Hengky Latan · Richard Noonan *Editors*

Partial Least Squares Path Modeling

Basic Concepts, Methodological Issues and Applications

Springer

Partial Least Squares Path Modeling

Hengky Latan • Richard Noonan
Editors

Partial Least Squares Path Modeling

Basic Concepts, Methodological Issues and Applications

Springer

Editors
Hengky Latan
Department of Accounting
STIE Bank BPD Jateng and Petra
Christian University
Semarang-Surabaya, Indonesia

Richard Noonan
Institute of International Education
Stockholm University
Stockholm, Sweden

ISBN 978-3-319-64068-6 ISBN 978-3-319-64069-3 (eBook)
DOI 10.1007/978-3-319-64069-3

Library of Congress Control Number: 2017955377

Mathematics Subject Classification (2010): 62H20, 62H25, 62H12, 62F12, 62F03, 62F40, 65C05, 62H30

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

We dedicate this book to the fond memory of Herman O. A. Wold (25 December 1908–16 February 1992), whose energy and creativity were an inspiration to all of us who had the good fortune to work with him. His pioneering work lives on and forms the base on which this book rests.

Foreword

I am immensely honored to have been asked to provide a foreword to this volume.

An aspiring high-school mathematics teacher, young Karl Jöreskog, is lured into graduate study by celebrated econometrician Herman Wold. In the course of his studies, Jöreskog turns from the mathematics that he loves to statistics. Under Wold's direction and encouragement, Jöreskog derives a maximum likelihood (ML) approach to estimating confirmatory factor analysis models which yields an inferential χ^2 distributed test statistic. Wold, a long-time advocate for least squares estimation methods, invents partial least squares (PLS) path modeling as a tool for approximating Jöreskog's results but without the heavy distributional, computing power and prior knowledge demands of ML estimation. From this origin story, some have concluded that PLS path modeling is inherently a lesser method, and now an anachronism which serves little purpose, given the cheap availability of computing resources and the development of factor analysis estimation methods that are increasingly robust to nonnormal distributions and that address an ever-expanding assortment of complex research situations. Simulation research, using populations defined by factor models, reaches the unsurprising conclusion that the factor-based approach to SEM performed better than composite-based approaches such as PLS path modeling.

This general theme, describing factor-based methods as "the real thing" and belittling composite-based alternatives as cheap imitations, runs through a great deal of the methods literature across the social sciences, and no wonder. In the earliest years of the twentieth century, the pioneering works of Charles Spearman had already bound together psychological measurement and factor analysis, making any other analytical method seem deficient. But this identification of "factor analysis" with "measurement" itself hinges on an anachronistic (and now somewhat quaint) philosophy of science. Spearman the empiricist argued that the common factor he extracted or fabricated from data, a common factor which Spearman labeled "general intelligence" or *g*, was in fact general intelligence itself. Intelligence, indeed, could be nothing else except the common factor resulting from this analysis of error-prone data, because the realm of legitimate scientific inquiry begins and ends with observable data. In philosophy of science circles, empiricism has long ago broadly given way to scientific realism, a perspective that takes unobservable conceptual variables—attributes like intelligence, customer satisfaction, or attitude—to be real entities with their own existence independent of data and statistical models. From a realist perspective, both the common factors in factor-based models and the composites in PLS path modeling are only proxies or empirical substitutes for the actual psychological attributes. Making inferences about the actual psychological attribute on the basis of a statistical model then requires that the researcher establish the validity of each proxy. The mathematics of the indeterminacy of factors and the unreliability of composites give reason enough for researchers to be cautious about the quality of both kinds of proxies.

I think it is very important to have alternatives. When there is only one way to do something, there is a tendency to just accept the limitations that come with that single path. It can be hard to even imagine a better way, even if the one available approach is actually rather weak. If there are alternatives, on the other hand, it can be easier to recognize the shortcoming of any one method by comparing it with the others. Moreover, if there are alternatives, then it may be possible to use the strengths of one method to offset or bypass the weaknesses of another method. For example, it was difficult to obtain ML estimates of the "interbattery factor model" until Michael Browne showed how to obtain them by transforming parameter estimates from canonical correlation, a composite-based method. More recently, Theo Dijkstra and colleagues have obtained consistent estimates of factor model parameters as a transformation of PLS path modeling parameter estimates, suggesting the possibility of combining factor-based and composite-based approaches within the same structural equation model.

We need to bring the composite-based approaches to SEM and the factor-based approach into the kind of relationship that can enable a true cross-pollination. Viewed as peers—mathematically different tools for accomplishing the same very challenging task of learning about the behavior of unobserved conceptual variables—the methods will be able to borrow from one another and to be inspired by one another, just as Wold and Jöreskog inspired each other.

In order to be a genuine participant in such a relationship, the PLS path modeling methodology must continue to grow, evolve, and mature. If PLS path modeling seems either moribund or stuck in the past, outside researchers will not expect to find new insights there and may not bother to look. Unfortunately, PLS path modeling endured a period of years which saw very little growth, even as the factor-based approach to SEM raced ahead. So it is doubly important that new research and advanced applications in this area be strongly encouraged.

These days, as it happens, it is hard to keep up with the pace of developments in PLS path modeling. Many exceptional packages are available as either commercial or open-source software, enabling researchers to learn by doing and to actively confront the limits of the known. And as researchers have become more familiar with these tools, they have been driven by need, by curiosity, and by competition to tackle new challenges.

For example, one of the major challenges for SEM is dealing with heterogeneity—different parameter values for different respondents—in all its many aspects. Researchers encounter heterogeneity as differences between known groups, as interaction or moderation effects, as the result of clustering of observations, and as mixtures of distributions within the same population. One by one, methods pioneers have stepped up and provided PLS path modeling procedures to address these issues, while still seeking to minimize distributional assumptions and thus honor the original spirit of Herman Wold's method.

Not every problem has a statistical solution. At the birth of PLS path modeling, one of the virtues claimed for PLS path modeling was its small sample size capabilities. In contrast to the large sample sizes required for maximum likelihood factor analysis, it was noted that PLS path modeling could yield parameter estimates and (jackknife or bootstrap) standard errors even when sample size was very small. This confidence in the small-sample performance of PLS path modeling drew upon a misunderstanding. Yes, PLS path modeling algorithms will function—will yield results and not simply quit or crash—even when sample size is quite small, but the quality and usefulness of such results will be poor. Simulation research, drawing data from correctly specified composite-based populations, has shown that researchers with small sample sizes are likely to be as well or better served by creating simple unit-weight composites, rather than seeking optimized weights through PLS path modeling. Bias in PLS path modeling parameter estimates can itself be a function of sample size, with bias shrinking toward 0 as sample size increases. No statistical method can turn a small amount of information into a large amount.

Undoubtedly, the different approaches to SEM—factor-based and composite-based—are encumbered with more than their share of controversy. Differing viewpoints are deeply entrenched within the factor method and PLS path modeling communities, and their roots run far back in the history of quantitative analysis. Still, I think the recently rising volume on this controversy could be a very positive sign for PLS path modeling. On the one hand, advocates for the factor-based approach to SEM find in PLS path modeling something worth opposing. That demonstrates the collective strength of the PLS path modeling community, even as it means that some users will face additional opposition when sharing their work in some academic journals. At the same time, old hands within the PLS path modeling community may wonder if they still recognize the method that they once knew. Unless PLS path modeling as a method is already perfect and complete, there must be a continuing evolution. And there is—a ferment, a liveliness, and a dissatisfaction with the status quo. That is the mark of a strong method, with both the opportunity and the energy to keep growing.

In June 2015, I was privileged to participate in a conference at the University of Sevilla in Spain. The conference ended with a panel session, and the session ended with Christian Ringle asking all the panelists whether they thought the future of PLS path modeling was bright—whether the future would see PLS path modeling widely accepted and embraced. At that moment, I must admit, I was full of doubt. The development of quantitative "measurement" procedures in the social sciences was a key step in those fields being recognized as true sciences, and factor analysis is still the core of psychometrics. So anything perceived as a threat to the dominance of the factor-based approach to SEM can still be perceived as an existential threat to an entire research establishment and something that must be fiercely resisted.

Today, however, I have reason to be more optimistic. First, some of the flawed ideas that have come down to us from the dawn of PLS path modeling are being left behind. Some of these ideas were easy targets for critics of PLS path modeling, allowing those critics to avoid dealing with the substance of the methodology. Leaving those notions behind will mean that critics will face a method with stronger foundations. Second, there is the diversity of contributions to the corpus of research, counting not only studies of PLS path modeling itself, coming from so many disciplines and so many parts of the world, but also work on related composite-based methods like generalized structured component analysis (GSCA) and regularized generalized canonical correlation analysis (RGCCA). This diversity only enriches the pool of ideas from which researchers can draw to build a better PLS path modeling. Given all of that, and given the usefulness of PLS path modeling in helping both researchers and practitioners to address the real-world problems confronting our planet, I can't help but see a bright future for composite-based approaches to SEM, including PLS path modeling.

Of course, the further development and growth of PLS path modeling is itself only a step toward a larger goal. The world faces a diversity of real problems. Many of these problems seem to involve causal elements that are best framed as unobserved conceptual variables. Structural equation modeling broadly offers an exceptional family of tools for addressing these problems. SEM as a family is stronger when all the members of that family are strong and robust, ready to challenge each other and to contribute ideas to the further development of all methods. Researchers facing the world's real problems need the best possible portfolio of tools to help them address these issues, to help them to make the world a better place. A stronger and more vibrant PLS path modeling makes for a stronger structural equation modeling, which in turn makes for a more hopeful tomorrow around the world.

J. Mack Robinson College of Business
Georgia State University, USA

Edward E. Rigdon

Editor's Preface

Partial least squares-path modeling (PLS-PM) is a multivariate statistical technique first introduced by Herman Wold in the late 1960s. The period from that time until the late 1980s can be seen as the "gestation period," which was followed by continued development and especially rapid development in the past decade. This can be seen by the increasing use of PLS-PM today for research in various fields, including accounting, business ethics, education, family business, information systems, international business, marketing, operations management, strategy management, sustainability, and tourism. Despite the rapid pace of development, PLS-PM has not escaped controversies. Some of the criticisms, such as its use with small sample sizes, the absence of a goodness of fit indices, bias in parameter estimation, and the problem of the measurement model, have spawned two groups debating among scientists, academics, and practitioners. The final conclusion to be addressed is whether the use of PLS-PM should be continued or abandoned.

In the past 3 years, loyal supporters of the PLS-PM have demonstrated that this method is feasible to maintain and continue to develop. New breakthroughs ranging from consistent partial least squares (PLSc), goodness of fit indices (i.e., SRMR and NFI), heterotrait–monotrait (HTMT) ratio, measurement invariance of composite (MICOM), PLS predict, importance-performance map analysis (IPMA), and new heterogeneity methods have led to a lively discussion of the emancipation of the PLS-PM method. On the other hand, the outdated discourse has now been left behind to build solid inferential statistics from the PLS-PM method. There is a glimmer of hope and an intersection between the two debating groups now, where consent and conclusion can be built. Thus, PLS-PM is no longer an alternative to covariance-based SEM but has transformed into a stand-alone method capable of solving real-world problems.

The purpose of this book is to introduce recent developments and techniques in the PLS-PM field. This book as a whole discusses the recent developments of the PLS-PM method and provides a comprehensive overview of the current state of the most advanced research related to PLS-PM. By focusing primarily on the advance of each PLS-PM technique, with example cases and situations, we hope that the chapters are both enlightening and instructional. Each chapter assumes that the reader has already mastered the equivalent of a multivariate statistics course that included coverage of most basic PLS-PM techniques.

Each chapter in this book contains an up-to-date description of a recent development or technique in PLS-PM and is often written by the author(s) who originally proposed or contributed substantially to its development. Each chapter also provides complete references to the pertinent literature on the topic. The decision regarding the selection and the organization of the chapters for this book was quite challenging. Obviously, within a book only a limited number of topics could be addressed. In the end, the choice of the material was governed by our own beliefs concerning the most important new developments within the PLS-PM field.

The book is divided into three main sections. The first section consists of six chapters emphasizing the basic concepts and extensions of the PLS-PM method. In Chap. 1, Noonan deals with the early history of the PLS-PM and some of the analytical context at the time. This chapter tells about some personal experiences and some events that are considered important in the journey of the PLS-PM method. It also discusses some important findings that occurred during that period. In Chap. 2, Henseler, Hubona, and Ray discuss the modern view of the PLS-PM method and provide rules of thumb covering several aspects and guidelines for the use of PLS-PM today. In this chapter, the readers can find new guidelines that can be useful in their research project.

In Chap. 3, Kock discusses the extension of the PLS-PM method into a factor-based SEM by introducing a new algorithm. In this chapter, PLS-PM is used for a common factor model using simulation. This chapter also provides practical steps to change the PLS-PM algorithm for factor-based SEM. In Chap. 4, Dijkstra develops a model that fully honors Wold's fundamental principle of soft modeling in terms of observable variables only. In this context, the principle states that "all information between the blocks of indicators is conveyed solely by observable composites." The latter may satisfy an interdependent, non-recursive system of linear equations. Now the model and mode B "like hand in glove," to use one of Wold's favorite phrases, as opposed to the latent variables model, where mode B can by necessity only yield an approximation. Here, it is an eminently natural approach, attaining consistency without invoking "consistency at large." In Chap. 5, Davino, Dolce, and Taralli present a new approach called the quantile composite-based path modeling (QC-PM). This approach can be used to explore the whole dependence structure and to highlight whether and how the relationships among variables (both observed and unobserved) change across quantiles. The approach is described from the point of view of methodology and application. A preliminary approach to handle observed heterogeneity is also provided. In Chap. 6, Schuberth and Cantaluppi present a new approach OrdPLSc combining ordinal PLS and consistent PLS in order to appropriately deal with ordinal categorical variables in the framework of PLS-PM. In doing so, OrdPLSc overcomes drawbacks from earlier approaches.

The second section of this book discusses the methodological issues that are the focus of the recent development of the PLS-PM method. This section consists of seven chapters. In Chap. 7, Dolce, Esposito Vinzi, and Lauro discuss predictive modeling with PLS-PM by introducing a new approach called nonsymmetrical composite-based path modeling (NSCPM). The authors demonstrate the capability of PLS-PM in predictive-orientated contexts and offer measures and evaluation criteria. This chapter tries to help readers understand how to evaluate the predictive performance of component-based path models. In Chap. 8, Cepeda-Carrión, Nitzl, and Roldán discuss the modern view for mediation analysis in PLS-PM. The authors offer a new approach in examining the mediation effects in PLS-PM. Guidelines for testing the effects of mediation are also available. Several examples are provided for simple and complex mediation models for real-world cases. In Chap. 9, Sarstedt, Ringle, and Hair propose a multi-method approach for identifying and treating unobserved heterogeneity. Bridging prior latent class methods in PLS-PM, their guideline contains many rules of thumb that researchers will find useful when analyzing their data.

In Chap. 10, Matthews deals with multigroup analysis (PLS-MGA) by providing examples of real-world applications. The author explains the PLS-MGA analysis step by step. This chapter also provides some practical guidelines for PLS-PM users in running PLS-MGA. In Chap. 11, Kock explains the common method bias and how to solve it. The full-collinearity approach is used to handle it. In Chap. 12, Petrarca, Russolillo, and Trinchera introduce a new algorithm to handle nonmetric variables in PLS-PM. This chapter shows how to include nonmetric variables (i.e., ordinal and/or nominal) in a PLS path model. It also discusses nonmetric PLS approach for handling these types of variables and how to integrate the logistic regression into the PLS path model for predicting binary outcomes via an application on real data. In Chap. 13, Sharma, Pohlig, and Kim evaluate the efficiency and accuracy of bootstrap parameter recovery in PLS-PM, CB-SEM, and the Bollen-Stine methods under various conditions of measurement and structural misspecifications, effect sizes, sample sizes, and data distributions. Model misspecifications are especially likely to arise in exploratory research where theories are uncertain and evolving. Their results suggest that PLS-PM is the method of choice in exploratory modeling when structural parameters are of interest, while CB-SEM and Bollen-Stine methods are favorable when measurement model parameters are the focus. They recommend a two-pronged strategy that appropriately utilizes the relative strengths of the two techniques when theoretical uncertainty exists at both measurement and structural levels.

The third part of this book discusses the real-world application of the PLS-PM method in various disciplines. This section consists of four chapters. In Chap. 14, Falk provides examples of applications in the field of psychology. In Chap. 15, Latan, Chiappetta Jabbour, and Lopes de Sousa Jabbour provide examples of applications from PLSc, PLS-MGA (MICOM), IPMA, and mediation-moderation analyses in the field of business ethics. In Chap. 16, Geladi, Grahn, and Esbensen provide examples of PLS applications in hyperspectral imaging where huge data sets can be collected. This requires special thinking about how dependent and independent variables are created and used for making a PLS model but also for dealing with the concept of replicate. One of the two examples shown uses PLS discriminant analysis, which is a very useful tool for handling a large amount of pixels in an image. Finally, in Chap. 17, Streuknes, Leroi-Werelds, and Willems discuss the application of IPMA and various usage guidelines for nonlinearity relationships.

This book could not have been completed without the assistance and support provided by many individuals. First, we would like to thank all the contributors for their time and effort in preparing chapters for this book. We would like to thank the referees who review each chapter on this book in the first and second rounds. We are also greatly indebted to Springer Publisher, which has been willing to publish this book. Thanks are also due to all the wonderful people on the editorial staff at Springer for their assistance and support in putting together this volume. Finally, we thank our families for their love and for continually enduring a seemingly endless list of projects.

Semarang-Surabaya, Indonesia — Hengky Latan
Stockholm, Sweden — Richard Noonan

Contents

Part I Basic Concepts and Extensions

1 Partial Least Squares: The Gestation Period 3
Richard Noonan
1.1 Introduction 3
1.2 A Prehistory: Large-Scale School Surveys 4
1.3 PLS Comes of Age 8
1.4 Some Concluding Personal Comments 15
1.4.1 Notes on Commonality Analysis 16
References 17

2 Partial Least Squares Path Modeling: Updated Guidelines 19
Jörg Henseler, Geoffrey Hubona, and Pauline Ash Ray
2.1 Introduction 20
2.2 The Nature of PLS Path Modeling 21
2.3 Model Specification 24
2.4 Assessing and Reporting PLS Analyses 27
2.5 Extensions 32
2.6 Discussion 33
References 35

3 Going Beyond Composites: Conducting a Factor-Based PLS-SEM Analysis 41
Ned Kock
3.1 Introduction 41
3.2 Measurement Error and the Attenuation Bias 43
3.3 Illustrative Model 44
3.4 PLSF-SEM's First Stage: Composites 45
3.5 PLSF-SEM's Second Stage: Factors 47
3.6 Monte Carlo Experiment 49
3.7 Discussion and Conclusion 51
References 52

4 A Perfect Match Between a Model and a Mode 55
Theo K. Dijkstra
4.1 Introduction 55
4.2 The Model: Composites as Factors 57
4.2.1 Fundamental Properties of the Model and Wold's Fundamental Principle 60
4.3 Estimation Issues 62
4.3.1 Estimation of Weights, Loadings, and Correlations 62
4.3.2 Mode A and Mode B 65
4.3.3 Estimation of the Structural Equations 67
4.3.4 Some Monte Carlo Results 71
4.4 Testing the Composites Model 73
4.4.1 Testing Rank Restrictions on Submatrices 74
4.4.2 Exploiting the Difference Between Different Estimators 74
4.4.3 Prediction Tests, via Cross-Validation 74
4.4.4 Global Goodness-of-Fit Tests 75
4.5 Some Final Observations and Comments 76
Appendix 77
References 78

5 Quantile Composite-Based Model: A Recent Advance in PLS-PM ... 81
Cristina Davino, Pasquale Dolce, and Stefania Taralli
5.1 Introduction 82
5.2 Measuring Equitable and Sustainable Well-Being 83
5.2.1 The Reference Framework 83
5.2.2 Data Description 85
5.3 A Hierarchical Composite Model for BES 87
5.3.1 Global Model PLS-PM Results 89
5.4 Observed and Unobserved Heterogeneity in PLS-PM 90
5.4.1 PLS-PM Results Across Geographic Areas 91
5.5 Quantile Composite-Based Path Modeling 96
5.5.1 Methodology 96
5.5.2 Results 99
5.6 Concluding Remarks 102
Appendix 103
References 105

6 Ordinal Consistent Partial Least Squares 109
Florian Schuberth and Gabriele Cantaluppi
6.1 Introduction 109
6.2 The Development from PLS Path Modeling to Consistent PLS Path Modeling 112
6.2.1 Partial Least Squares 113
6.2.2 Consistent Partial Least Squares 115
6.3 The Development from PLS to Ordinal PLS 117

6.3.1 Ordinal Partial Least Squares 118
6.3.2 Ordinal Categorical Variables According to Pearson 119
6.3.3 Polychoric and Polyserial Correlation 119
6.4 Ordinal Consistent Partial Least Squares 120
6.5 Evaluation of the Construct Scores in OrdPLS and OrdPLSc 122
6.6 Assessing the Results of OrdPLSc 124
6.6.1 Overall Model Evaluation 124
6.6.2 Measurement Model 125
6.6.3 Structural Model 126
6.7 An Empirical Example: Customer Satisfaction 126
6.8 Conclusion 135
Appendix 136
References 146

Part II Methodological Issues

7 Predictive Path Modeling Through PLS and Other Component-Based Approaches: Methodological Issues and Performance Evaluation 153
Pasquale Dolce, Vincenzo Esposito Vinzi, and Carlo Lauro
7.1 Introduction 153
7.2 Predictive Ability in Component-Based Path Modeling 156
7.3 Outer Weights Computation in PLS-PM 157
7.3.1 The Effects of the Different Outer Weighting Schemes on PLS-PM Predictive Ability 158
7.3.2 PLS Regression for Outer Model Regularisation and Multi-component Estimation 159
7.4 Performance Metrics for Component-Based Predictive Path Modeling 160
7.4.1 In-Sample Redundancy-Based measure 161
7.4.2 Out-of-Sample Prediction 162
7.5 Predictive Directions of the Relationships Among Blocks 164
7.5.1 A New Algorithm Coherent with the Predictive Directions 165
7.5.2 A Comparison Between Non-symmetrical Component-Based Path Modeling and PLS-PM 166
7.6 Conclusions 169
References 170

8 Mediation Analyses in Partial Least Squares Structural Equation Modeling: Guidelines and Empirical Examples 173
Gabriel Cepeda Carrión, Christian Nitzl, and José L. Roldán
8.1 Introduction 173
8.2 Advanced Procedure for Mediation Analysis in PLS-SEM 175
8.2.1 The Mediation Effect 175
8.2.2 Full Mediation 177

8.2.3 Partial Mediation ... 178
8.2.4 No Mediation ... 179
8.2.5 Multiple Mediation ... 180
8.3 Illustrative Examples ... 182
8.3.1 A Case of a Multiple Mediator Model ... 182
8.3.2 An Example of a Multistep Multiple Mediator Model ... 188
8.4 Conclusion ... 193
References ... 193

9 Treating Unobserved Heterogeneity in PLS-SEM: A Multi-method Approach ... 197
Marko Sarstedt, Christian M. Ringle, and Joseph F. Hair
9.1 Introduction ... 198
9.2 Guidelines for Uncovering and Treating Unobserved Heterogeneity in PLS Path Models ... 199
9.3 Step 1: Run FIMIX-PLS ... 201
9.3.1 Basic Concept ... 201
9.3.2 Considerations When Using FIMIX-PLS ... 202
9.4 Step 2: Follow-Up Latent Class Analysis ... 206
9.5 Step 3: Ex Post Analysis ... 208
9.6 Step 4: Run the MICOM Procedure ... 209
9.7 Step 5: Assess the Group-Specific Solutions ... 210
9.8 Summary ... 211
Appendix ... 212
References ... 213

10 Applying Multigroup Analysis in PLS-SEM: A Step-by-Step Process ... 219
Lucy Matthews
10.1 Introduction ... 219
10.2 Overview of Steps for Running MGA in PLS-SEM ... 221
10.3 Example Application of PLS-MGA ... 224
10.3.1 Step 1: Generate Data Groups ... 225
10.3.2 Step 2: Test for Invariance ... 227
10.3.3 Step 3: Analyze and Interpret Permutation Results ... 235
10.4 Summary ... 239
References ... 240

11 Common Method Bias: A Full Collinearity Assessment Method for PLS-SEM ... 245
Ned Kock
11.1 Introduction ... 245
11.2 What Is Common Method Bias? ... 246
11.3 Illustrative Model ... 247
11.4 Data Used in the Analysis ... 250
11.5 Path Coefficient Inflation ... 251
11.6 Confirmatory Factor Analysis ... 252

11.7 The Full Collinearity Test ... 253
11.8 Discussion and Conclusion ... 254
References ... 256

12 Integrating Non-metric Data in Partial Least Squares Path Models: Methods and Application ... 259
Francesca Petrarca, Giorgio Russolillo, and Laura Trinchera
12.1 Introduction ... 259
12.2 Methods ... 260
12.2.1 Partial Least Squares Path Modeling ... 260
12.2.2 PLS-PM with Binary Single-Indicator Endogenous Latent Variables ... 262
12.2.3 Non-metric Partial Least Squares Path Modeling ... 263
12.3 Data ... 266
12.4 Structural Paths and Constructs Definition ... 267
12.5 Results ... 271
12.5.1 STUDENTS Model ... 271
12.5.2 STUDENT WORKERS Model ... 275
12.6 Conclusions ... 276
References ... 278

13 Model Misspecifications and Bootstrap Parameter Recovery in PLS-SEM and CBSEM-Based Exploratory Modeling ... 281
Pratyush N. Sharma, Ryan T. Pohlig, and Kevin H. Kim
13.1 Introduction ... 282
13.2 The Naïve Bootstrap ... 283
13.2.1 Model-Based Bootstrap (Bollen-Stine) ... 284
13.3 Method ... 285
13.4 Results ... 287
13.4.1 Structural Misspecification Case ... 287
13.4.2 Measurement Misspecification Case ... 289
13.5 Discussion ... 290
Appendix ... 292
References ... 295

Part III Applications

14 Personality, Intellectual Ability, and the Self-Concept of Gifted Children: An Application of PLS-SEM ... 299
R. Frank Falk
14.1 The Theoretical Model ... 300
14.2 Hypotheses ... 302
14.3 Method ... 302
14.3.1 Subjects ... 302
14.3.2 Manifest Variables ... 303
14.4 Analysis ... 305

14.5 Results 305
14.6 Discussion 307
Appendix 308
References 308

15 Ethical Awareness, Ethical Judgment, and Whistleblowing: A Moderated Mediation Analysis 311
Hengky Latan, Charbel Jose Chiappetta Jabbour, and Ana Beatriz Lopes de Sousa Jabbour
15.1 Introduction 312
15.2 Literature Review and Hypothesis Development 314
15.2.1 The Ethical Decision-Making Model 314
15.2.2 Ethical Awareness, Emotions, and Ethical Judgment 317
15.2.3 Moderating Effect of Perceived Moral Intensity on Ethical Awareness and Ethical Judgment 317
15.2.4 Moderating Effect of Emotions on Ethical Judgment and Whistleblowing Intentions 318
15.2.5 Moderating Effect of Perceived Moral Intensity on Ethical Judgment and Whistleblowing Intentions 319
15.3 Research Method 320
15.3.1 Sample Selection and Data Collection 320
15.3.2 The Survey Structure 322
15.3.3 Data Analysis 325
15.4 Results 325
15.4.1 Multigroup Analysis (PLS-MGA) 327
15.4.2 Importance Performance-Map Analysis 330
15.4.3 Interaction Effect Analysis 331
15.5 Conclusion 332
15.6 Limitation and Future Research 333
References 334

16 Latent Variable Regression for Laboratory Hyperspectral Images 339
Paul Geladi, Hans Grahn, and Kim H. Esbensen
16.1 Multivariate and Hyperspectral Images and Their Relation to Regression 339
16.1.1 Multivariate Regression 339
16.1.2 Images, Multivariate, and Hyperspectral Images 341
16.2 Multivariate Analysis for Hyperspectral Images 343
16.2.1 The Example and Image Recording 343
16.2.2 Sampling Issues 346
16.2.3 Image Cleaning and Exploratory Analysis 347
16.3 Latent Variable (PLS) Regression for Hyperspectral Images 352
16.3.1 PLS Regression Models Between Mean Spectra and Data 352
16.3.2 Some Important PLS Regression Diagnostics 353
16.3.3 PLS Prediction Visualized in Images 359

16.3.4 PLS Discriminant Analysis PLSDA ... 361
References ... 364

17 Dealing with Nonlinearity in Importance-Performance Map Analysis (IPMA): An Integrative Framework in a PLS-SEM Context ... 367
Sandra Streukens, Sara Leroi-Werelds, and Kim Willems
17.1 Introduction ... 368
17.2 IPMA: The Basics ... 369
17.2.1 What Is IPMA? ... 369
17.2.2 IPMA and PLS-SEM ... 371
17.2.3 Performance Scores ... 371
17.2.4 Importance Scores ... 372
17.2.5 Defining the Quadrants ... 372
17.3 The SPC and Nonlinear Relationships ... 373
17.3.1 Introduction of the SPC ... 373
17.3.2 Nonlinear Relationships in the SPC ... 374
17.4 IPMA: An Integrative Framework ... 375
17.4.1 Stage 1: Research Design ... 378
17.4.2 Stage 2: The Functional Forms of the Relationships ... 380
17.4.3 Stage 3: Interpretation of Results ... 389
17.5 Empirical Illustration ... 394
17.5.1 Stage 1: Research Design ... 394
17.5.2 Stage 2: The Functional Forms of the Relationships ... 394
17.5.3 Stage 3: Interpretation ... 396
17.6 Conclusion ... 398
17.7 Limitations and Suggestions for Further Research ... 400
References ... 400

Author Index ... 405

Subject Index ... 409

About the Authors

Gabriele Cantaluppi is Associate Professor of Statistics. He graduated in Economics from the Catholic University of the Sacred Heart (Università Cattolica del Sacro Cuore) and got his Ph.D. in Methodological Statistics from the University of Trento, Italy. His main research interests include total quality management, in particular focusing on the analysis and the measurement of customer satisfaction by means of structural equation models with latent variables.

