TOWARDS A FRAMEWORK FOR OPERATIONAL RISK IN THE BANKING SECTOR

by

Mphekeleli Hoohlo

A dissertation submitted in partial fulfillment
of the requirements for the degree

of

DOCTOR OF PHILOSOPHY

in

Risk Theory (Finance)

Approved:

Eric Schaling, Ph.D. Supervisor	Thanti Mthanti, Ph.D. Co-supervisor
Odongo Kodongo, Ph.D. Panel Member (WBS)	Thabang Mokoaleli-Mokoteli, Ph.D. Panel Member (WBS)
Christopher Malikane, Ph.D. Panel Member (SEBS)	Paul Alagidede, Ph.D. Panel Member (WBS) and Academic Director of Wits Graduate School of Business

UNIVERSITY OF THE WITWATERSRAND
Johannesburg, Gauteng

2019

ABSTRACT

Towards a Framework for Operational Risk

in the Banking Sector

by

Mphekeleli Hoohlo

University of the Witwatersrand, 2019

Major Professor: Eric Schaling, Ph.D.
Department: Law, Commerce & Management

The purpose of this research is to provide clarity; based on theory and empirical evidence, on what the specific problems in the *operational risk* (OpRisk) literature are, given how its importance has increased significantly over the last decades. There have been a series of destructive events that have threatened the stability of the financial system due to (OpRisk). In most, if not all of these cases, human error is at the center of the chain of events that lead or may lead to (OpRisk) losses. There are many attitudes that can potentially infect organisational processes, the most persistent of these attitudes stem from human failings that are exploitable Barberis and Thaler (2003), thus forming a basis for the theoretical foundation of `OpRisk`.

Shefrin (2016) notes that people would rather incur greater risks to hold on to things they already have, than the risks they would taken to get into that position in the first place, thereby risking a banks' survival, rather than expose their trading losses by consciously deceiving senior management to hide unethical operational practices. In this paper the application of machine learning techniques on the observed data demonstrates how these issues can be resolved given their flexibility to different types of empirical data.

(128 pages)

PUBLIC ABSTRACT

Towards a Framework for Operational Risk
in the Banking Sector
Mphekeleli Hoohlo

The purpose of this research is to provide clarity; based on theory and empirical evidence, on what the specific problems in the *operational risk* (OpRisk) literature are, given how its importance has increased significantly over the last decades. There have been a series of destructive events that have threatened the stability of the financial system due to (OpRisk). In most, if not all of these cases, human error is at the center of the chain of events that lead or may lead to (OpRisk) losses. There are many attitudes that can potentially infect organisational processes, the most persistent of these attitudes stem from human failings that are exploitable Barberis and Thaler (2003), thus forming a basis for the theoretical foundation of `OpRisk`.

Shefrin (2016) notes that people would rather incur greater risks to hold on to things they already have, than the risks they would taken to get into that position in the first place, thereby risking a banks' survival, rather than expose their trading losses by consciously deceiving senior management to hide unethical operational practices. In this paper the application of machine learning techniques on the observed data demonstrates how these issues can be resolved given their flexibility to different types of empirical

DEDICATION

This work—the dissertation and all work associated with it—is dedicated to …. I will always be grateful to her and for her. I also dedicate this work to my child, who patiently loved a father who was often busy working, even when at home. Finally, I dedicate this work to my parents and siblings, who kept me level-headed throughout the process, providing wise and thoughtful advice.

This work is truly evidence of the love I am surrounded by.

Mphekeleli Hoohlo

ACKNOWLEDGEMENTS

Several individuals deserve recognition for their help in this work. First, I am greatly indebted to Dr. Ginger Lockhart for giving me the opportunity to find where my research passions are—especially quantitatively. She provided guidance in my budding career with the right amount of understanding and pressure to push me to be a more productive and capable researcher. I also owe a great deal to Dr. Jamison Fargo, for many, many reasons including helping me get into the program when my application was delayed. He has also proven to be an important mentor and friend. Relatedly, I am grateful to Dr. Karl White in allowing me to become his first student in many years, providing opportunities to develop as a researcher in an important field.

My friends and colleagues in the program also deserve serious recognition for their help and support. Specifically, Dr. Emily Brignone played an invaluable role as a leader and friend as I pursued my goals in the program. She provided timely and important help throughout the process. I appreciate her loyal, honest, genuine friendship. Another good friend that deserved mentioning is Sara Doutré, who also provided support and laughs when I needed it (and sometimes when I did not need it but still welcomed it!). My friends in the Prevention Science Lab continually offered support and collaboration that made my job easier and more enjoyable.

I also want to acknowledge my committee, in providing feedback, guidance, and support throughout my graduate student career. The department's and college's administrative staff and heads, although working in the background, were a huge help in getting me to where I am today. Finally, Utah State University as a school has proven to be a fantastic place to grow academically and to mature into a researcher.

CONTENTS

	Page
Abstract	ii
Public Abstract	iv
Dedication	v
Acknowledgments	vi
LIST OF TABLES	**viii**
LIST OF FIGURES	**ix**
1 INTRODUCTION	**1**
Purpose of the study	1
The theoretical foundation of OpRisk	1
Basel Committee quantitative risk management framework	3
OpRisk frameworks	4
Advanced Measurement Approach (AMA)	4
Standardised Measurement Approach (SMA)	5
Argument	6
Context of the study	7
Why `OpRisk?`	7
Analysis and interpretation issues with behavioral finance theory	8
Theoretical investigations for the quantification of moderm ORMF	9
A new class of ORMF models approach	10
Modeling	10
Problem statement	11
main problem	11
Objectives of the study	12
Exposure-based OpRisk (EBOR)	12
Modeling OpRisk depending on covariates	12
Interpretation Issues using cluster analysis	12
Significance of the study	13
Conclusion	13
2 AVERAGE MARGINAL EFFECTS	**15**
Introduction	15
Additive vs. Multiplicative Interpretations	16
Average Marginal Effect	18
Why Consider Average Marginal Effects?	18
Definition of the Average Marginal Effect	20

Confidence Intervals 21
The Counter-Factual Framework 21
Interpretation 21
Benefits and Limitations 22
Conclusions 23

3 MARGINAL MEDIATION ANALYSIS **25**
Introduction 25
Definition of Marginal Mediation Analysis 25
Interpretation 27
Principle 1: The individual paths are interpreted based on the corresponding endogenous variable's original metric. 28
Principle 2: The indirect effect, as a combination of the a and b paths, are interpreted based on the outcome's original metric. 28
Principle 3: Both the direct and total effects are interpreted based on the outcome's original metric. 28
Effect Sizes 29
Assumptions 30
Reproducibility and Interpretability 30

4 METHODS **32**
Introduction 32
Phase I: Development of Marginal Mediation Analysis 32
Properties of MMA 33
Software Development 33
Phase II: Monte Carlo Simulation Study of Marginal Mediation Analysis . 34
Literature Review 35
Simulations 35
Guideline Development 38
Phase III: Application of Marginal Mediation Analysis 38
Data 38
Substance Use 40
Religiosity 40
Respondent and Peer Attitudes 41
Psychological Well-being 41
Analyses 41
Conclusions 42

5 PHASES I & II: DEVELOPMENT AND TESTING OF MMA **44**
Introduction 44
Developmental Considerations 44
Moderation 44
The Decomposed Total Effect Equals The Total Effect 47
Standardized Effects 48
Software Development 49

Functions 49
Computation of the Marginal Effect 50
Standardization 52
Examples of Software Use 52
Monte Carlo Simulation Study 55
Literature Review 55
Simulations 56
Decomposed Total Effect Equals The Total Effect 56
Statistical Power 59
Estimation Accuracy 60
Confidence Interval Coverage 60
Conclusion 61

6 PHASE III: APPLICATION OF MMA **64**
Introduction 64
Results 65
Conclusions 69

7 DISCUSSION **70**
General Discussion 70
Findings from the Three Phases 71
Limitations 73
Future Research 74
Conclusions 75

APPENDICES **80**
Appendix A: R Code for Chapter 5 81
Examples from Chapter 5 82
Monte Carlo Simulation 83
Monte Carlo Simulation Data Analyses 87
Appendix B: R Code for Chapter 6 93
Data Preparation 94
Models 98
Tables and Figures 104

CURRICULUM VITA **107**

LIST OF TABLES

Table Page

2.1 The generalized linear model link functions with their associated units of interpretation. Note: This list is not exhaustive and there are likely more GLMs that are used within prevention research. . . . 22

4.1 The various experimental conditions of the Monte Carlo simulation study. 37

4.2 The effect sizes of the a and b paths in the Monte Carlo Simulation. . 37

5.1 Functions used in the `MarginalMediation` R package. 51

5.2 Average discrepancies between the decomposed total effect and the total effect for various sample sizes. 57

6.1 Descriptive statistics of the sample. 66

6.2 The percent of mediation (the percent of the total effect) by path for each outcome. 68

LIST OF FIGURES

Figure Page

2.1 a) Comparison of odds ratio and risk ratio with the average marginal risk (probability). b) Same comparison as a) but the y-axis is rescaled (log_{10}) to better show the negative marginal risk comparisons. Both highlight the discrepancy between odds and risk ratios at various levels of marginal risk and that neither approximate the marginal risk. 16

2.2 Demonstration of a non-linear relationship. a) The outcome is an exponential function of the predictor. b) When log transforming the outcome, the relationship becomes fairly linear. So the interpretation in the log transformed scale is additive, but once it is put back into the original units, it is multiplicative. 18

4.1 Path diagram of the replicated models from Ford and Hill (2012). The depression mediator and all of the outcomes are categorical. . . . 39

5.1 An example of moderated mediation, where the moderator (denoted W), moderates the effect of X on M (the a path). 45

5.2 An example of a common approach to visualizing a moderation effect where the effect of X on Y is shown for each level of W. 46

5.3 The structure of the `mma()` function. From left to right: (1) inputs inform the function of the model specifications, the indirect effects to be reported, and other arguments; (2) internal processes including checks/formatting and internal calculations, (3) the output with some truncated example output. Notably, the numbers along the solid arrows pointing at the `mma()` function show the order of the operations, namely checks, AMEs, indirect and direct effects, bootstrapped confidence intervals, and formatting of the output. 50

5.4 An example of the output from the `mma()` function. 54

5.5 The simulated differences between the decomposed total effect and the total effect. The discrepancies are higher for smaller sample sizes and larger effect sizes. 58

5.6 The simulated levels of power per tested sample size (x-axis) and effect size of the indirect path (color; a combination of the a path by the b path) stratified by the distribution of the mediator (binary or count). 59

5.7 The simulated accuracy per tested sample size (x-axis) and effect size of the indirect path (color; a combination of the a path by the b path). 61

5.8 The simulated confidence interval coverage per tested sample size (x-axis) and effect size of the indirect effect . The "ideal" level is set at 0.95. Panel a) shows that the confidence intervals from a broad perspective. Panel b) provides a closer look at the individual patterns of the confidence intervals around the ideal level of 0.95. 62

6.1 Results of the mediation models' individual paths regarding religiosity and substance use. Note: *** $p < .001$, ** $p < .01$, * $p < .05$ 67

CHAPTER 1

INTRODUCTION

Purpose of the study

The aim of this research is to apply a generalised linear model (GLM) that is suitable for exposure-based operational risk (EBOR) treatments within the operational risk management framework (ORMF), which effectively replaces historical loss severity curves obtained from historical loss counts, by forward-looking measures using event frequencies based on actual operational risk (OpRisk) exposures. Preliminary work on EBOR models was undertaken by (Einemann, Fritscher, and Kalkbrener, 2018). Secondly, the study provides comprehensive computational comparisons of various data-intensive techniques, versus classical statistical estimation methods for classification and regression performances between the two.

Cruz (2002) postulated that OpRisk, which focuses on the human side of risk management is difficult to manage with the reduced ability to measure it. The process of OpRisk, that is, the how manifests in conscious and/or unconscious states of the risk manager/s (Hemrit and Arab, 2012), and encompasses approaches and theories that focus on how one will choose when faced with a decision, based on how comfortable they are with the situation and the variables that are present. OpRisk is defined as: "**The risk of loss resulting from inadequate or failed internal processes, people and systems, and from external events. This definition includes legal risk, but excludes strategic and reputational risk.**" (Risk, 2001).

Our understanding of existing ORMF to date is limited to the assumption that financial institutions (FI's) are risk-neutral. In lieu of the afore-mentioned, this study finally seeks to invalidate the risk-neutral assumption, by means of various unsupervised learning techniques, by proposing that FI's are more risk-averse; this can be measured by analysing subtle patterns between data features and trends in the allocated risk capital estimates. In theory, a risk manager who experiences persistent/excessive losses due to particular risk events, would over-compensate cover for these particular risk types, and this would show in reduced losses in these types over time.

The theoretical foundation of OpRisk

Most banks' estimates for their risk are divided into credit risk (50%), market risk (15%) and OpRisk (35%). A major managerial concern for businesses is an inability to identify and account for their susceptibility to OpRisk events following a number of very costly and highly publicized operational losses. OpRisk became popular following a fraudulent trading incident, which was responsible for a catastrophic loss that lead to the collapse of Barings Bank (the UK's oldest bank) in 1995. The term OpRisk began to be used after the above-mentioned and similar types of OpRisk events became more common. In most, if not all of these cases, human error is at the center of the chain of events that lead or may lead to OpRisk losses. Shefrin (2016) notes that people would rather incur greater risks to hold on to things they already have, than the risks they would taken to get into that position in the first place.

A (rogue) trader (Nick Leeson), who risked the banks' survival rather than expose his trading losses, by consciously deceiving senior management to hide his unethical acts. He was found to have been responsible for unethical trading practices when he created illegal trades in his account, then used his position in the front and back offices of the bank to hide his trading losses. Worse still, he incurred a greater risk by lying in order to give a false impression of his profits. It was later discovered that he was placing illegal bets in the Asian-markets, and kept these contracts out of sight from senior management to cover up his illegal activity. When his fraudulent behaviour was discovered (after an earthquake hit at Kobe in Japan, that collapsed the Osaka Securities Exchange) he succumbed to unrecoverable losses due to trading positions he had accumulated, which resulted in a loss of around £1.3 billion to the bank, thus resulting in it's collapse.

Since then, there have been a series of destructive events that have threatened the stability of the financial system due to OpRisk. Large fines have been imposed on the culprits and regulatory scrutiny has been heightened as a result of a number of operational events e.g., the January 2016 "Dark Pool" trading penalties suffered by Barclays (\$70mn) and Credit Suisse (\$85mn), imposed by the United States (US) based securities exchange commision (SEC). These OpRisk loss events were due to fraudulent trading activity consisting of rogue traders dealing in illegally placed high frequency trades for private clients where prices were hidden.

In South Africa (SA), there is an upcoming case of price fixing and market allocation in trading foreign exchange (FX) currency pairs, reffered to the SA based competition tribunal for prosecution. Absa bank, Standard bank & Investec may be liable to payment of an admistrative penalty equal to 10% of their annual turnover in 2016, following accusations by the local based competition commission in February 2017, of rogue traders manipulating the price of the rand through buying and selling US dollars in exchange for the rand at fixed prices. According to the competition commission, it has been alleged that currency traders have been colluding or manipulating the price of the rand through these buy and sell orders to change supply of the currency.

This has compromised the quality and accuracy of risk management's advisory service and pedigree, and aroused huge interest as the value of the rand has implications on South African's. Furthermore, this kind of behaviour can lead to catastrophic operational losses, as with the case for the Barings event, resulting is a mismatch between business' expectations and the value the risk management practice was able to deliver, which is prevalent across FI's and remains unchanged. There are many attitudes that can potentially infect organisational processes, the most persistent of these attitudes stem from human failings that are exploitable (Barberis and Thaler, 2003); i.e. humans' propensity to be deceitful during periods of distress, thus forming a basis for a theoretical foundation of OpRisk management.

Basel Committee quantitative risk management framework

The Bank for International Settlements (BIS) is an organisation consisting of a group of central bank governors and heads of supervision of central banks around the world who represent an authority on good risk management in banking. More specifically, the BIS oversee the duties of the Basel Committee on Banking Supervision (BCBS)/Basel Commitee. The role of the BCBS is to set out guidelines on international financial regulation to cover risks in the banking sector. There have been three banking accords from the BCBS under the supervision of the BIS in dealing with financial regulation, viz., Basel I, Basel II & Basel III. These accords describe an overview of capital requirements for financial institutions (FI's) in order to create a level playing field, by making regulations uniform throughout the world.

The Capital Adequacy Accord (Basel I) was established in 1988. Basel I meant that FI's were required to assign capital for credit risk to protect against credit default. In 1996, an amendment to Basel I imposed additional requirements to cover exposure due to market risk as well as credit risks. Basel I effectively minimised rules that favoured local FI's over potential foreign competitors, by opening up global competition so that these banks could buffer against international solvency. In 2001, the Risk (2001) consultative package provided an overview of the proposed framework for regulatory capital (RC) charge for operational risk (OpRisk). An FI has an OpRisk component, which constitutes a substantial risk component other than credit and market risk. There are two types of OpRisk's viz., potential high severity risk where the probability of an extreme loss is very small but costly, and high frequency/low severity risk where frequency plays a major role in the OpRisk capital charge calculation.

The framework for a New Capital Adequacy Accord (Basel II) was implemented in June 2006. Basel II introduces more restrictive capital charge measures with specific emphasis on OpRisk. Under Basel II OpRisk loss events (i.e. a default in the credit risk jargon) are categorised under seven event types. e.g., Process Risk is one such category and is usually responsible for most OpRisk events, it is the risk that operational losses/problems would take place in the banks transactions.

Regarding the sequence Basel I and Basel II: Regulation begins as a qualitative recommendation which requires banks to have an assets-to-capital multiple of at least 20, then focuses on ratios in which both on-balance sheet and off-balance sheet items are used to calculate the bank's total risk-weighted assets (RWA's)[1], then on tail risk. In other words, auditors' discretion is replaced by market perception of capital, meaning there is a market risk capital charge for all items in the trading business line, then exciting new static risk management approaches which involve calculating a 99.9 percentile left tail confidence interval to measure OpRisk value-at-risk (VaR) and convert it into a RC charge.

Basel III establishes tougher capital standards through more restrictive capital definitions, higher RWA's, additional capital buffers, and higher requirements for minimum capital ratios (Mysis, 2013). Through Basel III, the BCBS is introducing a number of fundamental reforms grouped under three main headings (Committee and others, 2010): 1] A future of more capital through incremental trading book risk (credit items in trading book treated in the same way as if they were in banking book), 2] More liquidity through the introduction of a global liquidity risk standard (Basel III will push banks toward holding greater levels of liquid instruments, such as government bonds and more liquid corporate instruments), and 3] Lower risk under the new requirements of the capital base, i.e., establish more standardized risk-adjusted capital requirements.

The future regulatory environment requires OpRisk professionals, who are not only intelligent, creative and motivated but also have the courage to uphold the OpRisk advisory service standards. Businesses that want to successfuly manage risk, would be well advised to utilize new theoretical and empirical techniques, such that large and small scale experiments play an important role in risk analysis and regulatory research.

OpRisk frameworks

Advanced Measurement Approach (AMA)

Basel II describes three methods of calculating capital charge for OpRisk RC viz., the standardised approach (SA), the basic indicator approach (BIA) and the internal measurement approach (IMA). The basic indicator approach (BIA) sets the OpRisk RC equal to a percentage (15%) of the annual gross income of the firm as a whole to determine the annual capital charge. The SA is similar to the BIA except the firm is split into eight business lines and assigned a different percentage of a three year average gross income per business line, the summation of which is the capital charge Hoohlo (2015). In the IMA, the bank uses it's own internal models to calculate OpRisk loss.

The advanced measurement approach (AMA) is an IMA method which applies estimation techniques of OpRisk capital charge derived from a bank's internal risk measurement system Cruz (2002). Basel II proposed measurement of OpRisk to define capital requirements against unexpected bank losses whereas the unexpected loss (UL) is the quantile for the level α minus the mean. According to the

[1] Also reffered to as risk-weighted amount, it is a measure of the bank's total credit exposure

advanced measurement approach (AMA), thought to outperform the simpler standardised (SA) approach and (BIA), RC requirements are defined according to the UL limit in one year and the loss distribution at a 99.9% confidence level (α = 0.01%) aggegate loss distribution[2] used as a measure of RC. The BCBS proposes to define RC as $ RC = UL $. This involves simulations based on historical data to establish frequency and severity distributions for losses. In this case the RC is a VaR measure.

A FI's ability to determine the framework used for its regulatory OpRisk RC calculation, evolves from how advanced the FI is along the spectrum of available approaches used to determine capital charge (Risk, 2001). BCBS recognizes that a variety of potentially credible approaches to quantify OpRisk are currently being developed by the industry, and that these R&D activities should be incentivised. Increasing levels of sophistication of OpRisk measurement methodologies should generally be rewarded with a reduction in the regulatory OpRisk capital requirement.

The Basel III capital adequacy rules permit model-based calculation methods for capital, including the AMA for OpRisk capital. Under Basel III, standardised methods for OpRisk capital have been overhauled, however for a while there was no prospect of an overhaul of the AMA. Given the relative infancy of the field of OpRisk measurement, banks are mostly free to choose among various AMA principle-based frameworks to a significant degree of flexibility (Risk, 2016). A bank that undertakes an AMA should be able to influence their capital requirements through modeling techniques resulting in lowered pressure on OpRisk capital levels, which in turn has a positive impact on the bank.

Standardised Measurement Approach (SMA)

The flexibility of internal models was expected to narrow over time as more accurate OpRisk measurement was obtained and stable measures of RC were reached, ultimately leading to the emergence of best practice. Instead, internal models produced wildly differing results of OpRisk RC capital from bank to bank, contrary to the expectations of the BCBS. In March 2016, the BCBS published for consultation a standardised measurement approach (SMA) for OpRisk RC; that proposes to abandon the freedom of internal modelling (thus ending the AMA) approaches for OpRisk RC, in exchange for being able to use a simple formula to facilitate comparability across the industry. Under the SMA, RC will be determined using a simple method comprising of two components: A stylised systemic risk model (business indicator component), and an idiosyncratic risk model (loss component), which are combined via an internal loss multiplier (ILM), whose function is to link capital to a FI's operational loss experience to determine SMA capital.

[2] The aggregate loss distribution is obtained by convoluting a loss event frequency distribution and a loss severity distribution by means of the random sums method.

The SMA formula is thought to be consistent with regulators' intent for simplification and increased comparability across most banks. However, there is a feeling from some in the banking industry that the SMA is disadvantaged as it is not the same as measuring OpRisk. Mignola, Ugoccioni, and Cope (2016) and Peters, Shevchenko, Hassani, and Chapelle (2016) identified that the SMA does not respond appropriately to changes in the risk profile of a bank i.e., it is unstable viz., two banks of the same risk profile and size can exibit OpRisk RC differences exceeding 100%, and risk insensitive; that SMA capital results generally appear to be more variable across banks than AMA results, where banks had the option of fitting the loss data to statistical distributions.

Argument

Over the last twenty years, hard-won incremental steps to develop a measure for the size of OpRisk exposure along with the emergence of promising technologies presents a unique opportunity for bankers and treasurers - traditionally risk-averse players - to develop a novel type of way of looking at decision making under risk/uncertainty. New technologies have been introduced which make use of up to date technical solutions (such as homo heuristics developed by Gigerenzer and Brighton (2009), who mainatain their methods solve practical finance problems by simple rules of thumb, or Kahneman (2003)'s intuitive judgements and deliberate decision making), argued to more likely represent the true embedded OpRisk in financial organisations as these methods are designed to fit normal behavioral patterns in their formulation, which is consistent with how decisions are made under risk/uncertainty.

What are the important steps toward completing the post crisis reforms during the current year? Should the risk management fraternity follow the chartered[3] path followed in the Risk (2016) consultative document, scrapping away twenty years of internal measurement approaches (such as the AMA), or should the focus of financial regulators shift toward improving on what they see fit within current existing AMA frameworks. The question is should OpRisk managements' focus be on stimulating active discussions on practical approaches to quantify, model and manage OpRisk for better risk management and improved controls, or abandon the adoption of innovative measurement approaches, such as the AMA, in exchange for being able to use a simple formula across the whole industry?

This proposal begins with an account of significance and a commentary on the nature and scope of the practical problem. The next section of this chapter gives a background of the current issues when dealing with OpRisk measurement and research questions thereof. An overview of the loss distribution approach (LDA), an AMA technique used in the generation of `OpVaR` follows. The proposal concludes by

[3]Meaning as of the publication [@risk2016supporting] the methods brought forth in the consultative document have not been approved for the public, the ideas within an experimental (leased) phase for the exclusive use of BCBS and certain FI's

outlining the research methodology in which I explain the way that I combine the artificial neural network (ANN) management framework with a statistical theory.

