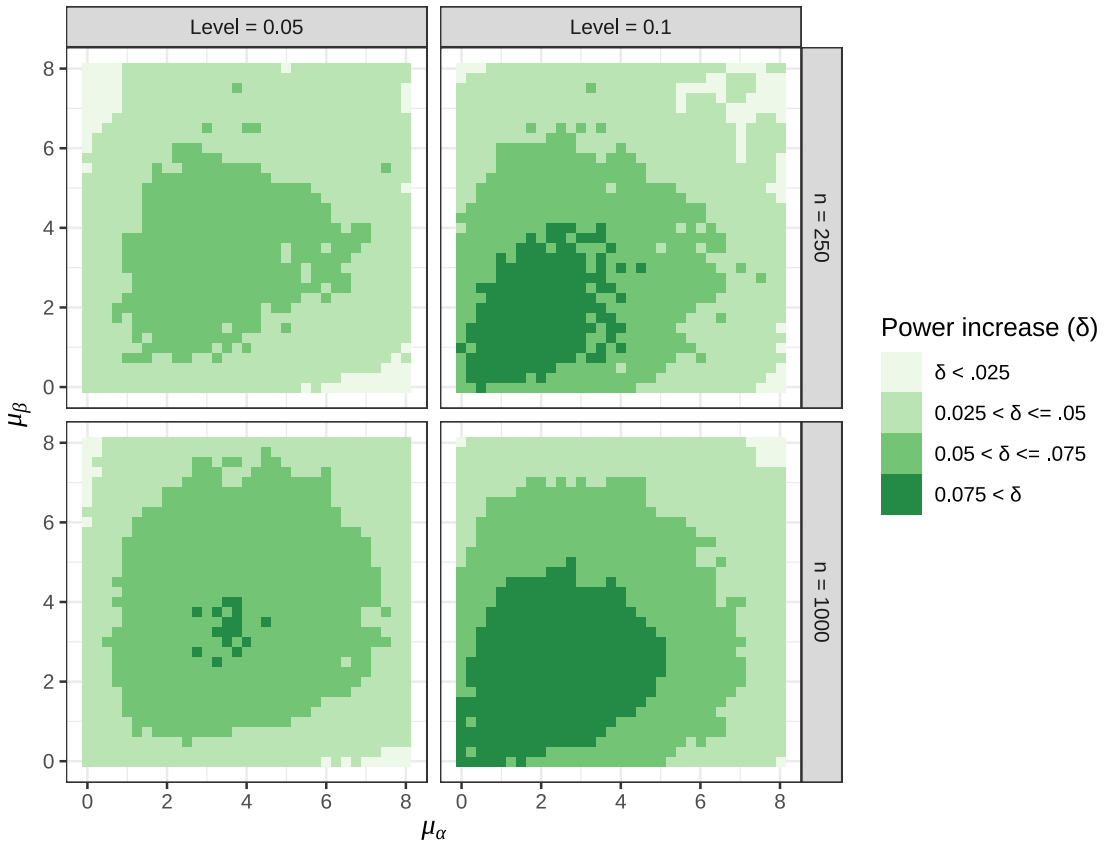


Figure 3.12: An simulation-based estimate of the power increase of the conditional tests over a confidence-interval based procedure. Results are shown over a grid of μ_α and μ_β values. Two test levels were considered (varying across columns) and two sample sizes were considered (varying by rows). Darker green regions indicate areas where the difference between the procedures' power functions was larger (in favor of the conditional procedure).

The largest increase in power occurs near the origin in all simulation settings (see Figures 3.12). As one moves away from the origin or either axis, δ eventually decreases. The power gain is larger when the level of the test is equal to 0.10, although the relative increase in δ is larger at the $\tilde{\alpha} = 0.05$ level. Additionally, the increase in power is larger when $n = 1,000$ versus $n = 250$.

Although the conditional procedure is not dramatically more powerful than existing methods, the increase is nontrivial, especially on a relative scale. As an example, consider the results of the simulations study when $\mu_\alpha = \mu_\beta = 3.5$ (Table 3.2).

In this setting, the confidence interval approach has low to moderate power, and the

Table 3.2: Simulation-based estimates of test power when $\mu_\alpha = \mu_\beta = 3.5$.

n	Level	Conf. Int. Power	Cond. Inf. Power	Relative power increase
250	0.05	0.19	0.24	30%
1000	0.05	0.18	0.26	42%
250	0.10	0.31	0.38	22%
1000	0.10	0.31	0.40	28%

relative increase in power is quite substantial. In many applied mediation analyses, one expects that associations between variables are weak, which is the region where the proposed method is most beneficial.

3.5.4 Impact of algorithm parameters on conditional procedure's performance

The final simulation study assesses the dependence of the procedure's performance on the choice of algorithm parameters. In particular, we would like to see whether the empirical level of the procedure depends on the ratio of outer to inner bootstrap repetitions. Additionally, we use this simulation study to again assess whether performance depends strongly on the effective sample size of the kernel weighted p-value.

The primary variable of interest, the number of non-parametric bootstrap repetitions used in Algorithm 3.1, will be varied over Monte Carlo simulation studies. The number of outer repetitions, denoted n_b, will take values in the set

$$\mathcal{N}_b = \{1, 2, 5, 10, 25, 50, 100, 250, 500, 1000, 5000\}.$$

For each $n_b \in \mathcal{N}_b$, a total of 10,000 bootstrap samples (λ, A) will be generated. The number of inner loops, $n_p = 10,0000/n_b$. For each synthetic dataset, a kernel-weighted p-value will be calculated for four different bandwidths chosen to give effective sample sizes of 100, 250, 500 and 1000. Again, these effective sample sizes correspond to using 1.0, 2.5,

5.0 and 10.0% of the total bootstrap samples to calculate the conditional p-value.

Both M and Y models will be Gaussian GLMs. The scale parameters for each model will be fixed equal to one. The simulation study will be conducted for two data-generating populations, each of which lives in the null parameter space. In the first, $\alpha = 3.0/\sqrt{250}$ and $\beta = 0$. In the second, the values of the regression coefficients are switched. These settings were chosen since they were settings for which the procedure appeared to be most anti-conservative in other simulation studies. 25,000 Monte Carlo trials will be performed for each data-generating population and value of $n_b \in \mathcal{N}_b$. To make the comparisons as informative as possible, the same 25,000 synthetic datasets were used for each choice of $n_b \in \mathcal{N}_b$.

Results are shown in Figures 3.13 and 3.14 for data-generating populations 1 and 2, respectively. The simulation results and findings are similar between both populations. Overall, the procedure becomes more conservative as the effective sample size increases. Effective samples sizes of 100 and 250 were more anti-conservative than larger effective sample sizes. Using a single outer bootstrap repetition produced tests that were more poorly calibrated than tests that used more than one outer bootstrap repetition. Performance did not vary substantially after 10 or more outer repetitions were used.

3.6 Discussion

Our proposed method addresses many of the issues that have often been debated about the appropriateness of conditional inference. Neither our application nor our observations fundamentally alter the debate, since they are issues which by nature cannot be definitely settled, but we believe the proposed procedure highlights conditional inference's utility and potential shortcomings.

First, our conditional test of the indirect mediation effect demonstrates the advantage of making inference more relevant to the observed data. Our analysis showed that the

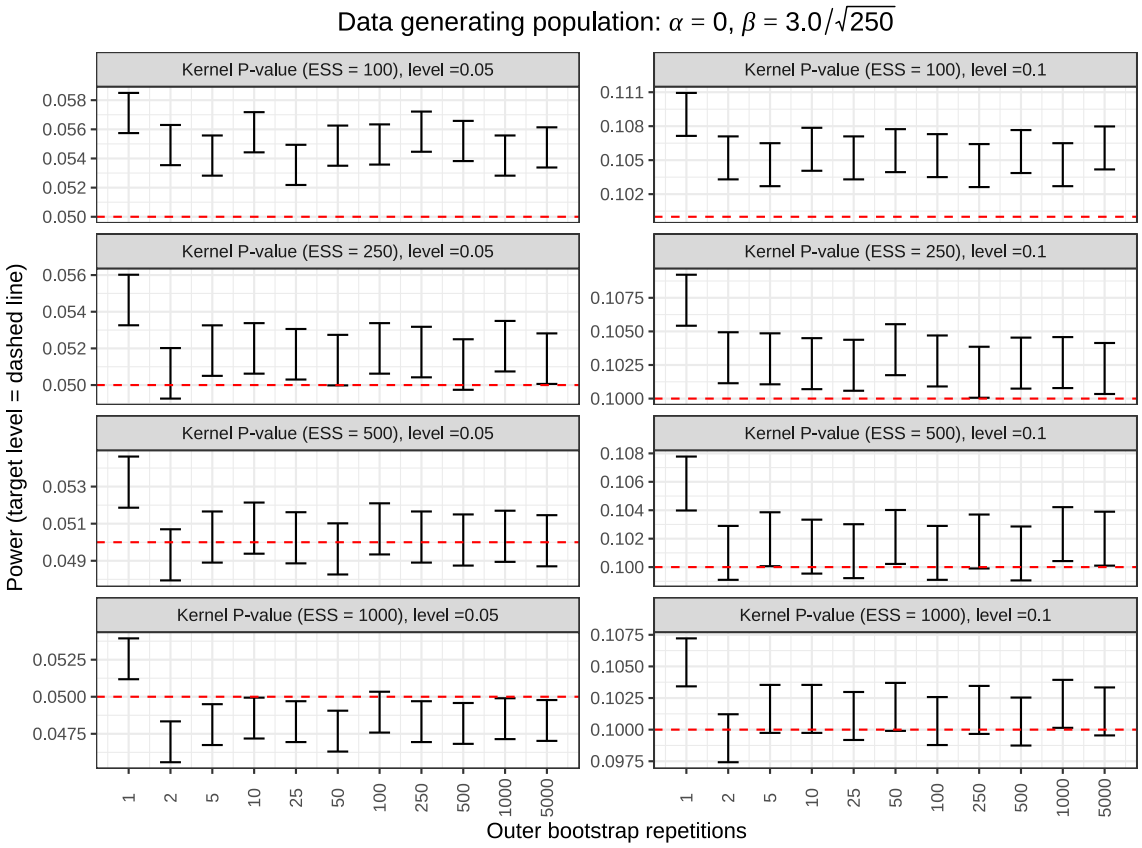


Figure 3.13: This figure plots the empirical level of the conditional test plotted against the number of outer bootstrap repetitions used in Algorithm 3.1. The limits of the error bars represent the estimate plus and minus two standard errors. Results are stratified by target significance levels (by column) and four effective sample sizes (by row). The dashed, horizontal line represents the nominal test level. The data-generating population set $\alpha = 0.0$ and $\beta = 3.0/\sqrt{250}$.

variation in the log-likelihood ratio's sampling distribution is explained by the value of a nuisance parameter. After conditioning on a nearly ancillary statistic, the conditional sampling distribution of the log-likelihood ratio was more uniform. We both derived an asymptotic approximation to the conditional sampling distribution and developed a bootstrap procedure for learning the conditional sampling distribution if one does not wish to use the limiting distribution. This lead to a moderate power increase (in the range of 7-10%) against a class of certain alternatives. Importantly, the variation in the conditional test's performance across the null parameter space was drastically reduced.

This brings us to the second oft debated topic: should one carry out conditional in-

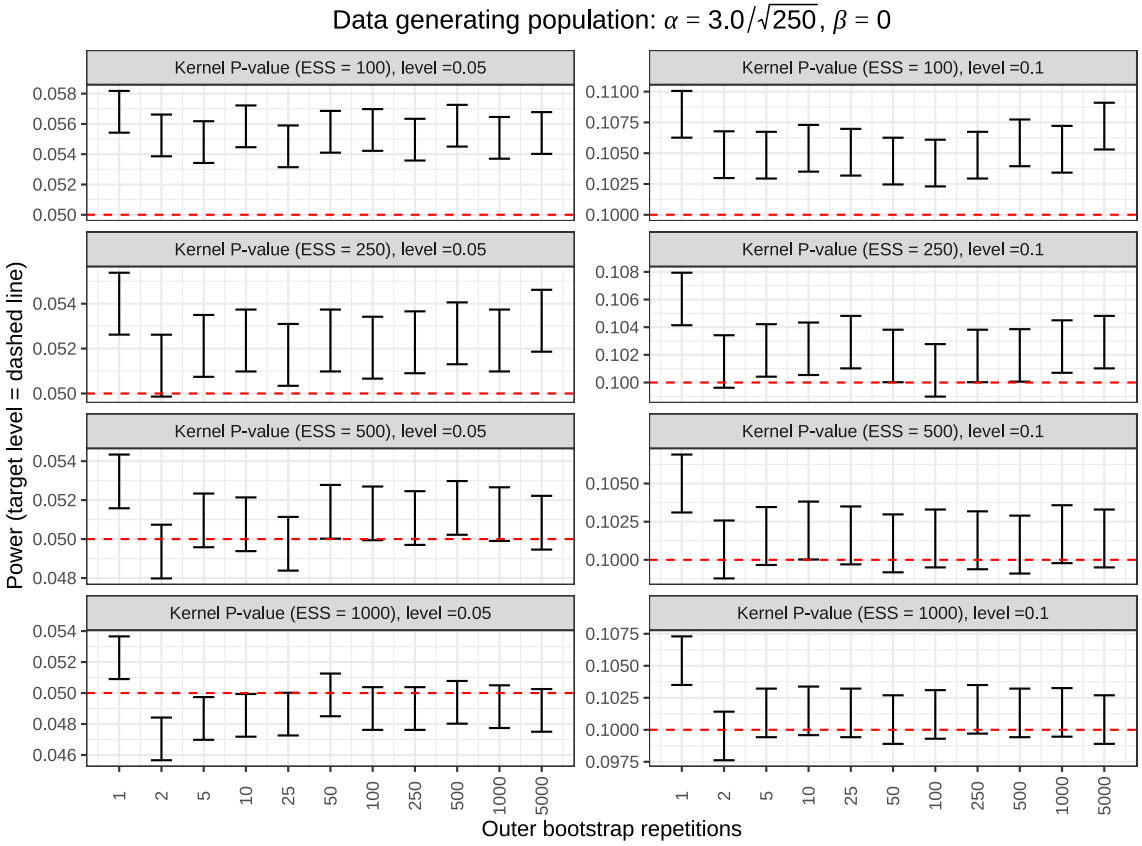


Figure 3.14: This figure plots the empirical level of the conditional test plotted against the number of outer bootstrap repetitions used in Algorithm 3.1. The limits of the error bars represent the estimate plus and minus two standard errors. Results are stratified by target significance levels (by column) and four effective sample sizes (by row). The dashed, horizontal line represents the nominal test level. The data-generating population set $\alpha = 3.0/\sqrt{250}$ and $\beta = 0.0$.

ference when an exactly ancillary statistic does not exist? The answer to this question likely depends on the application. In order to answer in an informed manner, one needs to understand the trade-offs in terms of the marginal and conditional tests' power and level.