Gabriel Cepeda Carrión is Full Professor at the Department of Management and Marketing of the University of Seville. His main research topics include knowledge management, dynamic capabilities, and organizational learning and unlearning. He is also an expert in qualitative (case study research) and quantitative (SEM and PLS) methods in management research. His research was published in several top-ranked journals including the *Journal of Knowledge Management*, *Journal of Business Research*, *British Journal of Management*, *International Journal for Quality in Health Care*, and *European Management Journal*. He has developed professional projects about knowledge management in industries, including banking, health care and professional sport.

Cristina Davino is Associate Professor of Statistics at the University of Macerata, Italy. Her fields of interest and areas of expertise are multidimensional data analysis, data mining, quantile regression, statistical surveys, evaluation of quality of life, evaluation of educational planning, and composite indicators. In 2014, she published a monograph on quantile regression in the *Wiley Series in Probability and Statistics*. Apart from serving as a reviewer for several international statistical journals, she has been member of the scientific and local organizing committee of several international conferences. She has been involved in many research projects, whose results have been disseminated in both international conferences and journals.

Ana Beatriz Lopes de Sousa Jabbour is Associate Professor at Montpellier Business School, France, and was an academic in operations management and supply chain at prestigious universities, including the University of Strathclyde, Glasgow, Scotland, UK, and UNESP-São Paulo State University, Brazil. She has expertise in sustainable operations management, including logistics and supply chain subjects. She has published in the *European Journal of Operational Research*, *Industrial Marketing Management*, *International Journal of Production Research*, *International Journal of Production Economics*, and *Journal of Business Ethics*, among others.

Theo K. Dijkstra graduated in 1977 in Econometrics (*cum laude*) at the University of Groningen, Netherlands, and obtained a Ph.D. there in 1981. In his thesis, with H.O.A. Wold as supervisor, he established numerical and stochastic convergence properties of partial least squares. He worked at the University of Groningen in econometrics and finance and at the School of Finance & Risk Management of the Free University of Amsterdam, Netherlands. For 12 years, he was also Head of Research at an institutional asset management company. As of 2010, he is part-time independent researcher as well as Honorary Professor of Econometrics at the University of Groningen. Dijkstra published in *Econometrica*, *International Economic Review, Journal of Econometrics*, *Psychometrika*, *Long Range Planning*, *Organizational Research Methods*, *European Journal of Information Systems*, *Journal of Flood Risk Management*, *Central European Journal of Operations Research*, *Insurance Mathematics and Economics*, *Information Systems Research*, *MIS Quarterly*, and others.

Pasquale Dolce received his Ph.D. in Statistics at the University of Naples "Federico II," Italy, and is a postdoctoral research fellow at Oniris Nantes, Research Unit of Statistics, Sensometrics and Chemometrics. His main research topics are focused on multidimensional data analysis, partial least squares methods, structural equation modeling (covariance-based and component-based approaches), nonsymmetrical data analysis, and three-way data analysis. He is a member of Nac Lab at the University of Naples.

Kim H. Esbensen Ph.D., Dr. (Hon) has been Research Professor in Geoscience Data Analysis and Sampling at GEUS, the National Geological Survey of Denmark and Greenland (2010–2015); Chemometrics and Sampling Professor at Aalborg University, Denmark (2001–2015); Professor of Process Analytical Technologies at Telemark Institute of Technology, Norway (1990–2000 and 2010–2015); and Associate Professor at the University of Québec in Chicoutimi, Canada (2013–2016). Being a member of six scientific societies, he phased out a more than 30-year academic career in his quest as an independent consultant from 2016: www.kheconsult.com. As he could not terminate his love for teaching completely, he is still very active as an international guest professor. Esbensen, a geologist/geochemist/data analyst by training, has been working over 20 years at the forefront of chemometrics, but since 2000 has devoted most of his scientific research to the theme of representative sampling of heterogeneous materials, processes, and systems (Theory of Sampling, TOS), PAT (Process Analytical Technology), and chemometrics. He has published over 250 peer-reviewed papers, has authored a widely used textbook in multivariate data analysis (33,000 copies), and is the editor of the magazine TOS forum (http://www.implublications.com/tos-forums).

Furthermore, he was chairman of the taskforce responsible for writing the world's first horizontal (matrix-independent) sampling standard (2013) http://webshop.ds.dk/product/M278012/ds-30772013.aspx.

R. Frank Falk Ph.D. is the Director of Research at the Institute for the Study of Advanced Development (ISAD), Westminster, CO, and Professor Emeritus at the University of Akron, OH. He chaired the Department of Sociology at the University of Denver, CO. At the University of Akron, he served as Head of Sociology and Chair of Classical Studies, Anthropology, and Archaeology and was Vice-President of Research. He specializes in research methodology with an interest in social psychology, gifted and talented, and adult emotional development. He is the senior author of *A Primer for Soft Modeling,* now available at www.researchgate.net.

Paul Geladi has a Ph.D. in Analytical Chemistry from the University of Antwerp, Belgium (1979). After finishing his Ph.D., he started specializing in chemometrics, mainly in Northern Europe and North America. This included teaching experimental design, multivariate data analysis, and multivariate image analysis. He has published 3 coauthored books and more than 150 research papers and has received more than 10,000 citations. His publications cover multivariate data analysis, multivariate regression, multiway analysis, and hyperspectral image analysis. He retired as Professor from the Swedish University of Agricultural Sciences in 2016 and is now active as emeritus.

Hans Grahn received his Ph.D. in organic chemistry from Umeå University, Sweden, on relaxation processes in Nuclear Magnetic Resonance (NMR) in 1986. He then started working on automated 2D and 3D NMR spectral interpretation using multivariate data analysis. He worked as a researcher at Syracuse University, NY, and at the University of Lausanne, Switzerland. He continued with MRI research and was appointed Associate Professor at the Department of Neuroscience, Karolinska Institute, Stockholm, Sweden, in 1992. Among his extensive entrepreneurial activities in medical imaging, he has continued publishing papers in multivariate data and image analysis with an emphasis on medical imaging applications. He has published more than 45 scientific papers and co-published 2 textbooks with Prof. Geladi. Dr. Grahn is currently Adjunct Professor at the School of Science and Technology, Örebro University, Sweden.

Joseph F. Hair is director of the DBA Doctor of Business Administration Degree Program at the Mitchell College of Business of the University of South Alabama, Mobile, AL, and holds the Cleverdon Chair of Business. He has authored over 50 books, including market leaders Multivariate Data Analysis, Prentice-Hall, seventh edition, 2010; A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM), Sage, second edition, 2017; and *MKTG*, Cengage, 2017. His numerous publications can be found in journals such as *Journal of Marketing Research*, *Journal of Academy of Marketing Science*, *Organizational Research Methods*, *Journal of Advertising Research*, and *Journal of Retailing*. His publications have been cited over 120,000 times.

Jörg Henseler holds the Chair of Product-Market Relations at the Faculty of Engineering Technology, University of Twente, Enschede, the Netherlands. He is also Visiting Professor at NOVA Information Management School, NOVA University of Lisbon, Portugal, and Distinguished Invited Professor in the Department of Business Administration and Marketing at the University of Seville, Spain. He holds a Ph.D. from the University of Kaiserslautern, Germany. His broad-ranging research interests encompass empirical methods of marketing and design research as well as the management of design, products, services, and brands. His work was published in *Computational Statistics and Data Analysis*, *European Journal of Information Systems*, *International Journal of Research in Marketing*, *Journal of the Academy of Marketing Science*, *Journal of Supply Chain Management*, *Long Range Planning*, *MIS Quarterly*, *Organizational Research Methods*, and *Structural Equation Modeling—An Interdisciplinary Journal*, among others. Prof. Henseler is a leading expert on partial least squares (PLS) path modeling, a variance-based structural equation modeling technique that is particularly useful in studies focused on the success factors for businesses. He authored dozens of academic articles, edited two books, and chaired two conferences on PLS. On a regular basis, Prof. Henseler provides seminars on PLS path modeling at the PLS School (www.pls-school.com), through which hundreds of scholars and practitioners have been trained in structural equation modeling. He furthermore chairs the Scientific Advisory Board of ADANCO, a software for variance-based structural equation modeling (www.composite-modeling.com). His personal website can be visited at http://www.henseler.com.

Geoffrey Hubona held full-time tenure-track, and tenured, assistant and associate professor faculty positions at three major state universities in the Eastern United States from 1993 to 2010. In these positions, he taught many various statistics, business information systems, and computer science courses to undergraduate, Master's, and Ph.D. students. He earned a Ph.D. in Business Administration (Information Systems and Computer Science) from the University of South Florida (USF), Tampa, FL (1993); an M.A. in Economics (1990) also from the USF; an MBA in Finance (1979) from the George Mason University, VA; and a B.A. in Psychology (1972) from the University of Virginia, Charlottesville, VA. He was full-time Assistant Professor from 1993 to 1996 at the University of Maryland Baltimore County, Catonsville, MD; a tenured Associate Professor in 1996–2001 at the Department of Information Systems at the Business College of the Virginia Commonwealth University, Richmond, VA; and Associate Professor in the Department of Computer Information Systems at the Robinson College of Business, of the Georgia State University in 2001–2010. He is the founder of the Georgia R School (2010–2014) and R-Courseware (2014–Present), online educational organizations that teach analytics, research methods, and quantitative analysis techniques.

Charbel Jose Chiappetta Jabbour is Associate Professor at Montpellier Business School, France, and was an academic at prestigious universities such as the University of Stirling, Scotland, UK; UNESP-São Paulo State University, Brazil; and USP-University of Sao Paulo, Brazil. He is experienced in technical and organizational aspects of sustainable management and has published highly cited articles in first-tier journals, including *Industrial Marketing Management*, *International Journal of Production Research*, *International Journal of Production Economics*, *Journal of Business Ethics*, *Journal of Cleaner Production*, and *Technovation*, among others.

Kevin H. Kim was Associate Professor in the Department of Psychology in Education at the University of Pittsburgh. He passed away in July 2014. His scholarship was in statistical methods as well as in the behavioral and social sciences. His work focused on structural equation modeling, multivariate statistics, multilevel modeling, and Asian-American mental health. He has published extensively in methodological and in several other academic disciplines and worked as co-investigator on numerous federally funded grants with faculty throughout the University of Pittsburgh and other universities. In 2007, he co-wrote the book entitled Univariate and Multivariate General Linear Models: Theory and Applications. He was also a highly regarded teacher of statistical methodology at the University of Pittsburgh, teaching numerous graduate level courses and supervising many Ph.D. students, including the two coauthors in this chapter.

Ned Kock is Killam Distinguished Professor and Chair of the Division of International Business and Technology Studies, A.R. Sanchez, Jr. School of Business, Texas A&M International University. He also serves as the main editor of the Health Correlator blog and as founding Editor-in-Chief of the International Journal of e-Collaboration (IJeC). He is the developer of WarpPLS, a widely used structural equation modeling (SEM) software. He holds a B.E.E. in Electronics Engineering from the Federal University of Technology—Paraná, Curitiba, Brazil; a M.Sc. in Computer Science from the Institute of Aeronautical Technology, Brazil; and a Ph.D. in Management with a concentration in information systems from the School of Management Studies, University of Waikato, New Zealand. Ned's Ph.D. research analyzed the impact of asynchronous e-collaboration technologies on business process improvement groups. Ned has authored and edited several books, book chapters, journal articles, conference proceedings papers, and book reviews. His articles were published in a number of journals including *Communications of the ACM*, *Decision Science Journal of Innovative Education*, *Decision Support Systems*, *European Journal of Information Systems*, *Defense Acquisition Review Journal*, *European Journal of Operational Research*, *Human Resource Management*, *IEEE Transactions (various)*, *Information & Management*, *Information Systems Journal*, *Information Technology & People*, *International Business Review*, *Journal of Evolutionary Psychology*, *Journal of the Association for Information Systems*, *Journal of Systems and Information Technology*, *MIS Quarterly*, and *Organization Science*.

Hengky Latan graduated in 2016 in Accounting (*summa cum laude*) at the Diponegoro University, Semarang, Indonesia. He worked at the STIE Bank BPD Jateng in the Department of Accounting, and at the Petra Christian University. He authored several books including multivariate statistics with IBM SPSS, Stata and PLS-SEM with SmartPLS, WarpPLS, and XLSTAT-PLSPM and GSCA. His articles were published in a number of journals including *Industrial Marketing Management*, *International Journal of Production Research*, *Journal of Business Ethics*, *Journal of Cleaner Production*, *the Technovation*, *Transportation Research Part E*, *Sustainability Accounting*, *Management and Policy Journal*, *Total Quality Management and Business Excellence*. He also attended conferences such as the Academy of Management (AoM), the British Academy of Management (BAM), and EurOMA.

Carlo Lauro graduated in Economics at the University of Naples "Federico II" with a dissertation in Statistics. Since 1984, he is full Professor of Statistics and Multivariate Analysis at the University of Naples "Federico II" and Chair of the Ph.D. course on computational statistics. He has been Director of the Department of Mathematics and Statistics for 12 years. Nowadays, he is President of the Nova University Consortium for Higher Education, consisting of seven Italian universities. He has published 165 papers and books. His research interests include multivariate statistical analysis, structural equation modeling, computational statistics and symbolic data analysis as well as applications in economics, impact evaluation, marketing research, Total Quality Measurement, and data mining. For his research activities, he received the following international awards: the Membership of the Chinese Academy of Complex Data Analysis (Bei Hang University, 2003), a Memorial Medal of the Polish Statistical Society (Cracow, 2003), and Charles University Memorial Medal (Prague, 2004). He has served as member of various editorial boards of statistical journals, such as *Journal of Classification*, *Applied Stochastic Models and Data Analysis*, and *Computational Statistics Quarterly*. He is currently Advisory Board member of *Computational Statistics and Data Analysis*, and *Advances in Data Analysis and Classification* and is director of the Italian *Journal of Applied Statistics*.

Sara Leroi-Werelds is Assistant Professor of Marketing at the University of Hasselt, Belgium. Her research mainly focuses on customer centricity and patient centricity. Topics include customer value, patient value, value co-creation, service ecosystems, and customer engagement. Leroi-Wereld's research was published in the *Journal of the Academy of Marketing Science*, *Journal of Service Management*, *European Management Journal*, and *Journal of Consumer Behaviour*, among others. She also teaches several marketing courses at the University of Hasselt such as Marketing, Strategic Marketing, Introduction to Marketing and Strategy, Retail Marketing, and Marketing Project.

Lucy M. Matthews is Assistant Professor of Marketing in the Jones College of Business of the Middle Tennessee State University. She has over 15 years of experience in sales and management roles. Her research interests include sales and sales management, methodology, autonomy, burnout, and engagement. Matthews published several articles in journals including *Journal of Business Research*, *Journal of Marketing Theory and Practice*, and *European Business Review*.

Christian Nitzl is a postdoctoral researcher at the University of the German Federal Armed Forces in Munich, Germany. His research interests include PLS path modeling with a special focus on the use of PLS in accounting, trust research, and accounting change in the governmental area. He has published in *Financial Accountability & Management*, *Journal of Accounting Literature*, *Journal of Accounting & Organizational Change*, *Journal of Public and Nonprofit Management*, and *Die Betriebswirtschaft*, among others. He has served many times as a reviewer for journals. He also regularly coaches researchers and practitioners in using PLS.

Richard Noonan studied Economics of Education at Columbia University, NY (Ed.D.), and Stockholm University, Sweden (Ph.D.). He served as Research Officer with the International Association for the Evaluation of Educational Achievement (IEA) and later as Associate Professor at the Institute of International Education, Stockholm University, Sweden. As an education economist, he served as consultant to UNESCO, World Bank, Asian Development Bank, the International Labour Organization, and many other multilateral and bilateral international development agencies. He published numerous articles and several books on education systems, with an emphasis on quantitative analysis, including the use of PLS in the analysis of education and training systems. He retired in 2012 and devotes his time to writing.

Francesca Petrarca is a researcher at the Italian National Statistical Institute (Istat) and at the Department of Methods and Models for Economics, Territory and Finance of the Sapienza University of Rome, Italy. She holds a Master's degree in statistics, information science, and decision support systems (2009) and a Ph.D. in Statistical Methods for Economics and Business (2014) from the University of Rome "Roma Tre", Italy. Her research focuses on longitudinal data, partial least squares (PLS) methods, data linkage, and spatial econometrics. Since 2013, she has been annual lecturer and lab teacher for the statistical course at the Department of Economics of the University of Rome "Roma Tre." She has experience as lab teacher for the statistics for economics course at LUISS, the Free International University of Social Sciences Guido Carli, as well as at the Catholic University of the Sacred Heart, both situated in Rome, Italy.

Ryan T. Pohlig is Sr. biostatistician for the College of Health Sciences at the University of Delaware, Newark. He holds a degree in research methodology, with a primary focus on statistics. He has participated in research from a wide range of fields including health sciences, psychology, biomechanics, social work, and law, leading to over 40 publications and funding from more than a dozen NIH grants. His interests include structural equation modeling, mixed modeling, mediation and moderation analyses, and study design.

Pauline Ash Ray is an Associate Professor at the Business Division of Thomas University, Thomasville, GA. She holds a Ph.D. from Trident University, Cypress, CA. The data analysis for her dissertation "Management of Change to Ensure IS Success: A Longitudinal Study" was performed using PLS. She holds a Bachelor's in Chemical Engineering from Mississippi State University, a Bachelor's in Accounting, and a Master's in Business from Mississippi University for Women as well as a Master level certificate in Accounting from Brenau University, GA.

She worked in process engineering and middle management for 20 years in the industry before teaching in graduate and undergraduate business programs. She was a coauthor on a PLS guidelines work: Henseler, Jörg, Hubona, Geoffrey & Ray, Pauline Ash (2016) "Using PLS Path Modeling in New Technology Research: Updated Guidelines," *Industrial Management & Data* Systems, and she has served as a reviewer for *Industrial Management & Data Systems* and *Business Research Journal*.

Christian M. Ringle is a Chaired Professor of Management at the Hamburg University of Technology (TUHH), Germany, and a Conjoint Professor at the Faculty of Business Law of the University of Newcastle, Australia. His research addresses the management of organizations, strategic and human resource management, marketing, and quantitative methods for business and market research. His research in these fields were published in well-known journals such as *Annals of Tourism Research, International Journal of Research in Marketing, Information Systems Research, Journal of Business Research, Journal of Service Research, Journal of the Academy of Marketing Science, Long Range Planning, MIS Quarterly*, and *Tourism Management*. He is a cofounder of SmartPLS (http://www.smartpls.com), a statistical software tool with a graphical user interface for partial least squares structural equation modeling (PLS-SEM). Prof. Ringle regularly teaches doctoral seminars on multivariate statistics, especially the PLS-SEM method, and the use of the statistical software such as SmartPLS. More information can be found at http://www.tuhh.de/hrmo/team/prof-dr-c-m-ringle.html.

José L. Roldán is Professor of Management in the Department of Business Administration and Marketing at the University of Seville, Spain. His current research interests include organizational agility, business intelligence, knowledge management, organizational culture, and partial least squares (PLS). His recent contributions were published in *European Journal of Operational Research*, *International Journal of Project Management*, *British Journal of Management*, *Journal of Business Research*, *International Business Review*, *European Journal of Information Systems*, *Computers in Human Behavior*, and *Industrial Marketing Management*, among others. He has also served as guest editor of the *European Journal of Information Systems* (EJIS), *Journal of Business Research*, and *European Management Journal*. Prof. Roldán has been co-chair of the Second International Symposium on Partial Least Squares Path Modeling: The Conference for PLS Users (Seville, Spain, 2015) and the Ninth International Conference on PLS and Related Methods (PLS'17) (Macau, China, 2017).

Giorgio Russolillo is Associate Professor (Maître de Conferences) at the National Conservatory of Arts and Crafts (CNAM) in Paris, France, where he teaches Statistics. He is a member of the Center for Studies and Research in Computer Science and Communications (CEDRIC) laboratory. He earned his Ph.D. in Statistics from the University of Naples "Federico II," Italy, in 2010 and joined CNAM in the same year as a postdoctoral fellow. From 2005 to 2006, he was biostatistician at the Italian National Research Council. His research interests involve multivariate data analysis models and methods applied to social and life sciences. Much of his current research focuses on partial least squares (PLS) algorithms for regularized regression and latent variable path modeling.

Marko Sarstedt is Chaired Professor of Marketing at the Otto-von-Guericke-University, Magdeburg, Germany, and Conjoint Professor of the Faculty of Business and Law at the University of Newcastle, Australia. He served previously as Assistant Professor of Quantitative Methods in Marketing and Management at the Ludwig-Maximilians-University Munich, Germany. His main research is in the application and advancement of structural equation modeling methods to further the understanding of consumer behavior and to improve marketing decision-making. His research was published in journals such as *Journal of Marketing Research*, *Journal of the Academy of Marketing Science*, *Organizational Research Methods*, *MIS Quarterly*, *International Journal of Research in Marketing*, *Long Range Planning*, *Journal of World Business*, and *Journal of Business Research*. Marko has coedited several special issues of leading journals and coauthored four widely adopted textbooks, including *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)* (together with Joseph F. Hair, G. Tomas M. Hult, and Christian M. Ringle).

Florian Schuberth graduated in Economics from the University of Würzburg, Germany. Currently, he is a Ph.D. student under the supervision of Prof. Dr. Martin Kukuk (University of Würzburg) and Prof. Dr. Jörg Henseler (University of Twente) at the Chair of Econometrics at the University of Würzburg. His main research interests are composite-based structural equation modeling and variance-based estimators.

Pratyush Nidhi Sharma is an Assistant Professor in the Alfred Lerner College of Business and Economics at the University of Delaware, Newark. He received his Ph.D. from the University of Pittsburgh in 2013. His research interests include online collaboration communities and networks, open-source software development, and research methods used in information systems, particularly partial least squares path modeling. He is also interested in issues surrounding technology use and adoption and human computer interaction. His research was published in journals such as *Journal of Information Systems*, *Journal of Business Research*, *Journal of International Marketing*, and *International Journal of Accounting Information Systems*, in book chapters, and in top conferences such as the *International Conference in Information Systems* (ICIS), *Americas Conference in Information System*s (AMCIS), *INFORMS*, and *PLS*.

Sandra Streukens is Associate Professor of Marketing at the University of Hasselt, Belgium. Her main research interests include return on marketing, customer engagement through gamification, service employee management, and PLS-SEM. Streukens' work was published in various journals such as *Journal of the Academy of Marketing Science*, *Journal of Service Research*, *European Management Journal*, *Industrial Marketing Management*, and *Long Range Planning*. Her teaching activities focus on quantitative research methods in business and marketing.

Stefania Taralli graduated in Political Sciences at the University of Perugia, Italy, in 1990. She started working as a freelance researcher for a number of public bodies and institutions, both national and European, focusing mainly on evaluation research in social and labor policies. Since September 2000, she is a researcher at Istat and, from October 2010, coordinates the "Provinces' Bes project" for measuring and monitoring the well-being for policy making at local levels (http://www.istat.it/it/benessere-e-sostenibilità/misure-del-benessere/il-bes-sul-territorio). She lives and works in the Marche region.

Laura Trinchera is Assistant Professor of Statistics at NEOMA Business School, France. She holds a Master's degree in Business and Economics (2004) and a Ph.D. in Statistics (2008) from the University of Naples "Federico II," Italy. Her research focuses on multivariate data analysis with an emphasis on structural equation modeling, partial least squares (PLS) methods, clustering, and classification algorithms. She is a member of the International Society of Business and Industrial Statistics (ISBIS), the Italian Statistical Society (SIS), and the French Society of Statistics (SFDS). She is co-chair of the "Component-based Methods for Predictive and Exploratory Path Modeling" specialized team of the working group on computational and methodological statistics in the European Research Consortium for Informatics and Mathematics (ERCIM). Since 2011, she has been an external lecturer for the ESSEC Ph.D. Program at ESSEC Business School, France. She was visiting researcher at University of California, Santa Barbara, CA; University of Michigan, Ann Arbor; University of Hamburg, Germany; Charles University, Prague, Czech Republic; and the HEC School of Management in Paris, France.

Vincenzo Esposito Vinzi Ph.D. is Full Professor of Statistics and Dean of the Faculty at ESSEC Business School, France. He is the former President of the International Society for Business and Industrial Statistics (ISBIS) and of the European Board of Directors of the International Association for Statistical Computing (IASC) and elected member of the ISI Council. He is Associate Editor of the international Statistics journals *Computational Statistics and Data Analysis*, *Computational Statistics*, and *Advances in Data Analysis and Classification*. Having published more than 70 papers and books, his research interests are focused on methodological developments in the field of multivariate statistical analysis with specific reference to factorial methods, Procrustes rotations, analysis of paired tables, structural equation modeling, partial least squares (PLS) regression, and predictive modeling and multiple table analysis. He has taken special care of the application relevance of his methodological contributions with specific reference to the domains of business and industry, sensory data, social surveys, customer satisfaction surveys with respect to both commodities and public services, Total Quality Measurement, and a focus on Big Data and business analytics. Among other awards, he was awarded the Pierre Vernimmen BNP Paribas "Teaching Awards 2005" (Visiting Professors Category) at HEC School of Management, Paris, France. Two of his papers were recognized as the Top-Cited 2005–2010 articles in the Journal of *Computational Statistics and Data Analysis*. He served as the

Editor-in-Chief of the "Handbook on Partial Least Squares: Concepts, Methods and Applications" (Springer, 2010).

Kim Willems is Assistant Professor at the Faculty of Economic and Social Sciences and Solvay Business School of the Free University Brussels, Belgium, and is affiliated with the University of Hasselt, Belgium. Her research focuses on retail marketing and consumer behavior, with particular interest in the transformative nature of smart services in customer relationship management. Her research was published in *Journal of Business Research*, *Psychology & Marketing*, *Journal of Service Management*, *and Technological Forecasting & Social Change*, among others. She teaches a foundational bachelor's course in marketing as well as courses at Master's level, such as strategic and advanced marketing.