Context of the study

Regulatory reforms are designed and fines imposed to protect against operational errors and other conduct costs connected with wrongdoing and employee misconduct. Despite the introduction and use of these seemingly robust strategies, regulations, processes and practices relating to managing risk in FI's, bank losses continue to occur at a rather distressing frequency. A cyclical pattern of OpRisk loss events still persists; as evidenced in the recent price fixing and collusion cases, defeating the explicit objectives of risk management frameworks. This demonstrates a scourge of reflexivity prevailing in financial markets emphasising that, there are theories that seem to work for a time only to outlive their use and become insufficient for the complexities that arise in reality.

Why OpRisk?

A forceful narrative in management theory is that an organisation running effective maintenance procedures combined with optimal team and individual performers i.e., the right balance of skills in the labour force and adequate technological advancements, means systems and services can be used to more efficiently produce material gains, enhance organisational effectiveness, meet business objectives and increase investment activity. Conversely, the risk of the loss of business certainty associated with lowered organisational competitiveness and inadequate systems technology that underpins operations and services is a key source leading to a potential breakdown in investment services activity (Hoohlo, 2015). In fact OpRisk control could set banks apart in competition. This serves as an incentive to support regulation, particularly Basel III recovery and resolution processes.

Consider the case of a regulator in a financial system, who assumes that he/she is consiously and accurately analysing an observed subject, trusting the validity and relying on the visual information that their sense of sight reveals. In the absence of visual confirmation they are hindered from extracting and/or analysing information about the system and their efforts to regulate could potentialy fail. The organisational methods and functioning of current information systems in this industry sector obscure the full extent of OpRisk challenges from the eyes of the risk practitioner.

When an attack such as an operational error occurs at a speed that the OpRisk agent (an individual legal entity or a group) is unable to react quickly enough, due to limitations of their processing speed, and they are not able to process all the information in the given time span, they could lose control/fail to comply with regulatory standards. The latter case is more often than not the most accurate reflection of current risk management practices. The agent represents one end of the spectrum of a risk management strategy, which mitigates risk and enforces regulation, dependent on the information recieved. The other end of the spectrum is one which does not react at all to changes in the system environment.

Current conventional financial systems where information processing is slow and have a tendency to rely on manual, uncertain, unpredictable and unrealistic controls, obscure risk management reporting and produce undesirable market conditions. The OpRisk management function should be able to assist the firms' ability to mitigate risks by acquiring and/or refining risk management solutions which deliver reliable and consistent benefits of improved control and management of the risks inherent in banking operations (Mysis, 2013). This proposal attempts to fill the gap in the current system where there is a risk management information lag or an obstruction from the eyes of the risk practitioner.

Analysis and interpretation issues with behavioral finance theory

Neoclassical theory has grown to become the primary take on modern-day economics formed to solve problems for decision making under uncertainty/risk. Expected Utility Theory (EUT) has dominated the analysis and has been generally accepted as the normative model of rational choice, and widely applied as a descriptive model of economic choice (Kahneman and Tversky, 2013). The traditional finance paradigm seeks to understand financial markets using models in which agents are "rational". According to Barberis and Thaler (2003), this means that agents update their beliefs on the onset of new information, and that given their beliefs, they make choices that are normatively acceptable, and that most people do this most of the time. Behavioral management theory is very much concerned with social factors such as motivation, support and employee relations. A critical component of behavioral finance is building models which better reflect actual behavior. Studies have revealed that these social factors are not easy to incorporate into finance models or to understand in the traditional framework.

A common thread running through the rational viz., the neoclassical take of modern-day economics vs the non-neoclassical schools of thought are findings of behavioral economics which tend to refute the notion that individuals behave rationally. Many argue that individuals are fundamentally irrational because they do not behave rationally giving rise to a literature and debates as to which heuristics and sociological and institutional priors are rational (Altman, 2008). In the real world there is a point of transition between the traditional (neoclassical) approach to decision making, based on data and data anaysis (logic and rational), by adding new parameters and arguments that are outside rational conventional thinking but are also valid. For example, that neoclassical theory makes use of the assumption that all parties will behave rationally overlooks the fact that human nature is vulnerable to other forces, which causes people to make irrational choices.

An essential ingredient of any model trying to understand trading behavior is an assumption about investor preferences (Barberis and Thaler, 2003), or how investors evaluate risky gambles. Investors systematically deviate from rationality when making financial decisions, yet as acknowledged by Kuhnen and Knutson

(2005), the mechanisms responsible for these deviations have not been fully identified. Some errors in judgement suggest distinct mental operations promote different types of financial choices that may lead to investing mistakes. Deviations from the optimal investment strategy of a rational risk neutral agent are viewed as risk-seeking mistakes and risk-aversion mistakes (Kuhnen and Knutson, 2005).

Theoretical investigations for the quantification of moderm ORMF

Kuhnen and Knutson (2005) explain that these risk-seeking choices (such as gambling at a casino) and risk-averse choices (such as buying insurance) may be driven by distinct neural[4] phenomena, which when activated can lead to a shift in risk preferences. Kuhnen and Knutson (2005) found that certain areas of the brain precede risk-seeking mistakes or risky choices and other areas precede risk-aversion mistakes or riskless choices. A risk-aversion mistake is one where a gamble on a prospect of a gain is taken by a risk-averse agent in the face of the chance of a prospective loss. The fear of losing prohibits one's urge to gamble, but people engage in gambling activity anyway. Barberis and Thaler (2003) show that people regularly deviate from the traditional finance paradigm evidenced by the extensive experimental results compiled by cognitive psycologists on how people make decisions given their beliefs.

Kahneman and Tversky (2013) maintains, preferences between prospects which violate rational behaviour demonstrates that outcomes which are obtained with certainty are overweighted relative to uncertain outcomes. This will contribute to a risk-averse preference for a sure gain over a larger gain that is merely probable or a risk-seeking preference for a loss that is merely probable over a smaller loss that it certain. As a psycological principle, overweighting of certainty favours risk-aversion in the domain of gains and risk-seeking in the domain of losses.

The present discussion replicates the common behavioral pattern of risk aversion, where people weigh losses more than equivalent gains. Furthermore, neuroeconomic research shows that this pattern of behavior is directly tied to the brain's greater sensitivity to potential losses than gains (Tom, Fox, Trepel, and Poldrack, 2007). This provides a target for investigating a more comprehensive theory of individual decision-making rather than the rational actor model and thus yield new insights relevant to economic theory[5] (Kuhnen and Knutson, 2005).

If people are reasonably accurate in predicting their choices, the presence of systematic violations of risk neutral behavior provides presumptive evidence against this i.e., people systematically violate EUT when choosing among risky gambles. This seeks to improve and adapt to reality and advance different interpretations of economic behaviour; viz., to propose a more adequately descriptive model, that can represent the basis for an alternative to the way the traditional model is built

[4] As recent evidence from human brain imaging has shown [@kuhnen2005neural] linking neural states to risk-related behaviours [@paulus2003increased].

[5] Representing ability of FI's financial market models to characterise the repeated decision-making process that applies to loss aversion

for decisions taken under uncertainty. This has led some influential commentators to call for an entirely new economic paradigm to displace conventional neoclassical theory with a psycologically more realistic preference specification (List, 2004).

A new class of ORMF models approach

A substantial body of evidence shows that decision makers systematically violate EUT when choosing between risky prospects. Indeed, people would rather satisfy their needs than maximise their utility, contravening the normative model of rational choice (i.e., EUT) which has dominated the analysis of decision making under risk. In recent work (Barberis and Thaler, 2003) in behavioral finance, it has been argued that some of the lessons learnt from violations of EUT are central to understanding a number of financial phenomena. In response to this, there has been several theories put forward advocating for the basis of a slightly different intepretation which describes how individuals actually make decisions under uncertainty/risk. Of all the non-EUT's, we focus on Prospect Theory (PT) as this framework has had most success matching most empirical facts[6].

Kahneman and Tversky (2013) list the key elements of PT, which are 1] a value function, and 2] a nonlinear transformation of the probability scale, that factors in risk aversion of the participants. According to Kahneman and Tversky (2013), the probability scale overweights small probabilities and underweights high probabilities. This feature is known as loss/risk aversion: It suggests that people have a greater sensitivity to losses than to gains (diminishing marginal utility for gains but the opposite for losses), and people are very sensitive about small losses unless accompanied by small gains.

PT looks for common attitudes in people (in FI's) with regard to their behaviour toward taking financial risks or gambles that cannot be captured by EUT. In light of this view, if an FI uses (internal) historical OpRisk loss data to model future events; say a historical case of fraud at the FI had occured and is incorporated in the model, the probability of making the same error in future is provided for in the model, versus risk events that haven't happened. The modelled future should over-provide for the loss events that have already occured, which fits normal patterns around individuals psycological make up and is consistent with risk-averse behavior. The idea at the basis of PT is that a better modeling method can be obtained which leads to a closer approximation of the over-all-value of OpRisk losses.

Modeling

In this study, an important new algorithm for ORMFs and is laid out coupled with data intensive estimation techniques; viz. Generalised Additive Models for locatin Scale & Shape (GAMLSS), Generalized Linear Models (GLDs), Artificial Neural Networks (ANNs), Random Forest (RF) & Decision Trees (DTs), which have

[6]OpRisk loss events in FI's are largely due to human failings that are exploitable e.g., fraudulent trading activity, and PT is based on the same behavioural element of how people make financial decisions about prospects

capabilities to tease out the deep hierarchies in the features of covariates irrespective of the challenges associated with the non-linear or multi-dimensional nature of the underlying problem, at the same time supporting the call from industry for a new class of EBOR models that capture forward-looking aspects. Machine Learning is used as a substitute tool for the traditional model based Autoregressive Moving Average (ARMA) used for analysing and representing stochastic processes. As opposed to the statistical tool, ML does not impose a functional relationship between variables, the functional relationship is determined by extracting the pattern of the training set and by learning from the data observed.

Using `ANN's` in OpRisk measurement and/or mixing historical data with a statistical theory is not a new approach for modeling and calculating OpRisk RC; as evidenced through Agostini, Talamo, and Vecchione (2010) in a study forecasting OpRisk RC via VaR, using advanced credibility theory (CT). The idea at the basis of CT is to advance the very recent literature that a better estimation of the OpRisk RC measurement can be obtained by integrating historical data and scenario analysis i.e., combining the historical simulations with scenario assessments through formulas that are weighted averages of the historical data entries and scenario assessments, advocating for the combined use of both experiences.

However, introducing PT and ANN's is an original way of looking at the approximation issue as opposed to advanced CT. The essential feature of PT are assumptions which are more compatible with basic principles of perception and judgement for decisions taken under uncertainty. According to Kahneman and Tversky (2013), the decision maker, who is a risk agent within the FI, constructs a representation of the losses and outcomes that are relevant to the decision, then assesses the value of each prospect and chooses according to the losses (changes in wealth), not the overall financial state of the FI. We wish to apply a method that tries to establish what accurately ascribes to decision rules that people wish to obey, to make a prediction about what operational loss events might result in the future, then use empirical data to test this idea in a way that is falsifyable.

Problem statement

main problem

The existing models of OpRisk VaR measurement frameworks assume FI's are risk neutral, and do not learn from past losses/mistakes:

* We address weaknesses in current OpRisk VaR measurement frameworks by assuming that FI's are more risk averse.
* Furthermore, we gain an understanding of how past losses affect risk attitudes using machine learning techniques. The calculated future losses are estimated via learning.

Objectives of the study

Exposure-based OpRisk (EBOR)

To quantify OpRisk losses by introducing generalised additive models for location, scale and shape (GAMLSS) in the framework for OpRisk management, that captures exposures to forward-looking aspects of the OpRisk loss prediction problem. EBOR treatments effectively replace historical loss severity curves obtained from historical loss counts, by looking into deep hierarchies in the features of covariates in investment banking (IB), and by forward-looking measures using event frequencies based on actual operational risk (OpRisk) exposures in the business environment and internal control risk factors (BEICF) thereof.

Modeling OpRisk depending on covariates

To investigate the performance of several supervised learning classes of data-intensive methodologies for the improved assessment of OpRisk against current *traditional* statistical estimation techniques. Three different machine learning techniques viz., DTs, RFs, and ANNs, are employed to approximate weights of input features (the risk factors) of the model. A comprehensive list of user defined input variables with associated root causes contribute to the *frequency* of OpRisk events of the underlying value-adding processes. Moreover, the *severity* of OpRisk is also borne out through loss impacts in the dataset . As a consequence of theses new mwthodologies, capital estimates should be able to adapt to changes in the risk profile of the bank, i.e. upon the addition of new products or varying the business mix of the bank providing sufficient incentives for ORMF to mitigate risk (Einemann et al., 2018).

Interpretation Issues using cluster analysis

To identify potential flaws in the mathematical framework for the loss distribution approach (LDA) model of ORM, which is based the derivation of OpRisk losses based on a risk-neutral measure $\mathbb{Q}$, by employing Cluster Analysis (CA). The study addresses weaknesses in the current *traditional* LDA model framework, by assuming managerial risk-taking attitudes are more risk averse. More precisely, CA learns the deep hierarchies of input features[7] that constitute OpRisk event *frequencies* & *severities* of losses during banking operations. In theory, a risk manager who experiences persistent/excessive losses due to particular risk events, would over-compensate cover for these particular risk types. This would show in reduced losses in those loss event types over time, subsequently determining whether risk adverse techniques over-compensate for persistent losses.

[7] A typical approach taken in the literature is to use an unsupervised learning algorithm to train a model of the unlabeled data and then use the results to extract interesting features from the data [@coates2012learning]

Significance of the study

This study fills a gap in that advancing OpRisk VaR measurement methods beyond simplistic and traditional techniques, new data-intensive techniques offer an important tool for ORMFs and at the same time supporting the call from industry for a new class of EBOR models that capture forward-looking aspects of ORM (Embrechts, Mizgier, and Chen, 2018). The current *traditional* approach consists of a loss data collection exercise (LDCE) which suffers from inadequate technologies at times relying on spreadsheets and manual controls to pull numbers together, and therefore do not support the use of data intensive techniques for the management of financial risks. In this study, a new dataset with unique feature characteristics is developed using an automated LDCE, as defined by Committee and others (2011) for internal data. The dataset in question is at the level of individual loss events, it is fundamental as part of the study to know when they happened, and be able to identify the root causes of losses arising from which OpRisk loss events.

This study will provide guidance on combining various supervised learning techniques with extreme value theory (EVT) fitting, which is very much based on the Dynamic EVT-POT model developed by Chavez-Demoulin, Embrechts, and Hofert (2016). This can only happen due to an abundance of larger and better quality datasets and which also benefits the loss distribution approach (LDA) and other areas of OpRisk modeling. In Chavez-Demoulin et al. (2016), they consider dynamic models based on covariates and in particular concentrate on the influence of internal root causes that prove to be useful from the proposed methodology. Moreover, EBOR models are important due to wide applicability beyond capital calculation and the potential to evolve into an important tool for auditing process and early detection of potential losses, culminating in structural and operational changes in the FI, hence releasing human capital to focus on dilemmas that require human judgement.

Conclusion

Mediation analysis is a powerful framework for understanding the processes by which one variable influences another. The assumptions are not much more than that of regression analysis. The interpretation, in linear models, is straightforward and simple. However, once the analysis ventures into non-normal, non-linear relationships, the interpretation becomes more difficult—particularly when it comes to the indirect and total effects.

In the end, Iacobucci (2012) is correct in saying this problem "lacks a strong solution" (pg. 583). Although important information can be obtained from the current methods, mediation analysis with categorical mediator(s) and/or outcome(s) still misses the mark on intuitive, meaningful effect sizes.

This project aims to alleviate these issues by integrating a post-estimation approach known as *Average Marginal Effects* (AMEs) within mediation analysis. This integration can allow simple and meaningful interpretation across variable types and combinations thus far shown to be problematic. The following chapter introduces AMEs, showing their benefit in interpretation and reporting when working with non-normal variables within generalized linear models.

CHAPTER 2

AVERAGE MARGINAL EFFECTS

Any fool can know. The point is to understand. — Albert Einstein

Introduction

When the outcome variable is not continuous and/or has a distribution far from normal, researchers in prevention science often rely on generalized linear models (GLMs). The power of GLMs is clear when you consider the broad range of situations it estimates with asymptotic consistency.[1] However, the problem with GLMs is that the estimates are not in an easily interpretable form. For example, in logistic regression (one type of GLM), the results are in "log-odds". To overcome this lack of interpretability, a simple exponentiation of the coefficient produces what is known as an odds ratio. Similarly, Poisson regression (another form of GLM), with an exponentiation, produces the risk ratio. Although some fields have adopted odds ratios (or relative risk, risk ratios, and other related metrics), these metrics have notable shortcomings.

1. Most of these metrics can be difficult to understand (i.e., many are not intuitive).
2. They cannot be combined with other metrics in a meaningful manner.

First, data have suggested that individuals, although with some variability, are able to intuitively grasp the meaning of phrases such as "highly probable" or "not likely" (Heuer Jr., 1999). Yet, this same intuition is not found in odds or odds ratios. For example, Montreuil, Bendavid, and Brophy (2005) found, of 84 articles in several epidemiology journals that used odds ratios, only 7 (8.3%) accurately

[1] Asymptotic consistency refers to the ability to, as the sample size increases, produce estimates that converge to the proper value.

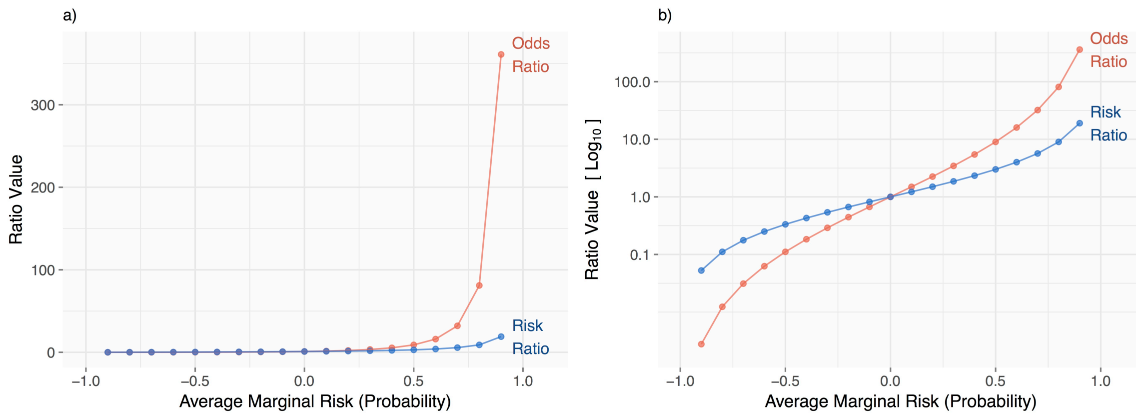

Figure 2.1: a) Comparison of odds ratio and risk ratio with the average marginal risk (probability). b) Same comparison as a) but the y-axis is rescaled (log_{10}) to better show the negative marginal risk comparisons. Both highlight the discrepancy between odds and risk ratios at various levels of marginal risk and that neither approximate the marginal risk.

interpreted the odds ratio and 22 (26%) interpreted odds as though they were probabilities ("risk"). Figure 2.1 highlights the large discrepancy between the odds ratio, the risk ratio, and the average marginal risk. Ultimately, reporting odds and interpreting them as risk is common.

The second is particularly important in the case of mediation analysis given the importance of the indirect effect (the combination of the a and b paths). If a is not in a unit that can be combined with b, then obtaining a meaningful indirect effect is not generally possible. Given its commonality and the limitation that it cannot be combined with other metrics, it is time to consider other strategies—at least in some situations.

Additive vs. Multiplicative Interpretations

The distinction between additive and multiplicative estimates is generally important but is particularly so in mediation analysis. In most quantitative research designs, the investigators are seeking information on the average effect in a population, whether this refers to an average difference across groups or an average change in the outcome for a given change in the predictor. Generally speaking, the average effect is referring to the marginal effect (i.e., the effect of a small change in the predictor in the outcome's units). In the linear regression framework, the average effect is the estimated coefficient and is interpreted *additively*—a one unit change in the predictor is associated with an X unit change in the outcome. Conversely, outcomes such as OR are *multiplicative*. Being multiplicative changes the interpretation to: a one unit change in the predictor is associated with an X times change in the outcome. Although subtle, the difference is important, especially for multi-part models (e.g., mediation analysis).

Being multiplicative indicates that the effect of the predictor changes based on the level of the predictor. For example, if the predictor is high, a small change in the predictor may have a big effect while if the predictor is low, a small change in the predictor has little effect. Figure 2.2 shows this phenomenon, where, in the outcomes original units there is an exponential function. In part a) of the figure, it is clear that a change from 2 to 3 in the predictor has a much larger effect than a change from 0 to 1. A regression would not work well here. If a log transformation is used, the relationship would be linear (and a regression can be used) but the interpretation becomes multiplicative (in this case a one unit increase in the predictor is associated with an $X * 100$ percent increase in the outcome).

Additive interpretations are generally the most intuitive and require less cognitive resources to understand the pattern being portrayed. In a multiplicative framework, simplicity in understanding the effect intuitively is somewhat lost (Iacobucci, 2008). Among others, this is one reason why Stata provides the `margins` command when dealing with two-part hurdle models.[2] These models break up the modeling into two parts: one for the binary part and one for the count part. In order to combine the two parts, Stata allows a transformation known as the *average*

[2]These models are often used for zero-inflated count outcomes.

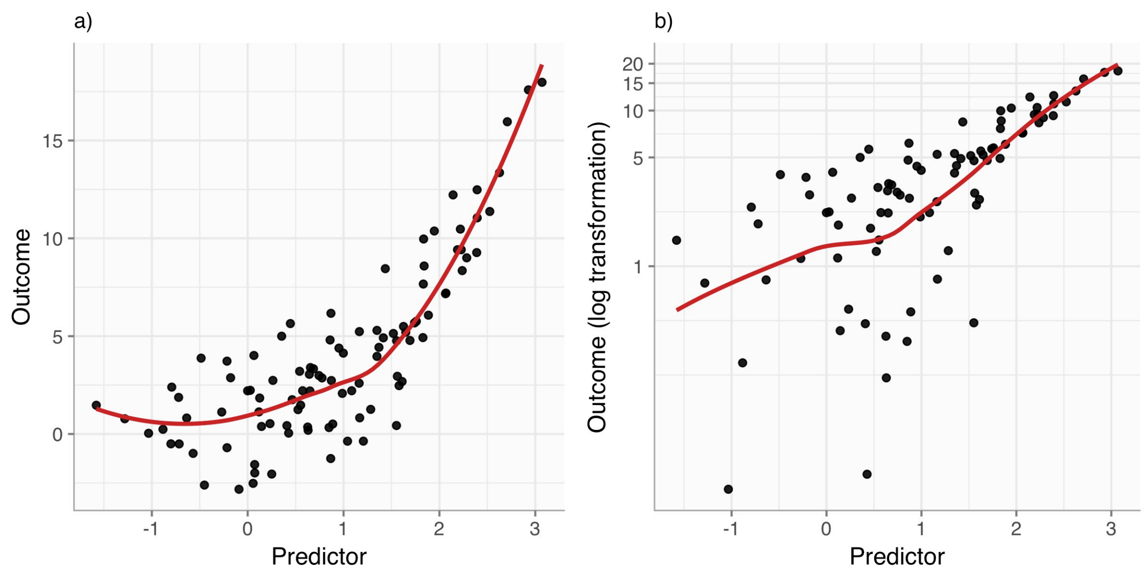

Figure 2.2: Demonstration of a non-linear relationship. a) The outcome is an exponential function of the predictor. b) When log transforming the outcome, the relationship becomes fairly linear. So the interpretation in the log transformed scale is additive, but once it is put back into the original units, it is multiplicative.

marginal effect (AME) that makes the two parts both additive, and therefore easily combined (as is also desired in mediation analysis).

Average Marginal Effect

Why Consider Average Marginal Effects?

When using GLMs, the model is fit with a link function (e.g., "logit", "probit", "log"). This change causes the marginal effect of a variable to rely on the values of the covariates in the model. This is well illustrated through an example. Say a logistic regression model was fit to the data, as shown in Equations 2.1 - 2.4, for p predictors.

$$logit(Y_i) = \beta_0 + \sum_{j=1}^{p} \beta_j X_{ij} + \epsilon_i \tag{2.1}$$

$$log(\frac{Prob(Y_i = 1)}{1 - Prob(Y_i = 1)}) = \beta_0 + \sum_{j=1}^{p} \beta_j X_{ij} + \epsilon_i \quad (2.2)$$

$$\frac{Prob(Y_i = 1)}{1 - Prob(Y_i = 1)} = e^{\beta_0 + \sum_{j=1}^{p} \beta_j X_{ij} + \epsilon_i} \quad (2.3)$$

$$Prob(Y_i = 1) = \frac{e^{\beta_0 + \sum_{j=1}^{p} \beta_j X_{ij} + \epsilon_i}}{1 + e^{\beta_0 + \sum_{j=1}^{p} \beta_j X_i + \epsilon_i}} \quad (2.4)$$

This implies that the marginal effect of, say, X_{i1} is:

$$\frac{\delta Y}{\delta X_1} = \frac{e^{\beta_0 + \sum_{j=1}^{p} \beta_j X_{ij} + \epsilon_i}}{(1 + e^{\beta_0 + \sum_{j=1}^{p} \beta_j X_{ij} + \epsilon_i})^2} \quad (2.5)$$

That is, the marginal effect of the predictor X_1 *depends* on the level of each covariate for each individual (all X_{ij}'s) and each estimate (all β_j's). This, understandably, complicates the interpretation.