We believe that the harm caused by conditioning on a non-ancillary statistic in our application is relatively minor, as shown through both theoretical analysis and extensive simulation studies. Marginal tests of the indirect effect are valid over the full null-parameter space, but are extremely conservative over a large region of the space. Our proposed conditional test is better calibrated over the full null-parameter space, but it is slightly anti-conservative over a region of the null-parameter space. Due to the better calibration, the

conditional procedure has greater statistical power over a large region of the alternative parameter space. Generally, statisticians are more comfortable with a procedure that is valid but conservative, even if the conservatism is an order of magnitude larger than the anti-conservatism of a competing procedure. In this setting, we believe that the relatively large increase in power is well worth the trade-off of having a slightly anti-conservative procedure for some data-generating populations.

To the best of knowledge, all widely used conditional inference methods either achieve an exact sampling distribution (e.g. Fisher's exact test) or simplify estimation of parameters (e.g. conditional logistic regression). Our application of conditional inference shows that the approach can be useful even when inference does not become exact. The development of conditional approaches also help to identified and highlighted when and why marginal methods underperform. Developing conditional tests or estimators for other non-standard or otherwise challenging inference and estimation problems may produce better statistical methods.

CHAPTER 4

Functional Random Effects in a Mechanistic Multilevel Analysis of a Biological System

4.1 Introduction

In many domains of research, quantitative deterministic laws, often stated using differential equations, are used to concisely represent the state of our scientific understanding. Such laws generally reflect well-established mechanisms, and to some extent may be seen as causal descriptions of how fundamental processes unfold in time and space. However, it can be challenging to rigorously calibrate and assess such mechanistic models against empirical measurements. A primary reason for this is that interesting real-world systems consist of many components, some of which are easier to measure than others. In response to this and other challenges, multilevel probability models have been successfully used in many domains to bridge the gap between idealized deterministic laws, and measurements taken in the real world. This chapter proposes a framework for enhancing what can be learned using this approach, focusing on gaining insight into components of the system that are described in terms of probability distributions of random functions.

Embedding a deterministic first-principles model into a probability model allows us to treat the observable and unobservable parts of the system on equal terms. Through numerous successes, it has been found that this approach can accommodate systematic and random measurement error, partially observed data, and other measurement challenges. At

a high level, such models take the form $P_\theta(Y, Z, |X)$, where Y is observable, Z is not observable, X is observable but we do not wish to consider its distribution, and θ is a parameter that governs the mechanistic process and measurement processes at hand.

Of particular interest here is the use of probability models to describe unobserved system components that exist in multiple realizations. For example, if we are studying a collection of exchangeable units that each have a distinct state for some characteristic, then we can use a probability distribution to describe the aggregate characteristics of the ensemble of such states. Let F denote this state for one unit, and extend the model above to $\prod_i P_\theta(Y_i, Z_i, F_i|X_i)$, where i indexes repeated observations, which here are taken to be independent for simplicity. Since $P_\theta(Y_i, Z_i, F_i|X_i) = P_\theta(Y_i, Z_i|F_i, X_i)P_\theta(F_i|X_i)$, we can focus our attention on $P_\theta(F_i|X_i)$ and specifically the components of θ that determine this distribution. This perspective is common to most applications of multilevel modeling. We note that mathematically, F_i could be taken to be part of Z_i, but we wish to distinguish it since F_i will have particular modeling goals that may differ from the other variables in Z_i.

Here we consider multilevel models in which the unit-level latent F_i are continuous functions of one variable (here that variable is time). These latent functions are treated as random, and are anticipated to share some common features, while easily diverging in other ways. Our goal is to learn about their underlying probability distribution, which captures both their stable and varying aspects. Since the latent variables are functions, we describe them using distributions that emit random functions. Ideally, these distributions will be able to encode the statistical characteristics that these functions must exhibit in order to be consistent with the data. Borrowing ideas from functional data analysis, we can consider the distribution of the degree of smoothness as a trait that can be learned from the data. Smoothness is only one of many behaviors that a distribution of random functions can exhibit.

Motivated by an application in pharmacokinetic modeling, we propose a modeling framework that is able to learn the extent to which tendencies toward either concavity

or convexity are likely to hold. We note that concavity, like smoothness, is commonly used in data analysis involving observable functions. For example, generalized additive models and isotonic regression are two techniques for modeling data that are observed as functions (perhaps observed incompletely and with additive noise) that are to some extent smooth or have restricted convexity. Our problem is somewhat more challenging, in that the functions of interest are only observed indirectly, through their role propagated through the mechanistic model, determining downstream observables.

This chapter presents a case study of ibuprofen pharmacokinetics motivated by drug bioequivalence studies. Typical bioequivalence studies are conducted with human subjects and are used to show that a new drug product (e.g. a generic drug product) delivers the product's active ingredient to the therapeutic site as well as the reference drug product. Due to the complexity of how a drug product interacts with the *in vivo* environment, pharmaceutical scientists have long sought alternative tests for evaluating the equivalence of drug products. Better tests minimize the exogenous variation that reduces the power of bioequivalence studies.

One such approach uses computers to simulate hypothetical subjects to assess whether a competitor drug product is equivalent to its name brand counterpart. This approach, sometimes called *in silico* experiments, uses pharmacokinetic and pharmacodynamic (PK/PD) models to describe a drug's transit through the human body. To conduct an informative *in silico* experiment, one must create a model that captures *in vivo* variation in drug concentrations. Principle sources of variation in rate constants are (a) inter-subject differences, (b) within-subject differences between trials and (c) temporal within-subject changes. Our aim is to describe and account for these potential sources of variation.

4.2 The motivating case study

The data motivating the work presented in this chapter was produce by a multi-site study lead by pharmaceutical scientists from the University of Michigan College of Pharmacy. The research program was funded through an FDA research contract. The project and the research it produced contributes to the FDA's efforts to develop better regulatory tests for pharmaceutical drugs.

The FDA regulatory framework for allowing generic drug products to be sold requires showing that the generic product will be statistically indistinguishable from the name brand drug product *in vivo*[1]. To show this, a drug company does not need to repeat the clinical trials that are used to establish the original product's safety and efficacy. Since the active ingredient in the two products is identical, bioequivalence tests establish whether a similar amount of drug reaches the drug's therapeutic target (e.g. blood or tissue). Showing this then establishes that the new drug product is efficacious and safe. For a drug like ibuprofen, which is used to treat pain, fever, and inflammation, this amounts to showing that the generic product has a similar maximum concentration of ibuprofen in the blood.

Currently, bioequivalence tests are divided into two categories of test: *in vitro* and *in vivo*. Each of these classes have deficiencies. *In vitro* tests, which often show that a drug dissolves sufficiently quickly in lab settings, generally fail to reflect the *in vivo* environment. This impacts their predictive power to describe how the new drug product will perform in human subjects. *In vivo* studies are often under-powered since they have small sample sizes. In these small n settings, exogenous variability often causes bioequivalence tests to fail when pharmaceutical scientists expect them to succeed.

The first task of the FDA funded research project was to better understand the *in vivo* sources of exogenous variability that affect systemic availability of an oral drug product

[1]It is important to clearly differentiate between a *drug* and a *drug product*. A drug product refers to the actual tablet that delivers the active ingredient, or drug, to a patient or subject. Drug products with identical active ingredients can be manufactured with different inactive ingredients. The choice of inactive ingredients may impact the drug product's ability to deliver the active ingredient, necessitating bioequivalence studies to show that a new product will behave similarly to its reference product.

[2]. Of particular interest were differences in gastrointestinal (GI) activity and conditions that might cause systemic differences in drug distribution. The first part of the research project involved conducting an intubation study that concurrently measured GI conditions and concentrations of ibuprofen in plasma.

4.2.1 Ibuprofen intubation study

The ibuprofen intubation study was designed to measure GI factors relevant to the distribution of ibuprofen *in vivo*. The full study was made up of 60 intubation studies. During each intubation study, a catheter was placed in a healthy subject's upper small intestine and stomach via their mouth and esophagus. The catheter was capable of measuring GI motility and taking aspiration samples of GI fluid. Each subject took an 800 milligram (mg) tablet of ibuprofen several hours after the intubation study began. After dosing, the GI pH and ibuprofen concentrations were measured by aspirating small amounts of fluid from the GI tract. Intravenous blood draws were used to record the blood or plasma concentration of ibuprofen.

Each study produces a sparse multivariate time series of GI pH, GI ibuprofen concentrations and plasma ibuprofen concentrations. GI conditions are monitored at between one and four locations along the GI tract. Table 4.1 shows timing of the aspiration and plasma draws relative to dosing.

In addition to the aspiration samples, the catheter measured GI motility through water-perfused manometry, which measures the contractive pressure of the GI tract. A grouping of three sensors was located distally to each aspiration port. Motility measurements were sampled at 10 Hertz.

The study protocol called for 60 subjects to undergo intubation. Of the 60 studies, 37 were successfully completed. The terminated intubation studies failed for a variety of reasons, including subject discomfort and vomiting or failing the pre-study screening test. During multiple studies, the physicians responsible for placing the catheter were unable to

Table 4.1: The ibuprofen intubation study sampling protocol.

Time (hours)	Plasma sample	Aspiration sample
0.000	✓	✓
0.167	✓	
0.250		✓
0.333	✓	
0.500	✓	✓
0.750	✓	✓
1.000	✓	✓
1.500	✓	✓
2.000	✓	✓
2.500	✓	✓
3.000	✓	✓
4.000	✓	✓
5.000	✓	✓
6.000	✓	✓
7.000	✓	✓
8.000	✓	
12.000 or 24.000	✓	

successfully reach the small intestine. For these studies, data from the small intestine are not available.

4.2.2 A compartmental model for the ibuprofen study

The foundation of our analysis is a four compartment model of the human body. During the intubation study, subjects ingested an 800 milligram ibuprofen tablet orally with a 240 milliliter glass of water. After reaching the stomach, the tablet disintegrates. Due to the stomach's low pH, the ibuprofen, which is an acid, does not rapidly dissolve until it reaches the high pH environment of the small intestine. Therefore, ibuprofen predominantly leaves the stomach as small particles suspended in gastric fluid. After leaving the stomach, the small ibuprofen particles dissolve readily in the small intestine. After dissolution, ibuprofen is able to passively diffuse across the intestinal wall and into the blood stream. The ibuprofen molecules circulates through the body in the blood until removed by the liver and excreted in the urine.

Our model includes four compartments, representing locations or organs where ibuprofen is held or processed during its *in vivo* transit. The first and the last compartments represent the stomach and plasma, while the second and third compartments represent the small intestine. The first small intestine compartment represents undissolved ibuprofen and the second represents dissolved ibuprofen (Figure 4.1).

We will denote the mass of ibuprofen in each compartment at time t using the following notation. Let:

- $G(t)$ denote the mass of undissolved ibuprofen in the *gastric* compartment at time t.
- $U(t)$ denote the mass of *undissolved* ibuprofen in the small intestine at time t.
- $D(t)$ denote the mass of *dissolved* ibuprofen in the small intestine at time t.
- $P(t)$ denote the mass of ibuprofen in the *plasma* (or circulation) at time t.

The system is evolved forward in time by solving the following system of first-order differential equations:

$$\begin{aligned} \frac{\partial G(t)}{\partial t} &= -k_s(t)G(t) \\ \frac{\partial U(t)}{\partial t} &= k_s(t)G(t) - k_d(t)U(t) \\ \frac{\partial D(t)}{\partial t} &= k_d(t)U(t) - k_{abs}D(t) \\ \frac{\partial P(t)}{\partial t} &= k_{abs}D(t) - k_{el}P(t) \end{aligned} \tag{4.1}$$

The logged dissolution rate $\log(k_d(t))$ depends linearly on $x_{ph}(t)$: $\log(k_d(t)) = k_{d0} + k_{d,ph}x_{ph}$, where x_{ph} is the small intestine pH at time t.