Abbreviations[1]

2SLS	Two-stage least squares
3SLS	Three-stage least squares
ACFE	Association of Certified Fraud Examiners
ADANCO	Advanced analysis of composites
ADF	Asymptotically distribution free
AIC	Akaike's information criterion
AIC_3	Modified AIC with factor 3
AIC_4	Modified AIC with factor 4
ALS	Alternating least squares
ANOVA	Analysis of variance
AVE	Average variance extracted
BCa	Bias-corrected and accelerated
BES	*Benessere Equo e Sostenibile*. Italian: Equitable and sustainable well-being
BFPI	Best fitting proper indices
BIC	Bayesian information criterion
BP	Bravais-Pearson [correlation]
CA	Correspondence analysis
CAIC	Consistent AIC
CAN	Consistent and asymptotically normal [estimator]
CBSEM	(also CB-SEM) Covariance-based SEM
CCA	Canonical correlation analysis
CI	Confidence interval
CIT	Critical incident technique
CMB	Common method bias
CSA	Covariance structure analysis
CV	Canonical variable

[1]*Note*: Common standard mathematical abbreviation, variable names, or other acronyms or abbreviations not specifically related to PLS methodology are not included here.

dG	Geodesic discrepancy
dULS	Unweighted least squares discrepancy
EAW	Ethical awareness
EDM	Ethical decision-making
EEOS	Equality of Educational Opportunity Survey
EJW	Ethical judgment
EM	Expectation-maximization
EMT	Emotion
EN	Entropy [statistic]
FIMIX-PLS	Finite mixture PLS
FIML	Full information maximum likelihood
FIMS	First International Mathematics Study
FP	Fix-point [approach]
GCCA	Generalized canonical correlation analysis
GLS	Generalized least squares
GoF	Goodness of fit
GoF_{rel}	Relative goodness of fit
GS	Goal specificity
GSCA	Generalized structured component analysis
HTMT	Heterotrait–monotrait [ratio of correlations]
IEA	International Association for the Evaluation of Educational Achievement
IPMA	1. Importance-performance map analysis. 2. Importance-performance matrix analysis
Istat	Italian National Institute of Statistics
LISREL	Linear structural relations
LV	Latent variable
MDL	Minimum description length
MDL_5	Minimum description length 5
MGA	Multigroup analysis
MICOM	Measurement invariance of composite models
MIMIC	Multiple indicators, multiple causes
ML	Maximum likelihood
MLR	Multiple linear regression
MV	Manifest variable
NFI	Normed fit index
NILES	Nonlinear iterative least squares
NIPALS	Nonlinear iterative partial least squares
NIR	Near infrared
NM-PLS	Nonmetric PLS
NM-PLSPM	Nonmetric PLS path modeling
NSCPM	Nonsymmetrical composite-based path modeling
NUTS	*Nomenclature des unitès territoriales statistiques*. French: Nomenclature of territorial units for statistics
OLS	Ordinary least squares

OrdPLS	Ordinal PLS
OrdPLSc	Ordinal consistent PLS
OrdPLS(c)	Reference to *both* OrdPLS *and* OrdPLSc
OS	Optimal scaling
PCA	Principal component analysis
PLS	Partial least squares
PLSc	Consistent PLS
PLSDA	PLS discriminant analysis
PLSF-SEM	Factor-based PLS-SEM
PLS-GAS	PLS genetic algorithm segmentation
PLS-IRRS	PLS iterative reweighted regressions segmentation
PLS-MGA	PLS multigroup analysis
PLS-PM	PLS path modeling
PLS-POS	PLS prediction-oriented segmentation
PLS-SEM	See PLS; SEM
PMI	Perceived moral intensity
PML-PLS	Partial maximum likelihood PLS
PRCA	Penalty–reward contract analysis
QC	Quantile correlation
QC-PM	Quantile composite-based path modeling
QR	Quantile regression
RC	Role clarity
REBUS-PLS	Response-based procedure for detecting unit segments
RGCCA	Regularized generalized canonical correlation analysis
RMS	Root mean square
RMSD	Root-mean-square difference
RMSEC	Root-mean-square error of calibration
RMSEP	Root-mean-square error of prediction
RMS_{theta}	Root-mean-square error correlation
SEM	Structural equation model
SIS	Statistical information system
SJER	Scandinavian Journal of Educational Research
SPC	Satisfaction-profit chain
SRMR	Standardized root mean [square] residual
ULS	Unweighted least squares
VAF	Variance accounted for
VIF	Variance inflation factor
WLS	Weighted least squares

Part I
Basic Concepts and Extensions

Chapter 1
Partial Least Squares: The Gestation Period

Richard Noonan

Abstract The aim of this chapter is to describe the context of some of the earliest applications of partial least squares in the analysis of large-scale school survey data. In the late 1960s, several large school surveys had been conducted, but the analytical methods available at the time were not capable of reflecting structural equation models covering these large data sets. Instead analysis proceeded more by analogical models than structural equation models. Such models had very limited usefulness for addressing significant policy issues. The development of partial least squares and its application in school survey research led not only to findings more relevant to policy issues of concern but also supported the development of the underlying theoretical models.

1.1 Introduction

A Personal Perspective Partial least squares (PLS) is a commonly used approach to data analysis in a wide range of fields today. I am privileged to be able to look back to the beginning of PLS and even a few years before and to some years of collaboration with Professor Herman Wold, at Uppsala University and later at Gothenburg University in Sweden. This chapter represents my personal "user" perspective on the origins of PLS, the statistical problems and issues researchers faced before PLS was developed, and the advantages derived from using PLS, as seen from the perspective of applied social research, especially in the field of survey research in education.

In 1969 I joined the small staff at the headquarters of the International Association for the Evaluation of Educational Achievement (IEA) in Stockholm. My main responsibility as Research Officer was to manage the analysis of the massive data set from the "Six-Subject Study" of educational achievement in 21 countries (see Peaker 1975). We had a massive amount of school survey data, but contemporary analytical approaches were inadequate for the kinds of research and policy issues we wanted to address.

R. Noonan (✉)
Institute of International Education, Stockholm University, Frescativägen 54, Stockholm, Sweden
e-mail: r.d.noonan@gmail.com

H. Latan, R. Noonan (eds.), *Partial Least Squares Path Modeling*,
DOI 10.1007/978-3-319-64069-3_1

It was in that context that I met Prof. Herman Wold. In 1973, when Wold invited me to collaborate with him in the application of nonlinear iterative partial least squares (NIPALS), a new world of data analysis methods opened before me. Wold needed real-world data applications to test, develop, and demonstrate the methods; we needed better methods for analyzing the massive IEA data set. We met often and had a pleasant and fruitful research collaboration that lasted more than a decade.

This chapter begins a few years before the emergence of PLS, with a brief prehistory of large-scale school survey research, some data analytic approaches used, and some problems and limitations faced when using these approaches. It continues with a brief account of some of the developments that led to the development of PLS, describes some early applications in educational research, and concludes with some personal observations.

1.2 A Prehistory: Large-Scale School Surveys

International Pilot Study of School Achievement In the late 1950s and early 1960s, a group of leading educational researchers, meeting under the auspices of the UNESCO Institute for Education in Hamburg, Germany, conducted a pilot study of school learning outcomes of 13-year-old students in 12 countries. Data were collected from students (nonverbal ability, gender, father's educational and occupational status, size of community, and school learning outcomes), teachers (rating of opportunity to learn the individual test items), and schools (use of streaming). Reporting was mainly in the form of simple means, standard deviations, cross tabulations, and breakdowns (Foshay 1962). Some simple probit regressions were computed relating selected test item scores to estimated opportunity to learn the tested item and student nonverbal ability.

Given the limitations of this pilot study, the research findings from a substantive perspective were little more than suggestive, as the authors freely admitted. Nevertheless, the results demonstrated that international comparative survey research involving school learning outcomes was feasible and could make important contributions to educational research and policy-making.

First International Mathematics Study (FIMS) Encouraged by the pilot study, the newly formed International Association for the Evaluation of Educational Achievement (IEA) launched the first full-scale international study of mathematics achievement in 1963 in 12 countries by Husén (1967). Samples were larger and more representative, and data were collected at four population levels. Tests and background instruments were much more extensive than in the pilot study. Reporting was mainly in the form of simple means, standard deviations, cross tabulations, and breakdowns for each of the four populations. However, regression analysis was used to study the relations between the achievement measures taken as dependent variables and the parental, teacher, school, and student background variables as independent variables. On the basis of exploratory analysis, the number of independent variables was reduced to 24, divided into 5 groups:

- Five parental variables (PAR_1 ... PAR_5)
- Five teacher variables (TEA_1 ... TEA_5)
- Four school variables related to teaching and learning processes ($SCH1_1$ ... $SCH1_4$)
- Six school variables related to the organization, management, and resources ($SCH2_1$ ... $SCH2_6$)
- Four student variables (STU_1 ... STU_4)

For each country and population, the total mathematics test score (corrected for guessing) was regressed on the 24 predictor variables in a single regression. The regression equation had the form

$$\begin{aligned} \mathrm{ACH} = f\,(&\mathrm{PAR}_1 \ldots \mathrm{PAR}_5, \mathrm{TEA}_1 \ldots \mathrm{TEA}_5,\\ &\mathrm{SCH1}_1 \ldots \mathrm{SCH1}_4, \mathrm{SCH2}_1 \ldots \mathrm{SCH2}_6, \mathrm{STU}_1 \ldots \mathrm{STU}_4) \end{aligned}$$

The results were presented first in the form of total variance accounted for (R^2), by country and for each of the four populations. Then in 20 tables (4 populations $\times$ 5 groups of variables), the correlations and standardized regression coefficients (beta-weights, β) were reported for each country and each variable in the respective variable groups. Finally in each of these tables and for each country, the contribution *of the respective variable group* to total variance explained (total R^2) was reported, calculated as

$$\text{Contribution to Variance Explained} = \sum\nolimits_{n=1}^{N} (\beta_n r_n)\,,$$

where N is the number of variables in each respective variable group.

Thus, for example, it was found in population 3b (students studying mathematics in final year of upper secondary school) that the total R^2 was 0.31 for England and 0.30 for Belgium (Husén op. cit., Vol. II, p. 264). However, the contribution of the second block of school variables to R^2 was 0.06 for England but -0.05 for Belgium (ibid., p. 280). Altogether more than 6% of all coefficients representing the contribution of a group of variables to the total R^2 were *negative*. Interpretation of the regression equations containing 24 predictor variables and negative contributions to R^2 was problematic.[1]

[1] Note that although the *tables* reported the correlations and regression coefficients for conceptually distinct groups of variables, the *regression equations* yielding those regression coefficients included *all 24 predictor variables*. On negative contributions to variance explained, see Note on Commonality Analysis at the end of this chapter.

Equality of Educational Opportunity Survey (EEOS) In the United States, the Civil Rights Act of 1964 mandated the conduct of a survey into the lack of equality of educational opportunity for individuals because of race, color, religion, or national origin in public educational institutions at all levels in the United States. That was a tall order. The survey addressed four issues:

- To what extent are racial and ethnic groups segregated in the public schools?
- Do the schools offer equal teaching and learning opportunities and financial, physical, and human resources?
- How much do students learn, as measured by standardized achievement tests?
- How are the learning outcomes related to the kinds of schools they attend?

In 1965 a massive school survey was conducted with a nationally representative probability sample of more than 500,000 students in some 780 schools covering the United States. Data were collected from school principals, teachers, and students. Five grade-level populations were defined, representing Grades 1, 3, 6, 9, and 12 (Coleman et al. 1966; Mayeske et al. 1969, 1972).

After preliminary exploratory analysis and aggregation of some 400 individual indicators, there remained 103 variables in 6 groups used in the regression analysis:

B	Student body social background
S	School variables
T	School personnel and financial expenditure
P	Pupil programs and policies
F	Facilities
O	Outcomes, educational expectations, attitudes, educational plans and desires, study habits, and achievement

The descriptive analysis comprised (a) means of the outcome variables, broken down by grade, region, racial or ethnic categories, and gender, and (b) correlations among these variables. The analytic objective of the study was twofold: "To find characteristics or attributes of the schools that seem to be related to school outcomes, and to suggest which of these characteristics may be most important in producing these outcomes" (Mayeske et al. 1972, p. 41). For that purpose regression analysis was used to partition explained variance among the predictor variables, following an approach referred to as "commonality analysis."[2]

In brief, the increment in R^2 when an independent variable or set of variables is entered *last* into the regression equation, as in stepwise multiple regression, is termed the "unique contribution" or the "first-order commonality" coefficient. A contribution shared by several independent variables or sets of variables is termed a "joint contribution" or a "higher-order commonality" coefficient.

[2]Commonality analysis was developed within the framework of the early large-scale surveys described in this chapter but is not widely used today. See Note on Commonality Analysis at the end of this chapter.

The data analytic work reported was comprehensive, covering many aspects of the issues addressed, but there were only two main forms of the regression equation, namely, O = f(B, S) and O = f(B, T, P, F), where O, B, S, T, P, and F represent *groups* or *blocks* of individual indicators. The results were reported in tables (1 table for each population and block of predictor variables, altogether 20 tables) showing the unique and higher-order commonality coefficients for each individual variable in the respective group of variables. No regression coefficients were reported. As the authors explained in the Foreword, "We are little interested in regression coefficients [...]. We *are* interested in the total amount of variation in the dependent variable that can be associated with the independent variable" (Mayeske et al. 1972, p. iii; emphasis in original).

IEA Six-Subject Study The successful completion of the mathematics study (FIMS) was followed in 1967 by the formal incorporation of the International Association for the Evaluation of Educational Achievement (IEA) as an independent research organization. The international headquarters was established in Stockholm, Sweden. In 1968 work began on the Six-Subject Study (SSS) in 21 countries. The subjects covered were:

- Science
- Reading Comprehension
- Literature
- English as a Foreign Language
- French as a Foreign Language
- Civic Education

Three populations were tested:

- All students in full-time schooling aged 10 years at the time of testing
- All students in full-time schooling aged 14 years at the time of testing
- All students in the terminal year in full-time secondary schooling

In brief, the data sets covered students (including home background, gender, attitudes, achievement), teachers (including educational background, experience, gender), and schools (including type of school, size of school, student/teacher ratio, expenditure). For the regression analysis, the indicators were organized into six blocks of variables:

1: Home background
2: Type of school
3: Learning conditions (teachers and teaching resources)
4: Kindred variables (attitudes and opinions)
5: Word knowledge (a proxy for verbal IQ)
6: Achievement

The purpose of this international study was not to produce "league tables" but to advance understanding of factors influencing school learning outcomes. There was a massive amount of data—hundreds of indicators covering students, teachers, and

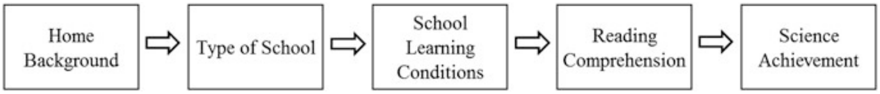


Fig. 1.1 General multiple regression model used in the IEA Six-Subject Study

schools—but no statistical methods adequate to enable a coherent analysis of the interrelationships among the variables. The main reports relied heavily on stepwise multiple regression with a focus on commonality analysis.

The explanatory analysis was based on *stepwise multiple regression using a predetermined order of entry of these blocks of variables*. The order of entry of the blocks into the regression was determined by a "chronological" conceptual model: The conditions of the home in which the child was raised represent the earliest and longest influence, the type of school the child attended (selective or general government school) had the second longest influence, etc. Thus, the home background block was entered first, followed by the type of school, etc. In some of the analyses, Reading Comprehension was entered as the final predictor, e.g., for Science, Literature, English as a Foreign Language, French as a Foreign Language, and Civic Education (Peaker 1975).

The general model followed is illustrated in Fig. 1.1 for science achievement. This kind of analysis was conducted for each of the subjects tested, for each population sampled, and for each participating country. Both between-student and between-school analyses were conducted.

The results of the analyses were reported mainly in the form of:

- Increments in R^2 as blocks of variables were entered into the regression in the predetermined order
- The unique and joint contributions of blocks of variables to R^2 (commonality analysis)
- Regression coefficients (betas) and partial regression coefficients (partial betas) as blocks of variables were entered into the regression in the predetermined order
- Increments in R^2 for alternative orders of entry of blocks of variables into the regression and in some cases
- Regression coefficients for the individual indicators making up the various blocks

It was a pioneering work and a significant contribution, and it was one of the foundations for the coming of age of partial least squares analysis.

1.3 PLS Comes of Age

Problems of Interpretation Quite apart from any discussion of the substantive results of the FIMS, the EEOS, and the IEA SSS, there was considerable discussion in the education research community about suitable methods for the analysis of large and complex data sets covering social systems. Percent of variance explained does not yield a satisfactory understanding of how things work, and the extension of variance explained to include unique and joint contributions can rapidly lead to incomprehensibility as the number of predictor variables increases.

Simple multiple regression analysis involving dozens of predictor is often beset by such multicollinearity problems as to render interpretation virtually impossible. What meaningful interpretation can be given to the frequently observed combination of positive and negative regression coefficients when mother's education, father's education, father's occupation, number of books in the home, type of school, teacher's education, teacher's years of experience, school resources, student attitudes, and a dozen other variables are used as predictors of science achievement when all these variables are positively correlated? What operational meaning can be given to "the effect of mother's education, holding father's education constant" when in the real world the two are highly correlated, especially when the regression coefficient for mother's education is *negative* (it happens)? No one can believe that student achievement would be higher if only mothers had lower levels of education, and yet a negative regression coefficient for mother's education is not uncommon in such simple multiple regressions. Reference to suppressor effects might be satisfying to statisticians but not to educators and education policy-makers.

Given these difficulties, why were variance-explained approaches, such as increments in R^2 and commonality analysis, used at all? Coleman (1975) noted that the use of such approaches was due partly to the fact that increments in R^2 were one of the few easily obtainable measures of the effect of a block of variables, as distinct from the effect of a single variable. As an alternative, Mayeske et al. (1972), in a reanalysis of the EEOS data set, formed composite variables using principal component weights with blocks of indicators.

Coleman (1975) recommended the use of composite variables representing each block. The composite variable for a given block was formed as a linear combination of the individual indicators in the block, weighted by the beta coefficients from the regression of the criterion variable (Achievement) on the predictor variables in the block. The blocks of variables could thus be reduced to individual composite variables, and the desired regression coefficients for the composites could be obtained in the usual manner. This approach could also be elaborated with other approaches (Noonan 1976). It was a step forward, but more was needed.

I met Herman Wold in 1973, and for the next decade, I would continue to pursue a research program to develop a *comprehensive and coherent* analysis of the IEA SSS data. The SSS data set covered large portions of the school system—characteristics of the student and the home, characteristics and resources of the school, teachers, teaching methods, and learning outcomes. Most of the studies based on the SSS data covered only small proportions of the massive data set, but my goal—inspired by PLS—was to achieve an analysis covering virtually all dimensions within a *single coherent analytical framework*.

The Dawn of Partial Least Squares The ideas on which PLS is based emerged from Herman Wold's work in econometrics. A crucial issue was the causal-predictive interpretation of a system of simultaneous equations. In the 1940s and 1950s, Wold focused on the causal-predictive issues, emphasizing a clear-cut causal interpretation of recursive systems. From the 1950s to the mid-1970s, his focus shifted from *causation* to *prediction*. This work led to the development of an iterative approach solving structural equation systems called fix-point (FP) approach, "fix-point" referring to the solution toward which the algorithm converges with successive iterations (Wold 1964, 1965, 1981, pp. 1–36).

One of the basic problems faced in applied social research is how to represent the dependence of a criterion Y on a set of explanatory variables X_i which are correlated and where neither variables X_i nor Y can be measured directly. Instead the variables X_i and Y can often be measured only by using some mathematical combination of indicators, x_{ij} and y_i. These variables X_i and Y are referred to as *latent variables* and are taken to represent underlying theoretical constructs ξ_i and η. The indicators x_{ij} and y_i are referred to as *manifest variables*. Wold developed the FP approach further with the introduction of nonlinear iterative least squares (NILES) in 1966 (Wold 1966), renamed nonlinear iterative partial least squares (NIPALS) by 1973 (Wold 1973). NIPALS was further developed with path modeling using latent variables in 1975 (Wold 1975).

NIPALS Applications in Educational Research The first application of NIPALS in educational research appeared in 1975, mainly as an illustration of its utility in the context of "complex situations with soft information" (Noonan et al. 1975). Our first joint article in the *Scandinavian Journal of Educational Research (SJER)*, published in 1977, used three relatively simple latent variables, as illustrated in Fig. 1.2a (Noonan and Wold 1977). It showed that with some relatively simple models, the NIPALS approach is equivalent to that proposed by Coleman (op cit). More importantly, it also showed that compared with the commonality approach referred to above, NIPALS was clearly superior, in terms of both interpretation and predictive power.

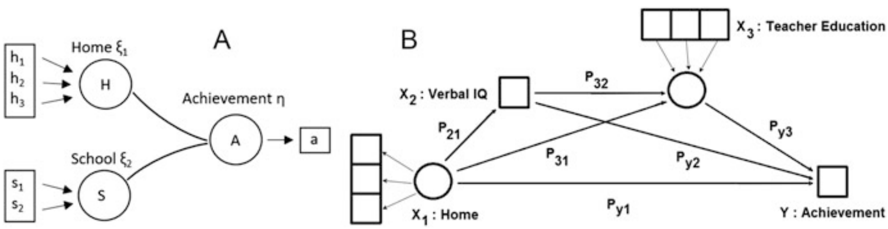


Fig. 1.2 NIPALS path models 1977 and 1978

A much more comprehensive application in 1978 based on the IEA data involved a path analysis model with hierarchically structured latent variables (Noonan 1978).[3] The purpose of that study was to investigate the effect of teacher education on student science achievement in developing countries. The IEA SSS data sets for Chile and India were used.

The IEA SSS data files contained a large number of items, but previous analyses used only a very small proportion of them in any one analysis. A new paradigm was needed for thinking about and analyzing the effects of student, home, school, and teacher factors on student learning outcomes. There was a need for methods for reducing the mass of data available to a smaller set of parsimonious descriptions of the underlying phenomena. There was also a need for a causal conceptualization of the educational situation in models elaborate enough to cover all major relevant aspects of schooling, within the limits of the data set.

In investigating the effects of teacher education on student achievement, it was necessary to control for other relevant variables. This was done with the help of a model of the schooling situation which covered as wide a scope as possible, within the limits of the IEA data bank. The analytical methods used enabled the calculation of both the direct effect of formal teacher education, holding all other relevant variables constant, and the indirect effect, through other variables that teacher education influences. A simplified version of the path model used is shown in Fig. 1.2b. It illustrates the "inner relations" among the variables representing the underlying theoretical concepts and the "outer relations" between the manifest variables and the latent variables. The term "inner variable" is used to cover theoretical constructs represented by either a single indicator (e.g., X_2, verbal IQ, and Y, achievement, in Fig. 1.2b) or a linear combination of several indicators (e.g., X_1, home, and X_3, teacher education).

Several weighting schemes had been proposed for combining the manifest variables to form the latent variables. In this article two of these weighting schemes were applied. For the outer relations, Fig. 1.2b illustrates the distinction then made between the two weighting schemes used in this study, "Mode A" and "Mode B." In Mode A, the manifest variables were weighted by simple correlations between the manifest variables and the latent variable (compare with principal components analysis). In Mode B, the manifest variables were weighted by multiple regression coefficients (compare with canonical variate analysis). These "modes" correspond to the "formative" and "reflective" outer relations discussed by Hauser (1973).

A large number of manifest variables were reduced to a total of 17 hierarchically structured latent variables covering the school science learning situation in as wide a scope as possible within the limits of the IEA data bank. In preparing this study, I

[3] This was a first and informal step toward the use of hierarchically structured latent variables within the NIPALS/PLS framework. There were two levels of latent variables. The "lower level" latent variables were formed using principal component weights. The resulting compounds were then treated as manifest variables in the usual manner.

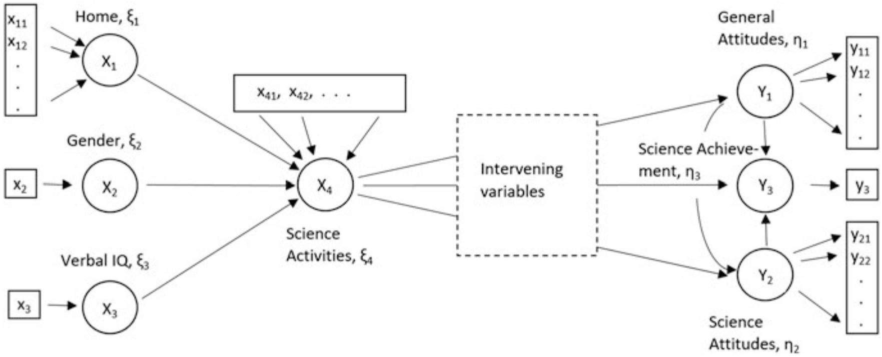


Fig. 1.3 PLS path model 1980

wrote a computer program with the capacity to handle an unlimited (in principle)[4] number of manifest variables and latent variables. Inner and outer relations were specified as input parameters.

From NIPALS to PLS By the end of the 1970s, Wold had dropped the terms "nonlinear" and "iterative" from the name and thereafter used the simpler name partial least squares, PLS (Wold 1980, 1981). Our second joint article in the SJER was published in 1980 (Noonan and Wold 1980). It represented a continuation of the same broad research program aiming at a comprehensive and coherent analysis of the SSS data for Sweden. Despite the name change from NIPALS to PLS, the article was titled "Part II," referencing the 1977 article using the name NIPALS. The first SJER article was mainly an illustration of the use of the NIPALS approach and a comparison with some other approaches used for the analysis of school survey data (stepwise multiple by blocks of variables, commonality analysis).

This second SJER article focused mainly on substantive issues but also tested five alternative forms for the specification of the latent variables as linear compounds of the manifest variables. The article assesses the magnitude of the effects of a variety of regional, home, student, teacher, and school factors on variation in student cognitive and affective outcomes in science learning. Fifty-nine indicators in 16 blocks were entered into a PLS analysis using the path model shown in Fig. 1.3.

Differences between national school systems are often reflected in analyses of hypothesized causal effects of home, school, and teacher variables on student learning outcomes. In 1982 a comparative study conducted of England, Scotland,

[4]Unlimited in principle, in practice limited by the maximum memory space allowed by the IBM mainframe computer we used, but never a problem in our analyses. This program was later developed to include hierarchically structured latent variables as a standard PLS option. The program was written in FORTRAN IV, later FORTRAN G, and comprised some 1500 lines of code. Around the same time, Lohmöller was writing his program (see Lohmöller 1989), which became the basis from which the most commonly used program today has evolved.

and Sweden using the IEA SSS science achievement data aimed to find *universal* patterns of *causal effects* (Noonan 1982). A distinction is made here between the *causal* effect and *allocation* effects, for example:

- *Causal effect*: Additional time on learning tasks *causes* additional learning, which results in a positive correlation between time on task and learning outcomes.
- *Allocation effect*: Students with special learning needs might receive more instructional time than other students, which could result in an observed *negative* correlation between time on task and learning outcomes for the population as a whole, even though the *causal* effect is positive for all students.

The school systems of these three counties differed considerably with respect to the relationship between home background, school resource allocation, and learning outcomes. Altogether more than 100 manifest variables were used, and these were distributed over 20 blocks of indicators used in the construction of the latent variables. In the search for universal patterns of causal effects, the study focused ultimately not on resources but on classroom practices. It was found, for example, that classroom activities that are most directly related to the "practice of science" (e.g., laboratory activities and field work) are more efficient in generating learning outcomes than such activities as discussion, questioning, and explication of the text.

With the 1982 search for universal patterns of causal effects, the "macro-focus" on school and teacher characteristics gave way of a "micro-focus" on teaching and learning activities. It appeared as if the use of path models for describing the influence of the education system on learning was itself somewhat limiting. In 1983 the use of PLS in the evaluation of school systems was described (Noonan and Wold 1983). It included the first explicit presentation of PLS with hierarchically structured latent variables.

Although model specification followed the standard path analytic approach, a broader and more comprehensive approach was needed in order to visualize the full range of significant factors influencing school learning. The analogue model shown in Fig. 1.4a represented a new beginning for the conceptualization of the scope of my own research program and my view of the analytical potential offered by PLS.

It might be said that PLS officially entered the world of educational research in 1986 with the publication of an article in the *International Encyclopedia of Education* (Noonan and Wold 1986) and an international handbook on educational research, methodology, and measurement (Noonan and Wold 1988).

My final contribution in the PLS story (Noonan 1989), published under my name but in reality very much a collaborative work, was the most comprehensive study of the IEA data to date. It covered a total of 191 manifest variables in 41 blocks. This was reduced, using hierarchically structured latent variables, to only 13 predictors in the inner relations, including 8 hierarchical latent variables which represented altogether 35 basic latent variables. Following the conceptual model shown in Fig. 1.4a, the results were presented as shown in Fig. 1.4b.