Importantly, this example makes it clear that each observational unit (e.g., individual) has a unique marginal effect given her observed levels of each variable. For example, if variable X is increased by one from wherever it was observed, each individual will have different effects (i.e., different marginal effects). One individual may have a large effect; another small. Each individual has a marginal effect of a given predictor associated with her set of characteristics (covariates). To understand the average effect in the population of interest, the mean of all such marginal effects can be calculated.

The AME, then, is the *averaged* marginal effect across all observational units in the data. In linear models, the AME is the same as the original model estimates. This is intuitive given the AME is the marginal effect in the outcome's original units—the exact interpretation of the estimates in a linear model. However, as stated previously, in GLMs the estimates are not in the original units and therefore must be estimated via a post-estimation calculation described below.

Definition of the Average Marginal Effect

In an instructive paper about the (now defunct) routine called “margeff” in Stata, Bartus (2005) highlights how the AME can be calculated—including the mathematical definition—and the benefits of AMEs compared to other related methods. The AME is a post-estimation calculation—it uses the model estimates and the data to provide the average effect. Bartus (2005) provided the definition of this post-estimation procedure of a continuous predictor as:

$$AME_k = \beta_k \frac{1}{n} \sum_{i=1}^{n} f(\beta X) \tag{2.6}$$

where f refers to the probability density function of F, βX is the linear combination of the predictors (i.e., the model predicted values for each observation), and AME_k is the average marginal effect for the kth variable. This definition provides the average change in the outcome for a one unit change in the continuous variable x_k across all n observations.

Relatedly, the AME of a dummy coded variable is:

$$AME_k = \frac{1}{n} \sum_{i=1}^{n} \Big[F(\beta X | x_{ki} = 1) - F(\beta X | x_{ki} = 0) \Big] \tag{2.7}$$

where $F(\beta X|x_{ki} = 1)$ is the predicted value of the ith observation when the dummy variable x_k equals one and $F(\beta X|x_{ki} = 0)$ is the predicted value when the dummy value of x_k equals zero holding all other variables constant. This, in effect, shows the discrete difference between the levels of the categorical variable in the outcome's original units.

Confidence Intervals

In general, two approaches are taken to estimate the confidence intervals of AMEs. The first approach is the delta method, which provides standard errors (StataCorp, 2015). Although beneficial, the second—bootstrapped confidence intervals—have proven accurate for understanding the variability in both the AME and mediation analysis. Therefore, this project uses bootstrapped confidence intervals.

The Counter-Factual Framework

As mentioned previously, this project—including the use of average marginal effects—fits within the counter-factual framework. Indeed, the definition of a dummy coded variable demonstrates this well—$\Big[F(\beta X|x_i = 1) - F(\beta X|x_i = 0)\Big]$. In essence, this answers the question: "What is the difference in the outcome when all observations of x_k are equal to one vs. when all observations of x_k are equal to zero?" The counter-factual framework strives to answer the same class of questions. When all assumptions are met, the AME is a direct, statistical answer to the causality conditions proposed by this framework.

Interpretation

Table 2.1 presents the various units that the AME will produce for the various GLM links. It is important to note that AMEs are in the outcome's original metrics whether they are probabilities, counts, or something else. The interpretation, then, is "for a one unit change in the predictor there is an associated [AME] change in the outcome."

Table 2.1: The generalized linear model link functions with their associated units of interpretation. Note: This list is not exhaustive and there are likely more GLMs that are used within prevention research.

Link Function	Average Marginal Effect
Identity	Original Continuous Unit
Log	Original Continuous Unit
Logit	Risk
Probit	Risk
Poisson	Count
Gamma	Count
Negative Binomial	Count

Benefits and Limitations

There are several benefits to using average marginal effects with GLMs.

- *Intuitive Interpretation and Few Assumptions.* The first, and most obvious, benefit to using AMEs is the simplicity of the interpretation. The effect is in the units used in the modeling; it is additive (i.e., the effect is the added increase or subtracted decrease in the outcome). It provides an interpretation that imitates that of ordinary least squares regression. Relatedly, there are no difficult modeling assumptions directly tied to AMEs. Instead, the underlying models' assumptions that are used to get the AME is what is important. The only additional assumption with AME is that the effect, for each individual, is linear enough to be represented by an additive value and that the average adequately reports this.
- *More Generalizable and Robust.* There is evidence suggesting that AMEs are more robust to problems associated with GLMs (including logistic regression) such as unobserved heterogeneity and model mis-specification (Mood, 2010; Norton, 2012). This allows the estimates to be more generalizable to individuals outside of the sample.
- *Low Computational Burden.* Given AMEs are a post-estimation calculation,

no new models need to be fit. Instead, using the estimates of the models, the average marginal effects can be calculated. The most computationally expensive aspect of the calculation is the bootstrapped confidence intervals.

- *Broadly Applicable.* The AME applies to any of the generalized linear models including logistic, Poisson, gamma, beta, negative binomial, and two-part hurdle models. This provides extensive flexibility in modeling, and, once applied to mediation, will allow flexibility in modeling based on the correct functional form.
- *Two-Part Models.* Particularly pertinent to this project is that the calculation of AMEs have been applied to two-part models, generally of the hurdle model types, as stated earlier. In fact, this is a common routine in the Stata statistical software. This provides valuable support for the proposed approach of using AMEs in mediation analysis.

Notably, the interpretation of AMEs hold to the assumption (as found in all regressions) that it is reasonable to adjust a single covariate while holding all others constant. This may not hold in reality, although it may be necessary to gain an understanding of the individual effect of a single variable. In data that are not representative of the population (e.g., non-random sample), AMEs may be biased because an over-representation of certain covariate values may be present. This is an overall modeling problem, since GLMs also assume a random sample. In this way, this problem is not specific to AMEs.

Conclusions

The Average Marginal Effect produces intuitive, interpretable, and additive estimates of an effect. They have been applied to two-part models, similar to mediation analysis, demonstrating their utility in difficult modeling situations. The following chapter discusses the integration of AME with mediation analysis—termed

Marginal Mediation Analysis— with its interpretation and assumptions, its benefits and limitations, and the basic procedures for its use.

CHAPTER 3

MARGINAL MEDIATION ANALYSIS

Without an interpretable scale, it is difficult to use effect size to communicate results in a meaningful and useful way. — Preacher and Hayes, 2011

Introduction

The proposed integration of average marginal effects and mediation analysis is designed to resolve two major obstacles currently found in mediation analysis:

1. The difficulty of performing mediation analysis with categorical mediators and/or outcomes, and
2. The lack of reliable and flexible effect size estimates in mediation analysis—particularly with categorical mediators and/or outcomes.

These issues are relatively common in prevention work (e.g., Ford and Hill, 2012; B. Hoeppner, Hoeppner, and Abroms, 2017; Wong and Brower, 2013) and the current approaches are not adequate—as was discussed at length in the previous chapters. In this chapter, the integration of Average Marginal Effects (AMEs) and mediation analysis—*Marginal Mediation Analysis* (MMA)—is discussed, including its interpretation and assumptions as well as its benefits and limitations. It is expected that this adjustment to both the modeling and the interpretation will help researchers in the health and prevention sciences to be able to model their data in the most properly-specified way and be able to communicate their findings clearly.

Definition of Marginal Mediation Analysis

The form of the general marginal mediation model, including the post-estimation step, are demonstrated in the following equations, where Equations 3.1

and 3.2 demonstrate the mediation estimation while Equations 3.3 and 3.7 show the post-estimation procedures.

$$M_{ij} = a_0 + \sum_{k=1}^{p} a_k x_{ki} + \epsilon_i \quad \text{for j = 1, ..., m mediators} \tag{3.1}$$

$$Y_i = \beta_0 + \sum_{j=1}^{m} b_j M_{ij} + \sum_{k=1}^{p} c_k^{'} x_{ki} + \epsilon_i \tag{3.2}$$

for the ith individual, for $k = 1, ..., p$ predictors, and $j = 1, ..., m$ mediators. The paths are all labeled with their common term (e.g., path a is labeled a). Combining these two equations provides the full mediation model. Using these models, we apply the post-estimation of the average marginal effects as presented by Bartus (2005). For a continuous x_k variable, the average marginal effect of path a is:

$$AME_k^a = a_k \frac{1}{n} \sum_{i=1}^{n} f(aX) \tag{3.3}$$

where f refers to the probability density function, aX is the linear combination of the predictors (i.e., the model predicted values for each observation), and AME_k^a is the average marginal effect of the a path for the kth variable. Ultimately, Equation 3.3 is identical to that of the following:

$$AME_k^a = \frac{1}{n} \sum_{i=1}^{n} \frac{f(aX_1) - f(aX_2)}{2h} \tag{3.4}$$

where

$$aX_1 = \begin{bmatrix} ax_{11} & ax_{12} & \dots & ax_{1k} + h & \dots & ax_{1p} \\ ax_{21} & ax_{22} & \dots & ax_{2k} + h & \dots & ax_{2p} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ ax_{n1} & ax_{n2} & \dots & ax_{nk} + h & \dots & ax_{np} \end{bmatrix} \tag{3.5}$$

and

$$aX_2 = \begin{bmatrix} ax_{11} & ax_{12} & \dots & ax_{1k} - h & \dots & ax_{1p} \\ ax_{21} & ax_{22} & \dots & ax_{2k} - h & \dots & ax_{2p} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ ax_{n1} & ax_{n2} & \dots & ax_{nk} - h & \dots & ax_{np} \end{bmatrix} \tag{3.6}$$

Both $f(aX_1)$ and $f(aX_2)$ are the model predicted value for the outcome given the small change due to h. Equations 3.4, 3.5, and 3.6 use a very small h value (default is 1×10^{-7}). This provides the change in the average predicted value for a very small increase and a very small decrease in in the x_k variable. This is also described in depth by Leeper (2017) since it is the strategy employed by `margins`. This approach is flexible—especially in the R statistical environment with the improvement of derivative computation that Leeper provided.

Similarly, the AME of a dummy coded variable in the a path is:

$$AME_k^a = \frac{1}{n} \sum_{i=1}^{n} \Big[F(aX|x_{ki} = 1) - F(aX|x_{ki} = 0) \Big] \tag{3.7}$$

where $F(aX|x_{ki} = 1)$ is the predicted value of the ith observation when the dummy variable equals one and $F(aX|x_{ki} = 0)$ is the predicted value when the dummy value equals zero. This is the same approach used by `margins`.

Notably, these same post-estimation equations (3.3 and 3.7) can be used for the b and c' paths as well.

Interpretation

The interpretation of Marginal Mediation is based on the original units of the mediator(s) and outcome(s). Because there are so many possible combinations of GLM types within mediation analysis, instead of outlining every combination, the basic principles are presented that apply to all situations.

Principle 1: The individual paths are interpreted based on the corresponding endogenous variable's original metric.

The individual paths have interpretations identical to those of AMEs, as discussed in Chapter 2. Therefore, the a path depends on the type of mediator and modeling approach chosen. For example, for a binary mediator, the a path is a probability (risk); a mediator representing a count has an a path that is in the same count units.

Principle 2: The indirect effect, as a combination of the a and b paths, are interpreted based on the outcome's original metric.

The indirect effect is joining two paths, possibly of different metrics. However, the interpretation is still straightforward: the entire effect will be in the outcome's original metric. An example may be beneficial to highlight this principle.

Suppose there are data with a hypothesized binary mediator (depression or no depression) and a continuous outcome (quality of life). A logistic regression is used to model path a and a linear regression is used for the b and c' paths. After calculating the AME of the paths, path a is in units of risk (of depression) while path b is the difference in the quality of life between depressed and not depressed individuals. By combining paths a and b through multiplication, we get the effect of the predictor on depression risk then what that depression risk does to quality of life; that is, the effect of the predictor on quality of life through depression.

Principle 3: Both the direct and total effects are interpreted based on the outcome's original metric.

Similar to Principle 2, the direct and total effects are in the outcome's original units. This is intuitive given that the AME of the direct effect is in the outcome's units, and it is not combined with any other path. For the total effect, the indirect and direct effect (which are both in the same units) are added together to get the complete effect. It was expected that, as in linear regression (but that is lacking in other situations), $a \times b + c' = c$. That is, the indirect and direct effects together equals the total effect. This expectation was tested, as is described in the next chapter.

Effect Sizes

A major advantage of this framework is the way effect sizes can be used intuitively.

> "First, virtually all effect size indices should be scaled appropriately, given the measurement and the question of interest. Without an *interpretable scale*, it is difficult to use effect size to communicate results in a meaningful and useful way…. Second, it should be emphasized that effect size estimates are themselves sample statistics and thus will almost certainly differ from their corresponding population values. Therefore, it is important to report confidence intervals for effect sizes…" (Preacher and Kelley, 2011, pg. 95, emphasis added).

The interpretation of MMA makes it clear that both of these aspects of proper effect size estimation and reporting are adequately represented. Of particular note, unlike many effect sizes that are only useful in certain research questions, the effect sizes produced by AMEs—and thus found in Marginal Mediation Analysis—are flexibly oriented to "be scaled appropriately" to best "communicate results in a meaningful and useful way" for a wide variety of situations.

Additionally, Preacher and Kelley (2011) continue: "it is important to develop a way to gauge the effect size of the product term *ab* itself," (pg. 95). That is, not only does the effect size of the individual paths need to be meaningful but the product of $a \times b$ must be as well. Marginal Mediation Analysis provides this comparability as the indirect effect can be compared without problems with the direct and total effects (they are all in the same units).

Assumptions

The assumptions inherent in MMA are the same as those presented in Chapter 1 regarding mediation analysis. The only additional assumption regards the ability of the effect to be represented additively (i.e., can the effect be represented linearly after accounting for the marginal effect for each observation?). In linear models, this is already included as an implicit assumption. For other models, although the relationship may not be linear in the outcome, taking the average of the effects across the observations is assumed to be representative of the relationship across the sample.

Reproducibility and Interpretability

As mentioned previously, categorical mediators/outcomes are generally not difficult to model using GLMs—only the interpretation is difficult. Generalized linear models, in conjunction with AMEs, allow researchers to use more correct functional forms, thereby reducing the justification to fit poorly specified models that have an easier interpretation.

With this framework, MMA can be applied across the GLM spectrum and essentially any combination of GLM types. For example, a marginal mediation model is defined when the mediator is binary and the outcome is continuous; when the mediator is a count and the outcome is ordinal; when the mediator is continuous and the outcome is binary. Each has a straightforward, yet informative, interpretation as outlined by the principles above. This attribute alone can increase the reproducibility of research using mediation analysis.

Finally, the interpretation across the paths and effects is straightforward and flexible. Other researchers, laypersons, lawmakers, and clinicians can assess the direction, magnitude, meaning, and utility of findings much easier—thus, increasing the reach and impact of research. In mediation analysis, this can prove largely beneficial given the already complex nature of the modeling scheme. By simplifying the interpretation, less cognitive resources are required to gain a basic understanding of the findings; instead, more resources can be used to understand how to apply it and assess future research questions based on the findings.

CHAPTER 4

METHODS

> *For precept must be upon precept, precept upon precept; line upon line, line upon line; here a little, and there a little.* — Isaiah 28:10, King James Bible

Introduction

As presented in the previous chapter, Marginal Mediation Analysis (MMA) has the ability to simplify the interpretation of mediated effects in a wide variety of situations, particularly in situations where an effect size otherwise does not exist (e.g., indirect effects when the mediator or outcome is categorical). In this chapter, methods fashioned to develop MMA and evaluate its performance are discussed via three phases:

1. Development of MMA
2. Monte Carlo Simulation Study of MMA
3. Application of MMA

These phases were designed to provide the theory and the software to perform MMA, assess the method's ability to accurately estimate the underlying effects, develop the guidelines of its use in finite samples, and apply it to real-world prevention data by replicating a recent study (Ford and Hill, 2012) that used a categorical mediator and categorical outcomes. Below, each phase is described in depth.

Phase I: Development of Marginal Mediation Analysis

To be useful to public health, psychological, and prevention researchers, the incorporation of average marginal effects within mediation analysis must happen in two ways: in theory and in software. This phase is focused on understanding the properties of MMA and on developing the software necessary to perform it.

Properties of MMA

Building on the mediation framework discussed by Hayes (2009) and by Edwards and Lambert (2007), MMA was established on linear regression—either ordinary least squares (OLS) for continuous outcomes/mediators or maximum likelihood (via GLMs) for categorical outcomes/mediators. In this framework, two or more regression equations are combined to provide the overall mediation model as discussed in Chapter 1. This method then adds a post-estimation step (Chapter 2) into this mediation framework.

The form of the general marginal mediation model, including the post-estimation step, were discussed in Chapter 3. Using this general framework, various considerations were made in the development of the method. First, an appropriate manner in which to integrate moderation (interaction effects) into the framework is important. Because of the work by Edwards and Lambert (2007), this included assessing the *reduced form*[1] of the models in addition to visualizations of the predicted values across levels of the moderator. Second, it has been noted by MacKinnon (2008) that in non-linear models the $a \times b + c'$ generally does not equal the c path as it does in linear models (Chapter 11 in his book). To assess whether these are equal within MMA, both a basic analysis and a Monte Carlo simulation (phase II) was used.

Software Development

A major aspect of this first phase is the development of the software for researchers to apply MMA. This software is provided via the R statistical environment given R is free, widely used by researchers in health and prevention, and ex-

[1]Reduced form refers to only having exogenous variables on the right-hand side of a regression equation (i.e., substituting the predictors of the mediators into the equation).

tensions to the software via "packages" are efficiently disseminated through the Comprehensive R Archive Network (CRAN). It consists of a number of functions to fit the model and assess the model's fit while efficiently producing the paths and effects in proper units.

Because of the flexibility of numerical derivation methods and the speed improvements by Thomas Leeper (2017), numerical derivatives were used to obtain the marginal effects for each observational unit. From here, means of the marginal effects were calculated for each variable in the model. To assess the model uncertainty, bootstrapping via the `boot` R package was applied. This approach relies on repeated resampling from the original sample *with replacement*. In general, this method does not rely on any distributional assumptions and works well with asymmetric distributions (as is found in indirect effects).

The package applies the best practices for both computational speed and user readability (Wickham, 2015), allowing other researchers to extend the package more easily. Additionally, several built-in tests will inform the functionality of the package before beginning Phase 2. These tests were performed on Linux, Mac, and Windows platforms. Finally, the package uses Git as the version control system. The necessary functions were developed first so that the package tests and simulations could begin. The usability of the package that is important in the disseminated version were developed afterwards.

Phase II: Monte Carlo Simulation Study of Marginal Mediation Analysis

The evaluation of MMA is an essential step in understanding its properties and robustness and further assess the performance of the software. The consistency of MMA, the statistical power at various sample sizes, and the accuracy of the bootstrapped confidence intervals were all tested via a Monte Carlo simulation study (Carsey and Harden, 2013; Paxton, Curran, Bollen, Kirby, and Chen, 2001).

In the simulation, data were simulated to come from a population of known parameters. A literature review of mediation analysis in prevention work highlighted the appropriateness of the population parameters chosen. The results of the simulation helped in the development of the guidelines for using MMA in practice.

Literature Review

Before performing the Monte Carlo simulations, a review of the literature is recommended (Paxton et al., 2001). This review focused on the use of mediation analysis in prevention research where the analyses contained categorical mediators and/or outcomes. This review included all recent articles (2008 - 2017) found that clearly identified a mediator or outcome that was categorical in nature. This search relied on terms such as “generalized linear models”, “logistic”, “dichotomous”, “polytomous”, and “count” in conjunction with “mediation analysis” across the Scopus database.

Simulations

Monte Carlo simulations, via the R statistical environment version 3.4.2, assessed the finite properties of MMA. Monte Carlo simulation was selected due to its simplicity in generating informative results and its high success in the literature (e.g., Graham, Olchowski, and Gilreath, 2007; Nylund, Asparouhov, and Muthen, 2007). Here, 500 data sets were simulated for each combination of experimental conditions (Carsey and Harden, 2013; Paxton et al., 2001). The data were simulated from a known population with a researcher specified causal model (i.e., the “population model”). The model consisted of either a binary mediator (0 = “No”, 1 = “Yes”) or a count variable (Poisson distribution), a continuous outcome, and a continuous predictor while also varying the sample size and the effect sizes for a total of 90 unique combinations of the conditions (see Table 4.1).

The a path population model is defined below for when the mediator is binary, where the $Prob(M_u = 1)$ is a latent continuous variable with a logistic relationship with the predictors and the ϵ_i is normally distributed with a mean of 0 and a standard deviation of 1.

$$log(\frac{Prob(M_u = 1)_i}{1 - Prob(M_u = 1)_i}) = a_0 + a_1 x + \epsilon_i \tag{4.1}$$

The observed variable, M_o, is defined as follows:

$$M_o = \begin{cases} 0, & \text{if } Prob(M_u = 1) < .5, \\ 1, & \text{otherwise.} \end{cases} \tag{4.2}$$

The a path population model for when the mediator is a count is shown below where M_u is a latent continuous variable with an exponential relationship with the predictors and the ϵ_i is normally distributed with a mean of 0 and a standard deviation of 1.

$$log(M_u) = a_0 + a_1 x + \epsilon_i \tag{4.3}$$

The observed variable, M_o, is defined as follows:

$$M_o = Po(\lambda = M_u) \tag{4.4}$$

This creates an observed, count variable that has λ values based on the latent mediator.

The b and c' paths population model is identical to Equation 3.2 with only one predictor and a single binary, or count, mediator, as shown below.

$$M_i = a_0 + a_1 x_1 + \epsilon_{mi} \tag{4.5}$$

Table 4.1: The various experimental conditions of the Monte Carlo simulation study.

Independent Variables	Conditions
Sample size	50, 100, 200, 500, 1000
Effect size of a path	small, moderate, large
Effect size of b path	small, moderate, large
Effect size of c' path	moderate
Type of mediator	binary, count
Total conditions	90

Table 4.2: The effect sizes of the a and b paths in the Monte Carlo Simulation.

Size	Odds Ratio (a path)	Risk Ratio (a path)	r (b path)
Small	1.58	1.34	0.10
Moderate	3.44	1.82	0.30
Large	6.73	3.01	0.50

$$Y_i = \beta_0 + bM_i + c^{'}x_i + \epsilon_{yi} \tag{4.6}$$

Table 4.1 highlights the conditions that were varied for each simulation. A distinct MMA model was applied to each of the 500 data sets for each possible combination of experimental conditions. This means 45,000 MMA models were fit. Using eight cores of powerful core i7 computers, these computations were finished over a span of several days.

Notably, the effect sizes for both the binary mediator and count mediator (a path) were the odds ratios and risk ratios corresponding to small, moderate, and large effect sizes. These are found in Table 4.2.

The focus of the simulations was to gauge the accuracy, power, and coverage of MMA at estimating the population effects while undergoing the experimental conditions. The dependent variables were:

1. bias (i.e., is the mean of the estimates at the population mean?),
2. power (i.e., how often does the null properly get rejected?),

3. confidence interval coverage (i.e., does the confidence interval cover the proper interval?), and
4. consistency regarding how closely $a \times b + c'$ is to c (i.e., does the indirect plus the direct effect equal the total effect?).

The effects of the conditions on these outcomes were assessed via visualizations and descriptive tables.

Guideline Development

Recommendations from the simulation study were documented, including necessary sample sizes, bias in various conditions, and the accuracy of bootstrapped confidence intervals in each condition. The documentation will be available in manual form online on the R website and GitHub.

Phase III: Application of Marginal Mediation Analysis

During the third phase, all important aspects of MMA discovered throughout the first two phases were used to replicate previous work regarding the relationship of adolescent religiosity with substance use (Figure 4.1). This study was selected given it used:

1. a large sample with a mix of binary and continuous mediators and outcomes,
2. one of the better and common statistical approaches, and
3. data that were publicly available and that had a more recent release to investigate.