At time 0, the subject takes the 800 milligram ibuprofen tablet, which we assume immediately enters the stomach compartment. The other compartments did not contain any ibuprofen at dosing since subjects were prohibited from taking ibuprofen for a week prior to their intubation study. Thus, initial condition of the compartmental model is

$$Z(0) = (G(0), U(0), D(0), P(0)) = (800, 0, 0, 0) \text{ mg}.$$

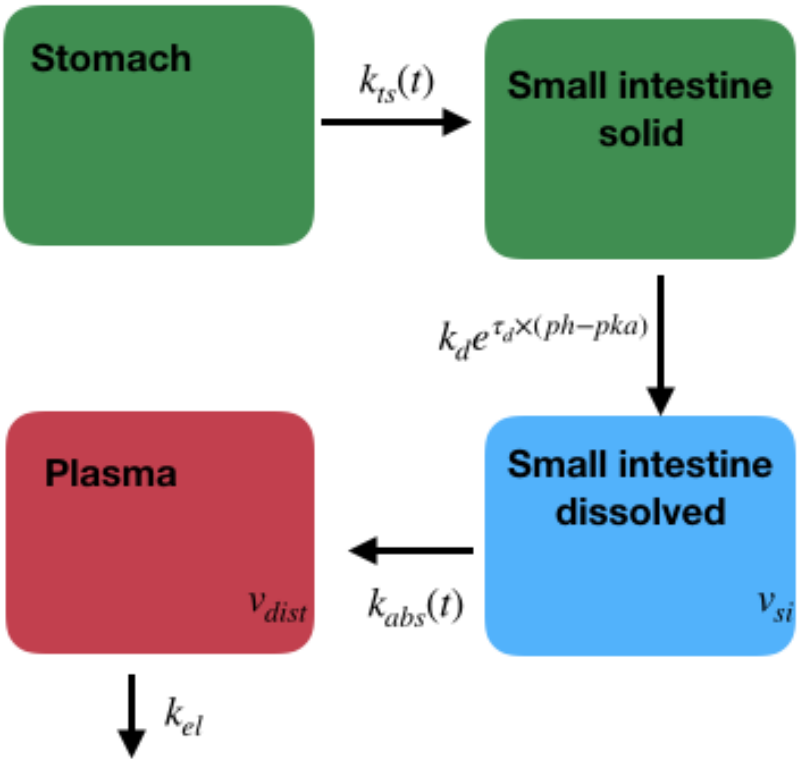


Figure 4.1: The four compartment model used to approximate ibuprofen's path through the human body. Green and blue compartments indicated when the drug is undissolved and dissolved, respectively. The red compartment represents the blood and any tissue that the ibuprofen can diffuse into.

For fixed $\theta = (k_{ts}, k_d, k_{abs}, k_{el})$, solving the system of equations given the initial condition calculates the mass of ibuprofen in each compartment at time t.

4.2.3 A first-principles statistical model of ibuprofen pharmacokinetics

As a starting point for our case study analysis, we fit a typical first-principles pharmacokinetics model to the case study data [37, 38]. We describe the model as a "first-principles" model since it assumes that the ibuprofen transfers between compartments at fixed and unchanging rates according to first-order differential equations. Rates are assumed to be constant across time within a subject but are allowed to vary between subjects through a hi-

Table 4.2: Population PK rate constant definitions

Parameter	Parameter role
k_s	Transfer rate constant from stomach compartment to small intestine
k_d	Dissolution rate constant in small intestine
k_{dph}	Multiplicative effect of small intestine pH on dissolution rate
k_{abs}	Permeability rate constant between small intestine and plasma compartments
k_{el}	Elimination rate constant from plasma compartment

erarchical structure. The model allows for both between-subject, and study-within-subject variation around the population mean and subject mean rates. In this section, we will describe the most complex random effects structure that we considered.

Our first-principles model assumes that at the study-within-subject level, the rate parameters are time-invariant. Definitions of the five PK rate constants are given in Table 4.2. These parameters represent the population average of the transfer and dissolution rates of ibuprofen in the four-compartment model.

Let $\theta = [k_s, k_d, k_{abs}, k_{el}]$ be the vector of pharmacokinetic parameters for the four compartment model. The statistical model assesses both (a) the degree to which subjects' individual ibuprofen kinetics vary between subjects around the population average, and (b) the within-subject variability of ibuprofen kinetics. We introduce the subject and visit-within-subject random effects θ_i and θ_{ij} with priors:

$$\theta_i|\theta \sim \mathcal{N}(\theta, \tilde{\Sigma}_s), \quad \theta_{ij}|\theta, \theta_i \sim \mathcal{N}(\theta_i, \text{diag}(\sigma_v)), \tag{4.2}$$

where $\tilde{\Sigma}_s := \text{diag}(\tau_s)\Sigma_s\text{diag}(\tau_s)$, with $\Sigma_s \in \mathbb{R}^{4\times4}$ is a symmetric positive definite matrix and $\tau_s, \tau_v \in \mathbb{R}^4$, $\tau_s > 0$, $\tau_v > 0$. The prior distribution for Σ_s will be Inverse-Wishart with 5 degrees of freedom and an identity scale matrix. This gives $\tilde{\Sigma}_s$ a *scaled Inverse-Wishart* prior, which has a nearly uniform prior on the correlations between-subject elements of θ_i [39]. At the visit-within-subject level, we assume that the pharmacokinetic parameters vary independently around their subject-specific means. See Table 4.3 for additional precise definitions of the first-principle model priors.

Table 4.3: Prior distributions of first-principle model

Parameter	Prior
$\log(\theta)$	$\mathcal{N}((0, 0, 2.5, 0), \text{diag}(1, 1, .1, 1))$
k_{dph}	$\mathcal{N}(0, 1)$
Σ_s	Inverse-Wishart(5, I)
$\log(\tau_s)$	$\mathcal{N}(0, 1)$
$\log(\tau_v)$	$\mathcal{N}(0, 1)$

Measurement model

For each subject-study combination, the predicted mean plasma and small intestine ibuprofen concentrations at time t are equal to the solution of the system of differential equations given in the previous section. We will denote the predicted ibuprofen mass in the plasma and small intestine compartments during subject i's j^{th} study at time t by $P_{ij}(t)$ and $D_{ij}(t)$. Note that each of these functions implicitly depend on population parameters and subject and study random effects.

Although the model makes predictions of the mass of ibuprofen in each compartment at time t, we observe the concentration of ibuprofen in the plasma and small intestines. As a result, we must divide $D_{ij}(t)$ and $P_{ij}(t)$ by a volume in order to connect study observations to the mechanistic model. The pharmaceutical sciences literature provides an estimate of ibuprofen's volume of distribution as a function of body mass. An estimate of an individual's volume of distribution is given by:

$$V_d = \text{Mass} * 0.1 \frac{\text{L}}{\text{Kg}}.$$

The volume of distribution represents an apparent volume rather than a physical volume. It represents the volume of the blood and any tissue that the ibuprofen may reside in while *in vivo* [40, 41].

We model the small intestine volumes as constant within a study and use a hierarchical model at the study level to estimate the population mean μ_{vol} and inter-study variation σ^2_{vol}

in small intestine volumes. Let V_s^{ij} be the small intestine fluid volume for subject i during study j. We model $V_s^{ij} \sim \mathcal{N}(\mu_{vol}, \sigma_{vol}^2)$.

Given an estimate of the volume of distribution and small intestine fluid volume, we can link the compartmental model to the study data. Let $y_{ij}(t)$ and $p_{ij}(t)$ represent the observed small intestine and plasma ibuprofen concentrations for subject i during study j at time t. Given the corresponding PK parameters $\tilde{\theta}_{ij}$, we use the following measurement models:

$$\log(y_{ij}(t))|\theta) \sim \mathcal{N}\left(\log\left(\frac{D_{ij}(t)}{V_s^{ij}}\right), \sigma_{si}^2\right)$$
$$p_{ij}(t)|\theta) \sim \mathcal{N}\left(\frac{P_{ij}(t)}{V_d^{ij}}, \alpha\left(\frac{P_{ij}(t)}{V_d^{ij}}\right)^{\kappa}\right).$$

4.3 Analysis of the case study data with the first-principles model

The first-principles model is unable to capture the rapid, short-term, and often delayed increase in plasma ibuprofen concentration. Predicted and observed plasma ibuprofen concentrations are shown in Figure 4.2. The time-invariant rates of the first-principles model are unable capture the peak concentration experienced during most studies. Instead, the first-principles model tends to overpredict plasma concentrations prior to T-max (the time at which we observe the maximum plasma concentration of ibuprofen) and to underpredict ibuprofen concentrations at T-max.

This deviation is made more clear by considering the residual ibuprofen concentrations in Figure 4.3. This figure shows the residual ibuprofen concentration against study time for 17 fasted-state intubation studies. The residual plots share a common feature that residuals tend to be negative after the beginning of the study, abruptly become positive at T-max, and

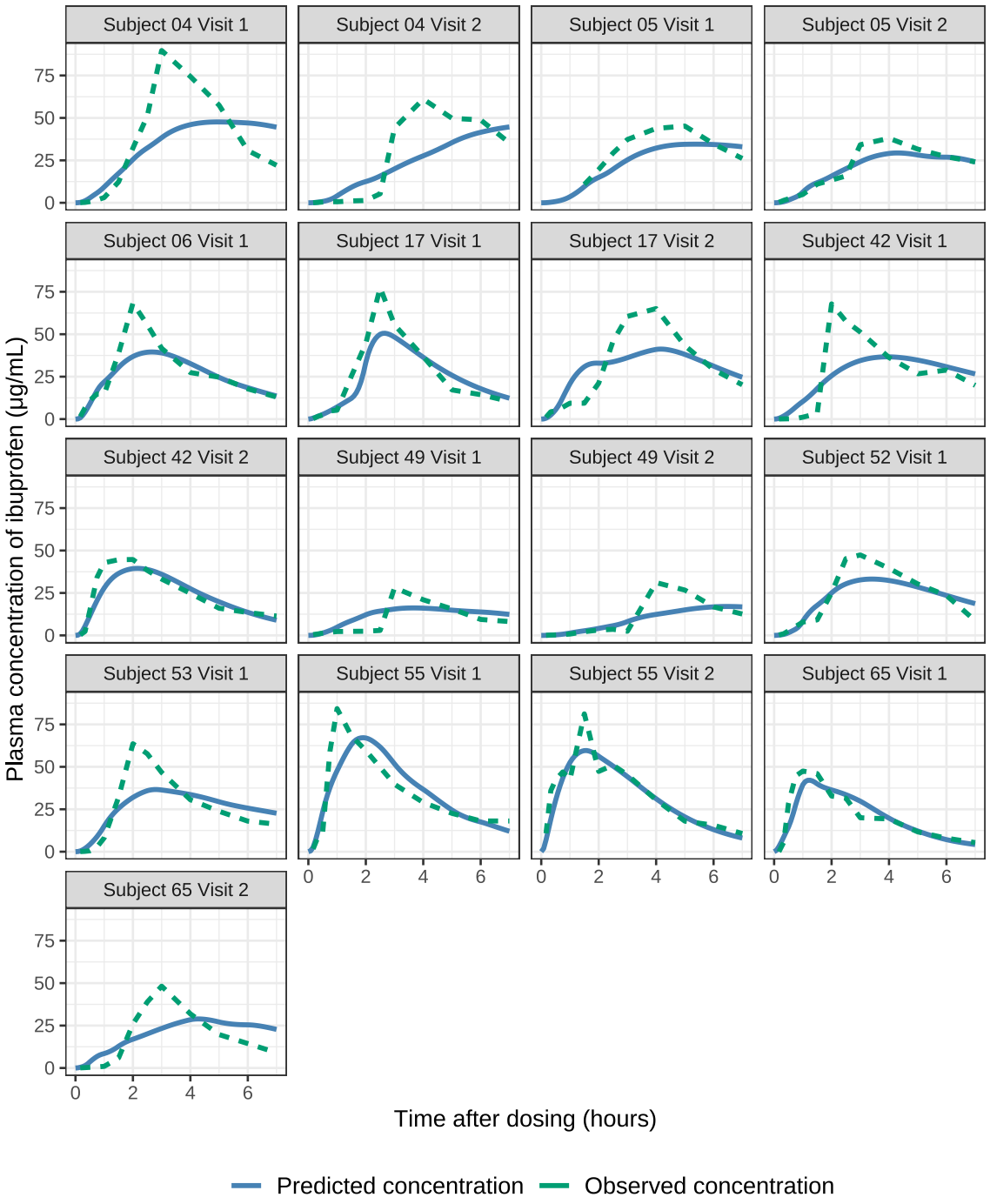


Figure 4.2: A comparison of observed ibuprofen concentrations and predicted ibuprofen concentrations from the first-principles model. Ibuprofen plasma concentrations (μg/mL) are plotted against study time for 17 fasted-state studies. 0 represents the time at which the subjects took the ibuprofen tablets. Solid blue and dashed green lines are used to distinguish between the fitted model's prediction of the subject's plasma concentration and the observed plasma ibuprofen concentration, respectively.