The visualization of the path model in Fig. 1.4b was by Herman Wold himself to correspond to the concentric circle diagram of the 1983 publication shown in

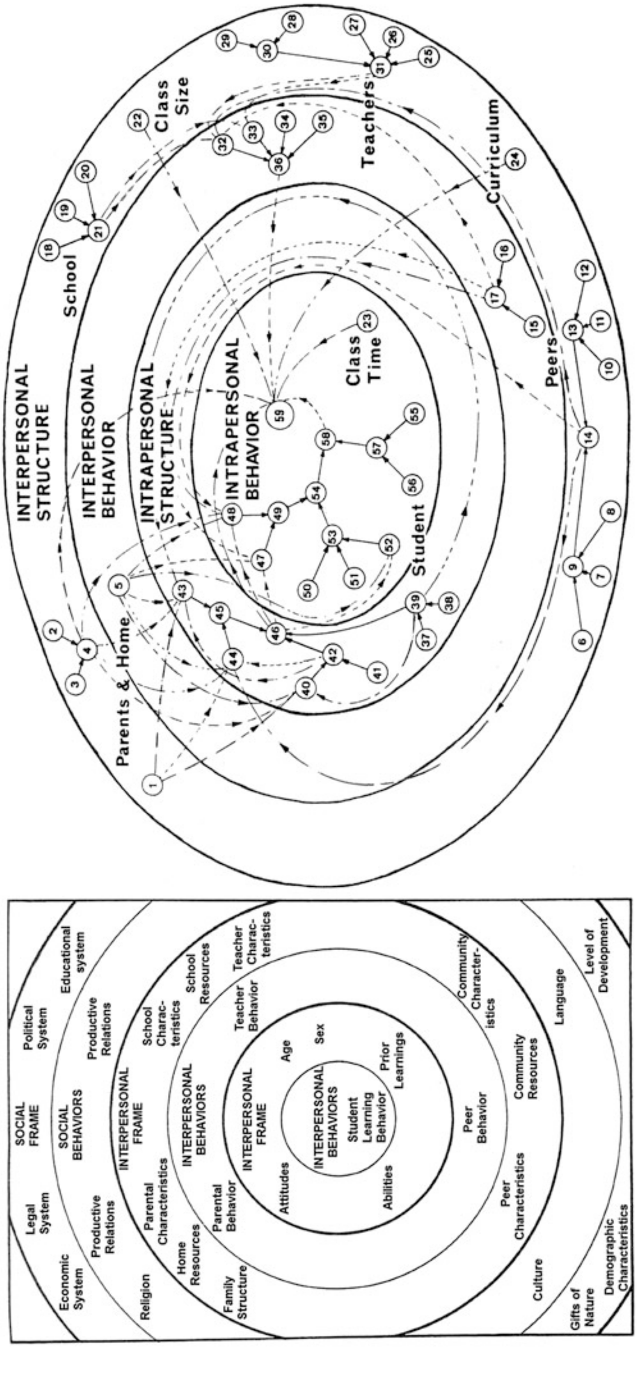


Fig. 1.4 An analogue model of school learning (theoretical and empirical)

Fig. 1.4a. I had not drawn a path diagram because I thought it was too complicated. At our next meeting over coffee, he showed me his visualization.

1.4 Some Concluding Personal Comments

As a young researcher struggling with the analysis of IEA's massive data set, I was excited to learn about NIPALS. It promised to solve some of the basic problems associated with the contemporary analytical methods for large-scale survey research data. My first forays into the world of NIPALS were small and simple, but the results were fruitful. With each succeeding publication, the aims became broader, and the path models became larger, as the limits of what was possible seemed to vanish. At some point I began to wonder where the limits lay. Were there any limits? And if so, where were the limits and what were the limiting factors?

In a complex, real-world social science context, regression coefficients become increasingly difficult to interpret as the number of predictors increases because of the limited validity of the *ceteris paribus*[5] assumption—the *ceteris* is often *not paribus*. For example, what is the effect of *mother's education* on student achievement, *holding constant* mother's occupation, father's education, father's occupation, number of books in the home, etc.? In reality these factors covary, and the use of these variables together as predictors can yield results very difficult to interpret without resorting to a discussion of "suppressor effects," as noted above. In such a situation, the formation of a compound variable would be the usual solution.

Even if the correlations among the predictor variables were low (so any suppressor effects would be small), interpretation would become increasingly difficult as the number of predictors increased. The main difficulty would be the ability of the mind to comprehend projections in multidimensional space of increasingly higher order. Thus, if the model is as comprehensive as that shown in Fig. 1.4a, even the use of latent variables can be unwieldy. I drew the conclusion, therefore, that the limit lay not in the method, the program, or computer power but in the ability of the mind to comprehend large and complex models. As the models got larger and larger, therefore, the introduction of multilevel hierarchically structured latent variables was a natural move.

Since those early days, the PLS family of methods has evolved and grown dramatically, as seen in the remaining chapters of this book. It has been a wonderful journey.

[5] Ceteris paribus: Latin. Other things being equal, holding other things constant

1.4.1 Notes on Commonality Analysis

Commonality analysis is not widely used today in the analysis of large data sets covering open systems. In the late 1960s and the 1970s, however, it was one of the most useful tools available, despite considerable shortcomings (which were often simply ignored). A brief review of the method is given here to illustrate the level at which we stood when NIPALS/PLS first appeared and opened vast new horizons.

Commonality analysis [developed and applied by Mood (1969, 1971), and Mayeske et al. (1969); see Kerlinger and Pedhazur (1973, pp. 297–305)] separates the explained variance of a dependent variable into *unique* and *joint* contributions of the predictor variables. The unique contribution of an independent variable is defined as the variance attributed to it (i.e., the increment in R^2) when it is entered last into the regression equation (as in a stepwise multiple regression analysis). In the simple case of a dependent variable Y regressed on two independent variables X_1 and X_2, the unique contributions of X_1 and X_2 are defined as:

$$U_1 = {R^2}_{y.12} - {R^2}_{y.2}$$

$$U_2 = {R^2}_{y.12} - {R^2}_{y.1}$$

The joint contribution of X_1 and X_2 is defined as:

$$J_{12} = {R^2}_{y.12} - U_1 - U_2$$

Thus,

$${R^2}_{y.12} = U_1 + U_2 + J_{12}$$

Commonality analysis can be extended to any number of predictor variables or blocks (groups) of predictor variables. For example, U_1 might represent the unique contribution of a block of student home background variables in predicting student learning outcomes, and U_2 might represent the unique contribution of a block of school variables predicting student learning outcomes. Then J_{12} would represent the joint (or "shared") contribution (or "the commonality") of the home and school blocks of variables in predicting student learning outcomes.

Commonality analysis is more useful in a predictive framework than in an explanatory framework, inasmuch as both the unique and joint contributions are defined in terms of variance explained. Because the unique and joint contributions are influenced by the intercorrelations among the independent variables, interpretation can be difficult, especially when there are many independent variables. For a regression with n predictors, there are $2^n - 1$ commonality coefficients, including n unique contribution coefficients and $2^n - 1 - n$ joint contribution coefficients. This can lead to considerable difficulties in interpretation of the higher-order joint contribution coefficients, especially when they are negative.

Finally, I argue that, contrary to the view expressed by Mayeske et al. (1972, p. iii), and regardless of the interpretational difficulties mentioned above, as scientists and policy-makers, we *are* interested in regression coefficients.

References

Coleman, J. S. (1975). Methods and results in the IEA studies of effects of school on learning. *Review of Educational Research, 45*(3), 355–386. Retrieved from https://www.jstor.org.

Coleman, J. S., et al. (1966). *Equality of educational opportunity*. Washington, DC: U.S. Department of Health, Education, and Welfare.

Foshay, A. W. (Ed.). (1962). *Educational achievement of 13-year olds in twelve countries*. Hamburg: UNESCO Institute for Education.

Hauser, R. M. (1973). Disaggregating a socio-psychological model of educational attainment. In A. S. Goldberger & O. D. Duncan (Eds.), *Structural equation models in the social sciences*. New York: Seminar Press.

Husén, T. (Ed.). (1967). *International study of achievement in mathematics: A comparison of twelve countries, 2 Vols*. New York: Wiley.

Kerlinger, F., & Pedhazur, E. (1973). *Multiple regression in behavioral research*. New York: Holt, Rinehart & Winston.

Lohmöller, J. B. (1989). *Latent variable path modelling with partial least squares*. Berlin: Springer.

Mayeske, G., et al. (1969). *A study of out nation's schools: A working paper*. Washington, DC: U.S. Department of Health, Education, and Welfare, Office of Education.

Mayeske, G., et al. (1972). *A study of out nation's schools*. Washington, DC: U.S. Department of Health, Education, and Welfare, Office of Education.

Mood, A. (1969). Macro-analysis of the American educational system. *Operations Research, 17*, 770–784.

Mood, A. (1971). Partitioning variance in multiple regression analysis as a tool for developing learning models. *American Educational Research Journal, 8*, 191–202.

Noonan, R. (1976). *School resources, social class, and student achievement: A comparative study of school resource allocation and the social distribution of mathematics achievement in ten countries*. Stockholm: Almqvist & Wiksell International.

Noonan, R. (1978). An empirical study of teacher training and student achievement in two countries: Chile and India, Part II. In T. Husén, L. Saha, & R. Noonan (Eds.), *Teacher training and student achievement in less developed countries. Staff Working Paper No. 310* (pp. 66a–699). Washington, DC: World Bank.

Noonan, R. (1982). School environments and school outcomes: An empirical study using the IEA data. In M. Niessen & J. Peschar (Eds.), *Comparative research in education 1975–1980* (pp. 190–202). Oxford: Pergamon.

Noonan, R. (1989). Evaluation of school systems using partial least squares (PLS): An application in the analysis of open systems, Chap. 3. In H. Wold (Ed.), *Theoretical empiricism: A general rationale for scientific model building*. Washington, DC: Paragon Press.

Noonan, R., & Wold, H. (1977). NIPALS path modelling with latent variables: Analyzing school survey data using non-linear iterative partial least squares. *Scandinavian Journal of Educational Research, 21*, 33–61.

Noonan, R., & Wold, H. (1980). PLS path modelling with latent variables: Analyzing school survey data using partial least squares – Part II. *Scandinavian Journal of Educational Research, 24*(1), 1–24.

Noonan, R., & Wold, H. (1983). Evaluating school systems using partial least squares. *Educational Evaluation: International Progress, 7*(3), 219–364.

Noonan, R., & Wold, H. (1986). Partial least squares path analysis. In T. Husén & T. N. Postlethwaite (Eds.), *The international encyclopedia of education: Research and studies* (Vol. 7, pp. 3769–3775). Oxford: Pergamon.

Noonan, R., & Wold, H. (1988). Partial least squares analysis. In J. P. Keeves (Ed.), *Educational research, methodology, and measurement: An international handbook* (pp. 710–716). Oxford: Pergamon.

Noonan, R., Abrahamsson, Å., Areskoug, B., Lorentzson, L. O., & Wallmyr, J. (1975). Applications of methods I-II (NIPALS clustering and path analysis) to the analysis of the IEA data bank. In H. Wold (Ed.), *Modelling in complex situations with soft information*. Group Report, Third World Congress of Econometrics, Toronto.

Peaker, G. F. (1975). *An empirical study of education in twenty-on countries: A technical report*. Stockholm: Almqvist & Wiksell.

Wold, H. (1964). A fix-point theorem with econometric background, Part I: The Theorem. *Arkiv för Matematik, 6*(12), 209–220.

Wold, H. (1965). A fix-point theorem with econometric background, Part II: Illustrations. Further developments. *Arkiv för Matematik, 6*(13), 221–240.

Wold, H. (1966). Nonlinear estimation by iterative least squares procedures. In F. N. David (Ed.), *Research papers in statistics: Festschrift for J. Neyman* (pp. 411–444). New York: Wiley.

Wold, H. (1973). Nonlinear iterative partial least squares (NIPALS) modelling: Some current developments. In P. R. Krishniah (Ed.), *Multivariate analysis – III*. New York: Academic Press.

Wold, H. (1975). Path models with one or two latent variables. In H. M. Blalock Jr. et al. (Eds.), *Quantitative sociology* (pp. 307–357). New York: Seminar Press.

Wold, H. (1980). Model construction and evaluation when theoretical knowledge is scarce: Theory and application of partial least squares, Chap. 3. In J. Kmenta & J. B. Ramsey (Eds.), *Evaluation of econometric models* (pp. 47–74). Cambridge, MA: Academic Press.

Wold, H. (Ed.). (1981). *The fix-point approach to interdependent systems. Contributions to economic analysis* (Vol. 132). Amsterdam: North Holland.

Chapter 2
Partial Least Squares Path Modeling: Updated Guidelines

Jörg Henseler, Geoffrey Hubona, and Pauline Ash Ray

Abstract Partial least squares (PLS) path modeling is a variance-based structural equation modeling technique that is widely applied in business and social sciences. It is the method of choice if a structural equation model contains both factors and composites. This chapter aggregates new insights and offers a fresh look at PLS path modeling. It presents the newest developments, such as consistent PLS, confirmatory composite analysis, and the heterotrait-monotrait ratio of correlations (HTMT). PLS path modeling can be regarded as an instantiation of generalized canonical correlation analysis. It aims at modeling relationships between composites, i.e., linear combinations of observed variables. A recent extension, consistent PLS, makes it possible to also include factors in a PLS path model. The chapter illustrates how to specify a PLS path model consisting of construct measurement and structural relationships. It also shows how to integrate categorical variables. A particularly important consideration is model identification: Every construct measured by multiple indicators must be embedded into a nomological net, which means that there must be at least one other construct with which it is related. PLS path modeling results are useful for exploratory and confirmatory research. The chapter provides guidelines for assessing the fit of the overall model, the reliability and validity of the measurement model, and the relationships between constructs.

This chapter is an updated reprint of Henseler, J., Hubona, G., Ray, P. A. (2016). Using PLS path modeling in new technology research: updated guidelines. *Industrial Management & Data Systems*, *116*(1), 2–20 doi:10.1108/IMDS-09-2015-0382. The first author acknowledges a financial interest in ADANCO and its distributor, Composite Modeling.

J. Henseler (✉)
University of Twente, Drienerlolaan 5, 7522 NB, Enschede, The Netherlands
e-mail: j.henseler@utwente.nl

G. Hubona
Georgia R School, 1307 Hixson Pike, Chattanooga, TN, 37405, USA
e-mail: ghubona@gmail.com

P.A. Ray
Business Division, Thomas University, 1501 Millpond Rd., Thomasville, GA, 31792, USA
e-mail: pash115@live.com

H. Latan, R. Noonan (eds.), *Partial Least Squares Path Modeling*,
DOI 10.1007/978-3-319-64069-3_2

Moreover, it provides a glimpse on various extensions of PLS, many of which will be described in more detail in later chapters of the book.

2.1 Introduction

Structural equation modeling (SEM) is a family of statistical techniques that have become very popular in business and social sciences. Its ability to model latent variables, to take into account various forms of measurement error, and to test entire theories makes it useful for a plethora of research questions.

Two types of SEM can be distinguished: covariance- and variance-based SEM. Covariance-based SEM estimates model parameters using the empirical variance-covariance matrix, and it is the method of choice if the hypothesized model consists of one or more common factors. In contrast, variance-based SEM first creates proxies as linear combinations of observed variables and then estimates the model parameters using these proxies. Variance-based SEM is the method of choice if the hypothesized model contains composites.

Among variance-based SEM methods, partial least squares (PLS) path modeling is regarded as the "most fully developed and general system" (McDonald 1996, p. 240) and has been called a "silver bullet" (Hair et al. 2011). PLS is widely used in information systems research (Marcoulides and Saunders 2006), strategic management (Hair et al. 2012a), marketing (Hair et al. 2012b), and beyond. Its ability to model both factors and composites is appreciated by researchers across disciplines and makes it a promising method particularly for new technology research and information systems research. Whereas factors can be used to model latent variables of behavioral research such as attitudes or personality traits, composites can be applied to model strong concepts (Höök and Löwgren 2012), i.e., the abstraction of artifacts such as management instruments, innovations, or information systems. A particularly interesting class of artifacts is success factors for businesses. Consequently, PLS path modeling is a preferred statistical tool for success factor studies (Albers 2010).

Not only has PLS and its use been subject of various reviews (c.f. Hair et al. 2012a, b), but just recently it has also undergone a series of serious examinations and has been the target of heated scientific debates. Scholars have discussed the conceptual underpinnings (Rigdon 2012, 2014; Sarstedt et al. 2014) and the strengths and weaknesses (Henseler et al. 2014; Rigdon et al. 2014) of PLS. As a fruitful outcome of these debates, substantial contributions to PLS emerged, such as bootstrap-based tests of overall model fit (Dijkstra and Henseler 2015a), consistent PLS to estimate factor models (PLSc, see Dijkstra and Henseler 2015b), and the heterotrait-monotrait ratio of correlations as a new criterion for discriminant validity (HTMT, see Henseler et al. 2015). All these changes render the extant guidelines on PLS path modeling outdated, if not even invalid. Consequently, Rigdon (2014) recommends breaking the chains and forging ahead, which implies an urgent need for updated guidelines on why, when, and how to use PLS.

The purpose of this chapter is manifold. First, it provides an updated view on what PLS actually is and which algorithmic steps it includes since the invention of consistent PLS. Second, it explains how to specify PLS path models, taking into account the nature of the measurement models (composite vs. factor), model identification, sign indeterminacy, special treatments for categorical variables, and determination of sample size. Third, it explains how to assess and report PLS results, including the novel bootstrap-based tests of model fit, the SRMR as a measure of approximate model fit, the new reliability coefficient rho_A, and the HTMT. Fourth, it sketches several ways of how to extend PLS analyses. Finally, it contrasts the understanding of PLS as presented in this chapter with the traditional view and discusses avenues for future developments.

2.2 The Nature of PLS Path Modeling

The core of PLS is a family of alternating least squares algorithms that emulate and extend principal component analysis as well as canonical correlation analysis. The method was invented by Herman Wold (c.f. 1974, 1982) for the analysis of high-dimensional data in a low-structure environment and has undergone various extensions and modifications. In its most modern appearance (c.f. Dijkstra and Henseler 2015a, b), PLS path modeling can be understood as a full-fledged structural equation modeling method that can handle both factor models and composite models for construct measurement, estimate recursive and non-recursive structural models, and conduct tests of model fit.

PLS path models are formally defined by two sets of linear equations: the measurement model (also called outer model) and the structural model (also called inner model). The measurement model specifies the relations between a construct and its observed indicators (also called manifest variables), whereas the structural model specifies the relationships between the constructs. Figure 2.1 depicts an example of a PLS path model.

PLS path models can contain two different forms of construct measurement: factor models or composite models (see Rigdon 2012, for a nice comparison of both types of measurement models). The factor model hypothesizes that the variance of a set of indicators can be perfectly explained by the existence of one unobserved variable (the common factor) and individual random error. It is the standard model of behavioral research. In Fig. 2.1, the exogenous construct ξ and the endogenous construct η_2 are modeled as factors. In contrast, composites are formed as linear combinations of their respective indicators. The composite model does not impose any restrictions on the covariances between indicators of the same construct, i.e., it relaxes the assumption that all the covariation between a block of indicators is explained by a common factor. The composites serve as proxies for the scientific concept under investigation (Ketterlinus et al. 1989; Maraun and Halpin 2008;

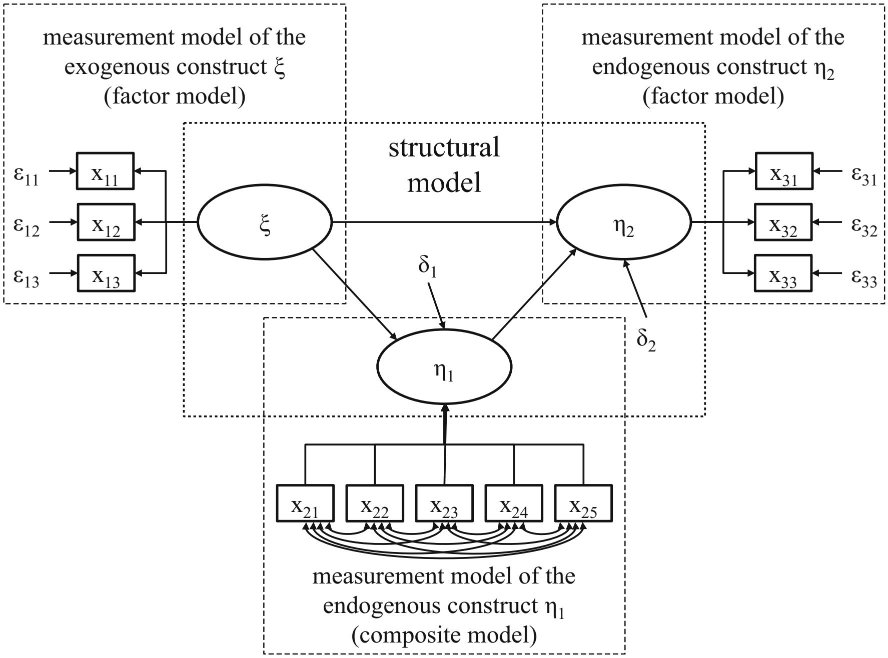


Fig. 2.1 PLS path model example

Rigdon 2012; M. Tenenhaus 2008).[1] The fact that composite models are less restrictive than factor models makes it likely that they have a higher overall model fit (Landis et al. 2000).

The structural model consists of exogenous and endogenous constructs as well as the relationships between them. The values of exogenous constructs are assumed to be given from outside the model. Thus, exogenous variables are not explained by other constructs in the model, and there must not be any arrows in the structural model that point to exogenous constructs. In contrast, endogenous constructs are at least partially explained by other constructs in the model. Each endogenous construct must have at least one arrow of the structural model pointing to it. The relationships between the constructs are usually assumed to be linear. The size and significance of path relationships are typically the focus of the scientific endeavors pursued in empirical research.

The estimation of PLS path model parameters happens in four steps: (1) an iterative algorithm that determines composite scores for each construct, (2) a correction for attenuation for those constructs that are modeled as factors (Dijkstra and Henseler 2015b), (3) parameter estimation, and (4) bootstrapping for inference testing.

[1] Note that also factors are nothing else than proxies (Rigdon 2012).

Step 1 For each construct, the iterative PLS algorithm creates a proxy as a linear combination of the observed indicators. The indicator weights are determined such that each proxy shares as much variance as possible with the proxies of causally related constructs. The PLS algorithm can be viewed at as an approach to extend canonical correlation analysis to more than two sets of variables; it can emulate several of Kettenring's (1971) techniques for the canonical analysis of several sets of variables (M. Tenenhaus et al. 2005). For a more detailed description of the algorithm, see Henseler (2010). The major outputs of the first step are the proxies (i.e., composite scores), the proxy correlation matrix, and the indicator weights.

Step 2 Correcting for attenuation is a necessary step if a model involves factors. As long as the indicators contain random measurement error, so will the proxies. Consequently, proxy correlations are typically underestimations of factor correlations. Consistent PLS (PLSc) corrects for this tendency (Dijkstra and Henseler 2015a, b) by dividing a proxy's correlations by the square root of its reliability (the so-called correction for attenuation). PLSc addresses the issue of what the correlation between constructs would be if there were no random measurement error. The major output of this second step is a consistent construct correlation matrix.

Step 3 Once a consistent construct correlation matrix is available, it is possible to estimate the model parameters. If the structural model is recursive (i.e., there are no feedback loops), ordinary least squares (OLS) regression can be used to obtain consistent parameter estimates for the structural paths. In the case of non-recursive models, instrumental variable techniques such as two-stage least squares (2SLS) should be employed. In addition to the path coefficient estimates, this third step can also provide estimates for loadings, indirect effects, total effects, and several model assessment criteria.

Step 4 Finally, the bootstrap is applied in order to obtain inference statistics for all model parameters. The bootstrap is a nonparametric inferential technique which rests on the assumption that the sample distribution conveys information about the population distribution. Bootstrapping is the process of drawing a large number of resamples with replacement from the original sample and then estimating the model parameters for each bootstrap resample. The standard error of an estimate is inferred from the standard deviation of the bootstrap estimates.

The PLS path modeling algorithm has favorable convergence properties (Henseler 2010). However, as soon as PLS path models involve common factors, there is the possibility of so-called Heywood cases (Krijnen et al. 1998), meaning that one or more variances implied by the model would be negative. The occurrence of Heywood cases may be caused by an atypical or too-small sample, or the common factor structure may not hold for a particular set of indicators.

PLS path modeling is not as efficient as maximum likelihood covariance-based SEM. One possibility is to further minimize the discrepancy between the empirical and the model-implied correlation matrix, an approach followed by efficient PLS (PLSe, see Bentler and Huang 2014). Alternatively, one could embrace the notion that PLS is a limited-information estimator and is less affected by model misspecification in some subparts of a model (Antonakis et al. 2010). Ultimately, there is no clear-cut resolution of the issues on this trade-off between efficiency and robustness with respect to model misspecification.

2.3 Model Specification

The analysts must take care that the specified statistical model complies with the conceptual model to be tested and further that the model complies with technical requirements such as identification and with the data conforming to the required format and statistical power.

Typically, the structural model is theory-based and is the prime focus of the research question and/or research hypotheses. The specification of the structural model addresses two questions: Which constructs should be included in the model? And how are they hypothesized to be interrelated? That is, what are the directions and strengths of the causal influences between and among the latent constructs? In general, analysts should keep in mind that the constructs specified in a model are only proxies and that there will always be a validity gap between these proxies and the theoretical concepts that are the intended modeling target (Rigdon 2012). The paths, specified as arrows in a PLS model, represent directional linear relationships between proxies. The structural model, and the indicated relationships among the latent constructs, is regarded as separate from the measurement model.

The specification of the measurement model entails decisions for composite or factor models and the assignment of indicators to constructs. Factor models are the predominant measurement model for behavioral constructs such as attitudes or personality traits. Factor models are strongly linked to true score theory (McDonald 1999), the most important measurement paradigm in behavioral sciences. If a construct has this background and random measurement error is likely to be an issue, analysts should choose the factor model. Composites help model emergent constructs, for which elements are combined to form a new entity. Composites can be applied to model strong concepts (Höök and Löwgren 2012), i.e., the abstraction of artifacts (man-made objects). Typical artifacts in new technology research would include innovations, technologies, systems, processes, strategies, management instruments, or portfolios. Whenever a model contains this type of construct, it is preferable to use a composite model.

Measurement models of PLS path models may appear less detailed than those of covariance-based structural equation modeling, but in fact some specifications are implicit and are not visualized. For instance, neither the unique indicator errors (nor their correlations) of factor models nor the correlations between indicators of composite models are drawn. Because PLS currently does not allow to either constrain these parameters or to free the error correlations of factor models, by convention these model elements are not drawn. No matter which type of measurement is chosen to measure a construct, PLS requires that there is at least one indicator available. Constructs without indicators, so-called phantom variables (Rindskopf 1984), cannot be included in PLS path models.

In some PLS path modeling software (e.g., SmartPLS and PLS-Graph), the direction of arrows depicted in the measurement model does not indicate whether a factor or composite model is estimated but whether correlation weights (Mode A, represented by arrows pointing from a construct to its indicators) or regression weights (Mode B, represented by arrows pointing from indicators to their construct) shall be used to create the proxy. In both cases PLS will estimate a composite model. Indicator weights estimated by Mode B are consistent (Dijkstra 2010), whereas indicator weights estimated by Mode A are not, but the latter excel in out-of-sample prediction (Rigdon 2012).

Some model specifications are made automatically and cannot be manually changed: Measurement errors are assumed to be uncorrelated with all other variables and errors in the model; structural disturbance terms are assumed to be orthogonal to their predictor variables as well as to each other[2]; correlations between exogenous variables are free. Because these specifications hold across models, it has become customary not to draw them in PLS path models.

Identification has always been an important issue for SEM, although it has been neglected in the realm of PLS path modeling in the past. It refers to the necessity to specify a model such that only one set of estimates exists that yields the same model-implied correlation matrix. It is possible that a complete model is unidentified, but also only parts of a model can be unidentified. In general, it is not possible to derive useful conclusions from unidentified (parts of) models. In order to achieve identification, PLS fixes the variance of factors and composites to one. An important requirement of composite models is a so-called nomological net. It means that composites cannot be estimated in isolation but need at least one other variable (either observed or latent) to have a relation with. Since PLS also estimates factor models via composites, this requirement extends to all factor models estimated using PLS. If a factor model has exactly two indicators, it does not matter which form of SEM is used—a nomological net is then required to achieve identification. If a construct is only measured by one indicator, one speaks of single-indicator measurement (Diamantopoulos et al. 2012). The construct scores are then identical to the standardized indicator values. In this case it is not possible to determine the amount of random measurement error in this indicator. If an indicator is error-prone, the only possibility to account for the error is to utilize external knowledge about the reliability of this indicator to manually define the indicator's reliability.