Data

To replicate this study, data from the 2014 (most recent) release of National Survey on Drug Use and Health (NSDUH; Ford and Hill, 2012) were used. As described by Ford and Hill (2012), the NSDUH is "an ongoing study sponsored by the

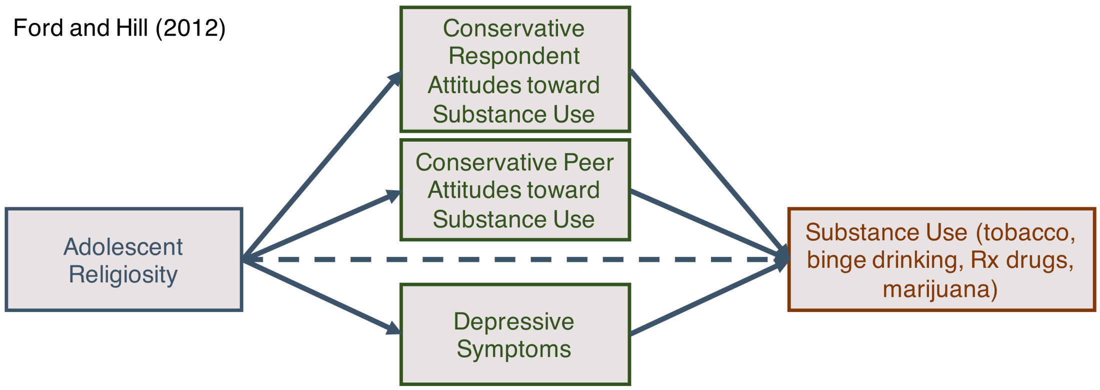

Figure 4.1: Path diagram of the replicated models from Ford and Hill (2012). The depression mediator and all of the outcomes are categorical.

U.S. Substance Abuse and Mental Health Services Administration that dates back to the 1970s" (page 4) and collects data on drug use of individuals 12 years and older across the United States. Ford's study used the 2007 data release while the replication uses the 2014. Several measures[2] were used to replicate the findings of Ford and Hill (2012):

1. Four substance use outcomes (tobacco use, prescription drug use, marijuana use, and illicit drug use).
2. Religiosity was based on the average response across four items relating to church attendance, the importance of religious beliefs to the individual, and participation in faith-based activities. Higher scores indicated more religiosity.
3. Respondent attitudes toward substance use was also the average response based on four items gauging the individual's response to someone their age using substances. Higher scores indicate more conservative attitudes.
4. Peer attitudes toward substance use is similar to the respondent attitudes except that it was asked how the individual's friends would feel about someone their age using substances. Again, the average response was used where

[2]Notably, this replication study omits the "heavy drinking" outcome because positive responses to it were very rare in the 2014 data.

higher scores indicate more conservative attitudes.

5. "Psychological well-being was indicated by major depression," (page 5) which was measured as at least five of the nine possible depression symptoms listed in the survey.

Substance Use

As stated previously, the substance use outcomes are tobacco use, prescription drug use, marijuana use, and illicit drug use.

- Tobacco use was defined as one of the following three items: 1) cigarette use within the last year, 2) smokeless tobacco use within the last year, and 3) cigar use within the last year.
- Prescription drug use consisted of four groups of drugs that are being used either without a prescription or for the sole purpose of obtaining a high within the last year: pain relievers, tranquilizers, stimulants, and sedatives.
- Marijuana use was a single item: marijuana use within the last year.
- Illicit drug use was defined as using any of the following drugs within the last year: cocaine, crack, heroin, hallucinogens, LSD, PCP, ecstasy, inhalants, or meth.

Each outcome was coded as dichotomous: use or no use within the last year.

Religiosity

Adolescent religiosity was the mean response across four items: 1) the number of times attended religious activities in past year, 2) religious beliefs are important, 3) religious beliefs influence decisions, and 4) the amount of participation in religious activities. The higher the average score the more the adolescent is considered to be religious.

Respondent and Peer Attitudes

The respondent's conservative views on drug use is the average of four items, answering the question "How do you feel about someone your age using [cigarettes daily, marijuana, marijuana monthly, drinking daily]?" Similarly, the peer's conservative views on drug use is the average of four items, answering the question "How do you think your close friends would feel about you using [cigarettes daily, marijuana, marijuana monthly, drinking daily]?"

Psychological Well-being

Finally, psychological well-being was defined as having had a major depressive episode in the past year. This was a binary (yes or no) variable based on "if they reported experiencing at least five of the following: felt sad, empty, or depressed most of the day or discouraged; lost interest or pleasure in most things; experienced changes in appetite or weight; sleep problems; other noticed you were restless or lethargic; felt tired or low energy nearly every day; felt worthless nearly every day; inability to concentrate or make decisions; any thoughts or plans of suicide," (Ford and Hill, 2012, pg. 5).

Analyses

The mediation analyses were replicated from Ford and Hill (2012). Although the mediation analysis is performed differently herein, the model specifications were identical to that employed there.

Importantly, Ford and Hill (2012) say: "we use the categorical data method outlined by MacKinnon (2008) to formally test the indirect effects," (pg. 5). This approach uses a significance test based on the estimates of both a and b and their standard errors. However, as stated throughout this project, the significance alone is insufficient information to provide for a mediation analysis; effect sizes are also necessary. Because of this, Ford and Hill (2012) continue by discussing the amount of the association between the predictor and outcome, in percentage units, that the mediator accounted for. This approach is useful but has some notable shortcomings. First, depending on the level of multi-collinearity in the models, the standard errors of the estimates can be inefficient which reduces the statistical power of this test. Second, it does not provide the effect size measures that would be most useful (e.g., the effect a one unit increase in the predictor has on the outcome through the mediator). Third, the measure is consistently too conservative with binary outcomes (Jiang and Vanderweele, 2015).

For the replication, then, each of the mediation models reported were run using MMA in place of the techniques employed by Ford and Hill (2012). Four distinct MMA models, one for each of the substance use outcomes, were assessed. These were all controlling for (adjusting for) parental attitudes towards substance use, age, race, sex, and income. The models included 500 bootstrapped samples to obtain 95% confidence intervals.

Further, using a variant of the "difference method," the amount of the total effect that was mediated was calculated using the following:

$$\text{Proportion mediated} = \frac{indirect}{indirect + direct}$$

Finally, the information provided through the use of MMA was also compared to that produced in the original paper.[3]

Conclusions

Ultimately, the goal of this project is to develop, evaluate and apply a method that can provide meaningful interpretation in mediation when the mediator and/or

[3] All analysis code is available on the Open Science Framework (osf.io/753kc) and in the Appendices of this document.

outcome is categorical. Each phase builds on this goal, as is discussed in the following chapters starting with the presentation of the results of Phase I and Phase II regarding the theory, software, and evaluation of MMA.

CHAPTER 5

PHASES I & II: DEVELOPMENT AND TESTING OF MMA

Exploring the unknown requires tolerating uncertainty. — Brian Greene

Introduction

The results of both Phase I (the development of the method and its software) and Phase II (the Monte Carlo simulations) regarding Marginal Mediation Analysis (MMA) are presented in this chapter.

Developmental Considerations

The general MMA framework was discussed in Chapter 3. This framework was extended with some important additional considerations, including integrating moderation and analytically assessing the relation between the decomposed total effect and the total effect.

Moderation

Moderation (interaction) is sometimes hypothesized to occur in conjunction with mediation. Moderation is any situation where the effect of a variable on another *depends* on the value of a third variable. This phenomenon, in conjunction with mediation, is often referred to as conditional process analysis, moderated mediation, or mediated moderation—depending on the source and situation.

An example of one of the many possible moderated mediation models is found in Figure 5.1 (for more examples, see Hayes, 2013). In this example, the moderator (denoted W in the figure), moderates that relationship between X and M. In other words, the effect of X on M depends on the value of W. This further suggests that

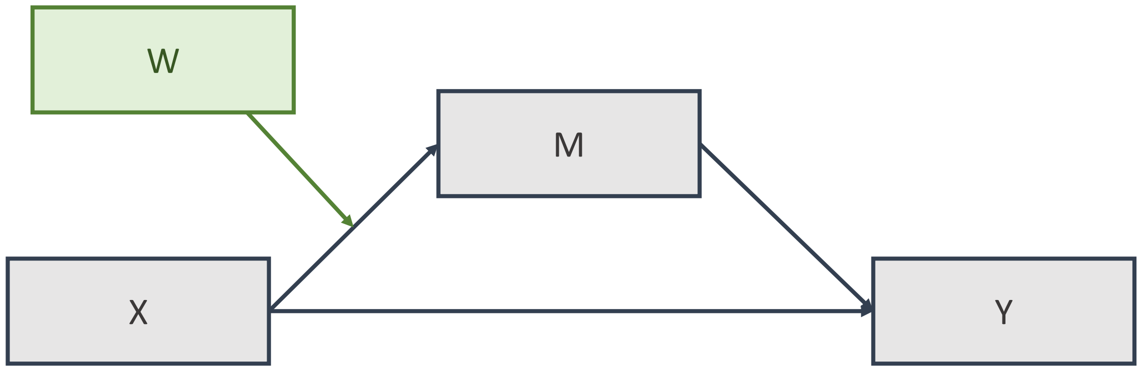

Figure 5.1: An example of moderated mediation, where the moderator (denoted W), moderates the effect of X on M (the a path).

the effect of X on Y, through M, also depends on the value of W.

In general, interactions make interpretation more difficult. In linear models, the interpretation of the interaction estimate becomes: "a one unit increase in X is associated with a $\beta_x + \beta_{int} \times W$ effect on the outcome." That is, to understand the size, and direction, of the effect of X, the level of W must be considered. For example, if W is a categorical variable with values of 0 and 1, and the following regression was estimated:

$$\hat{M}_i = 1.0 + 5.0X + 3.0W + .5X * W \tag{5.1}$$

then:

1. X is associated with a 5.0 increase in M when W is 0, and
2. X is associated with a 5.5 increase in M when W is 1.

The same logic holds for continuous moderators, although representative values of W must be chosen instead of using all possible values.

Yet, in non-linear situations, this becomes more strenuous. However, using average marginal effects, the interpretation can be like that of linear models. In general, this has been done by selecting various, representative values of W at which the average marginal effect is assessed (StataCorp, 2015). If W is categorical then

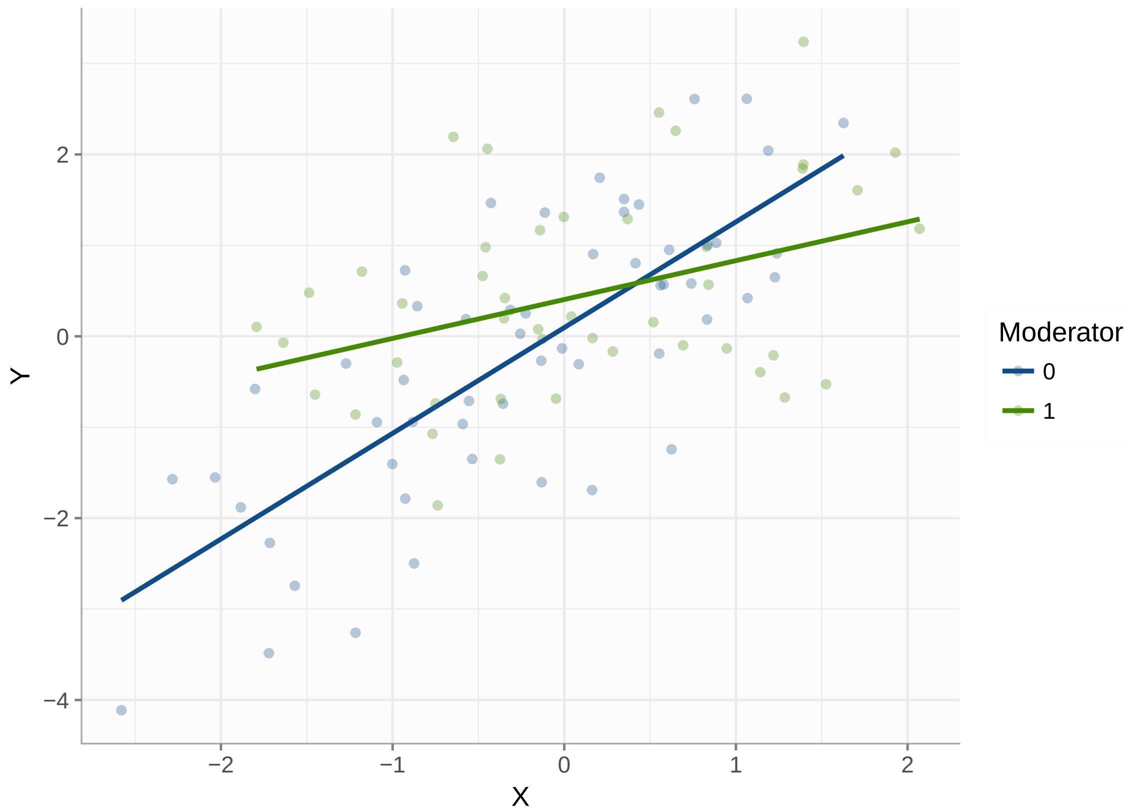

Figure 5.2: An example of a common approach to visualizing a moderation effect where the effect of X on Y is shown for each level of W.

all observed values can be used. Notably, in linear and generalized linear models, moderation is probably best understood using visualizations—showing the effect of X at various levels of W (see Figure 5.2).

In relation to MMA, moderation can be understood at both 1) an individual path level and 2) a complete model level. At an individual path level, moderation is understood as it is in non-mediating regression situations. For example, if path a is moderated, the effect of X on M can be understood via visualizations or representative values of W can be inserted into the regression equation, as in the example above.

To understand it in relation to the complete model, the framework discussed by Edwards and Lambert (2007) suggests using the *reduced form* of the mediation model to understand the moderation in the context of the indirect and direct effects. The reduced form refers to having only exogenous variables on the right-hand side of the equation (i.e., substituting the estimates of the a path into the b/c path model) as shown below.

Starting with the non-reduced form, we have M_i, an endogenous variable, on the right hand side.

$$Y_i = \beta_0 + b_1 M_i + c_1^{`} x_i + \epsilon_{yi} \tag{5.2}$$

Using the a path model, and assuming the same model specification as in Figure 5.1, we can substitute in the predictors of M_i.

$$Y_i = \beta_0 + b_1(a_0 + a_1 x_i + a_2 w_i + a_3 x_i * w_i) + c_1^{'} x_i + \epsilon_{yi} \tag{5.3}$$

This form is now reduced so that only exogenous variables are on the right-hand side. Using these estimates, it is now possible to assess the effect of x on Y when it depends on the level of w using the same approach with the individual paths. Importantly, using these estimates, the moderated effect of x can be visualized as well.

The Decomposed Total Effect Equals The Total Effect

When using the average marginal effect, the decomposed total effect $(a \times b + c')$ equals that of the original total effect (c). Winship and Mare (1983), demonstrated that, using calculus, an outcome variable Y can be decomposed by its total differential

$$dY = \frac{\delta Y}{\delta X} dX + \frac{\delta Y}{\delta M} dM \tag{5.4}$$

which implies the general formula

$$\frac{dY}{dX} = \frac{\delta Y}{\delta X} + \frac{\delta Y}{\delta M}\frac{dM}{dX} \tag{5.5}$$

(pg. 83, the symbols were altered to match that of the present project). That is, the total effect is equal to the direct plus indirect effects. If the average marginal effect is a good estimate of the derivative (or partial derivative), then:

$$\frac{dY}{dX} = \frac{\delta Y}{\delta X} + \frac{\delta Y}{\delta M}\frac{dM}{dX} = c' + b \times a = c \tag{5.6}$$

Therefore, it is expected that regardless of the distributions of the mediators or outcomes $a \times b + c' = c$.

This is further demonstrated with finite sampling properties in the Monte Carlo simulation in Phase II using both binary and count mediators.

Standardized Effects

It is often of considerable worth to understand standardized effects. These can be defined in numerous ways, depending on the situation and types of variables that are being used. In situations where the outcome is continuous, MMA can use a partial standardization approach discussed by Preacher and Hayes (2011) where the outcome is standardized using its own standard deviation. This produces interpretations that are based on the change in the outcome in standard deviation units. If using a dichotomous predictor, this essentially becomes a standardized mean difference (e.g., Cohen's D). It is also possible to standardize both the continuous outcomes and continuous predictors to obtain a partial correlation metric from these models as well.

As discussed below, the software to perform MMA includes the outcome standardization for continuous outcomes but does not provide standardization techniques for the predictors. Future iterations of the software will include this as well.[1]

Software Development

The software package developed for MMA is called `MarginalMediation` and is freely available via the R statistical environment. The software allows straightforward use of MMA across continuous, binary, and count mediators and/or outcomes (other distributions also work but have not been extensively tested). The computation is done in several steps:

1. Function and model checks
2. a path average marginal effect estimates
3. b and c' path average marginal effect estimates
4. Bootstrapped confidence intervals
5. Formatting and printing of the output

This strategy was undertaken to help in error-checking and allows the function to print informative output to the user during the modeling, which is especially useful for situations with large samples and many bootstrapped samples.

Functions

Using the package is based on a single function—`mma()`—that provides the main functionality (Figure 5.3). `mma()` is built on several other functions that perform specific duties that allow the simple syntax. The main functions of the package are shown in Table 5.1.

[1]The researcher can standardize the continuous predictors and outcomes before performing MMA, in which case they can obtain the partial correlation estimates with the current version of the software.

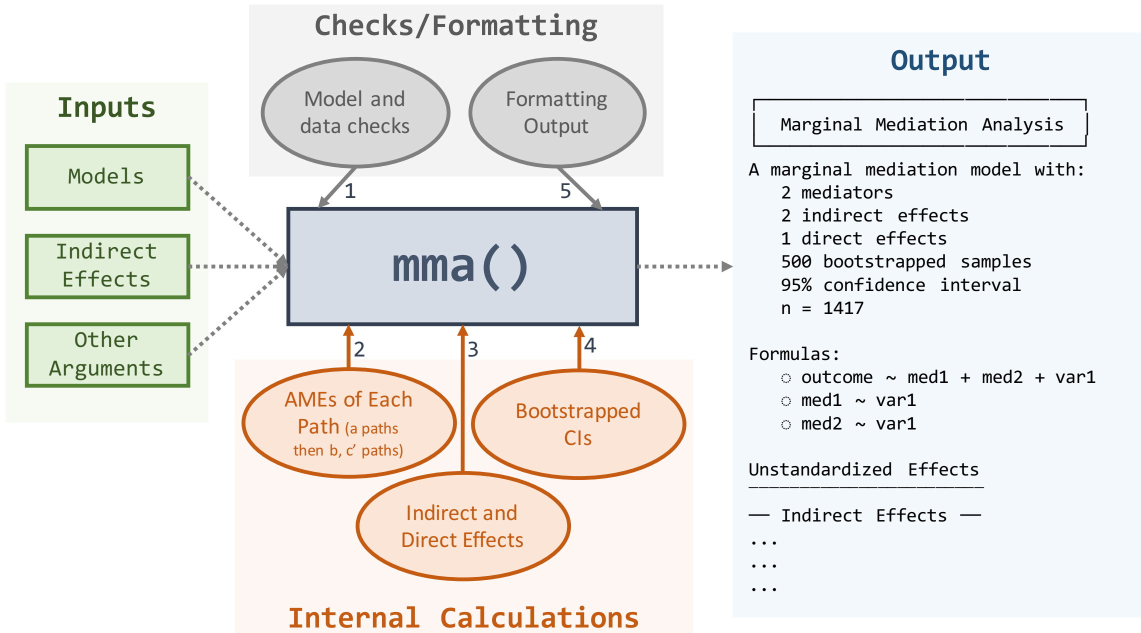

Figure 5.3: The structure of the `mma()` function. From left to right: (1) inputs inform the function of the model specifications, the indirect effects to be reported, and other arguments; (2) internal processes including checks/formatting and internal calculations, (3) the output with some truncated example output. Notably, the numbers along the solid arrows pointing at the `mma()` function show the order of the operations, namely checks, AMEs, indirect and direct effects, bootstrapped confidence intervals, and formatting of the output.

Computation of the Marginal Effect

`MarginalMediation` uses built-in R functionality that allows for relatively fast computation of the marginal effects. The approach taken here is identical to that of the `margins` R package (Leeper, 2017), as described in Chapter 3. This is repeated here. Specifically, for continuous predictors, the numerical derivative is used as shown below where a is the general symbol for the model estimates.

$$AME_k = \frac{1}{n} \sum_{i=1}^{n} \frac{f(aX_1) - f(aX_2)}{2h} \tag{5.7}$$

Table 5.1: Functions used in the `MarginalMediation` R package.

Type	Function	Behavior
Main Function	`mma()`	Performs the full MMA model
Marginal Function	`amed()`	Computes the average marginal effects of a given GLM model
Moderated Mediation	`mod_med()`	Computes the marginal effects at various levels of a moderator (still in testing)
Checks and formatting	These functions perform behind the scenes	Check model specification and function requirements
Other Functions	`mma_std_ind_effects()` and `mma_std_dir_effects()`	Obtain the standardized indirect and direct effects from the model
Other Functions	`mma_ind_effects()` and `mma_dir_effects()`	Obtain the unstandardized indirect and direct effects from the model
Other Functions	`perc_med()`	Obtain the percent of mediation for each specified path in the model

where

$$aX_1 = \begin{bmatrix} ax_{11} & ax_{12} & \dots & ax_{1k}+h & \dots & ax_{1p} \\ ax_{21} & ax_{22} & \dots & ax_{2k}+h & \dots & ax_{2p} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ ax_{n1} & ax_{n2} & \dots & ax_{nk}+h & \dots & ax_{np} \end{bmatrix} \tag{5.8}$$

and

$$aX_2 = \begin{bmatrix} ax_{11} & ax_{12} & \dots & ax_{1k}-h & \dots & ax_{1p} \\ ax_{21} & ax_{22} & \dots & ax_{2k}-h & \dots & ax_{2p} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ ax_{n1} & ax_{n2} & \dots & ax_{nk}-h & \dots & ax_{np} \end{bmatrix} \tag{5.9}$$

With a small h (default is 1×10^{-7}), this produces the average marginal effect across all the observations (e.g., the average change in the predicted value for a very small increase and a very small decrease in in the x_k variable).

For discrete predictors, the discrete difference is used as shown below,

$$AME_k = \frac{1}{n}\sum_{i=1}^{n}\Big[F(\beta X|x_{ki}=1) - F(\beta X|x_{ki}=0)\Big] \tag{5.10}$$

where $F(\beta X|x_{ki} = 1)$ is the predicted value of the ith observation when the dummy variable x_k equals one and $F(\beta X|x_{ki} = 0)$ is the predicted value when the dummy value of x_k equals zero holding all other variables constant. This, in effect, shows the discrete difference between the levels of the categorical variable in the outcome's original units.

These approaches are employed in `MarginalMediation` due to their flexibility across GLM types and model specifications. For example, it can handle many types of models (e.g., linear, GLM, multilevel) and can produce more interpretable estimates of the marginal effects of predictors that have quadratic terms (e.g., age and age^2).

Standardization

As was briefly noted earlier, partial standardization wherein the outcome is standardized is possible when the outcome is continuous. In these situations, the output of `MarginalMediation` will include both unstandardized and standardized effects (see Figure 5.4).

Examples of Software Use

To briefly demonstrate the use of `mma()`, fictitious data were first generated, where X, M, and Y are continuous. Using these data (called `df1`), the following `R` code demonstrates the use of `mma()` in the simplest case.

```r
library(MarginalMediation)

pathbc = glm(Y ~ X + M, data = df1)
patha  = glm(M ~ X, data = df1)

fit = mma(pathbc,
          patha,
          ind_effects = c("X-M"))
```

First, the individual sub-models are fit thereby creating `pathbc` and `patha` which are both `glm` objects. Then, the b and c paths (`pathbc`) model object is the first argument to `mma()`, followed by the a paths (in this case only a single a path but multiple—separated by commas—can be included). The necessary argument is the `ind_effects`. This argument expects a vector or list of quoted paths, where the paths are the form `"predictor-mediator"`. In this case, the predictor is called `X` and the mediator is called `M`.

The `fit` object as created by `mma()` contains a number of elements, including the indirect effects, the direct effects, the confidence interval, and the original data. Figure 5.4 provides an example of how the output could look if the `fit` object is printed. This output provides both unstandardized effects (both indirect and direct) that are in the units of the outcome and standardized effects—using the standard deviation of the outcome as recommended by MacKinnon (2008)—which are in the standard deviation units of the outcome.

Further, it can be assessed whether, in this case, the indirect plus the direct effects equal the total effect. Here, the total effect is 1.921 which is equal to the indirect effect (0.885) plus the direct effect (1.036). This suggests that comparisons between the effects can be confidently made.