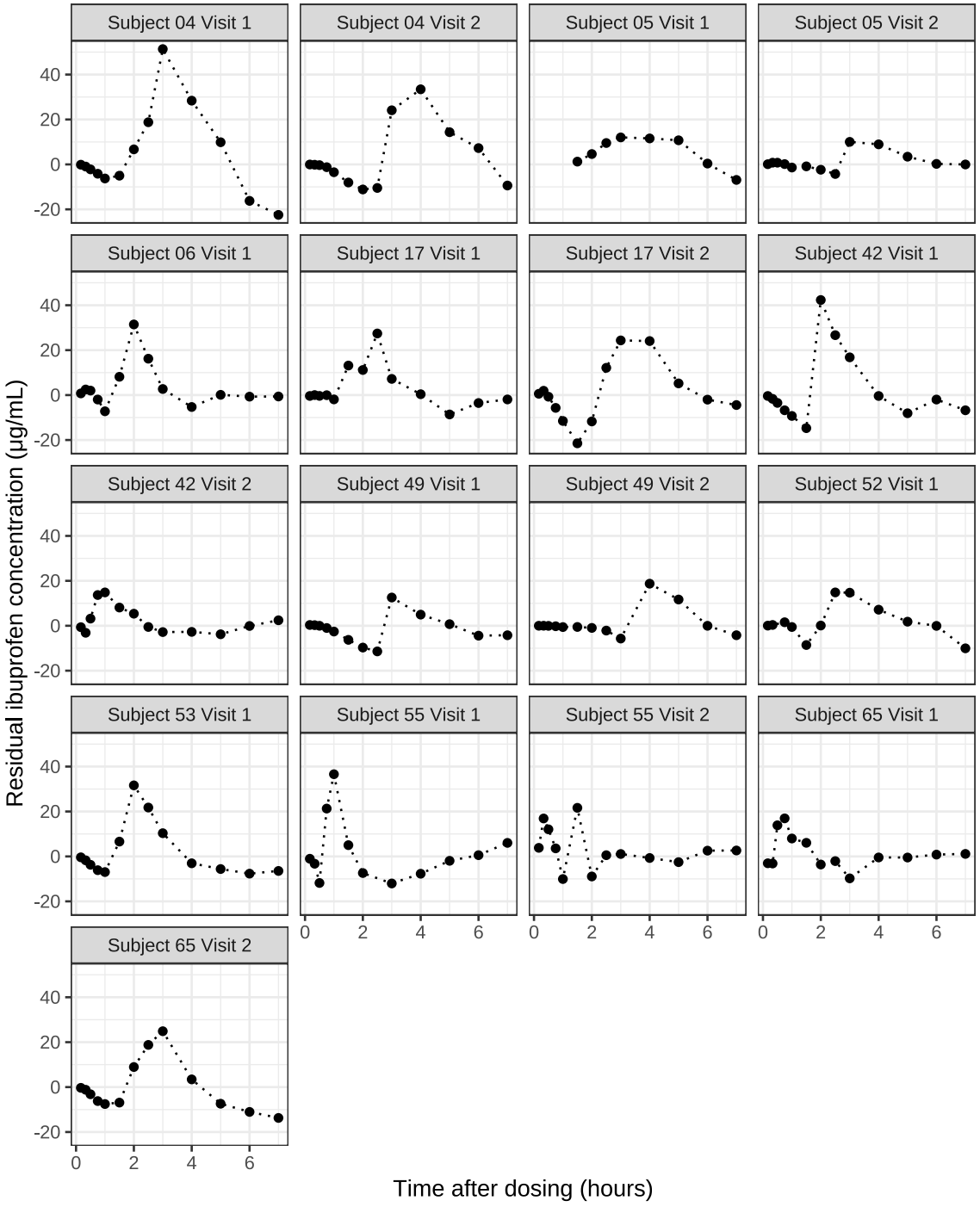


Figure 4.3: The residual ibuprofen plasma concentrations from the fitted first-principles model. The panels are stratified by subject and visit.

then tend back toward 0.

Figures 4.2 and 4.3 show that the first-principles model's assumption that its rate parameters are time-invariant is implausible. Although clear between-subjects heterogeneity exists in many of the rate constants, inter-subject and inter-study differences cannot account for the heterogeneity across the course of a study within a patient. Next, we present several possible mechanisms that might explain the within-study rate parameter heterogeneity

Possible biological mechanisms that underly the systems deviation from first-principles

Several plausible, biological mechanisms might underlie the observed rapid increases in plasma ibuprofen concentrations that the first-principles model failed to capture. We worked with our pharmaceutical science collaborators to identify mechanisms that the study data could potentially identify. We settled on three potential mechanisms that both violate the time-homogeneity assumption of the first-principles model.

First, the stomach-to-small intestine transfer rate k_s could be non-constant. It is known that the contents of the stomach leave in small packets rather than a constant stream. If the rate at which packets leaves fluctuates, then our time-homogeneous rate constant would be inappropriate. Allowing k_s to vary across time within a subject would allow one to approximate the underlying mechanism.

Next, the base dissolution rate k_d might change across time, even after controlling for the pH of the small intestine. Early in the research project, the study team believed that small intestinal motility would drive variation in plasma concentrations of ibuprofen. More small intestinal motility after dosing could cause ibuprofen dissolution to dramatically increase, causing the rapid appearance of ibuprofen in the plasma. To test this mechanism, one would allow k_d to vary across time within each subject.

Finally, the absorption rate constant of ibuprofen into the plasma could change over the course of the study. The absorption rate depends on the total small intestine surface area

that the ibuprofen is exposed to. Variation in motility changes the effective surface area of the small intestine. The absorption rate k_{abs} was allowed to vary across time in order to assess whether this mechanism was driving the rapid appearance of ibuprofen.

We found the greatest evidence in the data for variation in the stomach-to-small intestine transit rate. The second and third mechanisms do undoubtedly affect plasma concentrations [42], but were not well identified by the case study data. We were particularly disappointed to be unable to relate small intestinal motility to variation in either the ibuprofen dissolution or absorption rates. Ibuprofen dissolves readily in the higher pH environment of the small intestine and is able to passively diffuse across the intestinal wall once in solution.

A modified first-principles model of ibuprofen pharmacokinetics

The modification that we found to be most plausible and best supported by the data allows the gastric (stomach-to-small intestine) emptying rate $k_s(t)$ to vary across time. A straight forward way to let $k_s(t)$ change over time is to treat each study's k_s as a draw from a distribution that emits function-valued random variables. Because we did not have an *a priori* hypothesis about what type of variation exists between subjects' emptying rates, we began with a semi-parametric model for $\log(k_s(t))$. Let $k_s^{ij}(t)$ denote the latent emptying rate for subject i during study j at time t. Let the functions $\{\phi_1, \ldots, \phi_k\}$ be the B-spline basis over the interval $(0, 7)$ hours. Then, we model $\log(k_s^{ij})$ by:

$$\log(k_s^{ij}(t)) = \sum_{\ell=1}^{k} \phi_\ell(t)\beta_\ell^{ij}. \tag{4.3}$$

The vector of coefficients $\beta^{ij} = (\beta_1^{ij}, \ldots, \beta_k^{ij})^T$ is modeled hierarchically. We omit the details, but note that the selected prior model for β^{ij} penalized the curvature of $\log k_s^{ij}$ [43].

Figure 4.4 compares observed plasma concentrations to the modified model's predicted plasma concentrations. The modified model substantially improves upon the in-sample predictions compared to the first-principles model. In particular, the modified model is capable

of capturing the periods of rapid ibuprofen appearance in the plasma. Figure 4.5 shows the posterior estimates of the latent functions $k_s^{ij}(t)$ plotted against study time. Although the model for k_s^{ij} is capable of learning more complex shapes, each subject appeared to experience a single episode of faster emptying. This suggests that for the typical intubation study, most of the ibuprofen left the subject's stomach during a relatively short period of time. The short period of rapid emptying was often delayed after dosing.

4.4 Studying mechanisms in multilevel models using functional random effects

Our proposed approach to studying and identifying mechanisms in mechanistic analyses is useful in cases where there is reasonable uncertainty about exactly how the mechanism of interest impacts the biological system. If prior work has already studied and characterized the mechanism, then this approach is a less efficient way to estimating a model for the biological system. Additionally, this approach is most applicable when the underlying mechanism is assumed to cause one of the mechanistic system's rate parameters to vary across time. Furthermore, the method is capable of estimating both the population average rate and unit-level variation around the population mean trend.

Our approach parameterizes the time-varying rate parameter for unit i as a latent, function-valued random variable f^i. The function f^i is modeled as the additive composition of two independent function-valued random variables g^i and h^i

$$f^i = g^i + h^i,$$

where $g^i \sim G_\phi$ and $h^i \sim H_\psi$. The distribution functions G_ϕ and H_ψ are each members of families of probability distributions G and H indexed by the parameters $\phi \in \Phi$ and $\psi \in \Psi$, respectively. Here we assume that the functional space of G captures the mechanism of

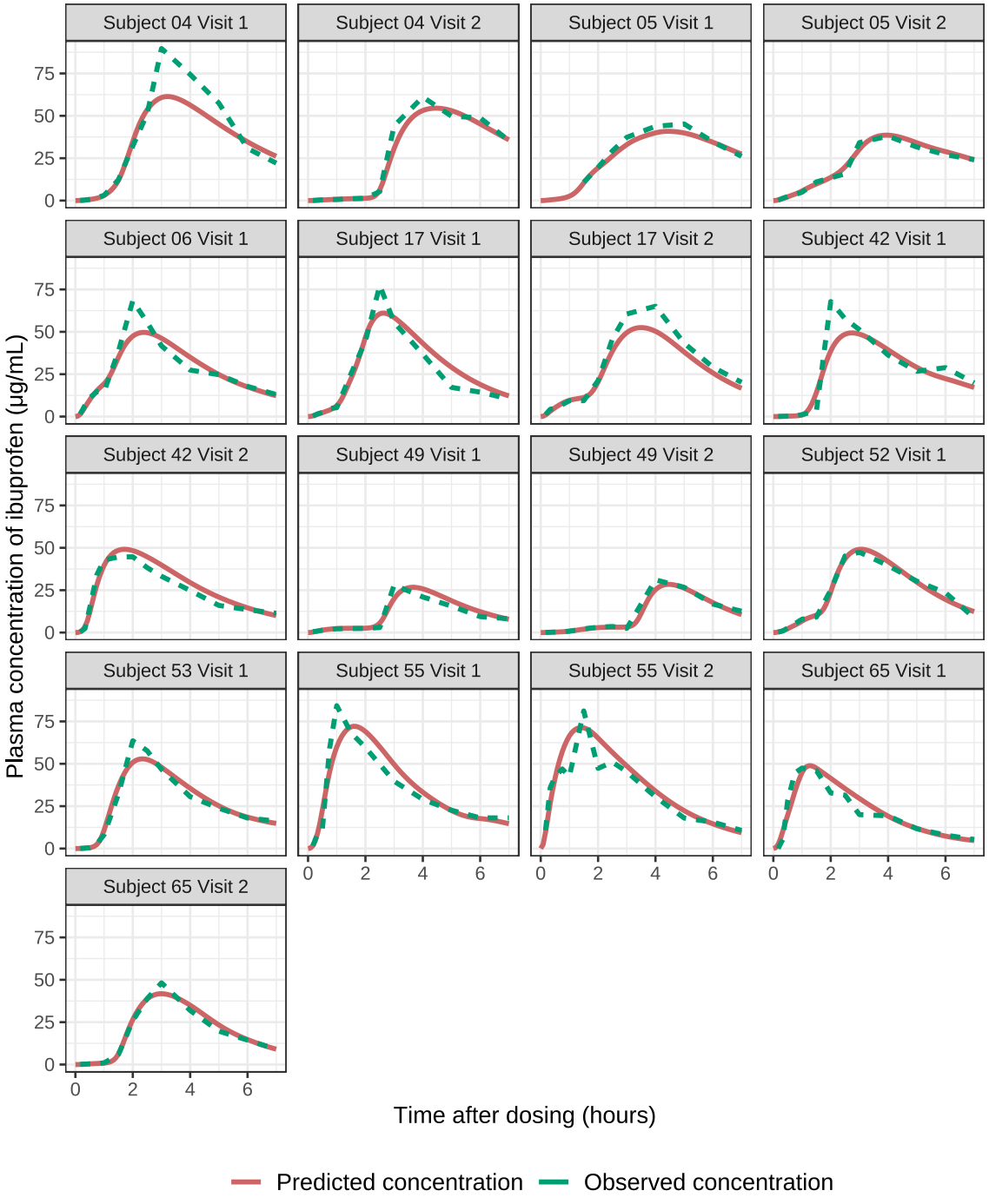


Figure 4.4: A comparison of observed ibuprofen concentrations and predicted ibuprofen concentrations from the modified first-principles model. Predictions are from the fitted first-principles model. Panels are stratified by subject and visit.

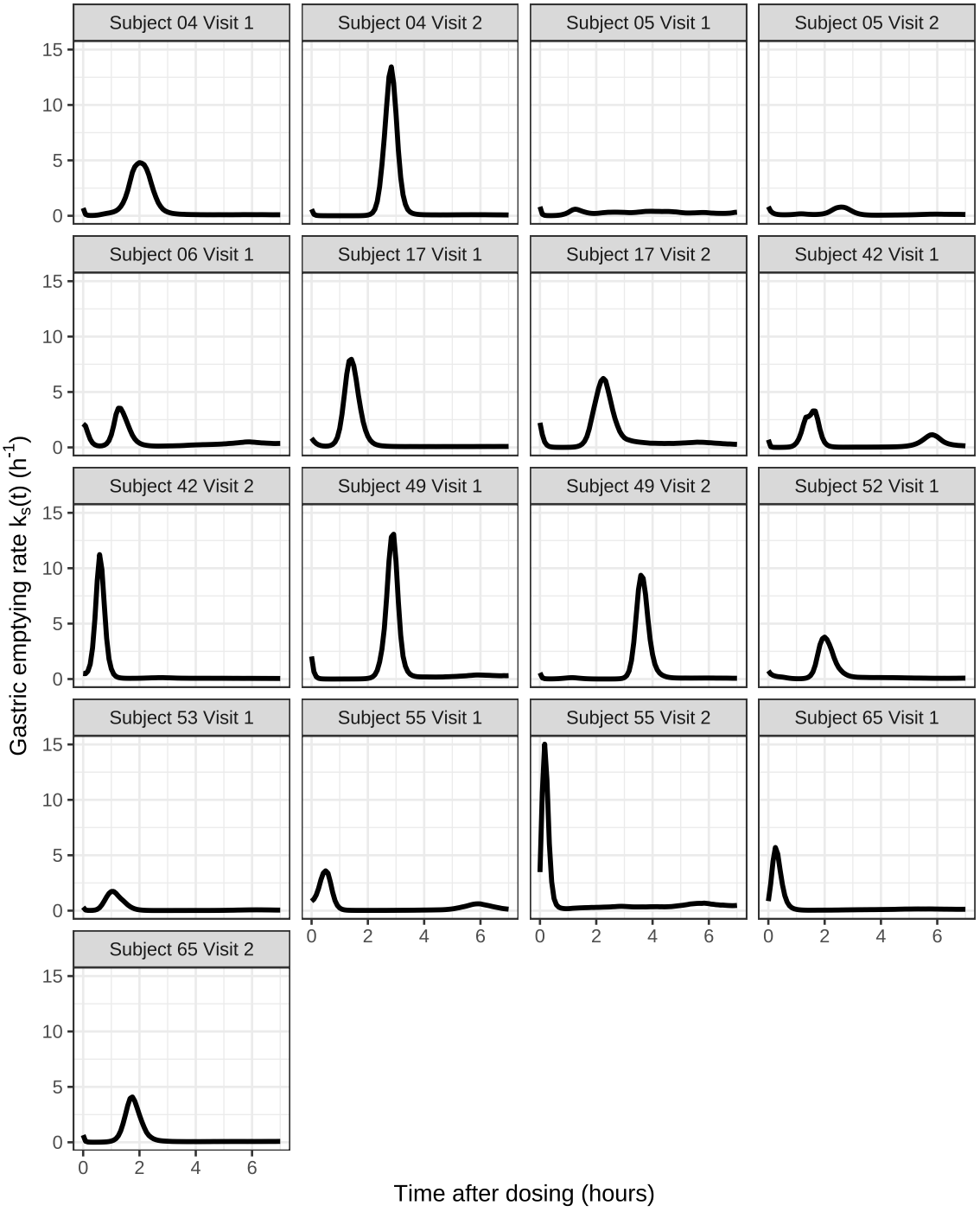


Figure 4.5: The estimated subject-by-visit gastric emptying rate $k_s^{ij}(t)$ plotted against time. Panels are stratified by subject and visit.

interest, while H is capable of capturing many time-varying patterns that deviate from G. In general, G will represent a much richer class of probability distributions.