A typical characteristic of SEM and factor-analytical tools in general is sign indeterminacy, in which the weight or loading estimates for a factor or a composite can only be determined jointly for their value but not for their sign. For example, if a factor is extracted from the strongly negatively correlated customer satisfaction indicators "How satisfied are you with provider X?" and "How much does provider X differ from an ideal provider?" the method cannot "know" whether the extracted

[2] This assumption should be relaxed in case of non-recursive models (Dijkstra and Henseler 2015a).

factor should correlate positively with the first or with the second indicator. Depending on the sign of the loadings, the meaning of the factor would either be "customer satisfaction" or "customer non-satisfaction." To avoid this ambiguity, it has become practice in SEM to determine one particular indicator per construct with which the construct scores are forced to correlate positively. Since this indicator dictates the orientation of the construct, it is called the "dominant indicator." While in covariance-based structural equation modeling this dominant indicator also dictates the construct's variance, in PLS path modeling, the construct variance is simply set to one.

Like multiple regression, PLS path modeling requires metric data for the dependent variables. Dependent variables are the indicators of the factor model(s) as well as the endogenous constructs. Quasi-metric data stemming from multipoint scales such as Likert scales or semantic differential scales are also acceptable as long as the scale points can be assumed to be equidistant and the number of scale points is sufficiently high (Rhemtulla et al. 2012). If these assumptions are not fulfilled, but the data are ordinal, researchers should rely on dedicated PLS-based approaches such as OrdPLSc (Schuberth et al. 2016; see also Chap. 6 of this book). To some extent it is also possible to include categorical variables in a model. Categorical variables (which are not necessarily ordinal) are particularly relevant for analyzing experiments (c.f. Streukens et al. 2010) or for control variables such as industry (Braojos-Gomez et al. 2015) or ownership structure (Chen et al. 2015). Figure 2.2 illustrates how a categorical variable "marital status" would be included in a PLS path model. If a categorical variable has only two levels (i.e., it is dichotomous), it can serve immediately as a construct indicator. If a categorical variable has more than two levels, it should be transformed into as many dummy variables as there are levels. A composite model is formed out of all but one dummy variable. The remaining dummy variable characterizes the reference level. Preferably, categorical variables should only play the role of exogenous variables in a structural model.

Sample size plays a dual role, namely, technically and in terms of inference statistics. Technically, the number of observations must be high enough that the regressions that form part of the PLS algorithm do not evoke singularities. It can

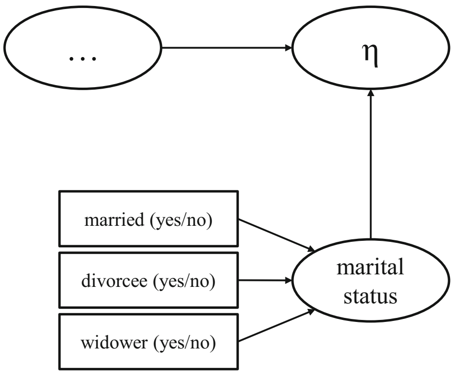


Fig. 2.2 Including a categorical control variable in a PLS path model (here marital status with the four categories "unmarried," "married," "divorcee," "widower"; the reference category is "unmarried")

thus be that the number of parameters or the number of variables in a model exceeds the number of observations. Inference statistics become relevant if an analyst wants to generalize from a sample to a population. The larger the sample size, the smaller the confidence intervals of the model's parameter estimates, and the smaller the chance that a parameter estimate's deviation from zero are due to sampling variation. Moreover, a larger sample size increases the likelihood to detect model misspecification (see Sect. 2.4 for PLS' tests of model fit). Hence, a larger sample size increases the rigor to falsify the model in the Popperian sense, but at the same time, the likelihood increases that a model gets rejected due to minor and hardly relevant aspects. The statistical power of PLS should not be expected to supersede that of covariance-based SEM.[3] Consequently, there is no reason to prefer PLS over other forms of SEM with regard to inference statistics. In research practice, there are typically many issues that have an impact on the final sample size. One important consideration should be the statistical power, i.e., the likelihood to find an effect in the sample if it indeed exists in the population. Optimally, researchers make use of Monte Carlo simulations to quantify the statistical power achieved at a certain sample size (for a tutorial, see Aguirre-Urreta and Rönkkö 2015).

2.4 Assessing and Reporting PLS Analyses

PLS path modeling can be used both for explanatory and predictive research. Depending on the analyst's aim—either explanation or prediction—the model assessment will be different. If the analyst's aim is to predict, the assessment should focus on blindfolding (M. Tenenhaus et al. 2005) and the model's performance with regard to holdout samples (Cepeda Carrión et al. 2016; Lancelot-Miltgen et al. 2016). However, since prediction-orientation still tends to be scarce in business research (Shmueli and Koppius 2011), in the remainder we will focus on model assessment if the analyst's aim is explanation.

PLS path modeling results can be assessed globally (i.e., for the overall model) and locally (for the measurement models and the structural model). For a long time, it was said that PLS path modeling did not optimize any global scalar and therefore did not allow for global model assessment. However, because PLS in the form as described above provides consistent estimates for factor and composite models, it is possible to meaningfully compare the model-implied correlation matrix with the empirical correlation matrix, which opens up the possibility for the assessment of global model fit.

[3] An allegedly higher statistical power of PLS (Reinartz et al. 2009) can be traced back to model misspecification, namely, making use of a composite model although the factor model would have been true (Goodhue et al. 2011).

The overall goodness of fit of the model should be the starting point of model assessment. If the model does not fit the data, the data contains more information than the model conveys. The obtained estimates may be meaningless, and the conclusions drawn from them become questionable. The global model fit can be assessed in two non-exclusive ways: by means of inference statistics, i.e., so-called tests of model fit, or through the use of fit indices, i.e., an assessment of approximate model fit. In order to have some frame of reference, it has become customary to determine the model fit both for the estimated model and for the saturated model. Saturation refers to the structural model, which means that in the saturated model all constructs correlate freely.

PLS path modeling's tests of model fit rely on the bootstrap to determine the likelihood of obtaining a discrepancy between the empirical and the model-implied correlation matrix that is as high as the one obtained for the sample at hand if the hypothesized model was indeed correct (Dijkstra and Henseler 2015a). Bootstrap samples are drawn from modified sample data. This modification entails an orthogonalization of all variables and a subsequent imposition of the model-implied correlation matrix. In covariance-based SEM, this approach is known as Bollen-Stine bootstrap (Bollen and Stine 1992). If more than 5% (or a different percentage if an alpha-level different from 0.05 is chosen) of the bootstrap samples yield discrepancy values above the ones of the actual model, it is not that unlikely that the sample data stems from a population that functions according to the hypothesized model. The model thus cannot be rejected. There is more than one way to quantify the discrepancy between two matrices, for instance, the maximum likelihood discrepancy, the geodesic discrepancy d_G, or the unweighted least squares discrepancy d_{ULS} (Dijkstra and Henseler 2015a), and so there are several tests of model fit. Monte Carlo simulations confirm that the tests of model fit can indeed discriminate between well-fitting and ill-fitting models (Henseler et al. 2014). More precisely, both measurement model misspecification and structural model misspecification can be detected through the tests of model fit (Dijkstra and Henseler 2014). Because it is possible that different tests have different results, a transparent reporting practice would always include several tests.

In addition to conducting the tests of model fit, it is also possible to determine the approximate model fit. Approximate model fit criteria help answer the question how substantial the discrepancy between the model-implied and the empirical correlation matrix is. This question is particularly relevant if this discrepancy is significant. Currently, the only approximate model fit criterion implemented for PLS path modeling is the standardized root mean square residual (SRMR, Hu and Bentler 1998, 1999). As can be derived from its name, the SRMR is the square root of the sum of the squared differences between the model-implied and the empirical correlation matrix, i.e., the Euclidean distance between the two matrices. A value of 0 for SRMR would indicate a perfect fit, and generally, an SRMR value less than 0.05 indicates an acceptable fit (Byrne 2013). A recent simulation study shows that even entirely correctly specified model can yield SRMR values of 0.06 and higher (Henseler et al. 2014). Therefore, a cutoff value of 0.08 as proposed by Hu and Bentler (1999) appears to be more adequate for PLS path models. Another useful approximate model fit criterion could be the Bentler-Bonett index or normed fit index (NFI, Bentler and Bonett 1980). The suggestion to use the NFI in connection with PLS path modeling can be attributed to Lohmöller (1989). For factor models, NFI values above 0.90 are considered as acceptable (Byrne 2013). For composite models, thresholds for the NFI are still to be determined. Because the NFI does not penalize for adding parameters, it should be used with caution for model comparisons. In general, the usage of the NFI is still rare.[4] Another promising approximate model fit criterion is the root mean square error correlation (RMS_{theta}, see Lohmöller 1989). A recent simulation study (Henseler et al. 2014) provides evidence that the RMS_{theta} can indeed distinguish well-specified from ill-specified models. However, thresholds for the RMS_{theta} are yet to be determined, and PLS software still needs to implement this approximate model fit criterion. Note that early suggestions for PLS-based goodness-of-fit measures such as the "goodness-of-fit" (GoF, see M. Tenenhaus et al. 2004) or the "relative goodness-of-fit" [GoF_{rel}, proposed by Esposito Vinzi et al. (2010)] are—contrary to what their name might suggest—not informative about the goodness of model fit (Henseler et al. 2014; Henseler and Sarstedt 2013). Consequently, there is no reason to evaluate and report them if the analyst's aim is to test or to compare models.

If the specified measurement (or outer) model does not possess minimum required properties of acceptable reliability and validity, then the structural (inner) model estimates become meaningless. That is, a necessary condition to even proceed to assess the "goodness" of the inner structural model is that the outer measurement model has already demonstrated acceptable levels of reliability and validity. There must be a sound measurement model before one can begin to assess the "goodness" of the inner structural model or to rely on the magnitude, direction, and/or statistical strength of the structural model's estimated parameters. Factor and composite models are assessed in a different way.

Factor models can be assessed in various ways. The bootstrap-based tests of overall model fit can indicate whether the data is coherent with a factor model, i.e., it represents a confirmatory factor analysis. In essence, the test of model fit provides an answer to the question "Does empirical evidence speak against the existence of the factor?" This quest for truth illustrates that testing factor models is rooted in the positivist research paradigm. If the test of overall model fit has not provided evidence against the existence of a factor,[5] several questions with regard to the factor structure emerge: Does the data support a factor structure at all? Is it clear that a factor can be extracted? How well has this factor been measured? Note that tests of overall model fit cannot answer these questions; in particular, entirely uncorrelated empirical variables do not necessarily lead to the rejection of the factor model. To

[4]For an application of the NFI, see Ziggers and Henseler (2016).

[5]Interestingly, the methodological literature on factor models is quite silent about what to do if the test speaks against a factor model. Some researchers suggest considering the alternative of a composite model, because it is less restrictive (Henseler et al. 2014) and not subject to factor indeterminacy (Rigdon 2012).

answer these questions, one should rather rely on several local assessment criteria with regard to the reliability and validity of measurement.

The amount of random error in construct scores should be acceptable, or in other words, the reliability of construct scores should be sufficiently high. Nunnally and Bernstein (1994) recommend a minimum reliability of 0.7. The most important reliability measure for PLS is rho_A (Dijkstra and Henseler 2015b); it currently is the only consistent reliability measure for PLS construct scores. Most PLS software also provides a measure of composite reliability (also called Dillon-Goldstein's rho, factor reliability, Jöreskog's rho, omega, or rho_c) as well as Cronbach's alpha. Both refer to sum scores, not construct scores. In particular, Cronbach's alpha typically underestimates the true reliability and should therefore only be regarded as a lower boundary to the reliability (Sijtsma 2009).

The measurement of factors should also be free from systematic measurement error. This quest for validity can be fulfilled in several non-exclusive ways. First, a factor should be unidimensional, a characteristic examined through convergent validity. The dominant measure of convergent validity is the average variance extracted (AVE, Fornell and Larcker 1981).[6] If the first factor extracted from a set of indicators explains more than one half of their variance, there cannot be any second, equally important factor. An AVE of 0.5 or higher is therefore regarded as acceptable. A somewhat more liberal criterion was proposed by Sahmer et al. (2006): They find evidence for unidimensionality as long as a factor explains significantly more variance than the second factor extracted from the same indicators. Second, each pair of factors that stand in for theoretically different concepts should also statistically be different, which raises the question of discriminant validity. Two criteria have been shown to be informative about discriminant validity (Voorhees et al. 2016): the Fornell-Larcker criterion (proposed by Fornell and Larcker 1981) and the heterotrait-monotrait ratio of correlations (HTMT, developed by Henseler et al. 2015). The Fornell-Larcker criterion says that a factor's AVE should be higher than its squared correlations with all other factors in the model. The HTMT is an estimate for the factor correlation (more precisely, an upper boundary). In order to clearly discriminate between two factors, the HTMT should be significantly smaller than one. Third, the cross-loadings should be assessed to make sure that no indicator is incorrectly assigned to a wrong factor.

The assessment of composite models is somewhat less developed. Again, the major point of departure should be the tests of model fit. The tests of model fit for the saturated model provide evidence for the external validity of the composites. Henseler et al. (2014) call this step a "confirmatory composite analysis." For composite models, the major research question is "Does it make sense to create this composite?" This different question shows that testing composite models

[6]The AVE must be calculated based on consistent loadings; otherwise, the assessment of convergent and discriminant validity based on the AVE is meaningless.

follows a different research paradigm, namely, pragmatism (Henseler 2015). Once confirmatory composite analysis has provided support for the composite, it can be analyzed further. One follow-up suggests itself: How is the composite made? Do all the ingredients contribute significantly and substantially? To answer these questions, an analyst should assess the sign and the magnitude of the indicator weights as well as their significance. Particularly if indicators weights have unexpected signs or are insignificant, this can be due to multicollinearity. It is therefore recommendable to assess the variance inflation factor (VIF) of the indicators. VIF values much higher than one indicate that multicollinearity might play a role.

Once the measurement model is deemed to be of sufficient quality, the analyst can proceed and assess the structural model. If OLS is used for the structural model, the endogenous constructs' R^2 values would be the point of departure. They indicate the % of variability accounted for by the precursor constructs in the model. The adjusted R^2 values take into account model complexity and sample size and are thus helpful to compare different models or the explanatory power of a model across different datasets.

If the analyst's aim is to generalize from a sample to a population, the path coefficients should be evaluated for significance. Inference statistics include the empirical bootstrap confidence intervals as well as one-sided or two-sided p-values. We recommend to use 4999 bootstrap samples. This number is sufficiently close to infinity for usual situations, is tractable with regard to computation time, and allows for an unambiguous determination of empirical bootstrap confidence intervals (for instance, the 2.5% [97.5%] quantile would be the 125th [4875th] element of the sorted list of bootstrap values). A path coefficient is regarded as significant (i.e., unlikely to purely result from sampling error) if its confidence interval does not include the value of zero or if the p-value is below the predefined alpha-level. Despite strong pleas for the use of confidence intervals (Cohen 1994), reporting p-values still seems to be more common in business research.

For the significant effects, it makes sense to quantify how substantial they are, which can be accomplished by assessing their effect size f^2. Values for f^2 above 0.35, 0.15, and 0.02 can be regarded as strong, moderate, and weak, respectively (Cohen 1988). The path coefficients are essentially standardized regression coefficients, which can be assessed with regard to their sign and their absolute size. They should be interpreted as the change in the dependent variable if the independent variable is increased by one and all other independent variables remain constant. Indirect effects and their inference statistics are important for mediation analysis (Zhao et al. 2010), and total effects are useful for success factor analysis (Albers 2010). Table 2.1 sums up the discussed criteria for model assessment.

Table 2.1 Assessment of PLS path modeling results in explanatory research settings

Assessment	Criterion
Overall model	
Test of model fit (estimated model)	SRMR < 95% bootstrap quantile (HI95 of SRMR) d_{ULS} < 95% bootstrap quantile (HI95 of d_{ULS}) d_{G} < 95% bootstrap quantile (HI95 of d_{G})
Approximate model fit (estimated model)	SRMR < 0.08
Measurement model	
Confirmatory composite and/or factor analysis (saturated model)	SRMR < 95% bootstrap quantile (HI95 of SRMR) d_{ULS} < 95% bootstrap quantile (HI95 of d_{ULS}) d_{G} < 95% bootstrap quantile (HI95 of d_{G})
Approximate model fit (saturated model)	SRMR < 0.08
Internal consistency reliability	Dijkstra-Henseler's $\rho_{\mathrm{A}} > 0.7$ Dillon-Goldstein's $\rho_{\mathrm{c}} > 0.7$ Cronbach's $\alpha > 0.7$
Convergent validity	AVE > 0.5
Discriminant validity	HTMT significantly smaller than 1 Fornell-Larcker criterion Loadings exceed cross-loadings
Structural model	
Endogenous variables	R^2, adjusted R^2
Direct effects	Path coefficient (absolute size, sign) Significance (p-value, confidence interval) Effect size
Indirect effects	Coefficient (absolute size, sign) Significance (p-value, confidence interval)
Total effects	Coefficient (absolute size, sign) Significance (p-value, confidence interval)

2.5 Extensions

PLS path modeling as described so far analyzes linear relationships between factors or composites of observed indicator variables. There are many ways how this rather basic model can be extended.

A first extension is to depart from the assumption of linearity. Researchers have developed approaches to include nonlinear relationships into the structural model. In particular, interaction effects and quadratic effects can be easily analyzed by means of some rudimentary extensions to the standard PLS path modeling setup (Dijkstra and Henseler 2011; Dijkstra and Schermelleh-Engel 2014; Henseler and Chin 2010; Henseler and Fassott 2010; Henseler et al. 2012). Interaction effects pay tribute to the fact that not all individuals function according to the same mechanism but that the strength of relationships depends on contingencies.

In addition to interaction effects, there are more comprehensive tools to take into account the heterogeneity between individuals. Heterogeneity can be observed, i.e., it can be traced back to an identified variable, or unobserved, i.e., there is no a priori explanation for why an individual's mechanism would differ from others. Because incorrectly assuming that all individuals function according to the same mechanism represents a validity thread (Becker et al. 2013b), several PLS-based approaches to discover unobserved heterogeneity have been proposed. Prominent examples include finite mixture PLS (Ringle et al. 2010a, c), PLS prediction-oriented segmentation (PLS-POS, Becker et al. 2013b), and PLS genetic algorithm segmentation (PLS-GAS, Ringle et al. 2010b, 2014). In order to assess observed heterogeneity, analysts should make use of multigroup analysis (Sarstedt et al. 2011). No matter whether heterogeneity is observed or unobserved, another concern for the analysts must be not to confound heterogeneity in the structural model with variation in measurement. Particularly in cross-cultural research, it has therefore become a common practice to assess the measurement model invariance before drawing conclusions about structural model heterogeneity. There is a plethora of papers discussing how to assess the measurement invariance of factor models (see, e.g., French and Finch 2006); there is only one approach for assessing the measurement invariance of composite models (Henseler et al. 2016).

2.6 Discussion

The plethora of discussions and developments around PLS path modeling called for a fresh look at this technique as well as new guidelines. As important aspect of this endeavor, we provide an answer to the question "What has changed?" This answer is given in Table 2.2, which contrasts traditional and modern perspectives on PLS. It is particularly helpful for researchers who have been educated in PLS path modeling in the past and who would like to update their understanding of the method.

The fact that PLS today strongly differs from how it used to be has also implications for the users of PLS software. They should verify that they use a software that has implemented the newest developments in the PLS field. One possibility would be ADANCO (Henseler and Dijkstra 2015), a new software for variance-based SEM, which also includes PLS path modeling.

The modularity of PLS path modeling as introduced in the second section opens up the possibility of replacing one or more steps by other approaches. For instance, the least squares estimators of the third step could be replaced by neural networks (Buckler and Hennig-Thurau 2008; Turkyilmaz et al. 2013). One could even replace the PLS algorithm in Step 1 by alternative indicator weight generators, such as principal component analysis (M. Tenenhaus 2008), generalized structured component analysis (Henseler 2012; Hwang and Takane 2004), regularized generalized canonical correlation analysis (A. Tenenhaus and Tenenhaus 2011), or even plain sum scores. Because in these instances the iterative PLS algorithm would not serve as eponym, one could not speak of PLS path modeling anymore. However, it still would be variance-based structural equation modeling.

Table 2.2 Contrasting traditional and modern perspectives on PLS

Traditional view on PLS	Modern view on PLS
PLS has some but not all abilities of structural equation modeling	PLS is a full-fledged structural equation modeling approach
PLS can estimate formative (using Mode B) and reflective measurement models (using Mode A)	PLS can consistently estimate composite models (using Mode B), formative models (MIMIC specification), and factor models (using consistent PLS for the latter)
Identification is not an issue for PLS	To ensure identification, analysts must provide a nomological net for each multi-item construct
PLS path models must be recursive	PLS path models can contain feedback loops or take into account endogeneity if an adequate estimator is used for the structural model. A sufficient number of exogenous variables must be available
PLS needs fewer observations than other SEM techniques	PLS does not need fewer observations than other techniques when it comes to inference statistics. Analysts should ensure sufficient statistical power and representativeness of data
In contrast to other SEM techniques, PLS does not rely on the assumption of normality	With regard to assumptions made for the estimation of parameters, PLS does not differ from other SEM techniques. For inference statistics, PLS applies a nonparametric technique, namely, bootstrapping, which can equally be applied by other SEM techniques
PLS only permits local model assessment by means of certain criteria	PLS path models can and should be assessed globally by means of tests of model fit and approximate measures of model fit. Models should be locally assessed, too
The reliability of PLS construct scores is indicated by Cronbach's alpha and/or composite reliability	The reliability coefficient rho_A is a consistent estimate of the reliability of PLS construct scores; composite reliability (based on consistent loadings) is a consistent estimate of the reliability of sum scores
Discriminant validity should be assessed by comparing each construct's average variance extracted with its squared construct correlations	Discriminant validity should be assessed by means of the heterotrait-monotrait ratio of correlations (HTMT) and by comparing each construct's average variance extracted (based on consistent loadings) with its squared consistent construct correlations
Bootstrapping should be conducted in combination with sign change correction in order to avoid inflated standard errors	For each construct, a dominant indicator should be defined in order to avoid sign indeterminacy

Finally, recent research confirms that PLS serves as a promising technique for prediction purposes (Becker et al. 2013a). Both measurement models and structural models can be assessed with regard to their predictive validity. Blindfolding is the standard approach used to examine if the model or a single effect of it can predict values of reflective indicators. It is already widely applied (Hair et al. 2012b; Ringle et al. 2012). Criteria for the predictive capability of structural models have been proposed (c.f. Chin 2010) but still need to disseminate. We anticipate that once business and social science researchers' interest in prediction becomes more pronounced, PLS will face an additional substantial increase in popularity.

References

Aguirre-Urreta, M., & Rönkkö, M. (2015). Sample size determination and statistical power analysis in PLS using R: An annotated tutorial. *Communications of the Association for Information Systems, 36*(3), 33–51.

Albers, S. (2010). PLS and success factor studies in marketing. In V. Esposito Vinzi, W. W. Chin, J. Henseler, & H. Wang (Eds.), *Handbook of partial least squares* (pp. 409–425). Berlin: Springer.

Antonakis, J., Bendahan, S., Jacquart, P., & Lalive, R. (2010). On making causal claims: A review and recommendations. *The Leadership Quarterly, 21*(6), 1086–1120. doi:10.1016/j.leaqua.2010.10.010.

Becker, J.-M., Rai, A., & Rigdon, E. E. (2013a). *Predictive validity and formative measurement in structural equation modeling: Embracing practical relevance*. Proceedings of the International Conference on Information Systems (ICIS), Milan, Italy.

Becker, J.-M., Rai, A., Ringle, C. M., & Völckner, F. (2013b). Discovering unobserved heterogeneity in structural equation models to avert validity threats. *MIS Quarterly, 37*(3), 665–694.

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin, 88*(3), 588–606. doi:10.1037/0033-2909.88.3.588.

Bentler, P. M., & Huang, W. (2014). On components, latent variables, PLS and simple methods: Reactions to Rigdon's rethinking of PLS. *Long Range Planning, 47*(3), 138–145. doi:10.1016/j.lrp.2014.02.005.

Bollen, K. A., & Stine, R. A. (1992). Bootstrapping goodness-of-fit measures in structural equation models. *Sociological Methods & Research, 21*(2), 205–229. doi:10.1177/0049124192021002004.

Braojos-Gomez, J., Benitez-Amado, J., & Llorens-Montes, F. J. (2015). How do small firms learn to develop a social media competence? *International Journal of Information Management, 35*(4), 443–458. doi:10.1016/j.ijinfomgt.2015.04.003.

Buckler, F., & Hennig-Thurau, T. (2008). Identifying hidden structures in marketing's structural models through universal structure modeling. An explorative Bayesian neural network complement to LISREL and PLS Market. *Marketing: Journal of Research and Management, 4*, 47–66.

Byrne, B. M. (2013). *Structural equation modeling with LISREL, PRELIS, and SIMPLIS: Basic concepts, applications, and programming*. Stanford, CA: Psychology Press.

Cepeda Carrión, G., Henseler, J., Ringle, C. M., & Roldán, J. L. (2016). Prediction-oriented modeling in business research by means of PLS path modeling. *Journal of Business Research, 69*(10), 4545–4551. doi:10.1016/j.jbusres.2016.03.048.

Chen, Y., Wang, Y., Nevo, S., Benitez-Amado, J., & Kou, G. (2015). IT capabilities and product innovation performance: The roles of corporate entrepreneurship and competitive intensity. *Information Management, 52*(6), 643–657.

Chin, W. W. (2010). Bootstrap cross-validation indices for PLS path model assessment. In V. Esposito Vinzi, W. W. Chin, J. Henseler, & H. Wang (Eds.), *Handbook of partial least squares: Concepts, methods and applications* (pp. 83–97). New York: Springer.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Mahwah, NJ: Lawrence Erlbaum.

Cohen, J. (1994). The earth is round (p<.05). *The American Psychologist, 49*(12), 997–1003. doi:10.1037/0003-066X.49.12.997.

Diamantopoulos, A., Sarstedt, M., Fuchs, C., Wilczynski, P., & Kaiser, S. (2012). Guidelines for choosing between multi-item and single-item scales for construct measurement: A predictive validity perspective. *Journal of Academy of Market Science, 40*(3), 434–449. doi:10.1007/s11747-011-0300-3.

Dijkstra, T. K. (2010). Latent variables and indices: Herman Wold's basic design and partial least squares. In V. Esposito Vinzi, W. W. Chin, J. Henseler, & H. Wang (Eds.), *Handbook of partial least squares: Concepts, methods and applications* (pp. 23–46). New York: Springer.

Dijkstra, T. K., & Henseler, J. (2011). Linear indices in nonlinear structural equation models: Best fitting proper indices and other composites. *Quality and Quantity, 45*(6), 1505–1518. doi:10.1007/s11135-010-9359-z.

Dijkstra, T. K., & Henseler, J. (2014). *Assessing and testing the goodness-of-fit of PLS path models*. Third VOC Conference, Leiden.

Dijkstra, T. K., & Henseler, J. (2015a). Consistent and asymptotically normal PLS estimators for linear structural equations. *Computational Statistics and Data Analysis, 81*(1), 10–23. doi:10.1016/j.csda.2014.07.008.

Dijkstra, T. K., & Henseler, J. (2015b). Consistent partial least squares path modeling. *MIS Quarterly, 39*(2), 297–316.

Dijkstra, T. K., & Schermelleh-Engel, K. (2014). Consistent partial least squares for nonlinear structural equation models. *Psychometrika, 79*(4), 585–604. doi:10.1007/S11336-013-9370-0.

Esposito Vinzi, V., Trinchera, L., & Amato, S. (2010). PLS path modeling: From foundations to recent developments and open issues for model assessment and improvement. Chap. 2. In V. E. Vinzi, W. W. Chin, J. Henseler, & H. Wang (Eds.), *Handbook of partial least squares: Concepts, methods and applications* (pp. 47–82). Berlin: Springer.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*(1), 39–50. doi:10.2307/3151312.

French, B. F., & Finch, W. H. (2006). Confirmatory factor analytic procedures for the determination of measurement invariance. *Structural Equation Modeling, 13*(3), 378–402.

Goodhue, D. L., Lewis, W., & Thompson, R. L. (2011). *A dangerous blind spot in IS research: False positives due to multicollinearity combined with measurement error*. Paper presented at the AMCIS 2011, Detroit, MI.

Hair, J. F., Ringle, C. M., & Sarstedt, M. (2011). PLS-SEM: Indeed a silver bullet. *Journal of Marketing Theory and Practice, 19*(2), 139–152.

Hair, J. F., Sarstedt, M., Pieper, T. M., & Ringle, C. M. (2012a). The use of partial least squares structural equation modeling in strategic management research: A review of past practices and recommendations for future applications. *Long Range Planning, 45*(5–6), 320–340. doi:10.1016/j.lrp.2012.09.008.