If a covariate, $X2$ is added to the data, this can be easily added to the model as shown below.

```
┌──────────────────────────────┐
|  Marginal Mediation Analysis  |
└──────────────────────────────┘
A marginal mediation model with:
   1 mediators
   1 indirect effects
   1 direct effects
   100 bootstrapped samples
   95% confidence interval
   n = 100

Formulas:
   ○ Y ~ X + M
   ○ M ~ X

Unstandardized Effects
──────────────────────

── Indirect Effects ──
      A-path  B-path Indirect   Lower   Upper
X-M 0.94825 0.93282  0.88454 0.64805 1.16212

── Direct Effects ──
  Direct  Lower   Upper
X 1.0358 0.8144 1.24912


Standardized Effects
────────────────────

── Indirect Effects ──
    Indirect   Lower   Upper
X-M  0.36097 0.26446 0.47425

── Direct Effects ──
  Direct   Lower   Upper
X 0.4227 0.33235 0.50976
-----
```

Figure 5.4: An example of the output from the `mma()` function.

```r
library(MarginalMediation)
pathbc = glm(Y ~ X + X2 + M, data = df1)
patha  = glm(M ~ X + X2, data = df1)

fit2 = mma(pathbc,
           patha,
           ind_effects = c("X-M",
                           "X2-M"))
```

It is also possible to access various aspects of these MMA model fit objects.

```r
perc_med(fit2, "X-M")
```

This informs the researcher that the indirect effect accounts for approximately 55%

of the total effect from X to Y in `fit2`.

Monte Carlo Simulation Study

With the software package `MarginalMediation`, the simulations were able to assess the package's functionality and the overall framework's ability to estimate the underlying effects accurately. First, to assess the appropriateness of the experimental conditions, a literature review was conducted.

Literature Review

Studies were sought that saliently reported results wherein both mediation analysis and generalized linear models were used. Since 2012, this produced 57 articles (via Scopus).[2] Among these, three general categories of articles were found:

1. Articles that were methodologically building on mediation analysis.
2. Articles that applied mediation where a mediator and/or outcome was categorical and the authors used the "difference method" (MacKinnon, 2008) to assess the amount of mediation.
3. Articles that applied mediation where a mediator and/or outcome was categorical and the authors either used the structural equation modeling approach or did not include the categorical mediator and/or outcome in the mediation.

Of these, number two was most prevalent. The literature suggested that the parameters selected for the simulations were relevant, particularly the small effect sizes and large sample sizes. Most studies used extant, large questionnaire data sets and the majority were cross-sectional.

Importantly, this search demonstrated the commonality of the "difference method" as discussed by MacKinnon (2008). This method relies on the following:

[2] The search terms included: "mediation analysis" and ["logistic" or "generalized linear models" or "GLM" or "poisson"].

$$a \times b + c' = c \tag{5.11}$$

$$a \times b = c - c' \tag{5.12}$$

In essence, this says that it is possible to estimate the indirect effect, that is $a \times b$ by assessing the difference $c - c'$. However, in situations where the decomposed total effect does not equal the total effect, this method may not be valid although many studies still used this approach in categorical data situations.

Simulations

The Monte Carlo simulation produced 45,000 marginal mediation models (although including the bootstrapped intervals there were 22.5 million models run). These simulated models were run on powerful Core i7 computers over the span of several days. The following subsections discuss the results of the simulations in regard to each outcome of interest.

Decomposed Total Effect Equals The Total Effect

One of the major questions about the performance of MMA regards whether the decomposed total effect equals the total effect ($a \times b + c' = c$). Table 5.2 highlights the average discrepancy between the decomposed total effect and the total effect divided by the total effect (thereby adjusting the discrepancy for the size of the total effect). Clearly, on an average level, deviations are extremely small, generally $< .5\%$ discrepancy, with the majority $< .1\%$ discrepancy. The discrepancies also decrease in size as the sample size increases.

Figure 5.5 presents the individual simulated differences between the decomposed total effect and the total effect. Once assessing the individual discrepancies, two patterns are of note:

1. There are larger discrepancies for smaller sample sizes and larger effect sizes.

Table 5.2: Average discrepancies between the decomposed total effect and the total effect for various sample sizes.

	Mediator	
Sample Size	Binary	Count
50	0.0017	-0.0080
100	0.0024	-0.0028
200	0.0007	0.0027
500	0.0001	-0.0036
1000	-0.0000	-0.0027

2. Besides a single outlier in the count condition (Panel b), most discrepancies are small.

First, the largest discrepancies are where the sample sizes are small (n = 50) and the effect sizes are larger. This is intuitive in that as the effect size is larger, the amount of discrepancy that is still considered small also increases (i.e., variability simply due to the estimates being on a larger scale). For both binary and count mediators, the discrepancies, even in the large effect sizes, are very small as the sample increases to n = 1000. Given the literature review, this is a sample size that is often possible in the health and prevention sciences. Further, most effect sizes in the literature were moderate or smaller. These conditions had low variability in the discrepancy.

Second, in the count condition there is a clear outlying value (>2 for the n = 50 and large/large effect size condition). Other than this value—across the binary and count mediators—all other values are relatively close to zero. For the binary mediator condition, the scale was in risk (probability) units. The discrepancies, here, in the n = 50 condition are notable in their size while the other conditions had discrepancies that are essentially within rounding error. For the count mediator condition, the scale of the total effect was in count units. The outlier is notable in its large discrepancy in these units; most other values were essentially within rounding error of the effect size.

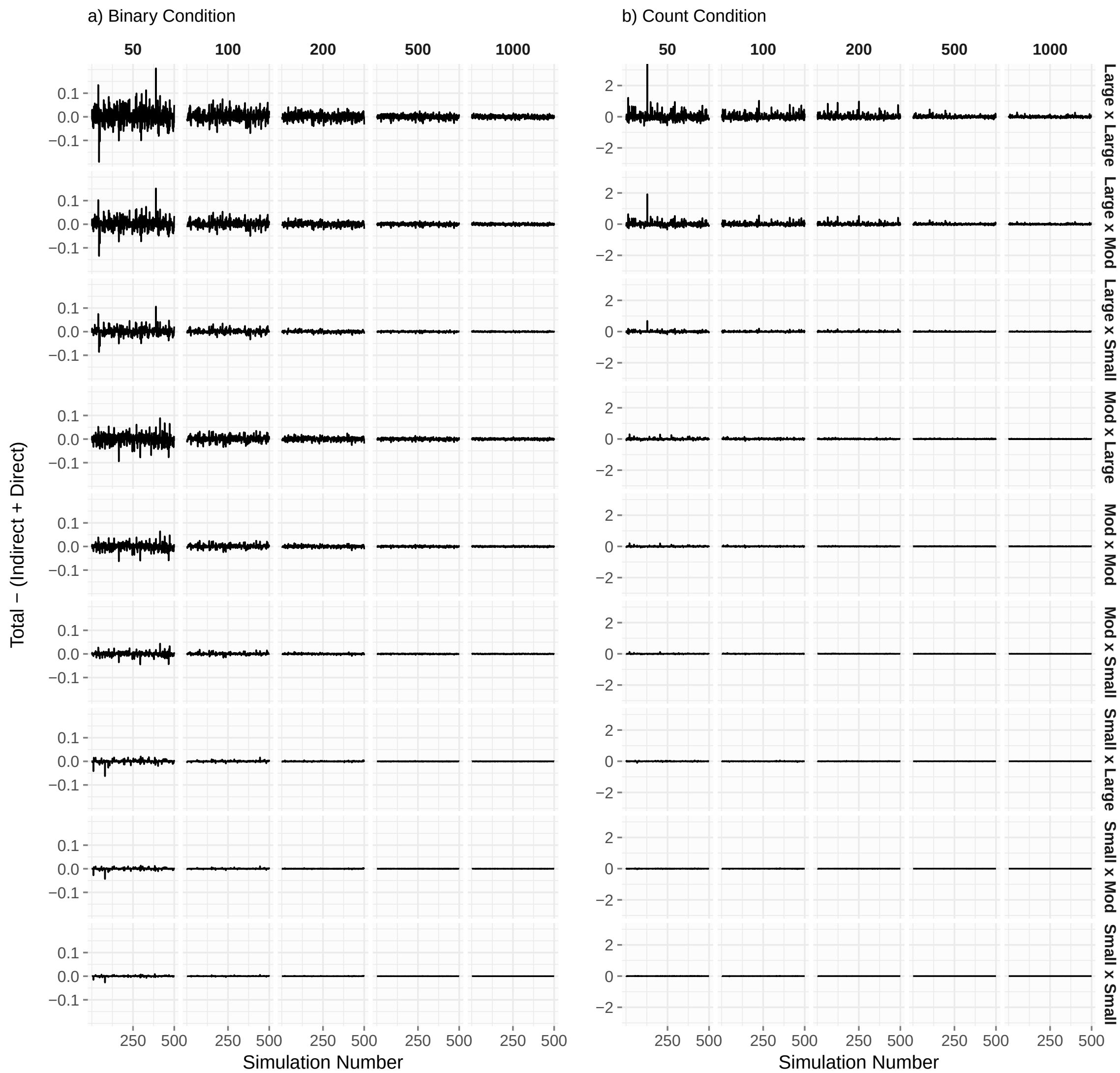

Figure 5.5: The simulated differences between the decomposed total effect and the total effect. The discrepancies are higher for smaller sample sizes and larger effect sizes.

Ultimately, this provides evidence of MMAs ability to estimate values that let the $a \times b + c' = c$ condition to hold, even in individual applications. This, however, is somewhat dependent on the sample size. As for differences across the effect sizes of the a and b paths, as an effect size increases so does the level of “rounding error.” That is, in large effects, larger discrepancies are still a minor deviation than if the effect was small. Therefore, the main aspect of this finding is that sample size is important in the accuracy of the indirect plus direct equaling the total effect.

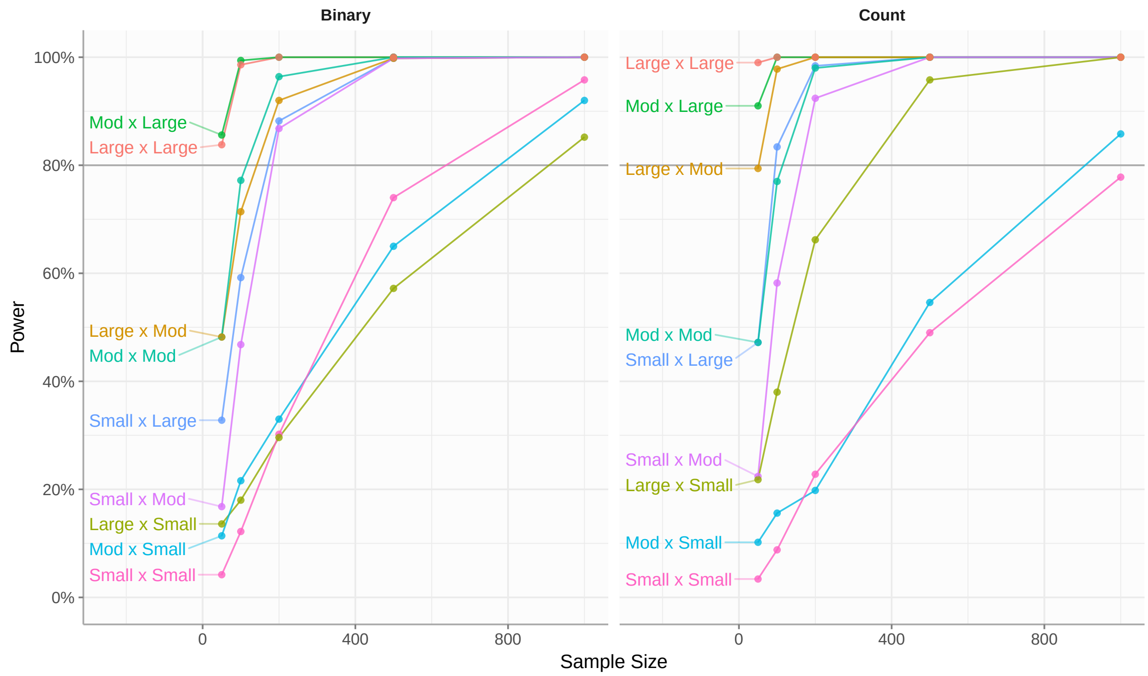

Figure 5.6: The simulated levels of power per tested sample size (x-axis) and effect size of the indirect path (color; a combination of the a path by the b path) stratified by the distribution of the mediator (binary or count).

Statistical Power

Figure 5.6 shows the statistical power of MMA across the various conditions. The figure shows the statistical power at each tested sample size for each combination of effect sizes (e.g., "Mod x Large" is a moderate a path effect size and a large b path effect size). Overall, most effect size combinations are adequately powered at a sample size of 200 across both binary and count mediator conditions. Interestingly, the "Small x Small" condition had more power at higher sample sizes than "Large x Small", which is contrary to intuition. However, the issue here was the issue of *complete separability* wherein the estimates and the standard errors are either biased or not estimable in logistic regression. With a large effect and a large sample size, this became common, thus reducing the statistical power in these conditions. The count mediator condition did not have this issue.

Overall, the method has the statistical power for even very small indirect effects with a sample size of 1000. As mentioned before, sample sizes greater than 1000 are common in the literature suggesting the method can be used even to detect small effect sizes.

Estimation Accuracy

It is also important for MMA to estimate the expected parameters. Figure 5.7 highlights that MMA is consistent in estimating the underlying effects for each combination of effect sizes across the various sample sizes. In the figure, which is stratified by the combination of effect sizes, shows the population parameter (vertical lines) and the estimated values (the density distributions). Overall, the distributions are centered at the true population parameter in each situation across the conditions.

As also seen in Figure 5.5, there is more variability in the estimation for larger effect sizes than for smaller. Again, this variability is likely due to the estimates being on a larger scale.

Confidence Interval Coverage

Finally, the confidence interval coverage is shown in Figure 5.8. Panel a) of the figure shows the overview—that the confidence interval coverage is around the 95% line for both the binary and count mediator conditions. However, looking at it much more closely in Panel b) it is clear that there is some deviation from the 95% line, particularly in the binary mediator condition. This is not a major deviation but an important one, nonetheless. Given the use of the percentile bootstrapping method herein, it may be important to apply other bootstrapping approaches such as the Bias-Corrected Bootstrap.

This finding of the indirect effect having confidence intervals that were too

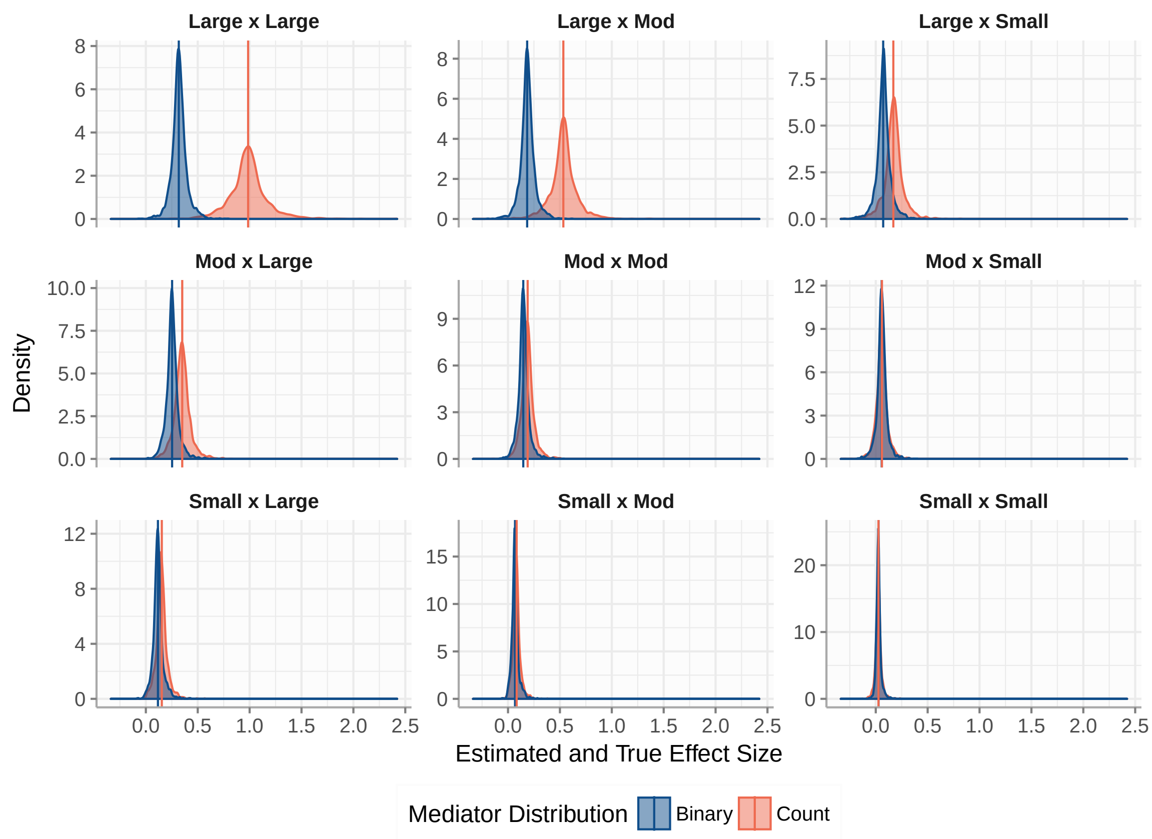

Figure 5.7: The simulated accuracy per tested sample size (x-axis) and effect size of the indirect path (color; a combination of the a path by the b path).

narrow has been found previously for the percentile bootstrapped (as applied in MMA; MacKinnon, Lockwood, and Williams, 2004). However, MacKinnon et al. (2004) also found the bootstrap methods, including the percentile approach, is among the best of the tested approaches. Other approaches, including the Monte Carlo confidence interval can be tested in future studies.

Conclusion

Marginal Mediation Analysis shows promise in its ability to accurately estimate models wherein the mediator is a binary or a count variable. Results regarding the decomposed total effect equaling the total effect are positive, although the estimation accuracy of this relationship depends on the sample and effect sizes. The

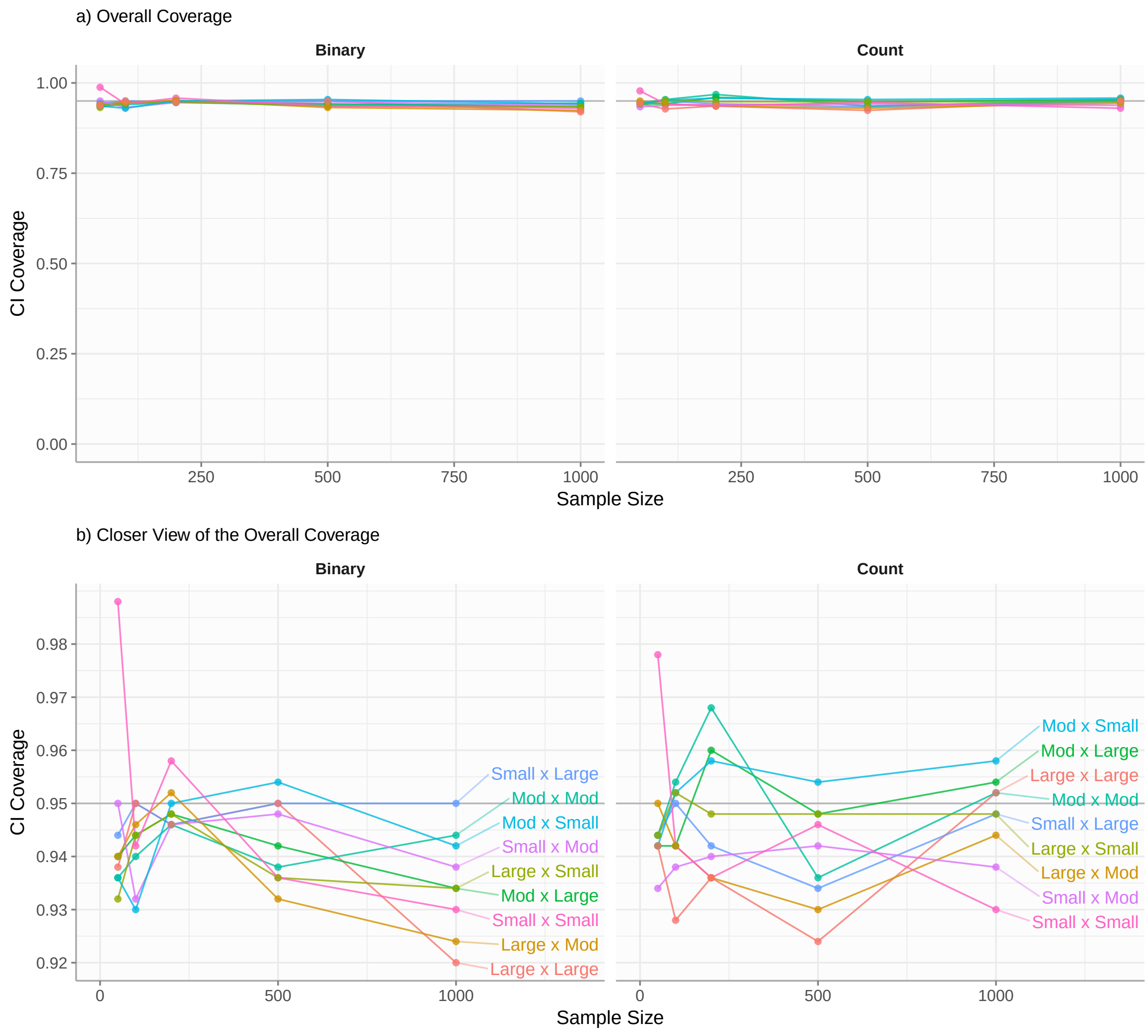

Figure 5.8: The simulated confidence interval coverage per tested sample size (x-axis) and effect size of the indirect effect . The "ideal" level is set at 0.95. Panel a) shows that the confidence intervals from a broad perspective. Panel b) provides a closer look at the individual patterns of the confidence intervals around the ideal level of 0.95.

statistical power is comparable to other modern mediation techniques wherein even small effect sizes can be estimated with a sample size of 1000. The estimation is consistent, ultimately averaging at the true population value. The confidence interval coverage was often too narrow for the binary mediator condition—sometimes having coverage of just above 92%—but it was approximately correct for the count mediator condition with some variability around 95%. Finally, the software for MMA is free to use in the R statistical environment in the `MarginalMediation` package. This allows researchers to begin to use the approach with little overhead. All in all, MMA appears to be a practical approach to difficult mediation situations.

CHAPTER 6

PHASE III: APPLICATION OF MMA

The power of intuitive understanding will protect you from harm until the end of your days. — Lao Tzu

Introduction

In 2012, Ford and Hill published an article that used some of the most common approaches to mediation when a mediator and/or outcome is categorical. Specifically, they used:

1. the difference method (MacKinnon, 2008),
2. the "categorical data method outlined by MacKinnon (2008)" (pg. 5) to assess the significance of the difference method, and
3. the percent of the total effect that was mediated.

These three approaches are not only common but likely some of the best approaches in this situation. However, as stated in Chapter 4, these have some notable shortcomings. First, the standard errors can be inefficient and biased if there is a high degree of multi-collinearity (or the degree to which there is perfect separability) in any of the models.[1] The significance of the difference method depends on these standard error estimates. Second, it does not provide the effect size measures that would be most useful [e.g., the effect of increasing the predictor on the outcome through the mediator(s)]. Third, the difference method is consistently too conservative with binary outcomes (Jiang and Vanderweele, 2015).

To build on the important findings from Ford and Hill (2012), this study replicates their work using more recent data from 2014 while using MMA to obtain ef-

[1] Perfect separability is where a predictor can perfectly predict the outcome in logistic regression.

fect sizes and confidence intervals for the indirect and direct effects.

Results

```
## Error in eval(expr, envir, enclos): object 'da36361.0001' not found
## Error in tolower(names(d)): object 'd' not found
## Error in eval(lhs, parent, parent): object 'd' not found
## Error in map(.x, .f, ...): object 'd1' not found
## Error in map_lgl(.x, .p, ...): object 'd1' not found
## Error in eval(lhs, parent, parent): object 'd1' not found
## Error in library(here): there is no package called 'here'
## Error in here("Data/NSDUH_2014_Results.rda"): could not find function "here"
## Error in library(survey): there is no package called 'survey'
## Error in svydesign(ids = ~1, strata = ~vestr, weights = ~analwt_c, data = d1):
## Error in svyglm(self ~ religious + age2 + irsex + newrace2 + irfamin3 + : could
## Error in svyglm(peer ~ religious + age2 + irsex + newrace2 + irfamin3 + : could
## Error in svyglm(dep ~ religious + age2 + irsex + newrace2 + irfamin3 + : could
## Error in coef(obj): object 'svy_a1' not found
## Error in rownames(est1) = c("Respondent", "Peer", "Depression"): object 'est1'
## Error in data.frame(est1): object 'est1' not found
## Error in svyglm(model, design = design, family = "quasibinomial"): could not fi
## Error in UseMethod("vcov"): no applicable method for 'vcov' applied to an objec
## Error in rownames(est2) = c("Tobacco", "Rx", "Marijuana", "Illicit"): object 'e
## Error in data.frame(est2): object 'est2' not found
```

The descriptive statistics are found in Table 6.1 for the 13,600 adolescents in the sample. Overall, these sample statistics were very similar to the 2007 sample used by Ford and Hill. However, the prevalence of drug use across each category dropped since 2007, although marijuana use did not drop substantially (13.85% in

2007 to 12.7% in 2014). Heavy drinking (10.4% in 2007) had only a single positive response in the entire sample of adolescents in 2014. Unfortunately, the number of major depressive episodes increased from 8.4% in 2007 to 11.3% in 2014. Attitudes regarding drug use were essentially identical as that in 2007 for the respondent, peer, and parent (and each had high reliabilities—all $\alpha \geq .80$—comparable to 2007). Finally, the attitudinal measures and the measure of religiosity had high reliabilities.