Our expectation is that if draws from G_ϕ for some $\phi \in \Phi$ are reasonable approximations to the true data-generating latent functions, then the posterior distribution of f^i will concentrate around G_ϕ. To achieve this, the posterior distribution will need to concentrate around a $\tilde{\psi} \in \Psi$ such that draws from $H_{\tilde{\psi}}$ are (nearly) constant.

Since G is a much richer class of probability distributions, one should consider whether the data will be capable of identifying the mechanism. Our expectation, which will be borne out through simulation studies, is that data will be capable of identifying the correct functional space.

To see why this is, first assume that the data-generating latent functions f^i are sampled from G_ϕ for some $\phi \in \Phi$. Let us also assume that there exists $\psi_0 \in \Psi$ such that $G_\phi = H_{\psi_0}$. Thus, it is seemingly impossible to determine whether f^i were sampled from G_ϕ or H_{ψ_0}. However, the likelihood function of G_ϕ is much greater than the likelihood function of H_{ψ_0} when both are evaluated at the f^i. Because MCMC maximizes the posterior likelihood, we expect that posterior will concentrate onto G_ϕ rather than onto H_{ψ_0}.

4.4.1 A model for the latent random function $\log(k_s)$

Our analysis of the motivating dataset suggests that the stomach-to-small intestine transit rate varies across time. We modified the structural model to allow the rate to be time-invariant by modeling the rate as a latent function-valued random variable. We expect that the latent $\log(k_s)$ tends to be concave for most subjects. However, we would like to use a methodology that is capable of identifying an alternative model for $\log(k_s)$ if the data are inconsistent with a concave model for the latent functions. One approach to learning this tendency toward a particular shape was outlined in Section 4.4. We will decompose each study's latent function into two components, a quadratic or parametric function g^{ij}, and a smooth function h^{ij}, so $\log(k_s^{ij}) = g^{ij} + h^{ij}$.

The g^{ij}'s will be sampled from a three parameter distribution that emits quadratic random functions. We will use the terms "quadratic" and "parametric" interchangeably to refer to the g^{ij}'s. The non-parametric component h^{ij} will be sampled from a mean zero Gaussian process with squared exponential covariance function. We chose to use a Gaussian process model for h^{ij} because this class of distributions should be able to capture many types of non-quadratic behavior.

We will use the same random function distributions for both our simulations and case study data analysis. In this section, we describe the hierarchical models for the simulation setting because it has a single layer of hierarchy, which makes the presentation simpler. The model specification for the case study analysis will include an additional layer of hierarchy, but the probabilistic structure for that additional layer will be identical.

Let $k_s^i(t)$ denote the i^{th} unit's realization of k_s. We model k_s^i as

$$\log(k_s^i(t)) = g^i(t) + h^i(t), \tag{4.4}$$

where $g^i(t) = \theta_1^i + \theta_2^i(t-\theta_3^i)^2$, $h^i \sim \mathcal{GP}\left(0, K_{\alpha,\rho}(t,t')\right)$, and $K_{\alpha,\rho}(t,t') = \alpha^2 \exp\left\{-\frac{(t-t')^2}{2\rho^2}\right\}$. Again, we model θ^i hierarchically: $\theta^i|\theta \sim \mathcal{N}(\theta, \Sigma_\theta)$. The vector θ denotes the population average and Σ_θ measures between-unit variation around θ. In order to conduct our analysis using Bayesian methods, the parameters θ, α, and ρ will have prior distributions depending on hyperparameters. Priors will be chosed so that the prior predictive distribution produces "reasonable" data.

Each component of θ describes a feature of the population curve f. Of particular interest for this application is the sign of θ_2, which will indicate whether the typical unit has convex or concave g^i. The (2,2) element of Σ_θ, which controls the spread of θ_2^i around the population mean θ_2, measures the strength of the tendency toward either convexity or concavity. The other two components θ_1 and θ_3 are best interpreted together. The function f achieves its maximum or minimum value of θ_1 at time $t = \theta_3$ (whether θ_1 is the maximum or minimum depends on the sign of θ_2).

The parameters of the Gaussian process, α^2 and ρ, control the marginal variance and smoothness of the latent functions h^i. For all $t \in \mathbb{R}$, Var$(g^i(t)|\rho, \alpha, t) = \alpha^2$. Larger α in the data-generating process leads to greater inter-subject variation in latent functions g^i. The parameter ρ controls the smoothness of the latent functions. When $|t_1 - t_2| > 2 * \rho$, then the correlation between $g^i(t_1)$ and $g^i(t_2)$ is negligible.

4.5 Latent function identification in a synthetic mechanistic model

4.5.1 Synthetic model description

The structural model used in this simulation study has two compartments. The transit of the compartments' contents is governed by a linear system of differential equations. The contents of the first compartment transit to the second according to a time-varying first-order rate $k_s(t)$. The contents of the second compartment empty at a constant rate k_e. In order to make the model easier to describe, for the remainder of the section we will assume that the mechanistic model describes the transit of a drug *in vivo*.

Let $y(t) = (y_1(t), y_2(t))' \in \mathbb{R}^2$ denote the mass of the drug at time t in the first and second compartments, respectively. The following system of differential equations describe the evolution of the $y(t)$:

$$\frac{\partial y}{\partial t} = \begin{bmatrix} -k_s(t) & 0 \\ k_s(t) & k_e \end{bmatrix} y(t). \tag{4.5}$$

In order to calculate the mass of drug in each compartment at time $t = t_0 + \delta$, $\delta > 0$, the system of differential equations is solved. Given initial conditions $y(t_0) = (y_1(t_0), y_2(t_0))'$,

the system has the following closed-form solution:

$$y(t) = \begin{bmatrix} y_1(t_0)e^{-\phi(t)\delta} \\ \frac{\phi(t)y_1(t_0)}{k_e-\phi(t)}e^{-\phi(t)\delta} - \left(y_2(t_0) - \frac{\phi(t)y_1(t_0)}{k_e-\phi(t)}\right)e^{-k_e\delta} \end{bmatrix}, \tag{4.6}$$

where $\phi(t) = \exp\left\{\int_{t_0}^{t_0+\delta} k_s(x)dx\right\}$. In practice, we will treat $k_s(t)$ as constant over the time interval $[t_0, t_0 + \delta]$. This choice simplifies solving the system of equations 4.5 whenever an analytical solution to $\int k_s(t)dt$ does not exist. We approximate the true solution to the system by using $\phi(t) = \exp\{k_s(t_0)\}$. In practice, δ will be small, and thus $k_s(t_0) \approx k_s(t')$ for all $t' \in [t_0, t_0 + \delta]$.

Since this exercise is used as a proof of concept, we wish to incorporate features of the case study dataset that make the case study analysis challenging. In most pharmacokinetic analyses, one directly models the mass of drug in all compartments, but observes the concentration of the drug in a subset of compartments (usually only one). In this example, the mass of drug in compartment one will be unobserved and the concentration of drug in compartment two will be measured at sampling times $t' \in \{t_0, \ldots, t_n\} \subset [0, 1]$. Let v be the volume of compartment two so that measurements consist of $c(t) := y_2(t)/v$ rather than $y_2(t)$.

In most analyses, the volume v is also treated as a latent variable and assumed to vary between subjects.We model subject i's volume hierarchically. We will treat v as the population mean volume and assume that $v^i|v \sim \mathcal{N}(v, \tau_v)$, where τ_v represents the inter-subject volume variance. Therefore, $c^i(t) = y_2^i(t)/v^i$ will denote subject i's drug concentration at time t.

Measurement model for $c^i(t)$

In order to estimate the two-compartment model's parameters, we must relate the observed data, concentrations in compartment two, to the compartmental model's predicted drug concentrations. Let $p^i(t)$ denote the concentration of drug in compartment two at time

t and let $c^i(t) = y_2^i(t)/v^i$ denote the model's prediction of subject i's drug concentration. We both generate and model $\log\left(p^i(t)\right)$ as a Gaussian random variable with the following structure:

$$\log(p^i(t)) = \log\left(c^i(t)\right) + \epsilon^i(t),\ \epsilon^i(t)|c^i(t) \overset{iid}{\sim} \mathcal{N}(0, \sigma^2). \tag{4.7}$$

The parameter σ represents the multiplicative error. On average, the observed concentration $p^i(t)$ will be $\exp(\sigma)\%$ away from the model's predicted concentration $c^i(t)$.

Choice of prior distributions

In most pharmacokinetic analyses, the analyst does not have strong prior beliefs concerning the parameters of the model. In practice, due to expectations of what "reasonable" data should look like, usually the parameters are constrained to be in a much smaller subspace of the full parameter space. We aim to use prior distributions that are weakly informative, meaning that they reflect our belief that the parameters likely live in the region of the parameter space that produces realistic data. For example, one would place little prior density on parameter combinations that generate data that routinely exceeds plasma drug concentrations of 200 μg/mL if such concentrations have not occurred during previous studies.

Table 4.4 lists each parameter, the probability model used for its prior distribution, and the values of the hyperparameters used during our simulations. Again, note that hyperparameter selection was done so that the predictive prior distribution produced reasonable synthetic data. Slightly different choices could have been used without greatly affecting the estimated posterior distribution.

Two parameters' prior models deserve additional comment. First, we use the scaled Inverse-Wishart (IW) prior for the covariance matrix Σ_θ. This prior model places nearly uniform density on the correlations between elements of θ_i. Given $\log(\tau_\theta) \sim \mathcal{N}(a_{\tau_\theta}, s^2_{\tau_\theta})$ and $\Psi_\theta \sim \text{IW}(4, I_3)$, then, $\Sigma_\theta = \text{diag}(\tau_\theta)\Psi_\theta\text{diag}(\tau_\theta)$. Second, we do not expect to be able

Table 4.4: Prior models for parameters of two-compartment model

Parameter	Prior Distribution	Hyperparameter selection
θ	$\mathcal{N}(a_\theta, s_\theta^2)$	$a_\theta = (0.0, 0.0, 0.5)'$, $s_\theta = (0.5, 1.50, 0.25)'$
$\log(\sigma)$	$\mathcal{N}(a_\sigma, s_\sigma^2)$	$a_\sigma = -3.0$, $s_\sigma = 0.75$
$\log(\alpha)$	$\mathcal{N}(a_\alpha, s_\alpha^2)$	$a_\alpha = -0.8$, $s_\alpha = 0.66$
ρ	Inv-Gamma(a_ρ, b_ρ)	$a_\rho = 2.93$, $b_\rho = 0.42$
$\log(k_e)$	$\mathcal{N}(a_{k_e}, s_{k_e}^2)$	$a_{k_e} = -0.5$, $s_{k_3} = 1.0$
v	$\mathcal{N}(a_v, s_v^2)$	$a_v = 4.0$, $s_v = 0.25$
$\log(\tau_v)$	$\mathcal{N}(a_{\tau_v}, s_{\tau_v}^2)$	$a_{\tau_v} = -1.0$, $s_{\tau_v} = 0.5$
$\log(\tau_\theta)$	$\mathcal{N}(a_{\tau_\theta}, s_{\tau_\theta}^2)$	$a_{\tau_\theta} = (-1.6, -1.6, -2.0)'$, $s_{\tau_\theta} = (0.5, 0.5, 0.1)'$
Ψ_θ	IW(ν, M)	$\nu = 4$, $M = I_3$

to estimate properties of the random function that occur at a scale smaller than our sampling frequency or larger than the total observation time. The hyperparameters $a_\rho = 2.93$ and $b_\rho = 0.42$ were selected so that less than 1% of the prior density for ρ fell below 0.05 and above 1.0.

Data-generating populations

Our simulation studies will involve two data-generating populations. For the first population, the quadratic latent function will be the dominant component of the latent function $\log(k_s^i)$. For the second, the Gaussian process component will be the dominant component. We will call these data-generating populations the *parametric data-generating population* and *Gaussian process data-generating population*, respectively. Here we present the data-generating parameters for each setting, beginning with the parameters that are common to both data-generating populations.

