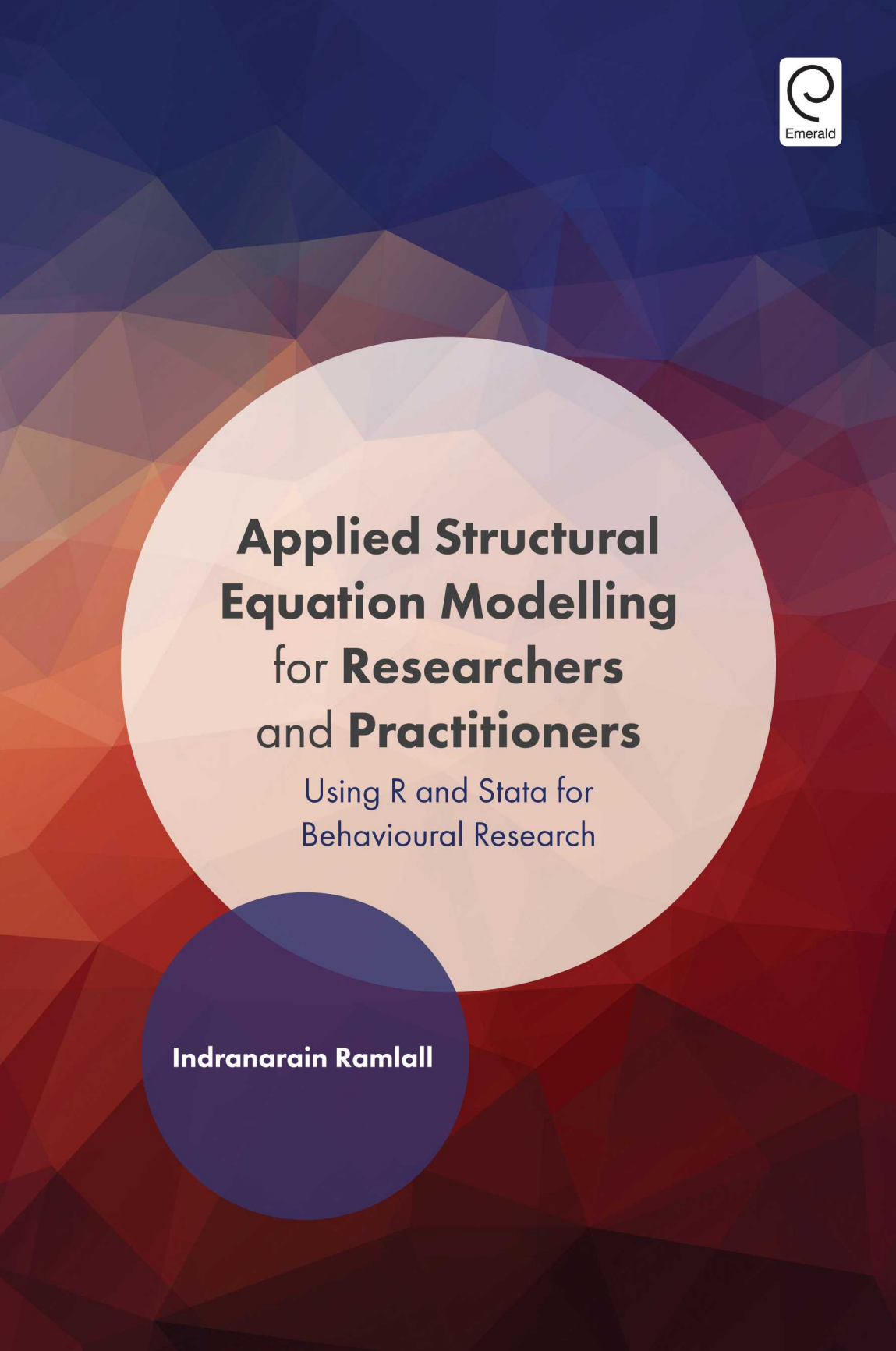

Emerald
Applied Structural Equation Modelling for Researchers and Practitioners
Using R and Stata for Behavioural Research
Indranarain Ramlall