Table 6.1: Descriptive statistics of the sample.

	Mean/ Percent (SD) n = 13,600
Drug use	
Prescription drug misuse (past year)	5.7%
Tobacco use (past year)	11.8%
Heavy drinking (past 30 days)	0%
Marijuana use (past year)	12.7%
Other illicit drug use (past year)	3.5%
Demographics	
Female	49.0%
Race (Non-White)	45.8%
Income (2x poverty level)	55.4%
Major Depression Episode	11.3%
Attitudinal measures	
Respondent (range 1-3)	2.6 ($\alpha = .86$)
Peer (range 1-3)	2.5 ($\alpha = .88$)
Parent (range 1-3)	2.9 ($\alpha = .84$)
Religiosity	$\alpha = .80$

Four MMA models were used to assess the pathways from adolescent religiosity to substance use, one for each outcome (any tobacco use, prescription drug misuse, marijuana use, and other illicit drug use). Each model controlled for parental conservative attitudes toward substance use, the adolescents' family income, and the adolescents' age, race, and sex. Figure 6.1 presents the individual paths in the

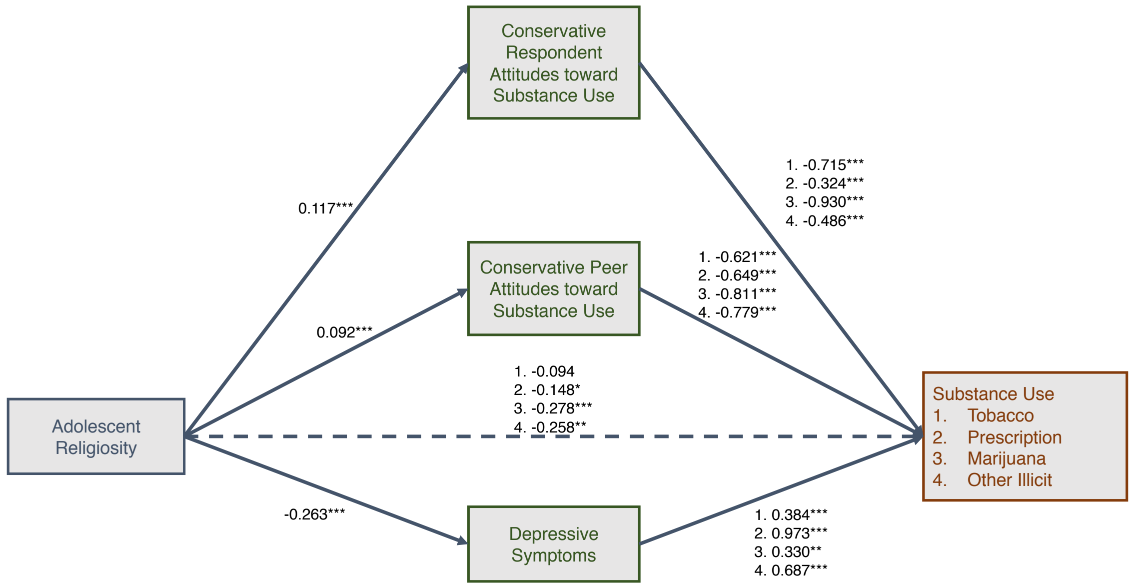

Figure 6.1: Results of the mediation models' individual paths regarding religiosity and substance use. Note: *** $p < .001$, ** $p < .01$, * $p < .05$

model units. Therefore, the paths leading to the conservative attitudes (both respondent and peer attitudes) are in the attitude metric with a range from 1 - 3. The paths leading to depression and the substance outcomes are all in log-odds. As the figure highlights, most paths were statistically significant at $p < .05$.

Because MMA provides information about each of the indirect effects naturally in the same units, it is possible to assess the amount mediated by each mediator while also controlling for the other mediators in a straightforward manner—without having to fit several other models and assess each $c - c'$. Table 6.2 presents the amount of the total effect of religiosity on substance use that is mediated through respondent conservative attitudes, peer conservative attitudes, and depression. Overall, the effect of religiosity on substance use is heavily mediated by the hypothesized mediators, more so for tobacco use than the others.[2]

In addition to this information, MMA provides information regarding the in-

[2]Using the approach used in Ford and Hill, the total mediated effects are slightly different than those estimated via the indirect and direct effects. This may be due to the weighting of the sample; an important area for further research into MMAs performance.

Table 6.2: The percent of mediation (the percent of the total effect) by path for each outcome.

	Outcome			
Mediator	Tobacco	Prescription	Marijuana	Illicit
Respondent Views	34.2	14.0	23.2	14.1
Peer Views	23.2	22.0	15.8	17.7
Depression	4.0	9.2	1.8	4.4
Total Mediation	61.4	45.2	40.8	36.1

direct and direct effects in the same units. Figure **??** highlights the indirect and direct effects with their associated 95% confidence intervals in the average marginal effects. All of the effects here are in risk (probability) units—i.e., risk of tobacco use, prescription misuse, marijuana use, or illicit drug use. Although all indirect effects and most direct effects are significant, the effect size estimates are particularly important here as the meaningfulness of these significant effects can be overlooked.

These resulting effects are all small, with most effects less than 0.01 (i.e., less than a single risk unit). That is, most effects show changes in the risk of the outcome by less than a single unit. For example, if adolescent religiosity is increased by one unit, its effect on the risk of tobacco use, through respondent attitudes, is a decrease of 0.007; through peer attitudes a decrease of 0.005; through depression a decrease of 0.001; and directly a decrease of 0.008. The total effect, then, is approximately -0.021. Therefore, if an individual has a risk of using tobacco at 50%, by increasing religiosity by one unit (holding the covariates constant), on average that individual's risk would decrease to 47.9%. About 0.013 of the effect of religiosity on tobacco use is mediated while 0.008 is direct from religiosity. Ultimately, these findings highlight the fact that the effect sizes, especially the indirect and direct effect sizes, are valuable companions to the p-values.

```
## Error in loadNamespace(name): there is no package called 'anteo'
```

Conclusions

The replication highlighted several important facets of the important work by Ford and Hill (2012). First, it simplifies the interpretation of the model by using average marginal effects. Second, it highlighted the effect sizes in terms of risk of substance use. This allowed the relatively small effects to be understood, not only in their significance, but in their meaning. Ultimately, MMA provided a more straightforward approach and substantially more information for each model than other mediation approaches.

CHAPTER 7

DISCUSSION

A model is a simplification or approximation of reality and hence will not reflect all of reality. ... While a model can never be "truth," a model might be ranked from very useful, to useful, to somewhat useful, to, finally, essentially useless. — Burnham and Anderson, 2002

General Discussion

Models are simply representations of reality. This is also true of mediation models even though they are used to model more complex relations. The value in using such models is generally seen in their ability to provide opportunities for intervention or prevention.[1]

As has been discussed throughout this project, mediation models are most useful when the model communicates both the significance (e.g., p-values) and meaningfulness (e.g., effect sizes). One without the other can be misleading, potentially resulting in faulty interventions and policies. Together, significance and meaning tell a more complete story of the data The significance helps researchers understand uncertainty; effect sizes communicate the potential for intervention to actually make meaningful changes in important outcomes.

However, some situations wherein mediation analysis is applied can provide a lack of interpretable information, particularly in terms of the effect sizes. This limits the usefulness of the model, whether or not it is an accurate representation

[1] Although not a major aspect of this project, it is important to note the predictive power a specific model has. For example, a model may show a significant effect in a logistic regression but it may not predict the outcome above chance. In this case, the question turns from "does X have an effect on Y?" to "does it predict Y?" This distinction has important implications when it comes to intervention and prevention but involves a number of important concepts that cannot be covered herein (e.g., model specification including penalty parameters, modeling approach such as tree based approaches). This general idea—that of the importance of predictive accuracy—is discussed in Hastie, Tibshirani, and Friedman (2009).

of reality. It is for this purpose that Marginal Mediation Analysis (MMA) was developed. It provides mediation analysis with the tools to communicate both significance and meaning. This is coming forth at an opportune time, as the American Psychological Association, among others, have called for more focus on effect sizes and less attention on p-values (Cumming, 2014).

Findings from the Three Phases

This project has presented the development of the approach and its software, the evaluation of its performance in possible real-world scenarios, and the application of it to health data regarding adolescent substance use. In its first phase, this project produced MMA with its accompanying software—the `MarginalMediation` R package. The software is freely available and allows for researchers to quickly apply it. The main function, `mma()`, is relatively quick even with the bootstrapping, and produces thorough output.

The next phase used Monte Carlo simulations to evaluate MMA and ways that it can possibly be improved. For example, given the results regarding the confidence interval coverage, it may be of benefit to try alternative approaches, either adaptations of bootstrapping (e.g., Bias-Corrected Bootstrap) or others. MacKinnon et al. (2004) found that Monte Carlo confidence intervals performed well and, therefore, may also be a valuable addition to MMA.

These simulations further demonstrated a trade-off regarding sample sizes and effect sizes: large effect sizes can be found in small samples but those same conditions provide much more variability in estimating the total effect accurately. Overall, these findings demonstrate the ability for the sample size, as it increases, to reduce bias and solidify relationships that should hold in mediation models (e.g., $a \times b + c' = c$).

The Monte Carlo simulations also allowed for the testing of the software.

Some situations, once simulated, demonstrated a need for change, often regarding the speed of the software, its accuracy, and necessary checks to avoid more serious problems. Ultimately, there was a natural feedback loop between the simulations and the software that were developed interactively. Once a stable version of the software was achieved, the reported simulations were all run based on that version (v0.5.0).

In the final phase, the application study highlighted important information regarding the MMA approach and adolescent health. The application study replicated work by Ford and Hill (2012), which was chosen to replicate for three major reasons:

1. the application study used a large sample with a mix of binary and continuous mediators and outcomes (common in the literature),
2. the statistical approach is one of the better approaches (also common in the literature), and
3. the data were open and a more recent release was available to investigate.

Although MMA can benefit the researcher in many situations, the benefits of using MMA are particularly clear within the context of this application study. Most importantly, MMA provided more information, in the form of effect size estimates, that help instruct on the meaningfulness of the results (Cumming, 2014). As Preacher and Kelley (2011) state: "it is important to develop a way to gauge the effect size of the product term *ab* itself," (pg. 95). That is, not only does the effect size of the individual paths need to be meaningful but the product of $a \times b$ must be as well. Although nearly all effects were significant, with such a large sample size significance tests alone can be misleading. For this study, the addition of the effect sizes are helpful to understand that each estimated effect was small. This provides a more complete view of the relationships tested herein.

Second, in terms of the substantive findings, there is strong evidence across many studies that adolescent religiosity is related to substance use. This is shown here as well. Consistently, religiosity was negatively related to the four substance use outcomes. About half of the total effect of religiosity on substance use was mediated by personal and peer attitudes about substance use. Depression also mediated the relationship, but to a much lesser degree.

Although not definitive, this study in conjunction with Ford and Hill (2012) presents evidence that religiosity may impact substance use outcomes through attitudes towards substance use. More research, particularly research with longitudinal data, are needed to further test and understand these relationships and their ability to inform intervention or policy.

Limitations

The MMA approach has two notable limitations. First, mediation analysis assumes no measurement error in the mediators. Although latent variable methods can help with this (Iacobucci, 2008; Lockhart, Mackinnon, and Ohlrich, 2011), the data necessary are not always available and the integration of average marginal effects within SEM is not clearly defined as of yet. Ultimately, the estimates are only as good as the measurements. Second, it may also be difficult for researchers to accept given the novelty of average marginal effects in the field. This is being alleviated through the use of various introductions to average marginal effects and its use in other fields (Barrett & Lockhart, in preparation).

Of course, the Monte Carlo simulation did not test for all conditions present in real-world modeling. Although it accounted for the main influences, there are other possible important influences that may impact the performance of the method, including missing values and model mis-specification. These are important influences to assess in future projects. Finally, the application study used cross-sectional data. This makes it difficult to demonstrate causality and puts additional pressure on the ability to control for confounding.

Future Research

Several foreseeable areas of investigation can prove useful for understanding and extending MMA. First, the application highlighted an important area for future inquiry—MMA with survey weighted data. The application study used data that were collected via a complex survey design and were therefore weighted. Further research is needed to understand MMAs behavior in these situations.

Second, this project specifically assessed binary and count mediators. Another important type of variable that could play an important role in mediation is "time-to-event" or survival data. Future research is needed to understand how this type of data with its accompanying statistical approaches can fit into MMA.

Third, MMA relies on the *sequential ignorability* assumption as described by Imai, Keele, and Tingley (2010). A sensitivity analysis is available to assess how important deviations from this assumption are on the estimates and conclusions (Imai et al., 2010; Imai, Keele, and Yamamoto, 2010). Integrating this sensitivity analysis would be a valuable addition to the approach. It likely would be a natural integration but this integration would need to be tested.

Relatedly, it could also be useful to look at using instrumental variables to help appease sequential ignorability. Although generally not applied in conjunction with mediation analysis, the approach could prove useful for MMA specifically and mediation analysis as a whole.

Lastly, the integration of latent variable approaches, including latent class analysis, is an important step in making this approach more broadly applicable. Work regarding average marginal effects, categorical data, and structural equation models would be an important contribution as well.

Conclusions

The results of the development, simulations, and application all show that MMA holds much promise in extending mediation analysis more fully to situations where the mediators and/or outcomes are categorical or non-normally distributed. Although further work is necessary to understand MMAs performance across more situations, the results of this project demonstrates its utility for common health and prevention research.

REFERENCES

Agostini, A., Talamo, P., and Vecchione, V. (2010). Combining operational loss data with expert opinions through advanced credibility theory. *The Journal of Operational Risk*, *5*(1), 3.

Altman, M. (2008). Behavioral economics, economic theory and public policy.

Barberis, N., and Thaler, R. (2003). A survey of behavioral finance. *Handbook of the Economics of Finance*, *1*, 1053–1128.

Bartus, T. (2005). Estimation of marginal effects using margeff. *Stata Journal*, *5*(3), 309–329. https://doi.org/The Stata Journal

Burnham, K., and Anderson, D. (2002). *Model selection and multimodel inference: A practical information-theoretic approach*. Springer-Verlag.

Carsey, T. M., and Harden, J. J. (2013). *Monte carlo simulation and resampling methods for social science*. SAGE.

Chavez-Demoulin, V., Embrechts, P., and Hofert, M. (2016). An extreme value approach for modeling operational risk losses depending on covariates. *Journal of Risk and Insurance*, *83*(3), 735–776.

Committee, B., and others. (2010). Basel iii: A global regulatory framework for more resilient banks and banking systems. *Basel Committee on Banking Supervision, Basel*.

Committee, B., and others. (2011). Operational risk–Supervisory guidelines for the advanced measurement approaches. *Basel: Bank for International Settlements*.

Cruz, M. G. (2002). *Modeling, measuring and hedging operational risk*. John Wiley & Sons New York,

Cumming. (2014). The new statistics: Why and how. *Psychological Science*, *25*(1), 7–29. https://doi.org/10.1177/0956797613504966

Edwards, J. R., and Lambert, L. S. (2007). Methods for integrating moderation and mediation: a general analytical framework using moderated path analysis. *Psychological Methods*, *12*(1), 1–22. https://doi.org/10.1037/1082-989X.12.1.1

Einemann, M., Fritscher, J., and Kalkbrener, M. (2018). Operational risk measurement beyond the loss distribution approach: An exposure-based methodology.

Embrechts, P., Mizgier, K. J., and Chen, X. (2018). Modeling operational risk depending on covariates. an empirical investigation.

Ford, J. A., and Hill, T. D. (2012). Religiosity and Adolescent Substance Use: Evidence From the National Survey on Drug Use and Health. *Substance Use &*

Misuse, *47*(7), 787–798. https://doi.org/10.3109/10826084.2012.667489

Gigerenzer, G., and Brighton, H. (2009). Homo heuristicus: Why biased minds make better inferences. *Topics in Cognitive Science*, *1*(1), 107–143.

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

Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The Elements of Statistical Learning* (Vol. 1). https://doi.org/10.1007/b94608

Hayes, A. F. (2009). Beyond Baron and Kenny: Statistical Mediation Analysis in the New Millennium. *Statistical Mediation Analysis in the New Millennium*, *76*(4), 408–420. https://doi.org/10.1080/03637750903310360

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. The Guilford Press.

Hemrit, W., and Arab, M. B. (2012). The major sources of operational risk and the potential benefits of its management. *The Journal of Operational Risk*, *7*(3), 71–92.

Heuer Jr., R. J. (1999). *Psychology of intelligence analysis*. Center for the Study of Intelligence.

Hoeppner, B., Hoeppner, S., and Abroms, L. (2017). How do text-messaging smoking cessation interventions confer benefit? A multiple mediation analysis of Text2Quit. *Addiction*, 673–682. https://doi.org/10.1111/add.13685

Hoohlo, M. (2015). *A new internal data measure for operational risk: A case study of a south african bank* (PhD thesis).

Iacobucci, D. (2008). *Mediation analysis*. SAGE [Kindle Version].

Iacobucci, D. (2012). Mediation analysis and categorical variables: The final frontier. *Journal of Consumer Psychology*, *22*(4), 582–594. https://doi.org/10.1016/j.jcps.2012.03.006

Iacobucci, D., Saldanha, N., and Deng, X. (2007). A mediation on mediation: Evidence that structural equation models perform better than regression. *Journal of Consumer Psychological*, *7*(2), 140–154. https://doi.org/10.1016/S1057-7408(07)70020-7

Imai, K., Keele, L., and Tingley, D. (2010). A general approach to causal mediation analysis. *Psychological Methods*, *15*(4), 309–34. https://doi.org/10.1037/a0020761

Imai, K., Keele, L., and Yamamoto, T. (2010). Identification, Inference and Sensitivity Analysis for Causal Mediation Effects. *Statistical Science*, *25*(1), 51–71.

https://doi.org/10.1214/10-STS321

Jiang, Z., and Vanderweele, T. (2015). When is the difference method conservative for assessing mediation? *American Journal of Epidemiology*, *182*(2), 105–108. https://doi.org/10.1093/aje/kwv059

Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist*, *58*(9), 697.

Kahneman, D., and Tversky, A. (2013). Prospect theory: An analysis of decision under risk. In *Handbook of the fundamentals of financial decision making: Part i* (pp. 99–127). World Scientific.

Kuhnen, C. M., and Knutson, B. (2005). The neural basis of financial risk taking. *Neuron*, *47*(5), 763–770.

Leeper, T. J. (2017). *Interpreting Regression Results using Average Marginal Effects with R's margins* (pp. 1–31). Retrieved from https://cran.r-project.org/web/packages/margins/vignettes/TechnicalDetails.pdf

List, J. A. (2004). Neoclassical theory versus prospect theory: Evidence from the marketplace. *Econometrica*, *72*(2), 615–625.

Lockhart, G., Mackinnon, D. P., and Ohlrich, V. (2011). Mediation analysis in psychosomatic medicine research. *Psychosomatic Medicine*, *73*(1), 29–43. https://doi.org/10.1097/PSY.0b013e318200a54b.Mediation

MacKinnon, D. P. (2008). *Introduction to statistical mediation analysis*. Routledge.

MacKinnon, D. P., Lockwood, C. M., and Williams, J. (2004). Confidence Limits for the Indirect Effect: Distribution of the Product and Resampling Methods. *Multivariate Behavioral Research*, *39*(1). https://doi.org/10.1207/s15327906mbr3901_4

Mignola, G., Ugoccioni, R., and Cope, E. (2016). Comments on the basel committee on banking supervision proposal for a new standardized approach for operational risk.

Montreuil, B., Bendavid, Y., and Brophy, J. (2005). What is so odd about odds? *J Can Chir*, *48*(5), 400–408.

Mood, C. (2010). Logistic regression: Why we cannot do what We think we can do, and what we can do about it. *European Sociological Review*, *26*(1), 67–82. https://doi.org/10.1093/esr/jcp006

Mysis. (2013). *Achieving Basel III compliance: how to tackle it and business issues mysis*. Retrieved from http://www.risktech-forum.com/research/achieving-basel-iii-compliance-how-to-tackle-it-and-business-issues

Norton, E. C. (2012). *Log Odds and Ends*. National Bureau of Economic Research. https://doi.org/10.1002/0471725331.ch16

Nylund, K. L., Asparouhov, T., and Muthen, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling*, *15*(1), 182. https://doi.org/10.1080/10705510701575396

Paxton, P., Curran, P. J., Bollen, K. a., Kirby, J., and Chen, F. (2001). Monte Carlo Experiments: Design and Implementation. *Structural Equation Modeling: A Multidisciplinary Journal*, *8*(2), 287–312. https://doi.org/10.1207/S15328007SEM0802_7

Peters, G., Shevchenko, P. V., Hassani, B., and Chapelle, A. (2016). Should the advanced measurement approach be replaced with the standardized measurement approach for operational risk?

Preacher, K. J., and Kelley, K. (2011). Effect Size Measures for Mediation Models : Quantitative Strategies for Communicating Indirect Effects, *16*(2), 93–115. https://doi.org/10.1037/a0022658

Risk, B. O. (2001). Supporting document to the new basel capital accord. *Consultative Document, January*, *200*.

Risk, B. O. (2016). Standardised measurement approach for operational risk. *Consultative Document, June*.

Shefrin, H. (2016). *Behavioral risk management: Managing the psychology that drives decisions and influences operational risk*. Springer.

StataCorp. (2015). *Stata Statistical Software: Release 14*. StataCorp LLC.

Tom, S. M., Fox, C. R., Trepel, C., and Poldrack, R. A. (2007). The neural basis of loss aversion in decision-making under risk. *Science*, *315*(5811), 515–518.

Wickham, H. (2015). *R Packages: Organize, Test, Document, and Share Your Code* (1st ed.). O'Reilly Media.

Winship, C., and Mare, R. (1983). Structural Equations and Path Analysis for Discrete Data. *The American Journal of Sociology*, *89*(1), 54–110.