Figure 4.6a shows 100 realizations of the latent function $\log(k_s^i)$ from each data-generating process. Functions drawn from the Gaussian process data-generating process are smooth and oscillate between being locally concave and convex. The quadratic functions are obviously either globally concave or convex. Both data-generating processes produce functions with roughly the same marginal variance across t. This, we hope will not unfairly bias the posterior estimates toward one component, and instead allow the data to select the compo-

Table 4.5: Data-generating values of parameters shared by both data-generating populations.

Parameter	Data-generating values
k_e	1
σ	0.05
v	4
τ_v	.25
Ψ_θ	$\begin{bmatrix} 1 & -.8 & 0 \\ -.8 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

Table 4.6: Data-generating values of latent function parameters.

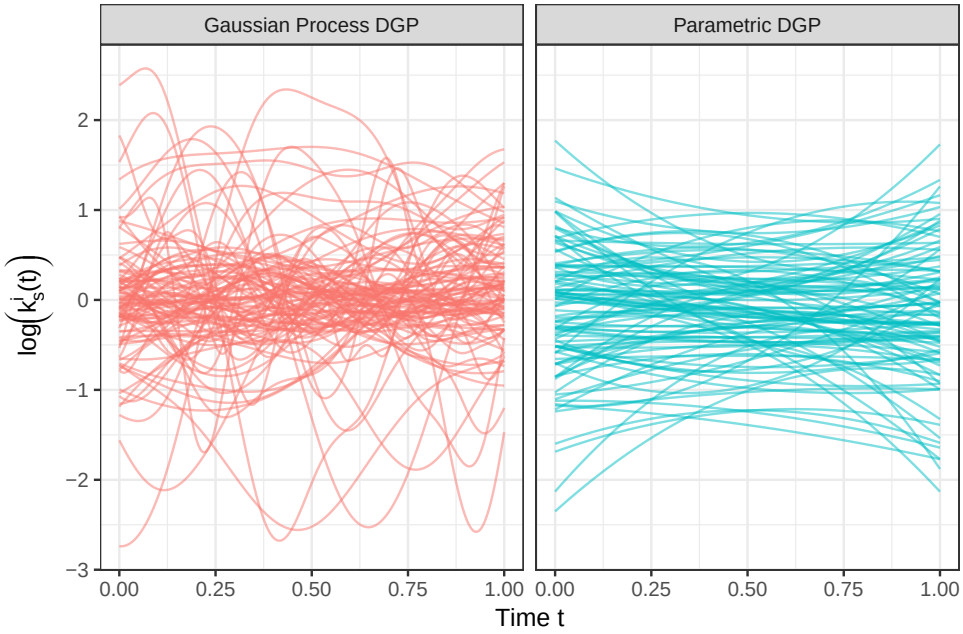
Parameter	Parametric generating values	Gaussian process generating values
α	.02	1.00
ρ	0.75	0.25
θ	$(1.5, -5, .25)'$	$(0.0, 0.0, 0.5)'$
τ_θ	$(1/2, 1, .2)'$	$(0.05, 0.05, 0.1)'$

nent that best describes the data.

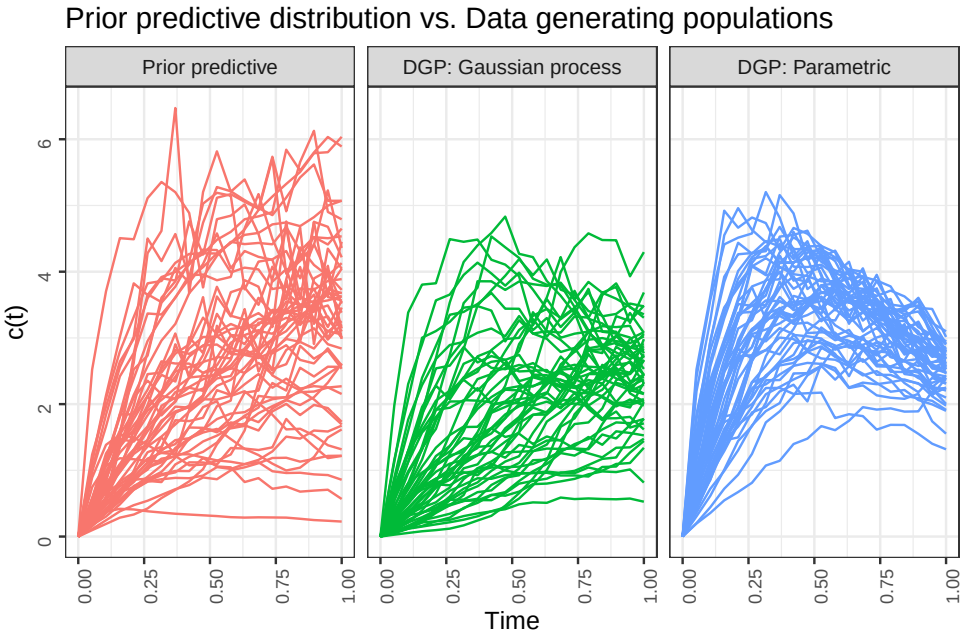
Figure 4.6b shows the data realized under these data-generating populations. We also show prior predictive draws that demonstrate that our prior model generates data that look reasonably similar to the true data-generating populations, albeit with more variation in rate of increase, shape, and spread. The parametric concentration curves increase more quickly than the Gaussian process concentration curves. Additionally, the parametric curves achieve their maximum plasma concentrations over a shorter time period.

4.5.2 Simulation results

We conducted a small simulation study in order to assess our ability to recover distributions of latent functions generated by the model described in Section 4.5.1. The simulation study generated data from both the parametric and Gaussian process data-generating populations with 20, 50, and 100 subjects or studies per synthetic dataset. For each setting, we generated 10 synthetic data sets and used Stan to estimate model parameters [44]. Multi-chain Monte



(a) Examples of latent functions from both the Gaussian process and parametric data-generating processes



(b) Examples of data generated by two compartment model. The left panel shows prior predictive draws, the center panel shows realizations from the Gaussian process data-generating process, and the right panel shows data from the parametric data-generating process.

Figure 4.6: Typical data from the synthetic data-generating populations.

Carlo (MCMC) was conducted using four chains to assess convergence and multi-modality. For each chain, 1000 warm-up and 1000 sampling iterations were used to estimate the parameters' joint posterior distribution. In most cases, the chains converged to the same region of the parameter space.

Overall, the results of the simulation study are a mixed success. Broadly speaking, the population mean latent function parameters are identified by the data, while the variance components of the latent function distributions were not. Figure 4.7 shows posterior mean estimates for the parameters α, ρ, $\theta \in \mathbb{R}^3$, and $\sqrt{(\text{diag}(\Sigma_\theta))} \in \mathbb{R}^3$, the inter-subject standard deviations in θ^i. When the data-generating population places more weight on the Gaussian process latent function, posterior mean estimates of α and ρ appear to be unbiased for all sample sizes (aqua triangles). When the latent function is approximately quadratic, these parameters are approximately identified. Importantly, posterior estimates of $\log(\alpha)$ are sufficiently negative so that the marginal standard deviation of the Gaussian process is very small. Thus, on average, the Gaussian process component impacts the latent function to a small degree. When α is small, posterior estimates of ρ are inconsequential.

We next turn to θ, the parameters of the quadratic latent function. When the latent function is approximately quadratic, posterior estimates of θ_1 appear to be approximately unbiased. For both θ_2 and θ_1, we see that the posterior estimates are converging to the true data-generating value as n increases. For these parameters, the data-generating value of the parameter was a low prior-density point in the parameter space, which makes convergence slower. Likewise, when the latent functions are predominantly generated by the Gaussian process, posterior means also appear to be approximately unbiased for all sample sizes.

The last 3 panels of Figure 4.7 show estimates for the variance components of θ^i. We specifically show the inter-subject standard deviations of the quadratic latent function parameters. Here, we are unable to estimate each population's data-generating values. The one exception is perhaps the standard deviation of θ_1^i for the parametric data-generating population. Overall, the variance components of the parametric latent function are not

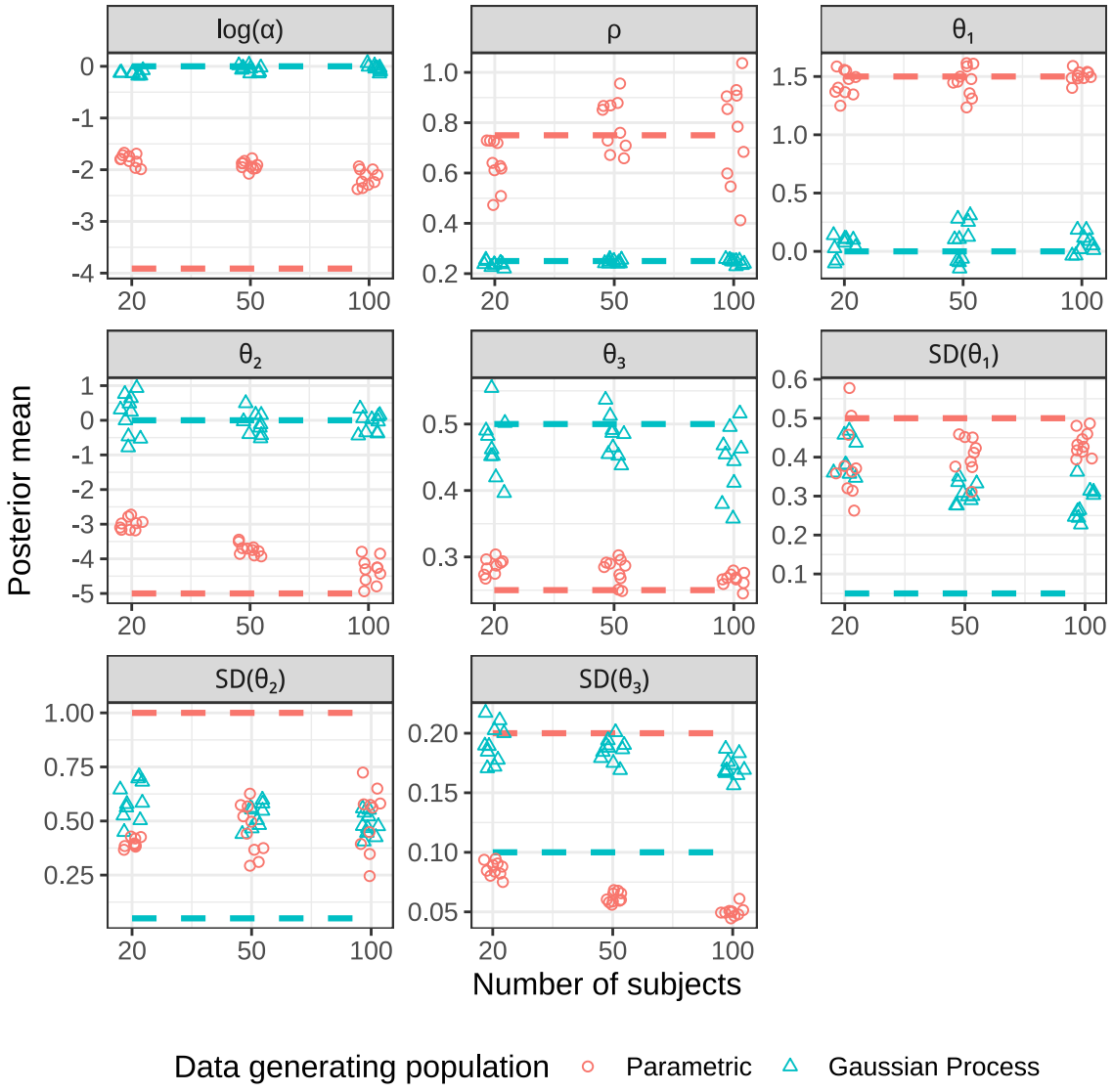


Figure 4.7: Posterior estimates from the two-compartment simulation study. Each panel presents posterior estimates for a single parameter and each point in a panel represents the posterior estimate from a single simulation trial. The x-axis indexes the number of subjects in the synthetic dataset used to estimate the parameter. The y-axis represents the posterior mean of the parameter. Points are colored based on whether the true data-generating population placed more weight on the parametric or Gaussian process latent functions. Horizontal dashed lines show the true value of the data-generating parameter, colored to correspond to each data-generating parameter. Posterior estimates converged to the true data-generating parameter when the posterior means (represented by points) have clustered around the dashed line of the same color.