Hair, J. F., Sarstedt, M., Ringle, C. M., & Mena, J. A. (2012b). An assessment of the use of partial least squares structural equation modeling in marketing research. *Journal of Academy of Market Science, 40*(3), 414–433. doi:10.1007/s11747-011-0261-6.

Henseler, J. (2010). On the convergence of the partial least squares path modeling algorithm. *Computational Statistics, 25*(1), 107–120. doi:10.1007/s00180-009-0164-x.

Henseler, J. (2012). Why generalized structured component analysis is not universally preferable to structural equation modeling. *Journal of Academy of Market Science, 40*(3), 402–413. doi:10.1007/s11747-011-0298-6.

Henseler, J. (2015). *Is the whole more than the sum of its parts? On the interplay of marketing and design research*. Inaugural Lecture, University of Twente.

Henseler, J., & Chin, W. W. (2010). A comparison of approaches for the analysis of interaction effects between latent variables using partial least squares path modeling. *Structural Equation Modeling, 17*(1), 82–109. doi:10.1080/10705510903439003.

Henseler, J., & Dijkstra, T. K. (2015). *ADANCO 2.0*. Kleve: Composite Modeling GmbH. Retrieved from http://www.compositemodeling.com.

Henseler, J., & Fassott, G. (2010). Testing moderating effects in PLS path models: An illustration of available procedures. In V. Esposito Vinzi, W. W. Chin, J. Henseler, & H. Wang (Eds.), *Handbook of partial least squares: Concepts, methods and applications* (pp. 713–735). Berlin: Springer.

Henseler, J., & Sarstedt, M. (2013). Goodness-of-fit indices for partial least squares path modeling. *Computational Statistics, 28*(2), 565–580. doi:10.1007/s00180-012-0317-1.

Henseler, J., Fassott, G., Dijkstra, T. K., & Wilson, B. (2012). Analysing quadratic effects of formative constructs by means of variance-based structural equation modelling. *European Journal of Information Systems, 21*(1), 99–112. doi:10.1057/ejis.2011.36.

Henseler, J., Dijkstra, T. K., Sarstedt, M., Ringle, C. M., Diamantopoulos, A., et al. (2014). Common beliefs and reality about PLS: Comments on Rönkkö & Evermann (2013). *Organizational Research Methods, 17*(2), 182–209. doi:10.1177/1094428114526928.

Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of Academy of Market Science, 43*(1), 115–135. doi:10.1007/s11747-014-0403-8.

Henseler, J., Ringle, C. M., & Sarstedt, M. (2016). Testing measurement invariance of composites using partial least squares. *International Marketing Review, 33*(3), 405–431. doi:10.1108/IMR-09-2014-0304.

Höök, K., & Löwgren, J. (2012). Strong concepts: Intermediate-level knowledge in interaction design research. *ACM Transactions on Computer-Human Interaction (TOCHI), 19*(3). doi:10.1145/2362364.2362371.

Hu, L.-T., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods, 3*(4), 424–453. doi:10.1037/1082-989X.3.4.424.

Hu, L.-T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1–55. doi:10.1080/10705519909540118.

Hwang, H., & Takane, Y. (2004). Generalized structured component analysis. *Psychometrika, 69*(1), 81–99. doi:10.1007/BF02295841.

Kettenring, J. R. (1971). Canonical analysis of several sets of variables. *Biometrika, 58*(3), 433–451. doi:10.1093/biomet/58.3.433.

Ketterlinus, R. D., Bookstein, F. L., Sampson, P. D., & Lamb, M. E. (1989). Partial least squares analysis in developmental psychopathology. *Development and Psychopathology, 1*(4), 351–371.

Krijnen, W. P., Dijkstra, T. K., & Gill, R. D. (1998). Conditions for factor (in)determinacy in factor analysis. *Psychometrika, 63*(4), 359–367. doi:10.1007/BF02294860.

Lancelot-Miltgen, C., Henseler, J., Gelhard, C., & Popovič, A. (2016). Introducing new products that affect consumer privacy: A mediation model. *Journal of Business Research, 69*(10), 4659–4666.

Landis, R. S., Beal, D. J., & Tesluk, P. E. (2000). A comparison of approaches to forming composite measures in structural equation models. *Organizational Research Methods, 3*(2), 186–207.

Lohmöller, J.-B. (1989). *Latent variable path modeling with partial least squares*. Berlin: Springer.

Maraun, M. D., & Halpin, P. F. (2008). Manifest and latent variates. *Measurement: Interdisciplinary Research and Perspectives, 6*(1–2), 113–117.

Marcoulides, G. A., & Saunders, C. (2006). PLS: A silver bullet? *MIS Quarterly, 30*(2), iii–iix.

McDonald, R. P. (1996). Path analysis with composite variables. *Multivariate Behavioral Research, 31*(2), 239–270. doi:10.1207/s15327906mbr3102_5.

McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah, NJ: Erlbaum.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.

Reinartz, W. J., Haenlein, M., & Henseler, J. (2009). An empirical comparison of the efficacy of covariance-based and variance-based SEM. *International Journal of Research in Marketing, 26*(4), 332–344. doi:10.1016/j.ijresmar.2009.08.001.

Rhemtulla, M., Brosseau-Liard, P. E., & Savalei, V. (2012). When can categorical variables be treated as continuous? A comparison of robust continuous and categorical SEM estimation methods under suboptimal conditions. *Psychological Methods, 17*(3), 354–373. doi:10.1037/a0029315.

Rigdon, E. E. (2012). Rethinking partial least squares path modeling: In praise of simple methods. *Long Range Planning, 45*(5–6), 341–358. doi:10.1016/j.lrp.2012.09.010.

Rigdon, E. E. (2014). Rethinking partial least squares path modeling: Breaking chains and forging ahead. *Long Range Planning, 47*(3), 161–167. doi:10.1016/j.lrp.2014.02.003.

Rigdon, E. E., Becker, J.-M., Rai, A., Ringle, C. M., et al. (2014). Conflating antecedents and formative indicators: A comment on Aguirre-Urreta and Marakas. *Information Systems Research, 25*(4), 780–784. doi:10.1287/isre.2014.0543.

Rindskopf, D. (1984). Using phantom and imaginary latent variables to parameterize constraints in linear structural models. *Psychometrika, 49*(1), 37–47. doi:10.1007/BF02294204.

Ringle, C. M., Sarstedt, M., & Mooi, E. A. (2010a). Response-based segmentation using finite mixture partial least squares: Theoretical foundations and an application to American customer satisfaction index data. *Annals of Information Systems, 8*, 19–49.

Ringle, C. M., Sarstedt, M., & Schlittgen, R. (2010b). Finite mixture and genetic algorithm segmentation in partial least squares path modeling: Identification of multiple segments in a complex path model. In A. Fink, B. Lausen, W. Seidel, & A. Ultsch (Eds.), *Advances in data analysis, data handling and business intelligence* (pp. 167–176). Berlin: Springer.

Ringle, C. M., Wende, S., & Will, A. (2010c). Finite mixture partial least squares analysis: Methodology and numerical examples. In V. Esposito Vinzi, W. W. Chin, J. Henseler, & H. Wang (Eds.), *Handbook of partial least squares: Concepts, methods and applications*, *Springer handbooks of computational statistics series* (Vol. II, pp. 195–218). New York: Springer.

Ringle, C. M., Sarstedt, M., & Straub, D. W. (2012). Editor's comments: A critical look at the use of PLS-SEM in MIS Quarterly. *MIS Quarterly, 36*(1), iii–xiv.

Ringle, C. M., Sarstedt, M., & Schlittgen, R. (2014). Genetic algorithm segmentation in partial least squares structural equation modeling. *OR Spectrum, 36*(1), 251–276.

Sahmer, K., Hanafi, M., & Qannari, M. (2006). Assessing unidimensionality within the PLS path modeling framework. In M. Spiliopoulou, R. Kruse, C. Borgelt, A. Nürnberger, & W. Gaul (Eds.), *From data and information analysis to knowledge engineering* (pp. 222–229). Berlin: Springer.

Sarstedt, M., Henseler, J., & Ringle, C. (2011). Multi-group analysis in partial least squares (PLS) path modeling: Alternative methods and empirical results. *Advances in International Marketing, 22*, 195–218. doi:10.1108/S1474-7979(2011)0000022012.

Sarstedt, M., Ringle, C. M., Henseler, J., & Hair, J. F. (2014). On the emancipation of PLS-SEM: A commentary on Rigdon (2012). *Long Range Planning, 47*(3), 154–160. doi:10.1016/j.lrp.2014.02.007.

Schuberth, F., Henseler, J., & Dijkstra, T. K. (2016). Partial least squares path modeling using ordinal categorical indicators. *Quality and Quantity*, 1–27. doi:10.1007/s11135-016-0401-7.

Shmueli, G., & Koppius, O. R. (2011). Predictive analytics in information systems research. *MIS Quarterly, 35*(3), 553–572.

Sijtsma, K. (2009). On the use, the misuse, and the very limited usefulness of Cronbach's alpha. *Psychometrika, 74*(1), 107–120. doi:10.1007/s11336-008-9101-0.

Streukens, S., Wetzels, M., Daryanto, A., & de Ruyter, K. (2010). Analyzing factorial data using PLS: Application in an online complaining context. In V. Esposito Vinzi, W. W. Chin, J. Henseler, & H. Wang (Eds.), *Handbook of partial least squares: Concepts, methods and applications* (pp. 567–587). New York: Springer.

Tenenhaus, M. (2008). Component-based structural equation modelling. *Total Quality Management and Business Excellence, 19*(7), 871–886. doi:10.1080/14783360802159543.

Tenenhaus, A., & Tenenhaus, M. (2011). Regularized generalized canonical correlation analysis. *Psychometrika, 76*(2), 257–284. doi:10.1007/S11336-011-9206-8.

Tenenhaus, M., Amato, S., & Esposito Vinzi, V. (2004). *A global goodness-of-fit index for PLS structural equation modelling* (pp. 739–742). Proceedings of the XLII SIS Scientific Meeting, CLEUP, Padova.

Tenenhaus, M., Esposito Vinzi, V., Chatelin, Y.-M., & Lauro, C. (2005). PLS path modeling. *Computational Statistics and Data Analysis, 48*(1), 159–205. doi:10.1016/j.csda.2004.03.005.

Turkyilmaz, A., Oztekin, A., Zaim, S., & Fahrettin Demirel, O. (2013). Universal structure modeling approach to customer satisfaction index. *Industrial Management & Data Systems, 113*(7), 932–949. doi:10.1108/IMDS-12-2012-0444.

Voorhees, C. M., Brady, M. K., Calantone, R., & Ramirez, E. (2016). Discriminant validity testing in marketing: An analysis, causes for concern, and proposed remedies. *Journal of Academy of Market Science, 44*(1), 119–134. doi:10.1007/s11747-015-0455-4.

Wold, H. (1974). Causal flows with latent variables: Partings of the ways in the light of NIPALS modelling. *European Economic Review, 5*(1), 67–86. doi:10.1016/0014-2921(74)90008-7.

Wold, H. (1982). Soft modeling: The basic design and some extensions. In K. G. Jöreskog & H. O. A. Wold (Eds.), *Systems under indirect observations: Part II* (pp. 1–54). Amsterdam: North-Holland.

Zhao, X., Lynch, J. G., & Chen, Q. (2010). Reconsidering Baron and Kenny: Myths and truths about mediation analysis. *Journal of Consumer Research, 37*(2), 197–206. doi:10.1086/651257.

Ziggers, G.-W., & Henseler, J. (2016). The reinforcing effect of a firm's customer orientation and supply-base orientation on performance. *Industrial Marketing Management, 52*(1), 18–26. doi:10.1016/j.indmarman.2015.07.011.

Chapter 3
Going Beyond Composites: Conducting a Factor-Based PLS-SEM Analysis

Ned Kock

Abstract There has been a long and ongoing debate, at points resembling an acrimonious dispute, among proponents and detractors of the use of the partial least squares (PLS) approach for structural equation modeling (SEM). The composite-factor estimation dichotomy has been the epicenter of this debate. In this chapter, we briefly discuss the implementation of a new method to conduct factor-based PLS-SEM analyses, which could be a solid step in the resolution of this debate. This method generates estimates of both true composites and factors, in two stages, fully accounting for measurement error. Our discussion is based on an illustrative model in the field of e-collaboration. A Monte Carlo experiment suggests that model parameters generated by the method are asymptotically unbiased. The method is implemented as part of the software WarpPLS, starting in version 5.0. This chapter provides enough details for the method's implementation in other venues such as R and GNU Octave.

3.1 Introduction

Structural equation modeling (SEM) is extensively used in many areas of research, including various business disciplines, as well as the social and behavioral sciences. The techniques underlying SEM are relevant for the incipient field of business data analytics. SEM employs latent variables, which are measured indirectly through "observed" or "manifest" variables, in sets associated with latent variables that are normally called "indicators."

The measurement of latent variables via indicators obtained from the administration of questionnaires includes error. Latent variables typically refer to perception-based constructs (e.g., satisfaction with one's job). Indicators normally store numeric answers to sets of questions in questionnaires, each set designed to refer to a latent variable and expected to measure it with a certain degree of imprecision.

N. Kock (✉)
Division of International Business and Technology Studies, Texas A&M International University, 5201 University Boulevard, Laredo, TX 78041, USA
e-mail: nedkock@gmail.com

H. Latan, R. Noonan (eds.), *Partial Least Squares Path Modeling*,
DOI 10.1007/978-3-319-64069-3_3

There has been an ongoing debate among proponents and detractors of the use of Wold's partial least squares (PLS) method (Adelman and Lohmöller 1994; Lohmöller 1989; Wold 1980) in the context of SEM. This debate has been at points acrimonious and has been going on for a long time. So far, it shows no signs of resolution. It arises from common factor model assumptions, which form the basis on which covariance-based SEM (CB-SEM) rests (Kline 2010; Mueller 1996). The debate is centered around two main issues.

The first issue is that Wold's original PLS design for "soft" SEM has a number of advantages over CB-SEM, such as minimal model identification demands, practically no data or model parameter distribution assumptions, virtually universal convergence to solutions, and estimation of "pseudo-factors." The latter, "pseudo-factors," provide a partial solution to the factor indeterminacy problem of CB-SEM.

The second issue fueling the debate is that the original PLS design does not base its model parameter estimation methods on the estimation of true factors. Estimation is based on "composites," which are exact linear combinations of indicators and are referred to above as "pseudo-factors." The composite estimates generated by the original PLS design can be conceptually seen as factors minus their corresponding measurement errors. Reliance on them leads to biased model parameter estimates (notably path coefficients and loadings) even as sample sizes grow to infinity (Kock 2014b).

In this chapter, we briefly describe what could be a solid step in the resolution of this debate, although it may open new avenues for debate on different issues. We show how researchers can implement what we refer to as "factor-based PLS-SEM" (PLSF-SEM). This new method generates estimates of both true composites and factors, in two stages, fully accounting for measurement error.

The PLSF-SEM method is implemented starting in version 5.0 of WarpPLS (Kock 2015). WarpPLS is an SEM software tool that is unique in that it enables nonlinear analyses where best-fitting nonlinear functions are estimated for each pair of structurally linked variables in path models and subsequently used (i.e., the nonlinear functions) to estimate path coefficients that take into account the nonlinearity. Moreover, WarpPLS provides a comprehensive set of model fit and quality indices that are compatible with both composite-based and factor-based SEM.

In our discussion all variables are assumed to be standardized, i.e., scaled to have a mean of zero and standard deviation of one. This has no impact on the generality of the method or of the discussion. All standardized variables can be rescaled back to their original scales. Also, our discussion builds on common factor model assumptions, a key one being that factors cause indicators. This assumption is strongly anchored in an epistemological argument regarding data collection via questionnaires: question-statements associated with indicators are developed by researchers based on mental representations of factors that exist prior to that development. Therefore, even with multidimensional factors (sometimes referred to as "formative"), the factors exist before the indicators. Given this, the idea that indicators may cause factors is questionable when data is obtained via questionnaires.

3.2 Measurement Error and the Attenuation Bias

Figure 3.1 shows two correlated factors F_1 and F_2 with three indicators each. Even though the indicators "reflect" their common factors (top part of figure), the factors can also be seen as aggregations of their respective indicators and measurement errors (bottom part of figure). In each factor the measurement error is uncorrelated with the factor's indicators.

Note that even though factor-to-indicator causality is assumed, weights do exist and factors can be seen as akin to "composites" that aggregate both indicators and measurement errors. The measurement error that is thus aggregated in each factor could be viewed as an "extra" indicator that (a) is uncorrelated with the actual indicators and (b) accounts for the variance in the factor that is not explained by the actual indicators.

The two correlated factors F_1 and F_2 can be expressed as weighted sums of their corresponding true composites C_1 and C_2 and measurement errors ε_1 and ε_2. These true composites are estimated in a specific way, as we will see later, and are generally not the same as the composites estimated via PLS algorithms. The weights ω_{1C} and ω_{2C} are applied to the true composites, and $\omega_{1\varepsilon}$ and $\omega_{2\varepsilon}$ are the measurement error weights. The true reliabilities α_1 and α_2 equal the corresponding true composite weights squared: ${\omega_{1C}}^2$ and ${\omega_{2C}}^2$. Since the standardized base measurement errors and true composites are uncorrelated, it follows that the measurement error weights and $\omega_{1\varepsilon}$ and $\omega_{2\varepsilon}$ equal $\sqrt{1-\alpha_1}$ and $\sqrt{1-\alpha_2}$, respectively. The factors F_1 and F_2 are correlated. Therefore their composites and measurement errors are cross-correlated, even though composites and measurement errors that refer to the same factor are uncorrelated. That is, even though $r(C_i, \varepsilon_i) = 0$, we have $r(C_i, \varepsilon_j) \neq 0$, $r(C_i, C_j) \neq 0$ and $r(\varepsilon_i, \varepsilon_j) \neq 0$. These nonzero cross-correlations are represented in Fig. 3.2.

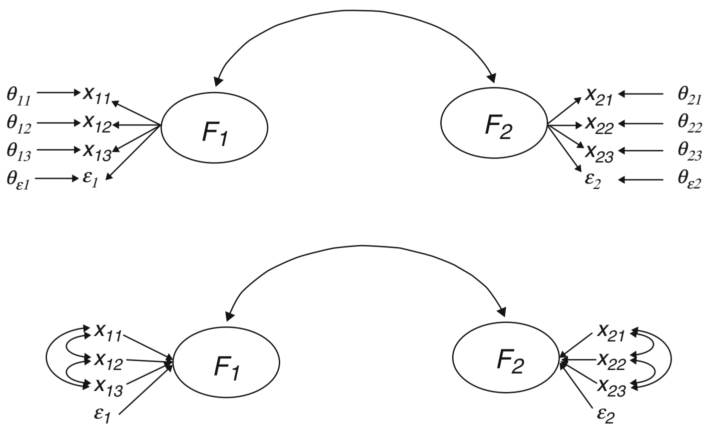


Fig. 3.1 Measurement errors for any pair of correlated factors. Note: factors (latent variables) are represented within *ovals*; the equivalent graph for composites would have the errors ε_1 and ε_2 removed

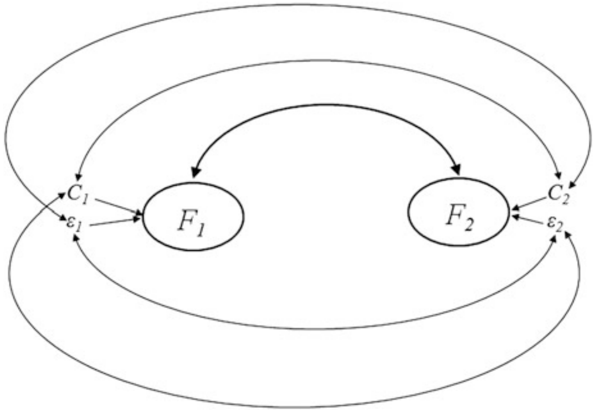


Fig. 3.2 Cross-correlation of measurement errors and composites. Note: factors, composites, and measurement errors are cross-correlated, but composites and measurement errors that refer to the same factor are not

The idea that measurement errors can give rise to an increase in the strength of the correlations between two factors is counterintuitive at first. Generally speaking, the presence of error tends to lead to a decrease in the strength of correlations. The discussion above, however, illustrates why the measurement errors associated with the factors F_1 and F_2 are important in making the strength of the correlation between the factors greater than the strength of the correlation between the corresponding composites.

The nonzero correlations $r(C_1, \varepsilon_2)$, $r(C_2, \varepsilon_1)$, and $r(\varepsilon_1, \varepsilon_2)$ contribute additively, together with $r(C_1, C_2)$, to the correlation between the factors $r(F_1, F_2)$. This is why the absolute correlation $|r(C_1, C_2)|$ between the true composites is lower than the absolute correlation $|r(F_1, F_2)|$ between the factors and ultimately why PLS-SEM tends to underestimate path coefficients. This is a fundamental problem, with obvious implications for practitioners, that our PLSF-SEM method aims to address.

3.3 Illustrative Model

Our discussion is based on the illustrative model depicted in Fig. 3.3, which builds on an actual empirical study in the field of e-collaboration (Kock 2005, 2008; Kock and Lynn 2012). This illustrative model incorporates the belief that *e-collaboration technology use* (F_1) by teams of workers tasked with the development of new products in organizations (e.g., a new consulting service, a new car part) increases both *team efficiency* (F_2) and *team performance* (F_3). *Team efficiency* (F_2) is related to the speed and cost at which teams operate. *Team performance* (F_3) is related to how well the new products developed by teams perform in terms of sales and profits.

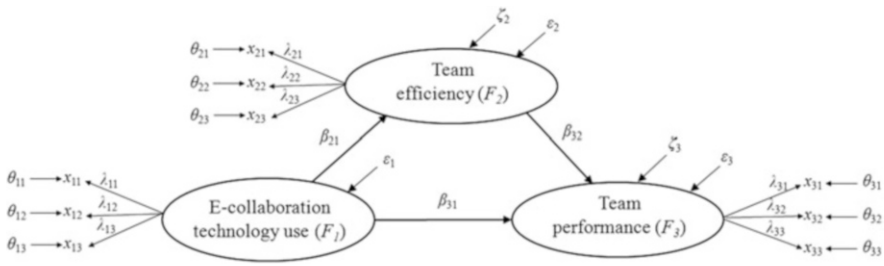


Fig. 3.3 Illustrative model

In this illustrative model, β_{ij} is the path coefficient for the link going from factor F_j to factor F_i; λ_{ij} is the loading for the jth indicator of factor F_i; θ_{ij} is the indicator error for the jth indicator of factor F_i; ε_i is the measurement error associated with F_i; and ζ_i is the structural error associated with F_i, which exists only for endogenous factors. An endogenous factor has at least one other factor pointing at it in the model.

Note that the measurement errors ε_i are not the same as the structural errors ζ_i. Measurement errors exist for any factors that are measured with a certain degree of imprecision, whether the factors are exogenous or endogenous. Structural errors exist only for endogenous factors. Analogously, the measurement errors ε_i should not be confused with the indicator errors θ_{ij}, even though these two types of errors are related. The former arise due to the existence of the latter and can be seen as "extra" indicators that account for the explained variances in their respective factors that are not accounted for by the actual factor indicators.

3.4 PLSF-SEM's First Stage: Composites

PLSF-SEM's first stage yields initial estimates of the composites. These estimates are used in the method's second stage, where factors and other model parameter estimates are produced. It starts by setting weights and loadings as 1 (reversed indicators must be properly adjusted) and initializing the composite estimates with a standardized vector of the summed indicators. Then measurement errors $\widehat{\varepsilon}_i$, reliabilities $\widehat{\alpha}_i$, measurement error weights $\widehat{\omega}_{i\varepsilon}$, and composite weights $\widehat{\omega}_{iC}$ are set as indicated in Eqs. (3.1)–(3.4).

$$\widehat{\varepsilon}_i := \mathrm{Stdz}\left[\mathrm{Rnd}(N)\right]. \tag{3.1}$$

$$\widehat{\alpha}_i := \frac{n_i \overline{\Sigma}_{x_i x_i}}{\left[1 + (n_i - 1)\, \overline{\Sigma}_{x_i x_i}\right]}. \tag{3.2}$$

$$\widehat{\omega}_{i\varepsilon} := \sqrt{1 - \widehat{\alpha}_i}. \tag{3.3}$$

$$\widehat{\omega}_{iC} := \sqrt{\widehat{\alpha}_i}. \tag{3.4}$$

In these equations, Rnd(N) is a function that returns an independent and identically distributed (i.i.d.) variable with N rows, with N being the sample size; Stdz(•) is a function that returns a standardized column vector; n_i is the number of indicators of factor F_i; x_i is a matrix with N rows and with each column referring to one of the indicators associated with F_i; and $\overline{\Sigma}_{x_i x_i}$ is the mean of the nonredundant correlation coefficients among the column vectors that make up x_i (e.g., the mean of the lower triangular version of $\Sigma_{x_i x_i}$).

Technical readers will notice that the reliability estimate $\widehat{\alpha}_i$ above is the Cronbach's alpha coefficient (Cronbach 1951; Kline 2010). We are aware that serious questions have been raised regarding Cronbach's alpha's psychometric properties. However, while the PLSF-SEM method uses the Cronbach's alpha coefficient as a basis for the estimation of measurement error and composite weights, it makes no assumptions about the coefficient's main purported psychometric properties that have been the target of criticism (Sijtsma 2009). This is an important distinction in light of measurement error theory (Nunnally and Bernstein 1994).

Moreover, we developed and tested a number of experimental versions of the PLSF-SEM method prior to writing this chapter, using various reliability estimates. The versions employing the Cronbach's alpha coefficient tended to yield the best results. Arguably employing Cronbach's alpha coefficients tended to yield good results in our simulations because the Cronbach's alpha coefficients provide good estimates of the true reliabilities when the degree of heterogeneity among the loadings in each latent variable is low. This appears to frequently be the case in practice. Two alternatives can be employed when this is not the case:

- To use the reliability estimates generated by Dijkstra's consistent PLS (a.k.a. PLSc) technique, which appear to be closer to the true reliabilities than Cronbach's alpha coefficients under high-loading heterogeneity conditions. Here Dijkstra's consistent PLS' reliabilities would be used in place of the Cronbach's alpha coefficients.
- To use Cronbach's alpha coefficients as initial reliability estimates and iterate across the two stages of the PLSF-SEM method, where the two stages are contained within an outermost loop. This outermost loop is responsible for convergence toward an asymptotically unbiased reliability measure. In each iteration the reliability estimates for each latent variable i would be adjusted to $1/2\left(\widehat{\omega}_i{}'\widehat{\lambda}_i + \widehat{\rho}_i\right)$, where $\widehat{\rho}_i$ is the composite reliability estimate for latent variable i.

PLSF-SEM's first stage then proceeds by iteratively estimating factors $\widehat{F}_i$, matrices $\widehat{\theta}_i$ with N rows and with each column storing one of the indicator error terms, column vectors of weights $\widehat{\omega}_i$, composites $\widehat{C}_i$, and column vectors of loadings $\widehat{\lambda}_i$ according to Eqs. (3.5)–(3.9). Convergence is achieved when the sum of the absolute differences between successive estimates of the matrix of loadings for the entire model $\widehat{\lambda}$ changes by less than a small fraction.

$$\widehat{F}_i := \text{Stdz}\left(\widehat{C}_i\widehat{\omega}_{iC} + \widehat{\varepsilon}_i\widehat{\omega}_{i\varepsilon}\right). \tag{3.5}$$

$$\widehat{\theta}_i := x_i - \widehat{F}_i\widehat{\lambda}_i{}'. \tag{3.6}$$

$$\widehat{\omega}_i := \Sigma_{x_i x_i}{}^{-1}\left[\Sigma_{x_i x_i} - \text{diag}\left(\Sigma_{x_i\widehat{\theta}_i}\right)\right]\widehat{\lambda}_i{}'^{+}. \tag{3.7}$$

$$\widehat{C}_i := \frac{1}{\widehat{\omega}_{iC}}\left(x_i\widehat{\omega}_i\right). \tag{3.8}$$

$$\widehat{\lambda}_i := \left(\widehat{C}_i{}^{+}x_i\right)'\widehat{\omega}_{iC}. \tag{3.9}$$

In these equations diag$\left(\Sigma_{x_i\widehat{\theta}_i}\right)$ is the diagonal matrix of *covariances* among the indicators and corresponding error terms, and the superscript + denotes the Moore-Penrose pseudoinverse transformation. It is useful to observe that diag$\left(\Sigma_{x_i\widehat{\theta}_i}\right)$ is a diagonal matrix because in the common factor model, $\Sigma_{x_{ij}\theta_{ij}} = 0$ for all $i \neq j$. That is, in the common factor model, indicator error terms are correlated with their corresponding indicators and uncorrelated with other indicators in the same factor.