Wong, M. M., and Brower, K. J. (2013). The prospective relationship between sleep problems and suicidal behavior in the National Longitudinal Study of Adolescent Health. *Journal of Psychiatric Research*, *46*(7), 953–959. https://doi.org/10.1016/j.jpsychires.2012.04.008.The

APPENDICES

Appendix A: R Code for Chapter 5

Required: R Packages from CRAN

```
if (!require(tidyverse)){
  install.packages("tidyverse")
  library(tidyverse)
}
if (!require(furniture)){
  install.packages("furniture")
  library(furniture)
}
if (!require(here)){
  install.packages("here")
  library(here)
}
if (!require(devtools)){
  install.packages("devtools")
  library(devtools)
}
```

Required: R Packages from GitHub

```
if (!require(MarginalMediation)){
  devtools::install_github("tysonstanley/MarginalMediation")
  library(MarginalMediation)
}
```

Examples from Chapter 5

Figure 5.2 on page 46

```
set.seed(843)
library(tidyverse)
tibble::data_frame(
  X = rnorm(100),
  W = rbinom(100, 1, .5),
  Y = X + .5*W + -.5*X*W + rnorm(100)
) %>%
  ggplot(aes(X, Y, group = factor(W), color = factor(W))) +
    geom_point(alpha = .3) +
    geom_smooth(method = "lm", se=FALSE) +
    scale_color_manual(values = c("dodgerblue4", "chartreuse4")) +
    labs(color = "Moderator") +
    anteo::theme_anteo_wh()
ggsave("figures/fig_interaction_effect.pdf",
       width = 7, height = 5, units = "in")
```

Monte Carlo Simulation

Notably, the code for both the binary mediator condition and the count mediator condition we run via the Terminal as, once the directory was where the R file was located:

```
Rscript Analyses_MMMC_scriptBinary.R 'c(1:45)'
```

and

```
Rscript Analyses_MMMC_scriptCount.R 'c(1:45)'
```

Binary Mediator

```
## Marginal Mediation: Monte Carlo Simulation Study
##   BINARY Mediator
## Tyson S. Barrett
##
## devtools::install_github("tysonstanley/MarginalMediation")

args <- commandArgs(TRUE)
args <- eval(parse(text = args))
library(MarginalMediation)
library(tidyverse)

## Create all combinations of independent variables
cond_binary = expand.grid(
  samplesize = c(50, 100, 200, 500, 1000),
  effecta    = c(.55, 1.45, 2.22),
  effectb    = c(.24, .62, 1.068),
  effectc    = c(.3)
)

## Population Models
## Binary Mediator
data_genB = function(ps, reps, samplesize, effecta, effectb, effectc){
  set.seed(84322)
  Xc = rnorm(ps)
  z  = effecta*Xc + rnorm(ps, 0, 1)
  pr = 1/(1+exp(-z))
  M  = rbinom(ps, 1, pr)
  Y  = effectb*M + effectc*Xc + rnorm(ps, 0, 1)
  M  = factor(M)
  df = data.frame(Y, M, Xc)
  bin = vector("list", reps)

  print(cbind(samplesize, effecta, effectb))
  print(lm(Y ~ M + Xc)$coefficients)
  print(lm(scale(Y) ~ M + Xc)$coefficients)
  med = amed(glm(M ~ Xc, df, family = "binomial"))
```

```
  for (i in 1:reps){
    d = df[sample(ps, samplesize), ]
    pathbc = glm(Y ~ M + Xc, data = d)
    patha  = glm(M ~ Xc, data = d, family = "binomial")
    bin[[i]] = mma(pathbc, patha,
          ind_effects = c("Xc-M"),
          boot = 500)
    bin[[i]] = list("IndEffects" = bin[[i]]$ind_effects,
                    "DirEffects" = bin[[i]]$dir_effects,
                    "Boot"       = bin[[i]]$boot,
                    "Total"      = lm(Y ~ Xc, d)$coefficients,
                    "MedSize"    = med)
    cat("\r", i)
  }
  print(exp(glm(M ~ Xc, family = "binomial")$coefficients))
  return(bin)
}

i = 0
for (j in args){
  set.seed(84322)
  i = i + 1
  cat("\nNumber:", j, "\n\n")

  out = data_genB(1e6, 500,
                  cond_binary[args[[i]],1],
                  cond_binary[args[[i]],2],
                  cond_binary[args[[i]],3],
                  cond_binary[args[[i]],4])

  save(out, file = paste0("Sims_Data/Binary2_",
                          cond_binary[args[[i]],1], "_",
                          cond_binary[args[[i]],2], "_",
                          cond_binary[args[[i]],3], "_",
                          cond_binary[args[[i]],4], ".rda"))

  cat("\nNumber:", j, "\n\n")
  cat("\nConditions Complete:\n",
      " Sample size =", cond_binary[args[[i]],1],
      "\n  A path      =", cond_binary[args[[i]],2],
      "\n  B path      =", cond_binary[args[[i]],3],
      "\n  C path      =", cond_binary[args[[i]],4], "\n")
}
```

Count Mediator

```
## Marginal Mediation: Monte Carlo Simulation Study
##   COUNT Mediator
## Tyson S. Barrett
##
## devtools::install_github("tysonstanley/MarginalMediation")

args <- commandArgs(TRUE)
args <- eval(parse(text = args))
library(MarginalMediation)
library(tidyverse)

## Create all combinations of independent variables
cond_count = expand.grid(
  samplesize = c(50, 100, 200, 500, 1000),
  effecta    = c(.3, .6, 1.1),
  effectb    = c(.084, .265, .49),
  effectc    = c(0, .3)
)

## Population Models
## Count Mediator
data_genC = function(ps, reps, samplesize, effecta, effectb, effectc){
  set.seed(84322)
  Xc = rnorm(ps)
  m1 = exp(effecta * Xc)
  M  = rpois(ps, lambda=m1)
  Y  = effectb*M + effectc*Xc + rnorm(ps, 0, 1)
  df = data.frame(Y, M, Xc)
  poi  = vector("list", reps)

  print(cbind(samplesize, effecta, effectb))
  print(lm(Y ~ M + Xc)$coefficients)
  print(lm(scale(Y) ~ M + Xc)$coefficients)
  med = amed(glm(M ~ Xc, df, family = "poisson"))

  for (i in 1:reps){
    d = df[sample(ps, samplesize), ]
    pathbc = glm(Y ~ M + Xc, data = d)
    patha  = glm(M ~ Xc, data = d, family = "poisson")
    poi[[i]] = mma(pathbc, patha,
                   ind_effects = c("Xc-M"),
                   boot = 500)
    poi[[i]] = list("IndEffects" = poi[[i]]$ind_effects,
                    "DirEffects" = poi[[i]]$dir_effects,
                    "Boot"       = poi[[i]]$boot,
                    "Total"      = lm(Y ~ Xc, d)$coefficients,
                    "MedSize"    = med)
```

```
    cat("\r", i)
  }
  print(exp(glm(M ~ Xc, family = "poisson")$coefficients))
  return(poi)
}

i = 0
for (j in args){
  set.seed(84322)
  i = i + 1
  cat("\nNumber:", j, "\n\n")

  out = data_genC(1e6, 500,
                  cond_count[args[[i]],1],
                  cond_count[args[[i]],2],
                  cond_count[args[[i]],3],
                  cond_count[args[[i]],4])

  save(out, file = paste0("Sims_Data/Count2_",
                          cond_count[args[[i]],1], "_",
                          cond_count[args[[i]],2], "_",
                          cond_count[args[[i]],3], "_",
                          cond_count[args[[i]],4], ".rda"))

  cat("\nNumber:", j, "\n\n")
  cat("\nConditions Complete:\n",
      " Sample size =", cond_count[args[[i]],1],
      "\n  A path      =", cond_count[args[[i]],2],
      "\n  B path      =", cond_count[args[[i]],3],
      "\n  C path      =", cond_count[args[[i]],4], "\n")
}
```

Monte Carlo Simulation Data Analyses

Data Preparations for tables and figures around page 57

```
options(na.rm=TRUE)

library(tidyverse)
library(furniture)
library(here)

es = read.csv("effect_sizes.csv") %>%
  data.frame(row.names = .$size) %>%
  select(-size)

filenames = list.files(here("Sims_Data/"),
                       pattern = ".rda")

tot = indc = indd =
  dirc = dird = vector("list", length(filenames))
for (i in filenames){
  cat("File:", i, "\n")
  load(paste0(here("Sims_Data/", i)))

  tot[[i]] = lapply(out, function(x) x$Total) %>%
    do.call("rbind", .) %>%
    data.frame %>%
    mutate(type = strsplit(i, "_")) %>%
    mutate(ss   = map_chr(type, ~.x[2])) %>%
    mutate(dist = map_chr(type, ~.x[1])) %>%
    mutate(boot = map_chr(dist, ~ifelse(grepl("2$", .x), 500, 100))) %>%
    mutate(dist = map_chr(dist, ~gsub("2", "", .x))) %>%
    mutate(ap   = map_chr(type, ~.x[3])) %>%
    mutate(bp   = map_chr(type, ~.x[4])) %>%
    mutate(cp   = map_chr(type, ~gsub("\\.rda", "", .x[5]))) %>%
    select(Xc, ss, dist, boot, ap, bp, cp)
  indc[[i]] = lapply(out, function(x) x$IndEffects[1, ]) %>%
    do.call("rbind", .) %>%
    data.frame %>%
    mutate(type = gsub(".rda", "", i)) %>%
    mutate(type = strsplit(type, "_")) %>%
    mutate(ss   = map_chr(type, ~.x[2])) %>%
    mutate(dist = map_chr(type, ~.x[1])) %>%
    mutate(boot = map_chr(dist, ~ifelse(grepl("2$", .x), 500, 100))) %>%
    mutate(dist = map_chr(dist, ~gsub("2", "", .x))) %>%
    mutate(ap   = ifelse(dist == "Count",
                     ifelse(map_chr(type, ~.x[3]) == "0.3", "Small",
                     ifelse(map_chr(type, ~.x[3]) == "0.6", "Mod",
                            "Large")),
```

```
                            ifelse(map_chr(type, ~.x[3]) == "0.55", "Small",
                            ifelse(map_chr(type, ~.x[3]) == "1.45", "Mod",
                                   "Large")))) %>%
    mutate(bp   = ifelse(map_chr(type, ~.x[4]) == "0.24", "Small",
                  ifelse(map_chr(type, ~.x[4]) == "0.62", "Mod",
                  ifelse(map_chr(type, ~.x[4]) == "1.068", "Large",
                  ifelse(map_chr(type, ~.x[4]) == "0.084", "Small",
                  ifelse(map_chr(type, ~.x[4]) == "0.265", "Mod",
                         "Large")))))) %>%
    mutate(cp   = map_chr(type, ~.x[5])) %>%
    mutate(power = ifelse(Lower > 0 & Upper > 0, 1, 0)) %>%
    mutate(ind_cat = paste(ap, "x", bp)) %>%
    mutate(true  = ifelse(dist == "Count",
                          es[ind_cat, "ind_count"],
                          es[ind_cat, "ind_binary"])) %>%
    mutate(ap_size = ifelse(dist == "Count",
                            es[ind_cat, "a_count"],
                            es[ind_cat, "a_binary"])) %>%
    mutate(bp_size = ifelse(dist == "Count",
                            es[ind_cat, "b_count"],
                            es[ind_cat, "b_binary"])) %>%
    mutate(ci    = ifelse(true < Upper & true > Lower, 1, 0)) %>%
    select(-type)
  dirc[[i]] = lapply(out, function(x) x$DirEffects[1, ]) %>%
    do.call("rbind", .) %>%
    data.frame %>%
    mutate(type = gsub(".rda", "", i)) %>%
    mutate(type = strsplit(type, "_")) %>%
    mutate(ss   = map_chr(type, ~.x[2])) %>%
    mutate(dist = map_chr(type, ~.x[1])) %>%
    mutate(boot = map_chr(dist, ~ifelse(grepl("2$", .x), 500, 100))) %>%
    mutate(dist = map_chr(dist, ~gsub("2", "", .x))) %>%
    mutate(ap   = map_chr(type, ~.x[3])) %>%
    mutate(bp   = map_chr(type, ~.x[4])) %>%
    mutate(cp   = map_chr(type, ~.x[5])) %>%
    mutate(ci   = ifelse(Lower > 0 & Upper > 0 & cp > 0, 1,
                  ifelse(Lower < 0 & Upper > 0 & cp == 0, 1, 0))) %>%
    mutate(power = ifelse(Lower > 0 & Upper > 0, 1, 0)) %>%
    mutate(true  = cp) %>%
    select(-type)
}

ind1 = do.call('rbind', indc) %>%
  mutate(var = "continuous") %>%
  select(Indirect, Lower, Upper, ss, dist,
         boot, ap, bp, cp, ci, true, power,
```

```
         var, ap_size, bp_size)
dir1 = do.call('rbind', dirc) %>%
  mutate(var = "continuous") %>%
  select(Direct, Lower, Upper, ss, dist,
         boot, cp, ci, var, power)
tot1  = do.call('rbind', tot) %>%
  select(Xc, ss, dist, boot)
```

Table 5.2 on page 57

```
total_total %>%
  group_by(dist, sample) %>%
  summarize(tot1 = mean(total),
            tot2 = mean(indirect + direct),
            true = mean(true)) %>%
  mutate(est = (tot1 - tot2) / tot1) %>%
  mutate(est2 = tot1 - tot2) %>%
  data.table::dcast(sample~dist, value.var = "est") %>%
  xtable::xtable(digits = 4) %>%
  xtable::print.xtable(include.rownames = FALSE)
```

Figure 5.5 on page 58

```
## Total = Total
total_total = cbind(tot1[,1], ind1[1:45000,1], dir1[1:45000,1]) %>%
  data.frame %>%
  set_names(c("total", "indirect", "direct")) %>%
  mutate(sample = ind1[1:45000,"ss"]) %>%
  mutate(sample = factor(sample,
                         levels = c("50", "100", "200", "500", "1000"))) %>%
  mutate(ap   = ind1[1:45000,"ap"],
         bp   = ind1[1:45000,"bp"],
         dist = ind1[1:45000,"dist"],
         true = ind1[1:45000,"true"]) %>%
  mutate(true = as.numeric(as.character(true))) %>%
  mutate(diff = (total - (indirect + direct))) %>%
  mutate(est  = (total - (indirect + direct))/true) %>%
  mutate(eff  = paste(ap, "x", bp)) %>%
  group_by(sample, eff, dist) %>%
  mutate(index = 1:n())
p1tot = total_total %>%
  filter(dist == "Binary") %>%
  ggplot(aes(index, diff)) +
    geom_path() +
    scale_x_continuous(breaks = c(250, 500),
                       labels = c("250", "500")) +
    scale_y_continuous(breaks = c(-.1,0,.1)) +
    facet_grid(eff~sample) +
    anteo::theme_anteo_wh() +
```

```
    theme(panel.spacing = unit(.1, "cm"),
          strip.text.y = element_blank()) +
    labs(x = "Simulation Number",
         y = "Total - (Indirect + Direct)\n",
         subtitle = "a) Binary Condition")
p2tot = total_total %>%
  filter(dist == "Count") %>%
  ggplot(aes(index, diff)) +
  geom_path() +
  scale_x_continuous(breaks = c(250, 500),
                     labels = c("250", "500")) +
  scale_y_continuous(breaks = c(-2,0,2)) +
  coord_cartesian(ylim = c(-2.9,3.1)) +
  facet_grid(eff~sample) +
  anteo::theme_anteo_wh() +
  theme(panel.spacing = unit(.1, "cm")) +
  labs(x = "Simulation Number",
       y = "",
       subtitle = "b) Count Condition")

plot_total = gridExtra::grid.arrange(p1tot, p2tot, ncol = 2)
ggsave("figures/fig_total_total.pdf",
       plot = plot_total,
       width = 10, height = 10, units = "in")
```

Figures 5.6, 5.7, and 5.8 on pages 59, 61, and 62, respectively.

```
## Accuracy, Power, Coverage
inds = ind1 %>%
  mutate(accuracy = (as.numeric(Indirect) -
                     (as.numeric(ap_size) *
                        as.numeric(bp_size)))) %>%
  group_by(ss, dist, boot = as.numeric(boot), ap_size,
           bp_size, cp, var, ap, bp) %>%
  summarize(Ind = mean(Indirect),
            Low = mean(Lower),
            Hi  = mean(Upper),
            ci  = mean(ci),
            power = mean(power),
            acc = mean(accuracy)) %>%
  ungroup
dirs = dir1 %>%
  group_by(ss, dist, boot = as.numeric(boot), cp, var) %>%
  summarize(Dir = mean(Direct),
            Low = mean(Lower),
            Hi  = mean(Upper),
            ci  = mean(ci),
            power = mean(power))
```

```
ggplot(inds, aes(x = as.numeric(ss), y = power,
                 color = paste(ap, "x", bp),
                 group = interaction(ap, bp, boot, dist, cp))) +
  geom_hline(yintercept = .8, color = "darkgrey") +
  geom_line(alpha = .8) +
  geom_point(alpha = .8) +
  facet_grid(~ dist, space = "free", scales = "free") +
  anteo::theme_anteo_wh() +
  theme(panel.spacing = unit(.25, "cm"),
        axis.line = element_line(color = "darkgrey"),
        legend.position = "none") +
  scale_y_continuous(breaks = c(0, .2, .4, .6, .8, 1),
                     labels = scales::percent) +
  labs(y = "Power",
       x = "Sample Size") +
  ggrepel::geom_text_repel(data = inds %>%
            filter(ss == 50),
          aes(label = paste(ap, "x", bp)),
          nudge_x = -150) +
  coord_cartesian(xlim = c(-250, 1000),
                  ylim = c(0,1))
ggsave("figures/sim_fig_power.pdf",
       width = 10, height = 6, units = "in")

ggplot(ind1, aes(x = Indirect, fill = dist,
                 color = dist,
                 group = interaction(ap, bp, boot, dist, cp))) +
  geom_density(alpha = .5) +
  geom_vline(aes(xintercept = true, color = dist)) +
  facet_wrap(~ paste(ap, "x", bp), scales = "free") +
  anteo::theme_anteo_wh() +
  theme(panel.spacing = unit(.25, "cm"),
        axis.line = element_line(color = "darkgrey"),
        legend.position = "bottom") +
  labs(y = "Density",
       x = "Estimated and True Effect Size",
       fill = "Mediator Distribution",
       color = "Mediator Distribution") +
  scale_fill_manual(values = c("dodgerblue4", "coral2")) +
  scale_color_manual(values = c("dodgerblue4", "coral2"))
ggsave("figures/sim_fig_acc.pdf",
       width = 8, height = 6, units = "in")

p1 = ggplot(inds, aes(x = as.numeric(ss), y = ci,
                      color = paste(ap, "x", bp),
                      group = interaction(ap, bp, boot, dist, cp))) +
```

```
  geom_hline(alpha = .8, yintercept = .95, color = "darkgrey") +
  geom_line(alpha = .8) +
  geom_point(alpha = .8) +
  facet_grid(~ dist, space = "free", scales = "free") +
  anteo::theme_anteo_wh() +
  theme(panel.spacing = unit(.25, "cm"),
        axis.line = element_line(color = "darkgrey"),
        legend.position = "none") +
  labs(y = "CI Coverage",
       x = "Sample Size",
       subtitle = "a) Overall Coverage") +
  coord_cartesian(ylim = c(0,1))
p2 = ggplot(inds, aes(x = as.numeric(ss), y = ci,
                      color = paste(ap, "x", bp),
                      group = interaction(ap, bp, boot, dist, cp))) +
  geom_hline(alpha = .8, yintercept = .95, color = "darkgrey") +
  geom_line(alpha = .8) +
  geom_point(alpha = .8) +
  facet_grid( ~ dist, space = "free", scales = "free") +
  anteo::theme_anteo_wh() +
  theme(panel.spacing = unit(.25, "cm"),
        axis.line = element_line(color = "darkgrey"),
        legend.position = "none") +
  labs(y = "CI Coverage",
       x = "Sample Size",
       subtitle = "b) Closer View of the Overall Coverage") +
  ggrepel::geom_text_repel(data = inds %>%
                             filter(ss == 1000),
                           aes(label = paste(ap, "x", bp)),
                           nudge_x = 250,
                           segment.alpha = .4) +
  coord_cartesian(xlim = c(0, 1350)) +
  scale_y_continuous(breaks = c(.92, .93, .94, .95, .96, .97, .98))
p3 = gridExtra::grid.arrange(p1,p2, ncol = 1)
ggsave("figures/sim_fig_ci.pdf",
       plot = p3,
       width = 10, height = 9, units = "in")
```

Appendix B: R Code for Chapter 6

Required: R Packages from CRAN

```
if (!require(tidyverse)){
  install.packages("tidyverse")
  library(tidyverse)
}
if (!require(furniture)){
  install.packages("furniture")
  library(furniture)
}
if (!require(here)){
  install.packages("here")
  library(here)
}
if (!require(devtools)){
  install.packages("devtools")
  library(devtools)
}
if (!require(survey)){
  install.packages("survey")
  library(survey)
}
```

Required: R Packages from GitHub

```
if (!require(MarginalMediation)){
  devtools::install_github("tysonstanley/MarginalMediation")
  library(MarginalMediation)
}
```

Data Preparation

Data preparation using the 2014 National Survey on Drug Use and Health, as described in Chapter 4.

```
library(tidyverse)
library(furniture)

## Load data
load("Data/NSDUH_2014_Results.rda")
d = da36361.0001
names(d) = tolower(names(d))
rm(da36361.0001)

## Variables
d1 = d %>%
  select(
    ## ---------------------------------- ##
    ##  Outcomes                          ##
    ##    (1,2,8,11,12 = within last year) ##
    ## ---------------------------------- ##
    ## Tobacco Outcome
    cigrec,    ## cig
    chewrec,   ## chew
    cigarrec,  ## cigar
    #pipe30dy,  ## pipe (30 days here instead)

    ## Heavy Drinking Outcome
    dr5day,  ## 1+ is within last 30 days

    ## Rx Outcome
    analrec, ## pain relievers
    tranrec, ## tranquilizers
    stimrec, ## stimulants
    sedrec,  ## sedatives

    ## Marijuana Outcome
    mjrec,   ## marijuana

    ## Other Illicit Outcome
    cocrec,  ## cocaine
    crakrec, ## crack
    herrec,  ## heroin
    hallrec, ## hallucinogens
    lsdrec,  ## LSD
    pcprec,  ## PCP
    ecsrec,  ## ecstacy
    inhrec,  ## inhalants
```

```
    methrec, ## meth

    ## ----------------------------------- ##
    ## Mediators                           ##
    ##   Mean response (higher = more cons) ##
    ## ----------------------------------- ##
    ## Self Views Mediator
    yegpkcig, ## someone your age cig
    yegmjevr, ## someone your age mj
    yegmjmo,  ## someone your age mj monthly
    yegaldly, ## someone your age drinking daily

    ## Peer Views Mediator
    yefpkcig, ## you cig
    yefmjevr, ## you mj
    yefmjmo,  ## you mj monthly
    yefaldly, ## you drinking daily

    ## Psychological Well-being (Major Depression)
    ymdeyr,  ## past year major depressive epidosde (MDE)

    ## ---------------------------------- ##
    ## Predictor                          ##
    ##   Cronbach's Alpha                 ##
    ##   Standardized mean level          ##
    ## ---------------------------------- ##
    ## Religiosity
    yerlgsvc, ## past 12, times at church (1-6
    yerlgimp, ## religious beliefs are important (1-4 strong dis to strong agree)
    yerldcsn, ## religious belief influence decisions (1-4)
    yefaiact, ## religious activities

    ## ---------------------------------- ##
    ## Control Variables                  ##
    ## ---------------------------------- ##
    ## Parental Attitudes
    yeppkcig, ## parents feel about cig
    yepmjevr, ## parents feel about mj
    yepmjmo,  ## parents feel about mj monthly
    yepaldly, ## parents feel about drinking daily

    ## Demographics
    age2,    ## age
    catage,  ## age category (1 = 12-17 year old)
    irsex,   ## gender (1 = male)
    newrace2, ## race (1 = White, 2-7 non-white)
    irfamin3, ## total family income (6 = 50,000 - 74,999)
```

```
    poverty2, ## not used in the study but could be for ours
              ##  (1 = poverty, 2 = low middle, 3 = middle class or more)

    ## ---------------------------------- ##
    ## Sampling Variables                 ##
    ## ---------------------------------- ##
    analwt_c,  ## sample weight
    vestr,     ## analysis stratum
    verep      ## analysis replicate
    ) %>%
  filter(catage == "(1) 12-17 Years Old")

## Data Cleaning
dich = function(x){
  x = ifelse(grepl("(01)|(02)|(08)|(11)", x), 1, 0)
  x
}
map_to = function(x){
  lbls = sort(levels(x))
  lbls = (sub("^\\([0-9]+\\) +(.+$)", "\\1", lbls))
  x = as.numeric(gsub("^\\(0*([0-9]+)\\).+$", "\\1", x))
  x
}
d1[, c(1:18)]  = map_df(d1[, c(1:18)], ~dich(.x))
d1[, c(19:36)] = map_if(d1[, c(19:36)], is.factor, ~map_to(.x))

## Creating final modeling variables
d1 = d1 %>%
  mutate(tobacco = ifelse(rowSums(cbind(cigrec, chewrec,
                                        cigarrec)) > 0, 1, 0),
         drink   = dr5day,
         rx      = ifelse(rowSums(cbind(analrec, tranrec,
                                        stimrec, sedrec)) > 0, 1, 0),
         mari    = ifelse(mjrec == 1, 1, 0),
         illicit = ifelse(rowSums(cbind(cocrec, crakrec,
                                        herrec, hallrec,
                                        lsdrec, pcprec,
                                        ecsrec, inhrec,
                                        methrec)) > 0, 1, 0)) %>%
  mutate(self = rowMeans(cbind(yegpkcig, yegmjevr, yegmjmo, yegaldly)),
         peer = rowMeans(cbind(yefpkcig, yefmjevr, yefmjmo, yefaldly))) %>%
  mutate(dep = washer(ymdeyr, 2, value = 0)) %>%
  mutate(religious = rowMeans(cbind(scale(yerlgsvc),
                                    scale(yerlgimp),
                                    scale(yerldcsn),
                                    scale(yefaiact)))) %>%
  mutate(parent = rowMeans(cbind(yeppkcig, yepmjevr,
```

```
                                          yepmjmo, yepaldly)))
```