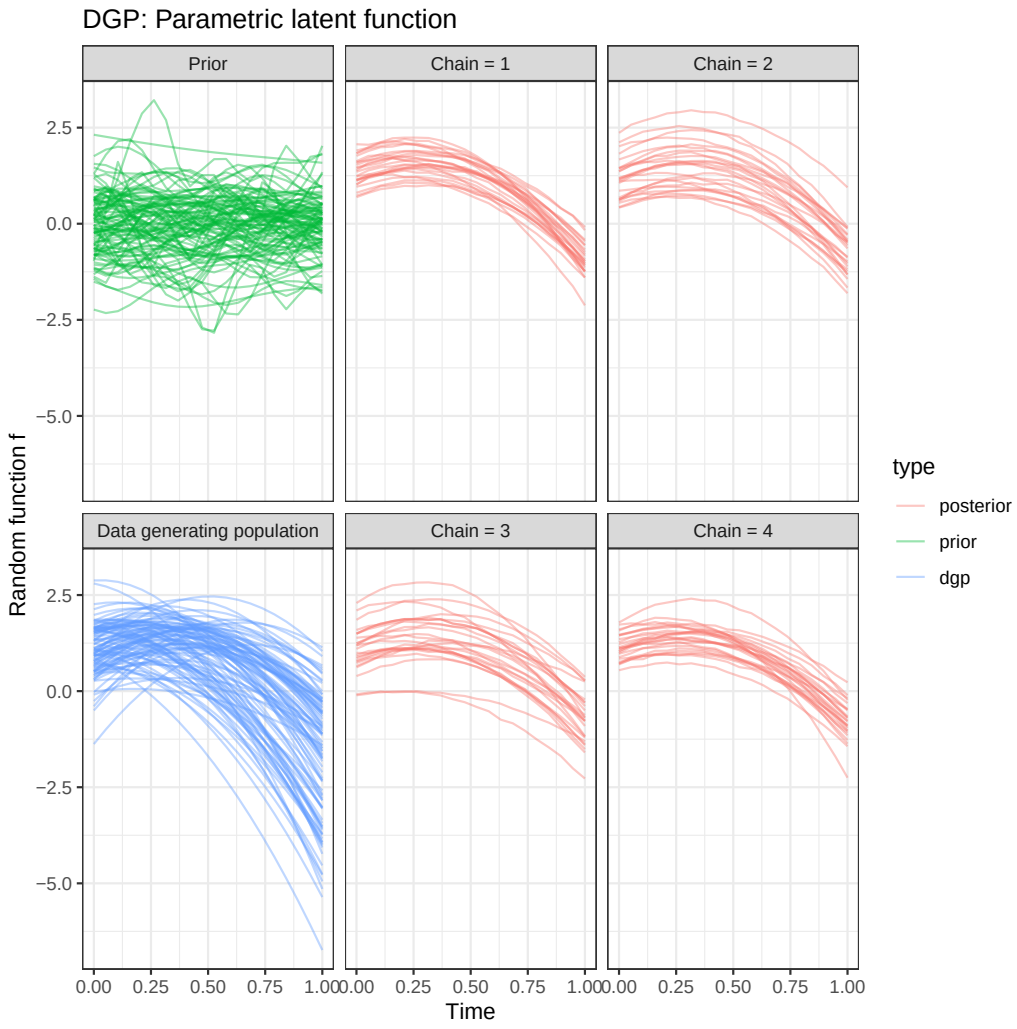


Figure 4.8: A comparison of latent functions drawn from the prior predictive (green), data-generating (blue), and posterior predictive distributions (red) for a population that placed more weight on the parametric latent function.

identified by the data.

Figure 4.8 shows results for one synthetic dataset with quadratic latent functions. The panels in the first and column show the prior predictive and data-generating distributions of latent $\log(k_s^i)$. The remaining panels show posterior predictive draws from the estimated distribution of $\log(k_s^i)$. The data-generating latent functions exhibit much greater spread than the posterior predictive estimates, which is driven by the underestimation of the inter-subject variation.

4.5.3 Synthetic data simulation discussion

It is unclear to us why the variance parameters are much more difficult to estimate. In part, we attribute the lack of identification to the nature of the data. Specifically, the data contain very little information about the latent function after most of the contents of the first compartment has transfered to the second compartment. After one-third of the total study time had passed, over 75% of the contents of the stomach had transfered to compartment two, averaged across the population. Since the contents of the compartments are indirectly what inform estimates of the parameters' values, this leaves the method with little information to estimate the emptying rate at the end of the study. For a quadratic latent function, much of the inter-subject variation is described by behavior far from the peak emptying rate, which is exactly where the data contain the least information.

Another possible explanation for the bias is the choice to model inter-subject differences on the log-scale. The latent function interacts with data on the linear scale after exponentiation. Thus, substantial inter-subject variation in latent k_s that falls below 0 is compressed after exponentiation. One can see how this plays out in Figure 4.8. For the data-generating distribution, substantial variation between latent functions occurs late in the study time-course and this variation largely falls below 0. This leads to an estimated distribution of latent functions that under estimates the true level of inter-subject variation.

Although estimates of the latent function variance components are biased, we do recover important features of concentration profiles. The maximum concentration (C-max) that a subject experiences is one of the relevant regulatory features, since C-max can affect both the safety and efficacy of the drug. In Figure 4.9, we compare the data-generating distribution of C-max to the posterior predictive distribution of C-max. For all samples sizes n, we approximately recover the population C-max distribution and the approximation improves as n grows. This result suggests that even if there is bias in the estimated model, a fitted model might be able to capture the population variation in interesting or relevant summaries of the observed data.

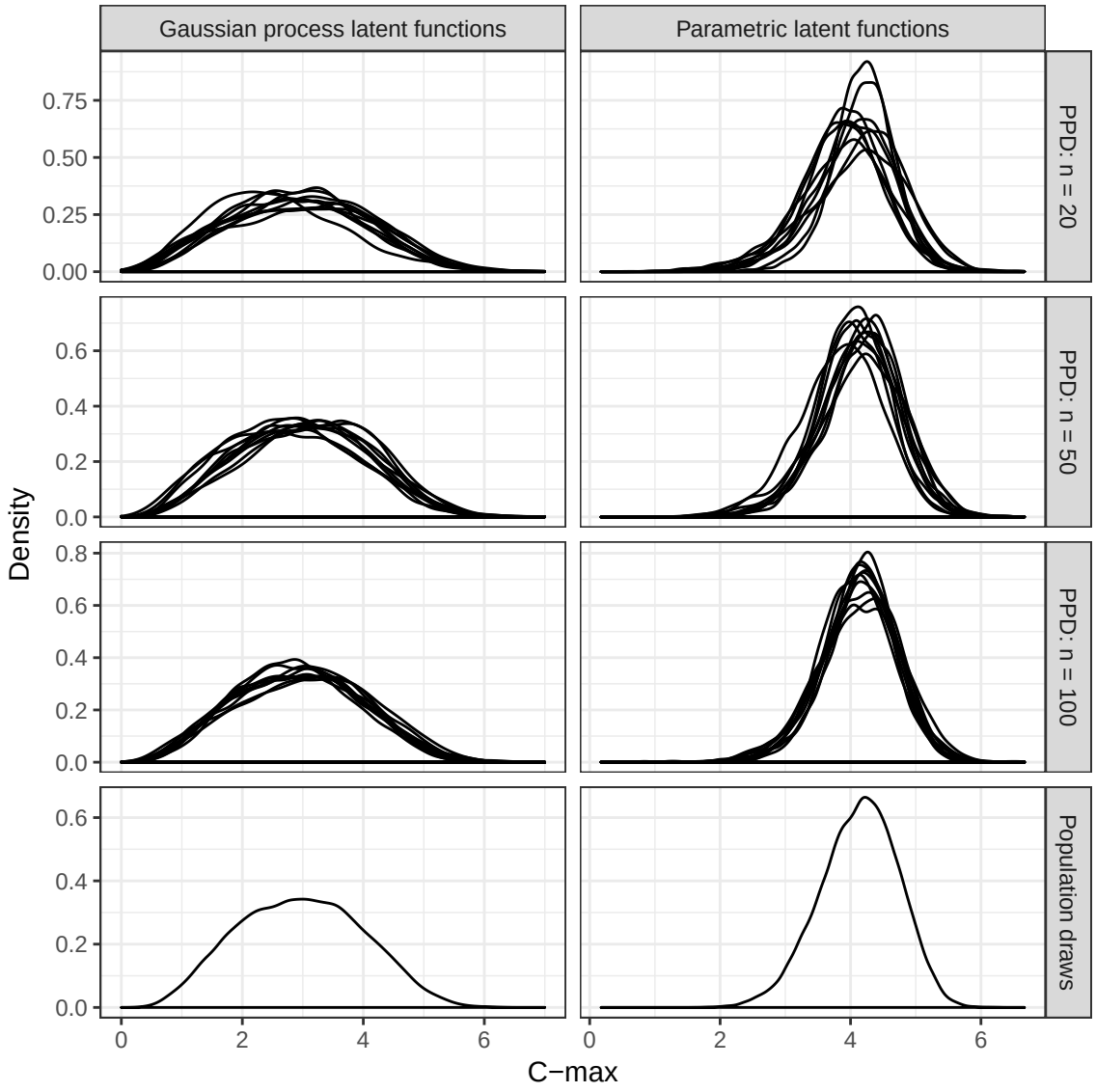


Figure 4.9: A comparison of the distribution of concentration maximums (C-max) between the data-generating population and posterior predictive distributions (PPDs). Simulation results are stratified by data-generating population and sample size n. Each line represents the estimate from a single synthetic data set.

4.6 Discussion

We were unable to successfully use our proposed approach to study mechanisms in the case study data. Attempts to fit a model with the latent function structure described in Section 4.4.1 were unsuccessful. The Monte Carlo chains converged to different modes (or failed to converge at all) and Stan's sampler diagnostics suggested that we had failed to take a valid sample from the parameter's posterior.

The analysis of the case study data was challenging for several reasons. First, the case study had an additional layer of hierarchy, since a subset of the study participants underwent two intubation studies. Additionally, there is reasonable evidence of stable within-subject kinetics. This finding necessitated the use of both subject and study-within-subject random effects, although the data to estimate the additional parameters was quite limited.

Second, the case study analysis was faced with partially-missing data problems that did not affect the synthetic data simulation studies. Roughly one third of the small intestine aspiration samples were not taken. This was commonly caused by an absence of fluid at the aspiration port during sampling. Additionally, the sampling times for both the plasma and aspiration samples were irregularly spaced. The study called for the irregular spacing because it was expected that most of the ibuprofen absorption would take place within the first few hours after dosing. In order to capture the shape of the plasma concentration trajectories, more frequent sampling was conducted during the first three hours of the seven-hour study (Table 4.1). However, many subjects experienced peak plasma concentrations later. This design choice caused the late-peak concentration data to contain less information about mechanistic parameters than early-peak concentration data.

Finally, the case study data had additional unobserved factors, both constant and time-varying, that impacted ibuprofen kinetics. The most troublesome of these factors was small intestine volume, which is known to vary dramatically over time and was unmeasured. Our models treated small intestine volume as a constant within-subject latent variable. Other time-varying latent factors can act as nuisance parameters when trying to estimate a sep-

arate latent factor. The model attempt to use the single flexible latent factor to explain variation caused by multiple mechanisms. This leads to latent factor estimates that are the sum total of multiple unobserved processes.

This project has clarified several other challenges with mechanistic modeling in biological systems. First, longitudinal, mechanistic biological data contains different amounts of information about parameters over the study time course. This arises because information about the system's parameters is dependent on data being observed in each part of the system. After the observable leaves one part of the system, the data are no longer informative about the parameters of that part of the system.

In PK/PD modeling, once a compartment and the compartments that proceed it no longer contain drug, the data no longer contain information about the rate parameters that govern drug kinetics in the empty compartments. In the case simulation examples, this issue arose because the identification of inter-subject variation was dependent on data from measurements late in the study. By halfway through the synthetic studies, for most subjects the majority of the drug has left compartment one. The second half of the study had little signal to contribute to the estimation of the distribution of the random latent functions. This ultimately led to biased estimates of the latent function distribution.

This also raises the question of whether one needs to correctly estimate features of the latent function distribution that do not impact the observed data. Despite the fact that our estimates of the latent function distribution were biased, we correctly recovered the distribution of maximum concentrations (see Figure 4.9). In any mechanistic analysis, the model will only approximate the biological system. If we recover a random effects structure that produces reasonable realizations, perhaps the fitted model can still be informative about variation between units, if not the mechanisms of interest.

Studying mechanisms in complex biological systems highlights the tension between model flexibility and identifiability. PK/PD analysis datasets rarely have large sample sizes, which requires compartmental and statistical models that roughly approximate the system's

true complexity. Conversely, the proposed latent function specification was capable of describing many different emptying behaviors. We used a flexible specification because we did not want to guide the analysis too strongly towards our expectations. Unfortunately, at least for our case study, the parameters of the latent function distributions were not identified by the study data.

When flexible models are not identified by the data, one generally chooses a more restrictive model. However, with a more restrictive model, the data are less likely to be able to choose an alternative mechanisms. Instead, the data will choose the best option among the available options. Likely a better approach for our purposes is *post-hoc* analysis of a semi-parametric model, followed by assessing sensitivity to choice of prior or other modeling decisions.

We proposed one approach to assessing whether data from biological systems are consistent with a mechanism, but there may be more efficient methods. In particular, the our approach introduced two potential models for the time-varying latent variable. Finding strong evidence for the mechanism requires that the posterior concentrates around one of the two components. In noisy, complex systems, the concentration onto one component might be extremely slow. There may be simpler, more parsimonious ways to identify mechanisms in biological systems.

During the course of our work, we found that semi-parametric specifications of latent functions were easier to work with than the quadratic functions. However, these fitted models are less interpretable and do not clearly represent the underlying mechanism or hypothesis. To assess whether a fitted model is consistent with the proposed mechanism, substantial *post-hoc* analysis is needed. Additionally, models with semi-parametric latent functions often overfit the observed data. In the PK/PD ibuprofen case study, posterior predictive draws from the fitted model had features not found in the observed study data.

We used standard specifications of semi-parametric latent functions found in the functional data analysis literature. These specifications implicitly link the smoothness and

marginal variance of the random function distribution. The smoothness-variance connection was partially responsible for unrealistic posterior predictive draws. When we attempted to decouple the latent smoothness-marginal variance, we found that the case study data were unable to identify the parameters that disconnected the two function traits.

APPENDIX A

Mediation analysis literature review

More recently, mediation analysis has been place in the counterfactual outcomes framework of causal analysis. Approaching mediation analysis from this perspective allowed [45, 15] to establish a minimum set of conditions such that the direct and indirect effects estimated by the coefficient-product method are valid. We now introduce the counterfactual framework of causal analysis, which closely follows the development of [15].