APPLIED STRUCTURAL EQUATION MODELLING FOR RESEARCHERS AND PRACTITIONERS

Using R and Stata for Behavioural Research

APPLIED STRUCTURAL EQUATION MODELLING FOR RESEARCHERS AND PRACTITIONERS

Using R and Stata for Behavioural Research

BY
INDRANARAIN RAMLALL
University of Mauritius, Mauritius

United Kingdom – North America – Japan
India – Malaysia – China

Emerald Group Publishing Limited
Howard House, Wagon Lane, Bingley BD16 1WA, UK

First edition 2017

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN: 978-1-78635-883-7 (Print)
ISBN: 978-1-78635-882-0 (Online)

Dedicated to my parents and GOD.

Contents

Preface

Structural equation models permeate every field of research in the world. Indeed, despite its deep-rooted origins in psychology, structural equation models gained considerable attention in different fields of study such as biology, engineering, environment, education, economics and finance. The main power ingrained in these types of models pertains to their ability to cater for various levels of interactions among variables such as bi-directional causality effects and, most importantly, catering for the effects of unobserved or latent variables.

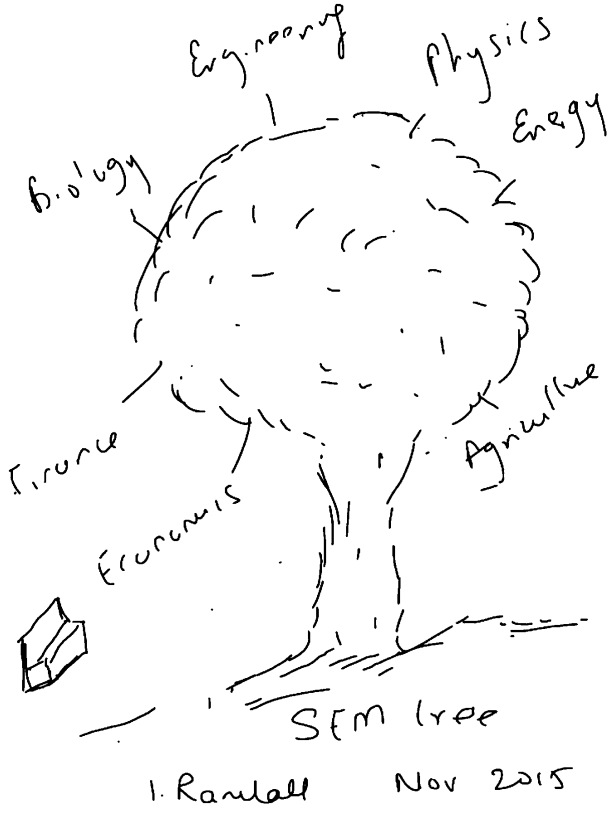
As at date, there exist some well-developed textbooks on structural equation models. However, most of them tend to address the subject mainly in a manner which may not really befit the needs of researchers who are new in this area. This is the main aim of this book, that is to explain, in a rigorous, concise and practical manner all the vital components embedded in structural equation modelling. The way the book is structured is to unleash all the vital elements in a smooth and quick to learn approach for the inquisitive readers. In essence, this book substantially leverages the learning curve for novice researchers in the area of structural equation models. Overall, this book is meant for addressing the needs of researchers who are found at the beginning or intermediate level of structural equation modelling learning curve.

LISREL and AMOS are now deemed as the workhorse for implementing structural equation models. Consequently, the book clings to two different software, namely R (a freeware) and STATA. R is used to explain the model in its lavaan package without going into too much sophistication. STATA implementation of

structural equation model is also explained. In fact, STATA 13 is now upgraded with enough power to implement structural equation models without being subject to much ado with respect to programming problems which usually characterize LISREL. In a nutshell, STATA 13 is powerful enough to perform different types of structural equation models.

This book can be used at graduate level for a one semester course on structural equation modelling. The way the book has been written is highly convenient as a self-learning tool to any interested reader.