Researchers familiar with the mathematics underlying PLS will see that the estimation steps above differ significantly from those employed in Wold's original PLS design (Adelman and Lohmöller 1994; Lohmöller 1989; Wold 1980). Particularly noteworthy is that the estimation steps above incorporate significantly more information in defining the relationships among weights and loadings, chiefly information about the relationships among indicators and their error terms.

In Wold's original PLS design and its variants, weights and loadings are typically assumed to be proportional to one another and thus linearly related. At the population level, our simulations suggest that usually they are not (i.e., the relationship between any factor's weights and loadings is usually nonlinear). As a result, the original PLS design does not yield estimates of the true composites. It is our contention that there is one unique true composite associated with each factor and that the PLSF-SEM method yields estimates of the true composites.

3.5 PLSF-SEM's Second Stage: Factors

PLSF-SEM's second stage starts with the estimation of the elements $\widehat{\Sigma}_{F_iF_j}$ of a target correlation matrix $\widehat{\Sigma}_{FF}$ via Eq. (3.10), which follows from the correlation attenuation notion of measurement error theory (Nunnally and Bernstein 1994). In this equation $\Sigma_{\widehat{C}_i\widehat{C}_j}$ is the correlation between composites estimated in the first stage, corresponding to the pair of factors F_i and F_j. Here $\widehat{\Sigma}_{F_iF_j}$ are the elements of the matrix of estimated correlations among factors $\widehat{\Sigma}_{FF}$, which can be seen as a population matrix estimate.

$$\widehat{\Sigma}_{F_iF_j} := \frac{\Sigma_{\widehat{C}_i\widehat{C}_j}}{\sqrt{\widehat{\alpha}_i\widehat{\alpha}_j}}. \tag{3.10}$$

In this second stage, the PLSF-SEM method will fit the matrix of correlations among estimated factors $\Sigma_{\widehat{F}\widehat{F}}$, which can be seen as a model-implied matrix estimate, to $\widehat{\Sigma}_{FF}$. To that end, the method proceeds by initializing factors as indicated in Eq. (3.14) and iteratively performing the assignments in Eqs. (3.11)–(3.15). Since factors and measurement errors are reestimated in each iteration, so must the correlation matrix elements $\Sigma_{\widehat{F}_i\widehat{F}_j}$, $\Sigma_{\widehat{F}_i\widehat{C}_i}$, and $\Sigma_{\widehat{F}_i\widehat{\varepsilon}_i}$. These are the elements of the correlation matrices among factors, factors and composites, and factors and measurement errors, respectively.

$$\widehat{\varepsilon}_i := \mathrm{Stdz}\left[\widehat{\varepsilon}_i + \left(\widehat{\Sigma}_{F_iF_j} - \Sigma_{\widehat{F}_i\widehat{F}_j}\right)\frac{\widehat{\Sigma}_{F_iF_j}}{\widehat{\omega}_{i\varepsilon}}\left(\widehat{C}_j\widehat{\omega}_{jC} + \widehat{\varepsilon}_j\widehat{\omega}_{j\varepsilon}\right)\right]. \tag{3.11}$$

$$\widehat{F}_i := \mathrm{Stdz}\left[\widehat{F}_i + \left(\widehat{\omega}_{iC} - \Sigma_{\widehat{F}_i\widehat{C}_i}\right)\widehat{C}_i\widehat{\omega}_{iC}\right]. \tag{3.12}$$

$$\widehat{\varepsilon}_i := \mathrm{Stdz}\left[\widehat{\varepsilon}_i - \Sigma_{\widehat{C}_i\widehat{\varepsilon}_i}\widehat{C}_i\widehat{\omega}_{iC} + \left(\widehat{\omega}_{i\varepsilon} - \Sigma_{\widehat{F}_i\widehat{\varepsilon}_i}\right)\widehat{F}_i\widehat{\omega}_{i\varepsilon}\right]. \tag{3.13}$$

$$\widehat{F}_i := \mathrm{Stdz}\left(\widehat{C}_i\widehat{\omega}_{iC} + \widehat{\varepsilon}_i\widehat{\omega}_{i\varepsilon}\right). \tag{3.14}$$

$$\widehat{\varepsilon}_i := \mathrm{Stdz}\left[\frac{1}{\widehat{\omega}_{i\varepsilon}}\left(\widehat{F}_i - \widehat{C}_i\widehat{\omega}_{iC}\right)\right]. \tag{3.15}$$

The assignments in Eqs. (3.11)–(3.13) are called “variation sharing” assignments and constitute a critical ingredient of the PLSF-SEM method. As a group they are akin to a “soft” version of the classic expectation-maximization algorithm (Dempster et al. 1977) used in maximum likelihood estimation, but with apparently faster convergence and nonparametric properties. Through these assignments factors and measurement errors obtain variation that they did not have at the end of PLSF-SEM’s first stage, but that is an integral part of the true measurement errors and factors. Ultimately all of this variation emanates from the true composites.

The assignments above are only carried out for the variables indexed by i where $\widehat{\omega}_{i\varepsilon} > 0$. That is, factors and measurement errors are only adjusted in those cases where measurement error is assumed to exist, which are also cases where $\widehat{\alpha}_i < 1$. Convergence is achieved when the sum of the absolute differences $\widehat{\Sigma}_{F_iF_j} - \Sigma_{\widehat{F}_i\widehat{F}_j}$ falls below a small fraction or when the sum of the absolute differences between successive estimates of $\Sigma_{\widehat{F}_i \widehat{F}_j}$ changes by less than a small fraction. Once convergence is achieved, final estimates of the composites, weights, and loadings are generated through Eqs. (3.16)–(3.18).

$$\widehat{C}_i := \mathrm{Stdz}\left[\frac{1}{\widehat{\omega}_{iC}}\left(\widehat{F}_i - \widehat{\varepsilon}_i\widehat{\omega}_{i\varepsilon}\right)\right]. \tag{3.16}$$

$$\widehat{\omega}_i := {x_i}^{+}\widehat{C}_i\widehat{\omega}_{iC}. \tag{3.17}$$

$$\widehat{\lambda}_i := x_i{}'{\widehat{F}_i}{}'^{+}. \tag{3.18}$$

The PLSF-SEM method then lastly proceeds to estimate path coefficients through a standard path analysis (Mueller 1996; Wright 1934, 1960) using the factor estimates. Standard errors for path coefficients and any other model parameter can be estimated via resampling or stable *P* value calculation methods (Kock 2014a, 2015), as is usually done in the original PLS design. The standard errors can subsequently be used to obtain chance probability estimates for hypothesis testing (Kock 2014c), for any model parameter.

3.6 Monte Carlo Experiment

We conducted a Monte Carlo experiment (Paxton et al. 2001) based on the true population model depicted in Fig. 3.4, whereby 300 samples were created for each of the following sample sizes: 50, 100, and 300. This Monte Carlo experiment was conducted as part of extensive internal tests of version 5.0 of WarpPLS.

A summarized set of results based on the analyses of simulated samples is shown in Table 3.1. True values, mean parameter estimates, and standard errors are shown next to one another. Results obtained through the PLSF-SEM method (under the

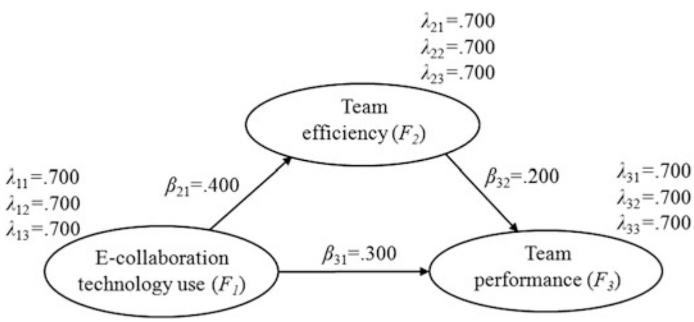


Fig. 3.4 True population model

Table 3.1 Summarized Monte Carlo experiment results

SEM method	PLSA	PLSF	PLSA	PLSF	PLSA	PLSF
Sample size	50	50	100	100	300	300
EU>TE(TruePath)	.400	.400	.400	.400	.400	.400
EU>TE(AvgPath)	.339	.380	.309	.385	.303	.394
EU>TE(SEPath)	.125	.161	.128	.127	.110	.070
EU>TP(TruePath)	.300	.300	.300	.300	.300	.300
EU>TP(AvgPath)	.260	.301	.248	.294	.234	.297
EU>TP(SEPath)	.135	.157	.108	.133	.085	.079
TE>TP(TruePath)	.200	.200	.200	.200	.200	.200
TE>TP(AvgPath)	.201	.234	.189	.225	.174	.203
TE>TP(SEPath)	.144	.163	.098	.132	.061	.079
EU3<EU(TrueLoad)	.700	.700	.700	.700	.700	.700
EU3<EU(AvgLoad)	.793	.692	.802	.695	.808	.699
EU3<EU(SELoad)	.129	.108	.113	.077	.112	.049

Notes: XX>YY = link from factor XX to YY; EU = e-collaboration technology use; TE = team efficiency; TP = team performance; XX1 … XXn = indicators associated with factor XX; TruePath = true path coefficient; AvgPath = mean path coefficient estimate; SEPath = standard error of path coefficient estimate; TrueLoad = true loading; AvgLoad = mean loading estimate; SELoad = standard error of loading estimate

"PLSF" columns) are contrasted with results obtained through the PLS Mode A algorithm (under the "PLSA" columns). PLS Mode A with the "path weighting" scheme was employed, the most widely used in analyses employing the original PLS design. We show results for all of the structural paths in the model but restrict ourselves to loadings for one indicator in one factor since all loadings are the same in the true population model used. This is also done to avoid repetition, as the same general pattern of results for loadings repeats itself for all indicators in all factors.

As we can see from the summarized results, the PLSF-SEM method yielded virtually unbiased estimates at $N = 300$, whereas PLS Mode A yielded significantly biased estimates at that same sample size. One of the reasons for these significantly biased estimates with PLS Mode A is the relatively low loadings in the true population model ($\lambda_{ij} = .7$, for all i and j), which tend to be a challenge for algorithms based on the original PLS design.

The relatively low loadings in the true population model apparently had little effect on PLSF-SEM's asymptotic convergence to the true values of the model parameters, although those loadings probably slowed down that convergence somewhat as sample sizes increased. In other simulations we conducted with higher loadings, convergence was achieved at smaller sample sizes. Generally speaking, high loadings are to be expected based on the common factor model, as they imply the use of psychometrically sound measurement instruments.

For several of the path coefficients and loadings, the PLSF-SEM method yielded lower standard errors, particularly as sample sizes increased. This is noteworthy because the PLSF-SEM method is clearly more computationally complex than PLS

Mode A and thus could have been expected to have a greater "cost" in terms of standard errors.

However, standard errors yielded at $N = 50$ were generally higher for the PLSF-SEM method. Apparently the difference was enough to have a negative effect on power, as the ratios of path coefficients to standard errors indicate. That is, at $N = 50$, one could argue based on the results that PLS Mode A has greater power than the PLSF-SEM method for this particular model, although the ratios of path coefficients to standard errors suggest that both methods may struggle to avoid type II errors at this small sample size, particularly for the paths whose true coefficients were lower than .400 (the path with the highest strength).

3.7 Discussion and Conclusion

While Wold's original PLS design offers several advantages over CB-SEM, it is largely incompatible with the common factor model (Kline 2010; Mueller 1996). Arguably the common factor model is the core foundation of CB-SEM. Given this, in Monte Carlo simulations where data is created based on common factor model assumptions, the original PLS design yields biased model parameters. Generally path coefficients are underestimated, and loadings are overestimated. This "advantages-with-costs" scenario has led to much debate over the years among proponents and detractors of the original PLS design.

In this chapter, we discussed what could be a solid step in the resolution of this debate. We showed how researchers can implement factor-based PLS-SEM (PLSF-SEM), a new method that generates estimates of both true composites and factors. The method does so in two stages and fully accounts for measurement error. Since it generates estimates of both true composites and factors, the PLSF-SEM method can potentially place researchers in a position where they can subsequently estimate a large number of model parameters.

At this point the reader may ask a reasonable question. Given that confirmatory factor analyses and hypothesis testing require primarily estimates of loadings and path coefficients, why would one want to generate factor estimates? The answer is that there are certain types of analyses that require factor scores and more will likely be developed in the future as estimates of true factor scores become available to methodological researchers.

For example, the recently developed full collinearity test concurrently assesses lateral and vertical collinearity among factors (Kock and Lynn 2012), providing the basis on which a number of methodological issues can be addressed (e.g., common method bias), but cannot be properly conducted without factor scores. Also, factor scores enable nonlinear analyses where best-fitting nonlinear functions are estimated for each pair of linked factors and subsequently used to estimate path coefficients that take into account the nonlinearity (Guo et al. 2011; Kock 2015; Moqbel et al. 2013).

It is our belief that the PLSF-SEM method is a solid step in the legitimization of modified versions of Wold's original "soft" PLS techniques for confirmatory factor and full-blown SEM analyses that are consistent with the common factor model. However, common sense suggests that the PLSF-SEM method has weaknesses that will be uncovered as time goes by. It is very unlikely that any new method will be problem-free.

As the PLSF-SEM method is refined and improved, it may serve as the basis for the development of novel statistical tests that could lead to new insights in the context of SEM. Users of WarpPLS, starting in version 5.0, will be able to test the PLSF-SEM method and variations for themselves. Also, we hope that this brief chapter will provide enough details for implementations, in numerical programming environments such as R and GNU Octave, to be developed and tested under various conditions. We welcome comments, suggestions, and corrections.

Acknowledgments The author is the developer of the software WarpPLS, which has over 7000 users in more than 33 different countries at the time of this writing, and moderator of the PLS-SEM e-mail distribution list. He is grateful to those users, and to the members of the PLS-SEM e-mail distribution list, for questions, comments, and discussions on topics related to SEM and to the use of WarpPLS.

References

Adelman, I., & Lohmöller, J.-B. (1994). Institutions and development in the nineteenth century: A latent variable regression model. *Structural Change and Economic Dynamics, 5*(2), 329–359.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297–334.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B (Methodological), 39*(1), 1–38.

Guo, K. H., Yuan, Y., Archer, N. P., & Connelly, C. E. (2011). Understanding nonmalicious security violations in the workplace: A composite behavior model. *Journal of Management Information Systems, 28*(2), 203–236.

Kline, R. B. (2010). *Principles and practice of structural equation modeling* (2nd ed.). New York: The Guilford Press.

Kock, N. (2005). What is e-collaboration. *International Journal of e-Collaboration, 1*(1), 1–7.

Kock, N. (2008). E-collaboration and e-commerce in virtual worlds: The potential of second life and world of warcraft. *International Journal of e-Collaboration, 4*(3), 1–13.

Kock, N. (2014a). Using data labels to discover moderating effects in PLS-based structural equation modeling. *International Journal of e-Collaboration, 10*(4), 1–14. doi:10.4018/ijec.2014100101.

Kock, N. (2014b). *One-tailed or two-tailed P values in PLS-SEM?* Laredo, TX: ScriptWarp Systems.

Kock, N. (2014c). Advanced mediating effects tests, multi-group analyses, and measurement model assessments in PLS-based SEM. *International Journal of e-Collaboration, 10*(3), 1–13.

Kock, N. (2015). *WarpPLS 5.0 user manual*. Laredo, TX: ScriptWarp Systems.

Kock, N., & Lynn, G. S. (2012). Lateral collinearity and misleading results in variance-based SEM: An illustration and recommendations. *Journal of the Association for Information Systems, 13*(7), 546–580.

Lohmöller, J.-B. (1989). *Latent variable path modeling with partial least squares*. Berlin: Springer.
Moqbel, M., Nevo, S., & Kock, N. (2013). Organizational members' use of social networking sites and job performance: An exploratory study. *Information Technology & People, 26*(3), 240–264.
Mueller, R. O. (1996). *Basic principles of structural equation modeling*. New York: Springer.
Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.
Paxton, P., Curran, P. J., Bollen, K. A., Kirby, J., & Chen, F. (2001). Monte Carlo experiments: Design and implementation. *Structural Equation Modeling, 8*(2), 287–312.
Sijtsma, K. (2009). On the use, the misuse, and the very limited usefulness of Cronbach's alpha. *Psychometrika, 74*(1), 107–120. doi:10.1007/s11336-008-9101-0.
Wold, H. (1980). Model construction and evaluation when theoretical knowledge is scarce: Theory and application of partial least squares, Chap. 3. In J. Kmenta & J. B. Ramsey (Eds.), *Evaluation of econometric models* (pp. 47–74). Cambridge, MA: Academic Press.
Wright, S. (1934). The method of path coefficients. *Annals of Mathematical Statistics, 5*(3), 161–215.
Wright, S. (1960). Path coefficients and path regressions: Alternative or complementary concepts? *Biometrics, 16*(2), 189–202.

Chapter 4
A Perfect Match Between a Model and a Mode

Theo K. Dijkstra

Abstract When the partial least squares estimation methods, the "modes," are applied to the standard latent factor model against which methods are designed and calibrated in PLS, they will not yield consistent estimators without adjustments. We specify a different model in terms of observables only, that satisfies the same rank constraints as the latent variable model, and show that now mode B is perfectly suitable without the need for corrections. The model explicitly uses composites, linear combinations of observables, instead of latent factors. The composites may satisfy identifiable linear structural equations, which need not be regression equations, estimable via 2SLS or 3SLS. Each time practitioners contemplate the use of PLS' basic design model the composites model is a viable alternative. The chapter is conceptual mainly, but a small Monte Carlo study exemplifies the feasibility of the new approach.

4.1 Introduction

Herman (H.O.A.) Wold (1908–1992) developed partial least squares (PLS) in a series of papers, published as well as privately circulated. The *seminal* published papers are Wold (1966, 1975, 1982). A key characteristic of PLS is the determination of composites, linear combinations of observables, by weights that are fixed

This chapter "continues" a sometimes rather spirited discussion with Wold, that started in 1977, at the Wharton School in Philadelphia, via my PhD thesis, Dijkstra (1981), and a paper Dijkstra (1983). There was a long silence, until about 2008, when Peter M. Bentler (UCLA) rekindled my interest in PLS, one of the many things for which I owe him my gratitude. Crucial also is the collaboration with Joerg Henseler (Twente), that led to a number of papers on PLS and on ways to get consistency without the need to increase the number of indicators, PLSc, as well as to a software program ADANCO for composites. I am very much in his debt too. The present chapter expands on Dijkstra (2010) by avoiding unobservables as much as possible while still adhering to Wold's fundamental principle of soft modeling.

T.K. Dijkstra (✉)
Faculty of Economics and Business, University of Groningen, Groningen, Netherlands
e-mail: t.k.dijkstra@rug.nl

H. Latan, R. Noonan (eds.), *Partial Least Squares Path Modeling*,
DOI 10.1007/978-3-319-64069-3_4

points of sequences of alternating least squares programs, called "modes." Wold distinguished three types of modes (not models!): mode A, reminiscent of principal component analysis, mode B, related to canonical variables analysis, and mode C, that mixes the former two. In a sense PLS is an extension of canonical variables and principal components analyses. While Wold designed the algorithms, great strides were made in the estimation, testing, and analysis of structured covariance matrices, as induced by linear structural equations in terms of latent factors and indicators (LISREL first, then EQS et cetera). Latent factor modeling became the dominant backdrop against which Wold designed his tools. One model in particular, the "basic design," became the model of choice in calibrating PLS. Here each latent factor is measured indirectly by a unique set of indicators, with all measurement errors usually assumed to be mutually uncorrelated. The composites combine the indicators for each latent factor separately, and their relationships are estimated by regressions.[1] The basic design embodies Wold's "fundamental principle of soft modeling": all information between the blocks of observables is assumed to be conveyed by latent variables (Wold 1982).[2] However, in this model PLS is not well-calibrated[3]: when applied to the true covariance matrix it yields by *necessity* approximations, see, e.g., Dijkstra (1981, 1983; 2010; 2014). For consistency, meaning that the probability limit of the estimators equals the theoretical value, Wold also requires the number of indicators to increase alongside the number of observations (consistency-at-large).

In this chapter we leave the realm of the unobservables, and build a model in terms of manifest variables that satisfies the fundamental principle of soft modeling, adjusted to read: *all information between the blocks is conveyed solely by the composites.* For this model, mode B is the perfect match, in the sense that estimation via mode B is the natural thing to do: when applied to the true covariance matrix it yields the underlying parameter values, not approximations that require corrections. A latent factor model, in contrast, would need additional structure (like uncorrelated measurement errors) and fitting it would produce approximations.

The chapter is structured as follows. The next section, Sect. 4.2, outlines the new model. We specify for a vector $\mathbf{y}$ of observable variables, "indicators," a structural model that generates via linear composites of separate blocks of indicators all the standard *basic design* rank restrictions on the covariance matrix, without invoking

[1] This includes simultaneous equations systems, which are generally not regressions. They were estimated by a Fix Point method, essentially iterations of 2SLS (two-stage-least-squares) regressions (Boardman et al. 1981). See below for 2SLS and Dijkstra and Henseler (2015a,b).

[2] "Soft modeling" indicates that PLS is meant to perform "substantive analysis of complex problems that are at the same time data-rich and theory-primitive" (Wold 1982).

[3] I am not saying here that methods that are not well-calibrated are intrinsically "bad." This would be ludicrous given the inherent approximate nature of statistical models. Good predictions typically require a fair amount of misspecification, to put it provocatively. But knowing what happens when we apply a statistical method to "the population" helps answering what it is that it is estimating. Besides, consistency, and to a much lesser extent "efficiency," was very important to Wold.

the existence of unobservable latent factors. They, the composites, are linked to each other by means of a "structural," "simultaneous," or "interdependent" equations system, that together with the loadings fully captures the (linear) relationships between the blocks of indicators.

Section 4.3 is devoted to estimation issues. We describe a step-wise procedure: first the weights defining the composites via generalized canonical variables,[4] then their correlations and the loadings in the simplest possible way, and finally the parameters of the simultaneous equations system using the econometric methods 2SLS or 3SLS. The estimation proceeds essentially in a non-iterative fashion (even when we use one of the PLS' traditional algorithms, it will be very fast), making it potentially eminently suitable for bootstrap analyses. We give the results of a Monte Carlo simulation for a model for 18 indicators; they are generated by six composites linked to each other via two linear equations, which are *not* regressions. We also show that mode A, when applied to the true covariance matrix of the indicators, can only yield the correct results when the composites are certain principal components. As in PLSc, mode A can be adjusted to produce the right results (in the limit).

Section 4.4 suggests how to test various aspects of the model, via tests of the rank constraints, via prediction/cross-validation, and via global goodness-of-fit tests.

Section 4.5 contains some final observations and comments. We briefly return to "the latent factors versus composites"-issue and point out that in a latent factor model the factors cannot fully be replaced by linear composites, no matter how we choose them: the regression of the indicators on the composites will not yield the loadings on the factors, *or* (inclusive) the composites cannot satisfy the same equations that the factors satisfy.

The Appendix contains a proof for a statement needed in Sect. 4.3.

4.2 The Model: Composites as Factors

Our point of departure is a random vector[5] $\mathbf{y}$ of "indicators" that can be partitioned into N subvectors, "blocks" in PLS parlance, as $\mathbf{y} = (\mathbf{y}_1; \mathbf{y}_2; \mathbf{y}_3; \ldots; \mathbf{y}_N)$. Here the semi-colon stacks the subvectors one underneath the other, as in MATLAB; $\mathbf{y}_i$ is of order $p_i \times 1$ with p_i usually larger than one. So $\mathbf{y}$ is of dimension $p \times 1$ with $p := \sum_{i=1}^{N} p_i$. We will denote cov($\mathbf{y}$) by $\boldsymbol{\Sigma}$, and take it to be positive definite (p.d.), so no indicator is redundant. We will let $\boldsymbol{\Sigma}_{ii}$:=cov($\mathbf{y}_i$). $\boldsymbol{\Sigma}_{ii}$ is of order $p_i \times p_i$ and it is of course p.d. as well. It is *not* necessary to have other constraints on $\boldsymbol{\Sigma}_{ii}$, in particular it need not have a one-factor structure. Each block $\mathbf{y}_i$ is condensed into a composite, a scalar c_i, by means of a conformable weight

[4] It should be pointed out that I see PLS' mode B as one of a family of generalized canonical variables estimation methods (Sect. 4.3.1), to be treated on a par with the others, without necessarily claiming that mode B is the superior or inferior method. None of the methods will be uniformly superior in every sensible aspect.

[5] Vectors and matrices will be distinguished from scalars by printing them in boldface.

vector $\mathbf{w}_i$: $c_i := \mathbf{w}_i^\top \mathbf{y}_i$. The composites will be normalized to have variance one: $\text{var}(c_i) = \mathbf{w}_i^\top \boldsymbol{\Sigma}_{ii} \mathbf{w}_i = 1$. The vector of composites $\mathbf{c} := (c_1; c_2; c_3; \ldots; c_N)$ has a p.d. covariance/correlation matrix denoted by $\mathbf{R}_c = (r_{ij})$ with $r_{ij} = \mathbf{w}_i^\top \boldsymbol{\Sigma}_{ij} \mathbf{w}_j$ where $\boldsymbol{\Sigma}_{ij} := \text{E}(\mathbf{y}_i - \text{E}\mathbf{y}_i)\left(\mathbf{y}_j - \text{E}\mathbf{y}_j\right)^\top$. A regression of $\mathbf{y}_i$ on c_i and a constant gives a loading vector $\mathbf{L}_i$ of order $p_i \times 1$:

$$\mathbf{L}_i := \text{E}\left(\mathbf{y}_i - \text{E}\mathbf{y}_i\right) \cdot \left(c_i - \text{E}c_i\right) = \text{E}\left(\mathbf{y}_i - \text{E}\mathbf{y}_i\right)\left(\mathbf{y}_i - \text{E}\mathbf{y}_i\right)^\top \mathbf{w}_i = \boldsymbol{\Sigma}_{ii} \mathbf{w}_i \tag{4.1}$$

So far all we have is a list of definitions but as yet no real model: there are no constraints on the joint distribution of $\mathbf{y}$ apart from the existence of moments[6] and a p.d. covariance matrix. We will now impose our version of Wold's fundamental principle in soft modeling:

all information between the blocks is conveyed solely by the composites

We deviate from Wold's original formulation in an essential way: whereas Wold postulated that all information is conveyed by unobserved, even unobservable, latent variables, we let the information to be fully transmitted by indices, by composites of observable indicators. So we postulate the existence of weight vectors such that for any two different blocks $\mathbf{y}_i$ and $\mathbf{y}_j$

$$\boldsymbol{\Sigma}_{ij} = r_{ij} \mathbf{L}_i \mathbf{L}_j^\top \tag{4.2}$$

$$= \mathbf{w}_i^\top \boldsymbol{\Sigma}_{ij} \mathbf{w}_j \cdot \boldsymbol{\Sigma}_{ii} \mathbf{w}_i \cdot \left(\boldsymbol{\Sigma}_{jj} \mathbf{w}_j\right)^\top$$

$$= \text{corr}\left(\mathbf{w}_i^\top \mathbf{y}_i, \mathbf{w}_j^\top \mathbf{y}_j\right) \cdot \text{cov}\left(\mathbf{y}_i, \mathbf{w}_i^\top \mathbf{y}_i\right) \cdot \left(\text{cov}\left(\mathbf{y}_j, \mathbf{w}_j^\top \mathbf{y}_j\right)\right)^\top \tag{4.3}$$

The cross-covariances between the blocks are determined by the correlation between their corresponding composites and the loadings of the blocks on those composites. Note that line (4.2) is highly reminiscent of the corresponding equation for the basic design, with latent variables. There it would read $\rho_{ij} \boldsymbol{\lambda}_i \boldsymbol{\lambda}_j^\top$ with ρ_{ij} representing the correlation between the latent variables, with $\boldsymbol{\lambda}_i$ and $\boldsymbol{\lambda}_j$ capturing the loadings. So the rank-one structure of the covariance matrices between the blocks is maintained fully, without requiring the existence of N additional unobservable variables.