Models

```
## Sampling Design
library(survey)
design = svydesign(ids = ~1,
                   strata = ~vestr,
                   weights = ~analwt_c,
                   data = d1)

## All a Path Models
## Unadjusted
svy_a1 = svyglm(self ~ religious, design = design)
svy_a2 = svyglm(peer ~ religious, design = design)
svy_a3 = svyglm(dep ~ religious, design = design,
                family = 'quasibinomial')

## Adjusted
svy_a12 = svyglm(self ~ religious + age2 +
                   irsex + newrace2 + irfamin3 + parent,
                 design = design)
svy_a22 = svyglm(peer ~ religious + age2 +
                   irsex + newrace2 + irfamin3 + parent,
                 design = design)
svy_a32 = svyglm(dep ~ religious + age2 +
                   irsex + newrace2 + irfamin3 + parent,
                 design = design,
                 family = 'quasibinomial')

## All b and c' Path Models (drink such low prevalence that it was not included)
svy_bc = svy_bc2 = list()
for (i in c("tobacco", "rx", "mari", "illicit")){

  ## Unadjusted Model
  model = as.formula(paste0(i, "~ self + peer + dep + religious"))
  svy_bc[[i]] = svyglm(model, design = design, family = "binomial")

  ## Adjusted Model
  model2 = as.formula(paste0(i, "~ self + peer + dep + religious + age2 +
                             irsex + newrace2 + irfamin3 + parent"))
  svy_bc2[[i]] = svyglm(model2, design = design, family = "binomial")

}

library(MarginalMediation)
## Tobacco
fit_tob = mma(svy_bc[["tobacco"]],
              svy_a1,
```

```
              svy_a2,
              svy_a3,
              ind_effects = c("religious-self",
                              "religious-peer",
                              "religious-dep"),
              boot = 500)
fit_tob2 = mma(svy_bc2[["tobacco"]],
               svy_a12,
               svy_a22,
               svy_a32,
               ind_effects = c("religious-self",
                               "religious-peer",
                               "religious-dep"),
               boot = 500)

## Rx
fit_rx = mma(svy_bc[["rx"]],
              svy_a1,
              svy_a2,
              svy_a3,
              ind_effects = c("religious-self",
                              "religious-peer",
                              "religious-dep"),
              boot = 500)
fit_rx2 = mma(svy_bc2[["rx"]],
               svy_a12,
               svy_a22,
               svy_a32,
               ind_effects = c("religious-self",
                               "religious-peer",
                               "religious-dep"),
               boot = 500)

## Marijuana
fit_mar = mma(svy_bc[["mari"]],
             svy_a1,
             svy_a2,
             svy_a3,
             ind_effects = c("religious-self",
                             "religious-peer",
                             "religious-dep"),
             boot = 500)
fit_mar2 = mma(svy_bc2[["mari"]],
              svy_a12,
              svy_a22,
              svy_a32,
              ind_effects = c("religious-self",
```

```
                               "religious-peer",
                               "religious-dep"),
              boot = 500)

## Illicit
fit_ill = mma(svy_bc[["illicit"]],
              svy_a1,
              svy_a2,
              svy_a3,
              ind_effects = c("religious-self",
                              "religious-peer",
                              "religious-dep"),
              boot = 500)
fit_ill2 = mma(svy_bc2[["illicit"]],
               svy_a12,
               svy_a22,
               svy_a32,
               ind_effects = c("religious-self",
                               "religious-peer",
                               "religious-dep"),
               boot = 500)

save(fit_tob, fit_tob2,
     fit_rx, fit_rx2,
     fit_mar, fit_mar2,
     fit_ill, fit_ill2,
     file = here("Data/NSDUH_2014_Results.rda"))
```

```
library(MarginalMediation)
library(tidyverse)
library(here)

load(file = here("Data/NSDUH_2014_Results.rda"))

## Extract direct effects
directs_fx = function(..., type){
  list(...) %>%
    map(~.x$dir_effects) %>%
    do.call("rbind", .) %>%
    data.frame(.) %>%
    select(Direct, Lower, Upper) %>%
    data.frame(., row.names =
                 gsub("religious", "Religiousity (Direct)", row.names(.))) %>%
    rownames_to_column() %>%
    mutate(Outcome = c(rep("Tobacco", 1), rep("Prescription", 1),
                       rep("Marijuana", 1), rep("Illicit", 1))) %>%
    select(Outcome, rowname, Direct, Lower, Upper) %>%
```

```
    set_names(c("Outcome", "Path", "Estimate", "Lower", "Upper")) %>%
    mutate(CI = paste0("(", round(Lower,4), ", ", round(Upper,4), ")")) %>%
    select(-CI) %>%
    mutate(type = type)
}
directs_un = directs_fx(fit_tob,
                        fit_rx,
                        fit_mar,
                        fit_ill,
                        type = "Unadjusted")
directs_adj = directs_fx(fit_tob2,
                         fit_rx2,
                         fit_mar2,
                         fit_ill2,
                         type = "Adjusted")

## Extract indirect effects and bind to directs
inds_fx = function(..., type){
  list(...) %>%
    map(~.x$ind_effects) %>%
    do.call("rbind", .) %>%
    data.frame(.) %>%
    select(Indirect, Lower, Upper) %>%
    data.frame(., row.names = gsub("religious-", "Religiousity Through",
                                   row.names(.))) %>%
    data.frame(., row.names = gsub("dep", "\nDepression",
                                   row.names(.))) %>%
    data.frame(., row.names = gsub("self", "\nRespondent Views",
                                   row.names(.))) %>%
    data.frame(., row.names = gsub("peer", "\nPeer Views",
                                   row.names(.))) %>%
    rownames_to_column() %>%
    mutate(Outcome = c(rep("Tobacco", 3),
                       rep("Prescription", 3),
                       rep("Marijuana", 3),
                       rep("Illicit", 3))) %>%
    select(Outcome, rowname, Indirect, Lower, Upper) %>%
    set_names(c("Outcome", "Path", "Estimate",
                "Lower", "Upper")) %>%
    mutate(CI = paste0("(", round(Lower,4), ", ",
                       round(Upper,4), ")")) %>%
    select(-CI) %>%
    mutate(type = type)
}
unadjusted = inds_fx(fit_tob,
                     fit_rx,
                     fit_mar,
```

```
                     fit_ill,
                     type = "Unadjusted") %>%
  rbind(directs_un)
adjusted = inds_fx(fit_tob2,
                   fit_rx2,
                   fit_mar2,
                   fit_ill2,
                   type = "Adjusted") %>%
  rbind(directs_adj)
inds = rbind(unadjusted, adjusted) %>%
  data.frame %>%
  mutate(type = factor(type,
                       levels = c("Unadjusted", "Adjusted"))) %>%
  mutate(Path = gsub("[0-9]","", Path)) %>%
  mutate(Outcome = factor(Outcome,
                          levels = c("Tobacco", "Prescription",
                                     "Marijuana", "Illicit")))
```

```
## Odds ratios and linear effects as done in Ford and Hill
## Sampling Design
library(survey)
design = svydesign(ids = ~1,
                   strata = ~vestr,
                   weights = ~analwt_c,
                   data = d1)

## All a Path Models
svy_a1 = svyglm(self ~ religious + age2 +
                  irsex + newrace2 + irfamin3 + parent,
                design = design)
svy_a2 = svyglm(peer ~ religious + age2 +
                  irsex + newrace2 + irfamin3 + parent,
                design = design)
svy_a3 = svyglm(dep ~ religious + age2 +
                  irsex + newrace2 + irfamin3 + parent,
                design = design,
                family = 'quasibinomial')

## path a
patha_fx = function(obj, row){
  cbind(coef(obj)[row],
          confint(obj)[row,1],
          confint(obj)[row,2])
}
est1 =
  rbind(
    patha_fx(svy_a1, "religious"),
```

```
    patha_fx(svy_a2, "religious"),
    patha_fx(svy_a3, "religious")
)
rownames(est1) = c("Respondent", "Peer", "Depression")
est1 = data.frame(est1) %>%
  set_names(c("Estimate", "Lower", "Upper"))

## All c and c' Path Models
svy_c = svy_c1 = list()
for (i in c("tobacco", "rx", "mari", "illicit")){

  model = as.formula(paste0(i, "~ religious + age2 +
                            irsex + newrace2 + irfamin3 + parent"))
  svy_c[[i]] = svyglm(model, design = design, family = "quasibinomial")

  model2 = as.formula(paste0(i, "~ self + peer + dep + religious + age2 +
                             irsex + newrace2 + irfamin3 + parent"))
  svy_c1[[i]] = svyglm(model2, design = design, family = "quasibinomial")

}

## Odds ratios of c and c' path models
pathc_fx = function(obj, row, drug){
  cbind(coef(obj[[drug]])[row],
            confint(obj[[drug]])[row,1],
            confint(obj[[drug]])[row,2])
}
est2 =
  rbind(
    cbind(pathc_fx(svy_c, "religious", "tobacco"),
          pathc_fx(svy_c1, "religious", "tobacco")),
    cbind(pathc_fx(svy_c, "religious", "rx"),
          pathc_fx(svy_c1, "religious", "rx")),
    cbind(pathc_fx(svy_c, "religious", "mari"),
          pathc_fx(svy_c1, "religious", "mari")),
    cbind(pathc_fx(svy_c, "religious", "illicit"),
          pathc_fx(svy_c1, "religious", "illicit"))
)
rownames(est2) = c("Tobacco", "Rx",
                   "Marijuana", "Illicit")
est2 = data.frame(est2) %>%
  set_names(c("c", "c_Lower", "c_Upper", "c1", "c1_Lower", "c1_Upper"))
```

Tables and Figures

Table 6.1 on page 66

```
## overall table1 (not adjusted for survey weights)
d1 %>%
  table1(factor(rx), factor(tobacco),
         factor(drink), factor(mari), factor(illicit),
         irsex, factor(ifelse(newrace2 == "(1) NonHisp White", 0, 1)),
         factor(
           ifelse(poverty2 == "(3) Income > 2X Fed Pov Thresh (See comment above)",
                  1, 0)),
         factor(dep), self, peer, parent, religious,
         type = c("simple", "condense"),
         var_names = c("Prescription", "Tobacco",
                       "Heavy Drinking", "Marijuana",
                       "Other Illicit",  "Sex",
                       "Race (Non-White)",
                       "Income (2x poverty)",
                       "Major Depression Episode",
                       "Respondent", "Peer",
                       "Parent", "Religiosity"),
         output = "latex2")

## Survey weighted
library(survey)
design = svydesign(ids = ~1,
                   strata = ~vestr,
                   weights = ~analwt_c,
                   data = d1)
svymean(~rx, design)
svymean(~tobacco, design)
svymean(~drink, design)
svymean(~mari, design)
svymean(~illicit, design)
svymean(~irsex, design)
svymean(~newrace2, design)
svymean(~poverty2, design)
svymean(~dep, design, na.rm=TRUE)
svymean(~self, design, na.rm=TRUE)
svymean(~peer, design, na.rm=TRUE)
svymean(~parent, design, na.rm=TRUE)
svymean(~religious, design, na.rm=TRUE)

## alpha of religiosity
with(d1,
  psych::alpha(cbind(scale(yerlgsvc),
                     scale(yerlgimp),
```

```
                      scale(yerldcsn),
                      scale(yefaiact))))

## alpha of respondent
with(d1,
  psych::alpha(cbind(yegpkcig, yegmjevr,
                     yegmjmo, yegaldly)))

## alpha of peer
with(d1,
  psych::alpha(cbind(yefpkcig, yefmjevr,
                     yefmjmo, yefaldly)))

## alpha of parent
with(d1,
  psych::alpha(cbind(yeppkcig, yepmjevr,
                     yepmjmo, yepaldly)))

## number of heavy drinking responses
sum(d1$drink)
```

Table 6.2 on page 68

```
perc_fx = function(obj){
  obj$ind_effects[,3]/(obj$dir_effects[,1] +
                         sum(obj$ind_effects[,3]))
}

percent_ind = cbind(perc_fx(fit_tob2),
                    perc_fx(fit_rx2),
                    perc_fx(fit_mar2),
                    perc_fx(fit_ill2)) %>%
  data.frame %>%
  set_names(c("Tobacco", "Prescription", "Marijuana", "Illicit")) %>%
  map_df(~.x*100) %>%
  mutate(Mediator = c("Respondent Views", "Peer Views", "Depression")) %>%
  select(Mediator, Tobacco, Prescription, Marijuana, Illicit)
percent_ind = percent_ind %>%
  rbind(data.frame(
    Mediator = "Total",
    Tobacco  = sum(percent_ind$Tobacco),
    Prescription = sum(percent_ind$Prescription),
    Marijuana = sum(percent_ind$Marijuana),
    Illicit = sum(percent_ind$Illicit)
  ))

library(xtable)
xtable(percent_ind, digits = 1)
```

Table ?? on page ??

```
est21 = est2 %>%
  rownames_to_column() %>%
  group_by(rowname) %>%
  summarize(perc = ((c - c1)/c)*100)
library(xtable)
xtable(est21, digits = 1) %>%
  print.xtable(include.rownames = FALSE)
```

Figure ?? on page ??

```
p = position_dodge(width = .2)
inds %>%
  filter(type == "Adjusted") %>%
  ggplot(aes(Path, Estimate, group = type, color = type)) +
    geom_hline(yintercept = 0, color = "darkgrey") +
    geom_point(position = p, alpha = .8) +
    geom_errorbar(aes(ymin = Lower, ymax = Upper),
                  position = p, alpha = .8,
                  width = .3) +
    facet_wrap(~Outcome) +
    coord_flip() +
    anteo::theme_anteo_wh() +
    theme(legend.position = "bottom",
          axis.line = element_line(color = "darkgrey"),
          panel.spacing = unit(.3, "in")) +
    scale_color_manual(values = c("chartreuse4", "coral2"),
                       guide = FALSE) +
    labs(x = "", y = "",
         color = "")
```

CURRICULUM VITA

Tyson S. Barrett

Address: 2810 Old Main Hill, Logan, UT 84322
Email: tyson.barrett@usu.edu
Website: tysonbarrett.com

EDUCATION

Ph.D. Quantitative Psychology | *Expected: Feb 2018*
Utah State University
Marginal Mediation Analysis: A New Framework for Interpretable Mediated Effects

B.S. Economics | *May 2014*
Utah State University
Cum Laude

B.S. Psychology | *May 2014*
Utah State University
Cum Laude

RESEARCH AREAS

Primary: Model Interpretability, Mediation Analysis, R Programming

Secondary: Health Behavior and Chronic Conditions, Machine Learning

RESEARCH EXPERIENCE

Research Assistant, Prevention Science Lab, Utah State University 2016-Present

Advisor: Ginger Lockhart, PhD

Statistical and Data Science Consultant, Utah State University 2016-Present

Advisors: Sarah Schwartz, PhD and Jamison Fargo, PhD

WOC Research Assistant, Veteran Affairs SLC 2016-Present

Advisors: Vanessa Stevens, PhD and Richard Nelson, PhD

Research Assistant, NCHAM, Utah State University 2016-Present

Advisor: Karl White, PhD

PUBLISHED WORKS IN REFEREED JOURNALS

* Denotes that I was the project methodologist or data scientist

Barrett, T.S. & Brignone, E. (2017). Furniture for Quantitative Scientists. *R Journal, 9*(2), 142-146.

*Borrie, S.A., Lansford, K.L., & **Barrett, T.S.** (2017). Generalized Adaptation to Dysarthric Speech. *Journal of Speech, Language, and Hearing Research, 60*, 3110-3117. doi: 10.1044/2017_JSLHR-S-17-0127.

Barrett, T.S. & White, K.R. (2017). Trends in Hearing Loss Among Adolescents. *Pediatrics, 140*(6), e20170619. doi: 10.1542/peds.2017-0619.

Fargo, J.D., Brignone, E., Metraux, S., Peterson, R., Carter, M.E., **Barrett, T.**, Palmer, M., Redd, A., Samore, M.H., & Gundlapalli, A.V. (2017). Homelessness Following Disability-Related Discharges from Active Duty Military Service in Afghanistan and Iraq. *Disability and Health Journal, 10*, 592-599. doi: 10.1016/j.dhjo.2017.03.003.

*Brown, L.T., Mohr, K.A.J., Wilcox, B.R., & **Barrett, T.S.** (2017). The effects of dyad reading and text difficulty on third-graders' reading achievement. *Journal of Educational Research.* doi: 10.1080/00220671.2017.1310711

Brignone, E., Gundlapalli, A.V., Blais, R.K., Kimerling, R., **Barrett, T.S.**, Nelson, R.E., Carter, M.E., Samore, M.A., & Fargo, J.D. (2017). Increased Health Care Utilization and Costs among Veterans with a Positive Screen for Military Sexual Trauma. *Medical Care, 55*, S70-S77. doi: 10.1097/MLR.0000000000000767

Long, E., **Barrett, T.S.**, & Lockhart, G. (2017). Network-behavior dynamics of adolescent friendships, alcohol use, and physical activity. *Health Psychology, 36*(6), 577-586. doi: 10.1037/hea0000483

*Borrie, S.A., Lansford, K.L., & **Barrett, T.S.** (2017). Rhythm perception and its role in perception and learning of dysrhythmic speech. *Journal of Speech, Language, and Hearing Research, 60*(3), 1-10. doi: 10.1044/2016_JSLHR-S-16-0094

Behl, D., White, K., **Barrett, T.**, Blaiser, K., Callow-Heusser, C., & Croshaw, T. (2017). A Multisite Study Evaluating the Benefits of Early Intervention via Telepractice. *Infant and Young Children, 30*(2), 147-161.

Barrett, T.S. & White, K. R. (2016). Prevalence and Trends of Childhood Hearing Loss Based on Federally-funded National Surveys: 1994–2013. *Journal of Early Hearing Detection and Intervention, 1*(2), 8-16.

Doutré, S. M., **Barrett, T.S.**, Greenlee, J. & White, K. R. (2016). Losing Ground: Awareness of Congenital Cytomegalovirus in the United States. *Journal of Early Hearing Detection and Intervention, 1*(2), 39-48.

Munoz, K., Rusk, S.E.P., Nelson, L., Preston, E., White, K.R., **Barrett, T.S.** & Twohig, M.P. (2016). Pediatric Hearing Aid Management: Parent Reported Needs for Learning Support. *Ear and Hearing Journal. 37*(6), 703-709.

Borgogna, N., Lockhart, G., Grenard, J, **Barrett, T.**, Shiffman, S. & Reynolds, K. (2015). Ecological Momentary Assessment of Urban Adolescents' Technology Use and Cravings for Unhealthy Snacks and Drinks: Differences by Ethnicity and Sex. *Journal of the Academy of Nutrition and Dietetics, 115*(5), 759-766.

Under Review/In Revision

Barrett, T.S. & Lockhart, G. Efficient exploration of many variables and interactions using regularized regression. *Prevention Science.* In Revision.

*Leopold, S., Healy, E. W., Youngdahl, C., **Barrett, T.S.**, & Apoux, F. Speech-material and talker effects in speech band importance. *Journal of the Acoustical Society of America.* In Revision.

OTHER PUBLICATIONS

R for the Health, Behavioral, and Social Sciences — A primer on using R for researchers in health, behavioral, and social sciences including importing data, data manipulation, data modeling, and data visualization. Currently published online at tysonstanley.github.io/Rstats/.

The furniture R package (Published on CRAN and GitHub) — Contains functions to help with several aspects of research in the health, behavioral, and social sciences. The main functions—`table1()` and `tableC()`—produce descriptive statistics and correlations in well-formatted tables (as commonly seen as the "Table 1" of academic journals). Over 8,000 downloads (see tysonbarrett.com/furniture).

The MarginalMediation R package (Published on GitHub) — Contains functions to perform Marginal Mediation. Papers discussing the method and software are forthcoming (see tysonbarrett.com/MarginalMediation).

The anteo R package (Published on GitHub) — Contains functions to help in interpreting machine learning models. Still under active development.

SELECTED PRESENTATIONS

Barrett, T.S., & Lockhart, G. (2017). Enhancing the Exploration and

Communication of Big Data in Prevention Science. Poster presented at the Annual Meeting of the Society of Prevention Research, Washington, DC. *Received "Distinguished Poster Award" and "Abstract of Distinction."*

Barrett, T.S., & Lockhart, G. (2017). Exploring the Predictors of Marijuana Use Among Adolescents with Asthma. Oral presentation at the Utah State University Research Symposium, Logan, UT.

Sanghavi, K., White, K., **Barrett, T.S.**, Wylie, A., Raspa, M., Cashman, D., Vogel, B. Caggana, M. & Bodurtha, J. (2017). Poster presented at the Early Hearing Detection and Intervention Conference, Atlanta, GA. *Received "Outstanding Poster Award."*

Brignone, E., Gundlapalli, A.V., **Barrett, T.S.**, Blais, R.K., Nelson, R.E., Carter, M.E., Kimerling, R., Samore, M.H., Fargo, J.D. (2016). Cost of Care among Male and Female Veterans with a Positive Screen for Military Sexual Trauma. Poster presented at the 2016 Annual Meeting of the International Conference of Psychology, Yokohama, Japan.

Barrett, T.S., Munoz, K. & White, K. (2016). How well do parent report hearing loss in their children? Poster presented at the Early Hearing Detection and Intervention Conference, San Diego, CA.

Barrett, T.S., Munoz, K. & White, K. (2016). Accounting for Temporary Loss in National Studies on Hearing Loss. Poster presented at the Early Hearing Detection and Intervention Conference, San Diego, CA.

Barrett, T.S., Munoz, K. & White, K. (2016). An Evaluation of Early Intervention delivered via Video Conferencing. Poster presented at the Early Hearing Detection and Intervention Conference, San Diego, CA.

Stevens, V., **Barrett, T.S.** & Nelson, R. (2016). Distribution and Daily Cost of Care in a Pediatric Hospital. Oral presentation to the Pediatric Guidance Council of Intermountain Healthcare, Salt Lake City, UT.

Barrett, T.S., Munoz, K. & White, K. (2015). Refinements to estimating prevalence of hearing loss in children. Poster presented at the Utah State University Research Symposium, Logan, UT.

Barrett, T.S., Prante, M., Peterson, R., Fargo, J.D., Pyle, N. (2014). Predictors of employability among homeless youth. Poster presented at the Psi-Chi Undergraduate Research Conference at Idaho State University, Pocatello, ID. *Best Undergraduate Poster Presentation Award.*

Barrett, T.S., Holland, D. (2014). Nascent Entrepreneurship, Impulsivity, and Self- Efficacy. Poster presented at the Research on Capitol Hill, Salt Lake City, UT.

Holland, D., **Barrett, T.S.** (2013). Impulsivity in young entrepreneurs. Round table discussion at the Babson Business Conference, Paris, France.

AWARDS

- Abstract of Distinction (Annual Meeting of the Society of Prevention Research)
- Distinguished Poster Award (Annual Meeting of the Society of Prevention Research)
- Outstanding Poster Award (Annual Meeting of EHDI)
- Dean's Scholarship (full tuition for two years)
- Best Poster Presentation (Psi-Chi Undergraduate Conference)

TEACHING INTERESTS

Primary Interest: Quantitative Methods

- Undergraduate and Graduate Statistics (ANOVA and Regression [OLS, GLM])
- Multilevel Modeling (Hierarchical Linear Modeling, GEE, Mixed Effects)
- Mediation Analysis (Marginal Mediation, Moderated Mediation)
- Reproducible Research (Research Methods, Open Science Framework)
- R for the Social Sciences (Undergraduate and Graduate Level)
- Exploratory Data Analysis
- Structural Equation Modeling (Psychometrics and Measurement Models, Mixture Modeling)
- Research Methods (Undergraduate and Graduate Level)

Secondary Interest: Public Health

- Research in Public Health
- Disabilities (Hearing Loss, Developmental Disabilities)

TEACHING EXPERIENCE

Instructor

- **R for the Health, Behavioral, Educational, and Social Sciences I and II**
  - 2016 - Present
  - Created, Developed, and Taught
  - Graduate Level
  - Most Recent Student-Responded Ratings:
    - Overall: 4.5 out of 5.0
    - Teaching: 4.9 out of 5.0
    - Course: 4.8 out of 5.0

Teaching Assistant

- **Econometrics I** (Graduate Level)
  - Fall 2016
- **Psychological Statistics** (Undergraduate Level)
  - 2014 – 2015

METHODOLOGICAL TRAINING

- Regression, Generalized Linear Models
- Mixed Effects, Generalized Linear Models
- Machine Learning
  - CART
  - Random Forest
  - Regularized Regression
  - Boosting/Bagging
  - Cross-Validation
- Social Network Analysis
- Data Visualization
  - Static and Dynamic Visuals

SOFTWARE AND PROGRAMMING EXPERIENCE

1. **R and RMarkdown: Expert (4 years of daily use)**
   - Database Queries, Creation, and Management
   - Data Analytics
   - Website Creation
   - Program Development and Deployment
2. **REDCap: Moderate Experience (1 year of weekly use)**
   - Database Creation

  - Data Entry and Survey Creation and Deployment
  - Data Management via REDCap API
3. **SQL: Some Experience (4 years of occasional use)**
  - Database Queries and Database Management Python: Minor Experience (1 year of occasional use)
  - Data Analytics