I hope this book to be particularly useful for all researchers who are new on the path of structural equation modelling.



1 Definition of SEM

1.1. Introduction

Known as causal models with a conspicuous presence in the field of consumer psychology, structural equation model (SEM) allows complex modelling of correlated multivariate data in order to sieve out their interrelationships among observed and latent variables. SEM constitutes a flexible and comprehensive methodology for representing, estimating and testing a theoretical model with the objective of explaining as much of their variance as possible. In simple terms, SEM is nothing more than an analysis of the covariance structure. SEM incorporates various statistical models such as regression analysis, factor analysis and variance/covariance analysis. Under SEM, a clear demarcation line is established between observed and latent variables. SEM can handle complex relationships as it can simultaneously factor in a measurement equation and a structural equation. Moreover, SEM represents a large sample technique, widely known under its rule of thumb, that is to have at least 10 observations per variable. SEM represents a vital multivariate data analysis technique, widely employed to answer distinct types of research questions in statistical analysis. Other names are associated with SEM such as simultaneous equation modelling, path analysis, latent variable analysis and confirmatory factor analysis. Technically speaking, SEM can be defined as a combination of two types of statistical technique, namely, factor analysis and simultaneous

equation models. SEM is so much coveted by researchers that there is even a journal in this area, namely, *Journal of Structural Equation Modelling*. In a nutshell, SEM can best be described as a powerful multivariate tool to study interrelationships among both observed and latent variables.

SEM = 2 types of statistical techniques:
Confirmatory Factor Analysis + Simultaneous Equation Models

While the objective of the measurement model is to relate the latent variables to the observed variables, the aim under structural equation focuses on the relationship between dependent and independent latent variables, the effects of explanatory (independent) latent variables on dependent (outcome) latent variables. Alternatively stated, the need to sieve observed variables to capture latent variables is effected in the measurement equation. In that respect, the measurement equation in SEM constitutes an exploratory tool, more specifically, a confirmatory factor analysis tool. Under SEM, there is the need to compare several structural equations via model comparison statistics to sieve out the most appropriate model. *Latent variables are also known as unobserved variables, intangible variables, 'directly unmeasured' variables, unknown variables or simply constructs, whereas manifest variables are called as observed variables, tangible variables, indicator variables or known variables.*

Latent variables are inherently linked to observed variables as they can only be captured by observed variables or indicators. Latent variables are inferred from observed variables. This can be best explained when it comes to LISREL software application whereby observed variables are input in such a way as to respect the order of the data input, while for latent variables, they can be defined in any order. Moreover, ellipses or circles are associated with latent variables while rectangles are inherently associated with the observed variables. Latent variables can be dependent or

independent variables. As a matter of fact, SEM comprises observed and latent variables, whether dependent or independent. Overall, SEM includes observed variables, latent variables and measurement error terms. Pure latent variables are those which are uncontaminated by measurement error. *Second-order latent variables are functions of other latent variables while first-order latent variables do not depend on any other latent variables*.

Examples of latent variables are intelligence, market psychology, achievement level and economic confidence. Examples of observed variables are economic performance, scores obtained and number of items sold.

Various versions of SEM prevail with the most basic version being the linear SEM. Other types of SEM consist of Bayesian SEM, non-linear SEM and hierarchical SEM. The main ingredient used in SEM application pertains to the covariance or the correlation matrices. It is of paramount significance to gain a proper insight into their inherent difference. In essence, covariance constitutes an unstandardised[1] form of correlation. SEM constitutes an exploratory factor analysis tool for gauging on the factors that underlie a set of variables. It assesses which items should be grouped together to form a scale. While exploratory factor analysis allows all the loadings to freely vary, confirmatory factor analysis constrains certain loadings to be zero. SEM is widely preferred to regression analysis by virtue of its powerful distinctive features such as the ability to incorporate multiple independent and dependent variables, inclusion of latent constructs and measurement errors being duly recognised. SEM is widely applied in non-experimental data (Figure 1.1).

1. Under STATA SEM estimation, unstandardized estimates pertain to covariance and standardized estimates for correlation coefficient.