In the causal inference literature, Rubin's counterfactual framework is widely known as a useful tool for establishing causal relationships between treatment and outcome measures [46]. In brief, the counterfactual framework posits that each observational unit has a potential outcome for every treatment level. Suppose we are studying the effect of a diet regime on weight loss. The binary variable X_i contains the experimental arm to which the i^{th} subject is assigned, with $X_i = 1$ if subject i is assigned a diet and $X_i = 0$ if subject i is assigned to the control group. We then have two potential outcomes, the weight that subject i gains or loses under control, $Y_i(X_i = 0)$, and treatment, $Y_i(X_i = 1)$. In practice we can only observe one of the two potential outcomes, $Y_i = Y_i(X_i)$ and never $Y_I(1 - X_i)$. Often the practitioner is interested in the average treatment effect, or $\mathbb{E}\left[Y(1) - Y(0)\right]$. Rubin showed that if treatment assignment is independent of the potential outcomes, then the difference in group means is an unbiased estimator of the treatment effect.

[15] extended the counterfactual framework to encompass the causal mediation analysis setting. To do so, two new potential outcomes must be defined. Once again, let X_i be a

binary variable storing experimental group assignment. Define M_i be the i^{th} individual's observed value of the mediating variable. The two potential mediators $M_i(X_i = 0)$ and $M_i(X_i = 1)$ are the values that the mediator takes if the i^{th} subject is assigned to the control or treatment group respectively. The potential outcomes $Y_i(X_i, M_i(X_i))$ for $X_i = 0$ and $X_i = 1$ are now a function of both the treatment assignment and mediator value, which itself is a function of treatment assignment. In our earlier dieting example, M_i could be the average number of minutes exercised each week by individual i over the course of the study. The scientist conducting the study might believe that following a proscribed diet might cause a participant to undertake additional exercise. The hypothesized outocme of this additional exercise is of course greater weight loss.

The counterfactual causal inference framework allows us to define and then estimate both the direct and indirect effects in terms of the potential outcomes. We use the notation of [15] to define each effect. First, the indirect effect for individual i is defined as

$$\delta_i(x) = Y_i(x, M_i(X_i = 1)) - Y_i(x, M_i(X_i = 0)). \tag{A.1}$$

The indirect effect represents the change in response if we were to hold the treatment fixed but allowed the mediator take potential values realized under treatment and control assignment. In our running example, $Y_i(x, M_i(X_i = 1))$ is the amount of weight loss under treatment assignment x and the exercise volume realized under treatment assignment. Conversely, $Y_i(x, M_i(X_i = 0))$ is the amount of weight loss under treatment assignment x, but the exercise volume realized under control assignment. The indirect effect takes the difference in these two potential outcomes, and, thus the indirect effect represents the change in outcome measure directly related to the mediating variable's change caused by a change in the treatment, while accounting for the treatment's effect. The direct effect is defined as

$$\xi_i(x) = Y_i(X_i = 1, M_i(x)) - Y_i(X_i = 0, M_i(x)). \tag{A.2}$$

The direct effect represents the change in response if the treatment assignment changes from treatment to control, but the mediator value is held constant at the value realized under treatment x. The quantity $\xi_i(x)$ is the amount of weight loss caused by the diet holding the amount of exercise constant at the level realized under treatment assignment x.

Using this notation, we can then express the LSEM in terms of the potential outcomes, which gives us a method of estimating both the direct and indirect effects $\xi_i(x)$ and $\delta_i(x)$.

- $Y_i(X_i, M_i(X_i)) = \mu_0 + \mu_1 X_i + \mu_2 C_i + \epsilon_{1i}(X_i, M_i(X_i)$,
- $M_i(X_i) = \alpha_0 + \alpha_1 X_i + \alpha_2 C_i + \epsilon_{2i}(X_i$, and
- $Y(X_i, M_i(X_i)) = \beta_0 + \gamma X_i + \beta_1 M_i + \beta_2 C_i + \epsilon_{3i}(X_i, M_i(X_i))$.

Furthermore, [45] showed that if the *no interaction* and *sequential ignorability* assumptions are met, then these estimates are consistent. First, the *no interaction* assumption says that the indirect and direct mediation effects do no depend on x, or $\delta_i(0) = \delta_i(1)$ and $\xi_i(0) = \xi_i(1)$. The *sequential ignorability* assumption is more nuanced.

Assumption 1 (Imai, Keele, Yamamoto, 2010)**.** *Assume that the following statements of conditional independence hold:*

1. $\{Y_i(x', m), M_i(t)\} \perp\!\!\!\perp X_i \mid C_i = c$,
2. $Y_i(x', m) \perp\!\!\!\perp M_i(x) \mid C_i = c, X = x$, *where* $0 < Pr(X_i = x \mid C_i = c)$ *and* $0 < Pr(M_i(x) = m \mid X_i = x)$ *for* $x = 0, 1$ *and all* $c \in \mathbb{R}^p$, $m \in \mathbb{R}$.

For a thorough discussion of Assumption 1, see [45]. Briefly, what Assumption 1 says is that conditional on any pretreatment confounders C_i, the potential outcomes and treatment assignment are independent. This assumption is met through randomized treatment assignment and therefore reasonable in practice. The second assumption cannot be empirically verified, however. It says that the potential outcomes are $M_i(x)$ and $Y_i(x', m)$ are independent, conditional on the pretreatment confounders C_i and the treatment assignment

X_i. We cannot verify this assumption because values of the mediator are not randomly assigned through experimental design, in general.

BIBLIOGRAPHY

[1] J. Savage *et al.*, "Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence," *Nature Genetics*, vol. 50, no. 7, pp. 912–919, 2018.

[2] B. Hens *et al.*, "Formulation predictive dissolution (fPD) testing to advance oral drug product development: An introduction to the US FDA funded '21st Century BA/BE' project," *International Journal of Pharmaceutics*, vol. 548, no. 1, pp. 120 – 127, 2018.

[3] T. VanderWeele, *Explanation in causal inference: methods for mediation and interaction*. Oxford University Press, 2015.

[4] S. Wright, "The relative importance of heredity and environment in determining the piebald pattern of guinea-pigs," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 6, no. 6, pp. 320–332, 1920.

[5] S. Wright, "The method of path coefficients," *The Annals of Mathematical Statistics*, vol. 5, pp. 161–215, 09 1934.

[6] M. Sobel, "Asymptotic confidence intervals for indirect effects in structural equation models," *Sociological Methodology*, vol. 13, pp. 290–312, 1982.

[7] M. Sobel, "Some new results on indirect effects and their standard errors in covariance structure models," *Sociological Methodology*, vol. 16, pp. 159–186, 1986.

[8] D. MacKinnon, C. Lockwood, J. Hoffman, S. West, and V. Sheets, "A comparison of methods to test mediation and other intervening variable effects," *Psychological Methods*, vol. 7, no. 1, p. 83, 2002.

[9] D. MacKinnon, *Introduction to statistical mediation analysis*. Routledge, 2008.

[10] H. Hyman, *Survey design and analysis: Principles, cases, and procedures*. Free Press, 1955.

[11] C. Judd and D. Kenny, "Process analysis: Estimating mediation in treatment evaluations," *Evaluation Review*, vol. 5, no. 5, pp. 602–619, 1981.

[12] R. Baron and D. Kenny, "The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations," *Journal of Personality and Social Psychology*, vol. 51, no. 6, pp. 1173–1182, 1986.

[13] K. Preacher and A. Hayes, "SPSS and SAS procedures for estimating indirect effects in simple mediation models," *Behavior Research Methods, Instruments, & Computers*, vol. 36, no. 4, pp. 717–731, 2004.

[14] K. Preacher and A. Hayes, "Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models," *Behavior Research Methods*, vol. 40, no. 3, pp. 879–891, 2008.

[15] K. Imai, L. Keele, and D. Tingley, "A general approach to causal mediation analysis," *Psychological Methods*, vol. 15, no. 4, pp. 309–326, 2010.

[16] O. Chén, C. Crainiceanu, E. Ogburn, B. Caffo, T. Wager, and M. Lindquist, "High-dimensional multivariate mediation with application to neuroimaging data," *Biostatistics*, vol. 19, no. 2, pp. 121–136, 2018.

[17] A. Derkach, R. Pfeiffer, T. Chen, and J. Sampson, "High dimensional mediation analysis with latent variables," *Biometrics*, vol. 75, no. 3, pp. 745–756, 2019.

[18] V. Zhao, M. Lindquist, and B. Caffo, "Sparse principal component based high-dimensional mediation analysis," *Computational Statistics & Data Analysis*, vol. 142, p. 106835, 2020.

[19] Y. Huang, "Joint significance tests for mediation effects of socioeconomic adversity on adiposity via epigenetics," *Annals of Applied Statistics*, vol. 12, pp. 1535–1557, 09 2018.

[20] H. Zhang, Y. Zheng, Z. Zhang, T. Gao, B. Joyce, G. Yoon, W. Zhang, J. Schwartz, A. Just, E. Colicino, P. Vokonas, L. Zhao, J. Lv, A. Baccarelli, L. Hou, and L. Liu, "Estimating and testing high-dimensional mediation effects in epigenetic studies," *Bioinformatics*, vol. 32, pp. 3150–3154, 06 2016.

[21] A. Yurtsever, M. Udell, J. Tropp, and V. Cevher, "Sketchy decisions: Convex low-rank matrix optimization with optimal storage," *arXiv preprint arXiv:1702.06838*, 2017.

[22] W. Burns, E. Peters, and P. Slovic, "Risk perception and the economic crisis: A longitudinal study of the trajectory of perceived risk," *Risk Analysis: An International Journal*, vol. 32, no. 4, pp. 659–677, 2012.

[23] T. Jernigan, S. Brown, *et al.*, "Introduction," *Developmental Cognitive Neuroscience*, vol. 32, pp. 1 – 3, 2018. The Adolescent Brain Cognitive Development (ABCD) Consortium: Rationale, Aims, and Assessment Strategy.

[24] M. Luciana *et al.*, "Adolescent neurocognitive development and impacts of substance use: Overview of the adolescent brain cognitive development (ABCD) baseline neurocognition battery," *Developmental Cognitive Neuroscience*, vol. 32, pp. 67 – 79, 2018. The Adolescent Brain Cognitive Development (ABCD) Consortium: Rationale, Aims, and Assessment Strategy.

[25] D. Hinkley, "Fisher's development of conditional inference," in *R.A. Fisher: An Appreciation* (S. Fienberg and D. Hinkley, eds.), (New York, NY), pp. 101–108, Springer New York, 1980.

[26] R. Fisher, "The logic of inductive inference," *Journal of the Royal Statistical Society*, vol. 90, no. 1, pp. 39–82, 1935.

[27] R. Fisher, "Theory of statistical estimation," in *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 22, pp. 700–725, Cambridge University Press, 1925.

[28] J. Kalbfleisch, "Sufficiency and conditionality," *Biometrika*, vol. 62, no. 2, pp. 251–259, 1975.

[29] R. Little, "Testing the equality of two independent binomial proportions," *The American Statistician*, vol. 43, no. 4, pp. 283–288, 1989.

[30] L. Choi, J. Blume, and W. Dupont, "Elucidating the foundations of statistical inference with 2 x 2 tables," *PloS One*, vol. 10, no. 4, 2015.

[31] N. Breslow, N. Day, K. Halvorsen, R. Prentice, and C. Sabai, "Estimation of multiple relative risk functions in matched case-control studies," *American Journal of Epidemiology*, vol. 108, no. 4, pp. 299–307, 1978.

[32] N. Breslow, N. Day, and J. Schlesselman, "Statistical methods in cancer research. Volume 1—The analysis of case-control studies," *Journal of Occupational and Environmental Medicine*, vol. 24, no. 4, pp. 255–257, 1982.

[33] H. Chernoff, "On the distribution of the likelihood ratio," *The Annals of Mathematical Statistics*, pp. 573–578, 1954.

[34] M. Drton, "Likelihood ratio tests and singularities," *The Annals of Statistics*, vol. 37, no. 2, pp. 979–1012, 2009.

[35] J. Pearl, *Causality*. Cambridge University Press, 2009.

[36] R. W. Keener, *Theoretical statistics: Topics for a core course*. Springer, 2011.

[37] K. Albert and C. Gernaat, "Pharmacokinetics of ibuprofen," *The American Journal of Medicine*, vol. 77, no. 1, Part 1, pp. 40 – 46, 1984.

[38] M. Davidian and D. Giltinan, *Nonlinear models for repeated measurement data*, vol. 62. CRC press, 1995.

[39] A. O'Malley and A. Zaslavsky, "Domain-level covariance analysis for multilevel survey data with structured nonresponse," *Journal of the American Statistical Association*, vol. 103, no. 484, pp. 1405–1418, 2008.

[40] L. Mazaleuskaya, K. Theken, L. Gong, C. Thorn, G. FitzGerald, R. Altman, and T. Klein, "Pharmgkb summary: ibuprofen pathways," *Pharmacogenetics and Genomics*, vol. 25, no. 2, pp. 96–106, 2015.

[41] A. Mansoor and N. Mahabadi, "Volume of distribution," 2019.

[42] Y. Wang and J. Brasseur, "Enhancement of mass transfer from particles by local shear-rate and correlations with application to drug dissolution," *AIChE Journal*, vol. 65, no. 8, p. e16617, 2019.

[43] M. Wand, D. Ruppert, and C. Crainiceanu, "Bayesian analysis for penalized spline regression using WinBUGS," *Journal of Statistical Software*, vol. 14, 01 2004.

[44] B. Carpenter, A. Gelman, M. Hoffman, D. Lee, B. Goodrich, M. Betancourt, M. Brubaker, J. Guo, P. Li, and A. Riddell, "Stan: A probabilistic programming language," *Journal of Statistical Software*, vol. 76, no. 1, 2017.

[45] K. Imai, L. Keele, and T. Yamamoto, "Identification, inference and sensitivity analysis for causal mediation effects," *Statistical Science*, vol. 25, no. 1, pp. 51–71, 2010.

[46] D. B. Rubin, "Estimating causal effects of treatments in randomized and nonrandomized studies," *Journal of Educational Psychology*, vol. 66, no. 5, p. 688, 1974.