We now have:

$$\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_{11} & r_{12}\mathbf{L}_1\mathbf{L}_2^\top & r_{13}\mathbf{L}_1\mathbf{L}_3^\top & \cdot & r_{1N}\mathbf{L}_1\mathbf{L}_N^\top \\ & \boldsymbol{\Sigma}_{22} & r_{23}\mathbf{L}_2\mathbf{L}_3^\top & \cdot & r_{2N}\mathbf{L}_2\mathbf{L}_N^\top \\ & & \cdot & \cdot & \cdot \\ & & & \boldsymbol{\Sigma}_{N-1,N-1} & r_{N-1,N}\mathbf{L}_{N-1}\mathbf{L}_N^\top \\ & & & & \boldsymbol{\Sigma}_{NN} \end{bmatrix} \tag{4.4}$$

[6] A random sample of indicator-vectors and the existence of second order moments is sufficient for the consistency of the estimators to be developed below; with the existence of fourth-order moments we also have asymptotic normality.

The appendix contains a proof of the fact that $\mathbf{\Sigma}$ is positive definite when and only when the correlation matrix of the composites, $\mathbf{R}_c$, is positive definite. Note that in a Monte Carlo analysis we can choose the weight vectors (or loadings) and the values of $\mathbf{R}_c$ independently.

We can add more structure to the model by imposing constraints on $\mathbf{R}_c$. This is done most conveniently by postulating a set of simultaneous equations to be satisfied by $\mathbf{c}$. We will call one subvector of $\mathbf{c}$ the *exogenous* composites, denoted by $\mathbf{c}_{\text{exo}}$, and the remaining elements will be collected in $\mathbf{c}_{\text{endo}}$, the *endogenous* composites. There will be conformable matrices $\mathbf{B}$ and $\mathbf{C}$ with $\mathbf{B}$ invertible such that

$$\mathbf{B}\mathbf{c}_{\text{endo}} = \mathbf{C}\mathbf{c}_{\text{exo}} + \mathbf{z} \tag{4.5}$$

It is customary to normalize $\mathbf{B}$, i.e., all diagonal elements equal one (perhaps after some re-ordering). The residual vector $\mathbf{z}$ has a zero mean and is uncorrelated with $\mathbf{c}_{\text{exo}}$. In this type of (econometric) model the relationships between the exogenous variables are usually not the main concern. The research focus is on the way they drive the endogenous variables and the interplay or the feedback mechanism between the latter as captured by a matrix $\mathbf{B}$ that has nonzero elements both above and below the diagonal. A special case, with no feedback mechanism at all, is the class of *recursive* models, where $\mathbf{B}$ has only zeros on one side of its diagonal, and the elements of $\mathbf{z}$ are mutually uncorrelated. Here the coefficients in $\mathbf{B}$ and $\mathbf{C}$ can be obtained directly by consecutive regressions, given the composites. For general $\mathbf{B}$ this is not possible, since $\mathbf{c}_{\text{endo}}$ is a linear function of $\mathbf{z}$ so that z_i will typically be correlated with every endogenous variable in the ith equation.[7]

Even when the model is not recursive, the matrices $\mathbf{B}$ and $\mathbf{C}$ will be postulated to satisfy certain zero constraints (and possibly other types of constraints, but we focus here on the simplest situation). So some B_{ij}'s and C_{kl}'s are zero. We will assume that the remaining coefficients are *identifiable* from a knowledge of the so-called reduced form matrix $\mathbf{\Pi}$

$$\mathbf{\Pi} := \boldsymbol{B}^{-1}\mathbf{C} \tag{4.6}$$

Note that

$$\mathbf{c}_{\text{endo}} = \mathbf{\Pi}\boldsymbol{c}_{\text{exo}} + \mathbf{B}^{-1}\mathbf{z} \tag{4.7}$$

so $\mathbf{\Pi}$ is a matrix of regression coefficients. Once we have those, we should be able to retrieve $\mathbf{B}$ and $\mathbf{C}$ from them. Identifiability is equivalent to the existence of certain rank conditions on $\mathbf{\Pi}$, we will have more to say about them later on. We could have additional constraints on the covariance matrices of $\mathbf{c}_{\text{exo}}$ and $\mathbf{z}$ but we will not develop that here, taking the approach that demands the least in terms of knowledge

[7] See Pearl (2009) for an in-depth causal analysis of simultaneous equations systems (based on and extending (Haavelmo 1944), probably the best apologia of econometrics ever).

about the relationships between the composites. It is perhaps good to note that granted identifiability, the free elements in **B** and **C** can be interpreted as regression coefficients, provided we replace the "explanatory" endogenous composites by their regression on the exogenous composites. This is easily seen as follows:

$$\mathbf{c}_{\text{endo}} = (\mathbf{I}-\mathbf{B})\,\mathbf{c}_{\text{endo}} + \mathbf{C}\mathbf{c}_{\text{exo}} + \mathbf{z} \tag{4.8}$$

$$= (\mathbf{I}-\mathbf{B})\left(\boldsymbol{\Pi}\boldsymbol{c}_{\text{exo}} + \mathbf{B}^{-1}\mathbf{z}\right) + \mathbf{C}\mathbf{c}_{\text{exo}} + \mathbf{z} \tag{4.9}$$

$$= (\mathbf{I}-\mathbf{B})\left(\boldsymbol{\Pi}\boldsymbol{c}_{\text{exo}}\right) + \mathbf{C}\mathbf{c}_{\text{exo}} + \mathbf{B}^{-1}\mathbf{z} \tag{4.10}$$

where $\mathbf{B}^{-1}\mathbf{z}$ is uncorrelated with $\boldsymbol{\Pi}\boldsymbol{c}_{\text{exo}}$ and $\mathbf{c}_{\text{exo}}$. So the free elements of $(\mathbf{I}-\mathbf{B})$ and $\mathbf{C}$ can be obtained by a regression of $\mathbf{c}_{\text{endo}}$ on $\boldsymbol{\Pi}\boldsymbol{c}_{\text{exo}}$ and $\mathbf{c}_{\text{exo}}$, equation by equation.[8] Identifiability is here equivalent to invertibility of the covariance matrix of the "explanatory" variables in each equation. A necessary condition for this to work is that we cannot have more coefficients to estimate in each equation than the total number of exogenous composites in the system.

We have for $\mathbf{R}_c$

$$\mathbf{R}_c = \begin{bmatrix} \operatorname{cov}(\mathbf{c}_{\text{exo}}) & \operatorname{cov}(\mathbf{c}_{\text{exo}})\cdot\boldsymbol{\Pi}^{\intercal} \\ & \boldsymbol{\Pi}\operatorname{cov}(\mathbf{c}_{\text{exo}})\,\boldsymbol{\Pi}^{\intercal} + \mathbf{B}^{-1}\operatorname{cov}(\mathbf{z})\,(\mathbf{B}^{\intercal})^{-1} \end{bmatrix} \tag{4.11}$$

Thanks to the structural constraints, the number of parameters in $\mathbf{R}_c$ could be (considerably) less than $\frac{1}{2}N(N-1)$, potentially allowing for an increase in estimation efficiency.

As far as $\boldsymbol{\Sigma}$ is concerned, the model is now completely specified.

4.2.1 *Fundamental Properties of the Model and Wold's Fundamental Principle*

Now define for each i the measurement error vector $\mathbf{d}_i$ via

$$\mathbf{y}_i - \operatorname{mean}(\mathbf{y}_i) = \mathbf{L}_i\,(c_i - \operatorname{mean}(c_i)) + \mathbf{d}_i \tag{4.12}$$

where $\mathbf{L}_i = \boldsymbol{\Sigma}_{ii}\,\mathbf{w}_i$, the loadings vector obtained by a regression of the indicators on their composite (and a constant).

By construction $\mathbf{d}_i$ has a zero mean and is uncorrelated with c_i. In what follows it will be convenient to have all variables de-meaned, so we have $\mathbf{y}_i = \mathbf{L}_i c_i + \mathbf{d}_i$. It is easy to verify that:

[8] The estimation method based on these observations is called 2SLS, two-stage-least-squares, for obvious reasons, and was developed by econometricians in the 1950s of the previous century.

The measurement error vectors are mutually uncorrelated, and uncorrelated with all composites:

$$\mathrm{E}\mathbf{d}_i\mathbf{d}_j^{\mathsf{T}} = 0 \text{ for all different } i \text{ and } j \tag{4.13}$$

$$\mathrm{E}\mathbf{d}_i c_j = 0 \text{ for all } i \text{ and } j \tag{4.14}$$

It follows that $\mathrm{E}\mathbf{y}_i\mathbf{d}_j^{\mathsf{T}} = 0$ for all different i and j. In addition:

$$\mathrm{cov}\,(\mathbf{d}_i) = \boldsymbol{\Sigma}_{ii} - \mathbf{L}_i\mathbf{L}_i^{\mathsf{T}} \tag{4.15}$$

The latter is also very similar to the corresponding expression in the basic design, but we cannot in general have a diagonal $\mathrm{cov}(\mathbf{d}_i)$, because $\mathrm{cov}(\mathbf{d}_i)\,\mathbf{w}_i$ is identically zero (implying that the variance of $\mathbf{w}_i^{\mathsf{T}}\mathbf{d}_i$ is zero, and therefore $\mathbf{w}_i^{\mathsf{T}}\mathbf{d}_i = 0$ with probability one). The following relationships can be verified algebraically using regression results, or by using conditional expectations formally (so even though we use the formalism of conditional expectations and the notation, we do just mean regression).

$$\mathrm{E}\,(\mathbf{y}_1|c_1) = \mathbf{L}_1 c_1 \tag{4.16}$$

because $\mathrm{E}(\mathbf{y}_1|c_1) = \mathrm{E}(\mathbf{L}_1 c_1 + \mathbf{d}_1|c_1) = \mathbf{L}_1 c_1 + 0$. Also note that

$$\mathrm{E}\,(c_1|\mathbf{y}_2, \mathbf{y}_3, \ldots, \mathbf{y}_N) \tag{4.17}$$

$$= \mathrm{E}(\mathrm{E}\,(c_1|\mathbf{y}_2, \mathbf{y}_3, \ldots, \mathbf{y}_N, \mathbf{d}_2, \mathbf{d}_3, \ldots, \mathbf{d}_N)\,|\mathbf{y}_2, \mathbf{y}_3, \ldots, \mathbf{y}_N) \tag{4.18}$$

$$= \mathrm{E}(\mathrm{E}\,(c_1|c_2, c_3, \ldots, c_N, \mathbf{d}_2, \mathbf{d}_3, \ldots, \mathbf{d}_N)\,|\mathbf{y}_2, \mathbf{y}_3, \ldots, \mathbf{y}_N) \tag{4.19}$$

$$= \mathrm{E}(\mathrm{E}\,(c_1|c_2, c_3, \ldots, c_N)\,|\mathbf{y}_2, \mathbf{y}_3, \ldots, \mathbf{y}_N) \tag{4.20}$$

$$= \mathrm{E}\,(c_1|c_2, c_3, \ldots, c_N) \tag{4.21}$$

We use the "tower property" of conditional expectation on the second line. (In order to project on a target space, we first project on a larger space, and then project the result of this on the target space.) On the third line we use $\mathbf{y}_i = \mathbf{L}_i c_i + \mathbf{d}_i$ so that conditioning on the $\mathbf{y}_i$'s and the $\mathbf{d}_i$'s is the same as conditioning on the c_i's and the $\mathbf{d}_i$'s. The fourth line is due to zero correlation between the c_i's and the $\mathbf{d}_i$'s, and the last line exploits the fact that the composites are determined fully by the indicators. So because $\mathrm{E}(\mathbf{y}_1|\mathbf{y}_2, \mathbf{y}_3, \ldots, \mathbf{y}_N) = \mathrm{E}(\mathbf{L}_1 c_1 + \mathbf{d}_1|\mathbf{y}_2, \mathbf{y}_3, \ldots, \mathbf{y}_N) = \mathbf{L}_1\mathrm{E}(c_1|\mathbf{y}_2, \mathbf{y}_3, \ldots, \mathbf{y}_N)$ we have

$$\mathrm{E}\,(\mathbf{y}_1|\mathbf{y}_2, \mathbf{y}_3, \ldots, \mathbf{y}_N) = \mathbf{L}_1\mathrm{E}\,(c_1|c_2, c_3, \ldots, c_N) \tag{4.22}$$

In other words, the best (least squares) predictor of a block of indicators given other blocks is determined by the best predictor of the composite of that block given the composites of the other blocks, together with the loadings on the composite. This contrasts rather strongly with the model Wold used, with latent factors/variables $\mathbf{f}$. Here instead of $\mathbf{L}_1\mathrm{E}(c_1|c_2, c_3, \ldots, c_N)$ we have

$$\mathrm{E}\left(\mathbf{y}_1|\mathbf{y}_2, \mathbf{y}_3, \ldots, \mathbf{y}_N\right) = \boldsymbol{\lambda}_1\mathrm{E}(\mathrm{E}\left(f_1|f_2, f_3, \ldots, f_N\right)|\mathbf{y}_2, \mathbf{y}_3, \ldots, \mathbf{y}_N) \tag{4.23}$$

Basically, we can follow the sequence of steps as above for the composites except the penultimate step, from (4.20) to (4.21). I would maintain that the model as specified answers more truthfully to the fundamental principle of soft modeling than the basic design.

4.3 Estimation Issues

We will assume that we have the outcome of a Consistent and Asymptotically Normal (CAN-)estimator for $\boldsymbol{\Sigma}$. One can think of the sample covariance matrix of a random sample from a population with covariance matrix $\boldsymbol{\Sigma}$ and finite fourth-order moments (the latter is sufficient for asymptotic normality, consistency requires finite second-order moments only). The estimators to be described are all (locally) smooth functions of the CAN-estimator for $\boldsymbol{\Sigma}$, hence they are CAN as well.

We will use a step-wise approach: first the weights, then the loadings and the correlations between the composites, and finally the structural form coefficients. Each step uses a procedure that is essentially non-iterative, or if it iterates, it is very fast. So no explicit overall fit-criterion, although one could interpret the approach as the first iterate in a block relaxation program that aims to optimize a positive combination of target functions appropriate for each step. The view that a lack of an overall criterion to be optimized is a major flaw is ill-founded. Estimators should be compared on the basis of their distribution functions, the extent to which they satisfy computational desiderata, and the induced quality of the predictions. There is no theorem, and their cannot be one, to the effect that estimators that optimize a function are better than those that are not so motivated. For composites a proper comparison between the "step-wise" (partial) and the "global" approaches is still open. Of the issues to be addressed two stand out: *efficiency* in case of a proper, correct specification, and *robustness* with respect to distributional assumptions and specification errors (the optimization of a global fitting function that takes each and every structural constraint seriously may not be as robust to specification errors as a step-wise procedure).

4.3.1 Estimation of Weights, Loadings, and Correlations

The only issue of some substance in this section is the estimation of the weights. Once they are available, estimates for the loadings and correlations present themselves: the latter are estimated via the correlation between the composites, the former by a regression of each block on its corresponding composite. One could devise more intricate methods but in this stage there seems little point in doing so.

To estimate the weights we will use generalized Canonical Variables (CV's) analysis.[9] This includes of course the approach proposed by Wold, the so-called mode B estimation method. Composites simply *are* canonical variables. Any method that yields CV's matches naturally, "perfectly," with the model. We will describe some of the methods while applying them to $\mathbf{\Sigma}$ and show that they do indeed yield the weights. A continuity argument then gives that when they are applied to the CAN-estimator for $\mathbf{\Sigma}$ the estimators for the weights are consistent as well. Local differentiability leading to asymptotic normality is not difficult to establish either.[10]

For notational ease we will employ a composites model with three blocks, $N = 3$, but that is no real limitation. Now consider the covariance matrix, denoted by $\mathbf{R}(\mathbf{v})$, of $\mathbf{v}_1^\top \mathbf{y}_1$, $\mathbf{v}_2^\top \mathbf{y}_2$, and $\mathbf{v}_3^\top \mathbf{y}_3$ where each $\mathbf{v}_i$ is normalized $(\mathrm{var}(\mathbf{v}_i^\top \mathbf{y}_i) = 1)$. So

$$\mathbf{R}(\mathbf{v}) := \begin{bmatrix} 1 & \mathbf{v}_1^\top \mathbf{\Sigma}_{12}\mathbf{v}_2 & \mathbf{v}_1^\top \mathbf{\Sigma}_{13}\mathbf{v}_3 \\ \mathbf{v}_1^\top \mathbf{\Sigma}_{12}\mathbf{v}_2 & 1 & \mathbf{v}_2^\top \mathbf{\Sigma}_{23}\mathbf{v}_3 \\ \mathbf{v}_1^\top \mathbf{\Sigma}_{13}\mathbf{v}_3 & \mathbf{v}_2^\top \mathbf{\Sigma}_{23}\mathbf{v}_3 & 1 \end{bmatrix}. \tag{4.24}$$

Canonical variables are composites whose correlation matrix has "maximum distance" to the identity matrix of the same size. They are "collectively maximally correlated." The term is clearly ambiguous for more than two blocks. One program that would seem to be natural is to maximize with respect to $\mathbf{v}$

$$z(\mathbf{v}) := \mathrm{abs}(R_{12}) + \mathrm{abs}(R_{13}) + \mathrm{abs}(R_{23}) \tag{4.25}$$

subject to the usual normalizations. Since

$$\mathrm{abs}(R_{ij}) = \mathrm{abs}(r_{ij}) \cdot \mathrm{abs}(\mathbf{v}_i^\top \mathbf{\Sigma}_{ii}\mathbf{w}_i) \cdot \mathrm{abs}\left(\mathbf{v}_j^\top \mathbf{\Sigma}_{jj}\mathbf{w}_j\right) \tag{4.26}$$

we know, thanks to Cauchy–Schwarz, that

$$\mathrm{abs}(\mathbf{v}_i^\top \mathbf{\Sigma}_{ii}\mathbf{w}_i) = \mathrm{abs}\left(\mathbf{v}_i^\top \mathbf{\Sigma}_{ii}^{\frac{1}{2}} \mathbf{\Sigma}_{ii}^{\frac{1}{2}} \mathbf{w}_i\right) \leq \sqrt{\mathbf{v}_i^\top \mathbf{\Sigma}_{ii}^{\frac{1}{2}} \mathbf{\Sigma}_{ii}^{\frac{1}{2}} \mathbf{v}_i \cdot \mathbf{w}_i^\top \mathbf{\Sigma}_{ii}^{\frac{1}{2}} \mathbf{\Sigma}_{ii}^{\frac{1}{2}} \mathbf{w}_i} \tag{4.27}$$

$$= \sqrt{\mathbf{v}_i^\top \mathbf{\Sigma}_{ii}\mathbf{v}_i \cdot \mathbf{w}_i^\top \mathbf{\Sigma}_{ii}\mathbf{w}_i} = 1 \tag{4.28}$$

with equality if and only if $\mathbf{v}_i = \mathbf{w}_i$ (ignoring irrelevant sign differences). Observe that the upper bound can be reached for $\mathbf{v}_i = \mathbf{w}_i$ for all terms in which $\mathbf{v}_i$ appears,

[9] Kettenring (1971) is *the* reference for generalized canonical variables.

[10] These statements are admittedly a bit nonchalant if not cavalier, but there seems little to gain by elaborating on them.

so maximization of the sum of the absolute correlations gives $\mathbf{w}$. A numerical, iterative routine[11] suggests itself by noting that the optimal $\mathbf{v}_1$ satisfies the first order condition

$$0 = \text{sgn}\,(R_{12}) \cdot \mathbf{\Sigma}_{12}\mathbf{v}_2 + \text{sgn}\,(R_{13}) \cdot \mathbf{\Sigma}_{13}\mathbf{v}_3 - l_1\mathbf{\Sigma}_{11}\mathbf{v}_1 \tag{4.29}$$

where l_1 is a Lagrange multiplier (for the normalization), and two other quite similar equations for $\mathbf{v}_2$ and $\mathbf{v}_3$. So with arbitrary starting vectors one could solve the equations recursively for $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$, respectively, updating them after each complete round or at the first opportunity, until they settle down at the optimal value. Note that each update of $\mathbf{v}_1$ is obtainable by a regression of a "sign-weighted sum"

$$\text{sgn}\,(R_{12}) \cdot \mathbf{v}_2^{\mathsf{T}}\mathbf{y}_2 + \text{sgn}\,(R_{13}) \cdot \mathbf{v}_3^{\mathsf{T}}\mathbf{y}_3 \tag{4.30}$$

on $\mathbf{y}_1$, and analogously for the other weights. This happens to be the classical form of PLS' mode B.[12] For $\mathbf{\Sigma}$ we do not need many iterations, to put it mildly: the update of $\mathbf{v}_1$ is already $\mathbf{w}_1$, as straightforward algebra will easily show. And similarly for the other weight vectors. In other words, we have in essentially just one iteration a fixed point for the mode B equations that is precisely $\mathbf{w}$.

If we use the correlations themselves in the recursions instead of just their signs, we regress the "correlation weighted sum"

$$R_{12} \cdot \mathbf{v}_2^{\mathsf{T}}\mathbf{y}_2 + R_{13} \cdot \mathbf{v}_3^{\mathsf{T}}\mathbf{y}_3 \tag{4.31}$$

on $\mathbf{y}_1$ (and analogously for the other weights), and end up with weights that maximize

$$z\,(\mathbf{v}) := R_{12}^2 + R_{13}^2 + R_{23}^2 \tag{4.32}$$

the simple sum of the squared correlations. Again, with the same argument, the optimal value is $\mathbf{w}$.

Observe that for this $z\,(\mathbf{v})$ we have

$$\text{tr}\left(\mathbf{R}^2\right) = 2 \cdot z\,(\mathbf{v}) + 3 = \sum_{i=1}^{3} \gamma_i^2 \tag{4.33}$$

where $\gamma_i := \gamma_i\,(\mathbf{R}\,(\mathbf{v}))$ is the ith eigenvalue of $\mathbf{R}\,(\mathbf{v})$. We can take other functions of the eigenvalues, in order to maximize the difference between $\mathbf{R}\,(\mathbf{v})$ and the identity matrix of the same order. Kettenring (1971) discusses a number of alternatives. One

---

[11]With $\mathbf{\Sigma}$ one does not really need an iterative routine of course: $\mathbf{\Sigma}_{ij} = r_{ij}\mathbf{\Sigma}_{ii}\mathbf{w}_i\mathbf{w}_j^{\mathsf{T}}\mathbf{\Sigma}_{jj}$ can be solved directly for the weights (and the correlation). But in case we just have an estimate, an algorithm comes in handy.

[12]See chapter two of Dijkstra (1981).

of them minimizes the product of the γ_i's, the determinant of $\mathbf{R}(\mathbf{v})$, also known as the generalized variance. The program is called GENVAR. Since $\sum_{i=1}^{N} \gamma_i$ is always N (three in this case) for every choice of $\mathbf{v}$, GENVAR tends to make the eigenvalues as diverse as possible (as opposed to the identity matrix where they are all equal to one). The determinant of $\mathbf{R}(\mathbf{v})$ equals $\left(1 - R_{23}^2\right)$, which is independent of $\mathbf{v}_1$, times

$$1 - \begin{bmatrix} R_{12} & R_{13} \end{bmatrix} \begin{bmatrix} 1 & R_{23} \\ R_{23} & 1 \end{bmatrix}^{-1} \begin{bmatrix} R_{12} \\ R_{13} \end{bmatrix} \tag{4.34}$$

$$= 1 - \left(\mathbf{v}_1^\intercal \boldsymbol{\Sigma}_{11} \mathbf{w}_1\right)^2 \begin{bmatrix} r_{12}\mathbf{v}_2^\intercal \boldsymbol{\Sigma}_{22}\mathbf{w}_2 & r_{13}\mathbf{v}_3^\intercal \boldsymbol{\Sigma}_{33}\mathbf{w}_3 \end{bmatrix} \begin{bmatrix} 1 & R_{23} \\ R_{23} & 1 \end{bmatrix}^{-1} \begin{bmatrix} r_{12}\mathbf{v}_2^\intercal \boldsymbol{\Sigma}_{22}\mathbf{w}_2 \\ r_{13}\mathbf{v}_3^\intercal \boldsymbol{\Sigma}_{33}\mathbf{w}_3 \end{bmatrix}$$

where the last quadratic form does not involve $\mathbf{v}_1$ either and we have with the usual argument that GENVAR produces $\mathbf{w}$ also. See Kettenring (1971) for an appropriate iterative routine (this involves the calculation of ordinary canonical variables of $\mathbf{y}_i$ and the $(N-1)$-vector consisting of the other composites).

Another program is MAXVAR, which maximizes the largest eigenvalue. For every $\mathbf{v}$ one can calculate the linear combination of the corresponding composites that best predicts or explains them: the first principal component of $\mathbf{R}(\mathbf{v})$. No other set is as well explained by the first principal component as the MAXVAR composites. *There is an explicit solution here, no iterative routine is needed for the estimate of* $\boldsymbol{\Sigma}$, if one views the calculation of eigenvectors as non-iterative, see Kettenring (1971) for details.[13] One can show again that the optimal $\mathbf{v}$ equals $\mathbf{w}$ when MAXVAR is applied to Σ, although this requires a bit more work than for GENVAR (due to the additional detail needed to describe the solution).

As one may have expected, there is also MINVAR, the program aimed at minimizing the smallest eigenvalue (Kettenring 1971). The result is a set of composites with the property that no other set is "as close to linear dependency" as the MINVAR set. We also have an explicit solution, and $\mathbf{w}$ is optimal again.

4.3.2 Mode A and Mode B

In the previous subsection we recalled that mode B generates weight vectors by iterating regressions of certain weighted sums of composites on blocks. There is also mode A (and a mode C which we will not discuss), where weights are found iteratively by reversing the regressions: now blocks are regressed on weighted sums of composites. The algorithm generally converges, and the probability limits of

[13]This is true when applied to the estimate for $\boldsymbol{\Sigma}$ as well. With an estimate the other methods will usually require more than just one iteration (and all programs will produce different results, although the differences will tend to zero in probability).

the weights can be found as before by applying mode A to $\boldsymbol{\Sigma}$. If we denote the probability limits (plims) of the (normalized) mode A weights by $\widetilde{\mathbf{w}}_i$, we have in the generic case that $\mathbf{y}_i$ is regressed on $\sum_{j\neq i}\text{sgn}(\text{cov}(\widetilde{\mathbf{w}}_i^\top \mathbf{y}_i, \widetilde{\mathbf{w}}_j^\top \mathbf{y}_j))\cdot\widetilde{\mathbf{w}}_j^\top \mathbf{y}_j$ so that

$$\widetilde{\mathbf{w}}_i \propto \sum_{j\neq i} \text{sgn}(\text{cov}(\widetilde{\mathbf{w}}_i^\top \mathbf{y}_i, \widetilde{\mathbf{w}}_j^\top \mathbf{y}_j)) \cdot \boldsymbol{\Sigma}_{ij}\widetilde{\mathbf{w}}_j \tag{4.35}$$

$$= \sum_{j\neq i} \text{sgn}(\text{cov}(\widetilde{\mathbf{w}}_i^\top \mathbf{y}_i, \widetilde{\mathbf{w}}_j^\top \mathbf{y}_j)) \cdot r_{ij}\mathbf{L}_i\mathbf{L}_j^\top\widetilde{\mathbf{w}}_j \tag{4.36}$$

$$= \mathbf{L}_i \cdot \left(\sum_{j\neq i} \text{sgn}(\text{cov}(\widetilde{\mathbf{w}}_i^\top \mathbf{y}_i, \widetilde{\mathbf{w}}_j^\top \mathbf{y}_j)) \cdot r_{ij}\mathbf{L}_j^\top\widetilde{\mathbf{w}}_j\right) \tag{4.37}$$

and so

$$\widetilde{\mathbf{w}}_i \propto \mathbf{L}_i, \text{in fact} \widetilde{\mathbf{w}}_i = \frac{1}{\sqrt{\mathbf{L}_i^\top \boldsymbol{\Sigma}_{ii}\mathbf{L}_i}}\mathbf{L}_i \tag{4.38}$$

An immediate consequence is that the plim of mode A's correlation, $\widetilde{r}_{ij}$, equals

$$\widetilde{r}_{ij} = \widetilde{\mathbf{w}}_i^\top \left(r_{ij}\mathbf{L}_i\mathbf{L}_j^\top\right)\widetilde{\mathbf{w}}_j = r_{ij} \cdot \frac{\mathbf{L}_i^\top\mathbf{L}_i}{\sqrt{\mathbf{L}_i^\top \boldsymbol{\Sigma}_{ii}\mathbf{L}_i}}\frac{\mathbf{L}_j^\top\mathbf{L}_j}{\sqrt{\mathbf{L}_j^\top \boldsymbol{\Sigma}_{jj}\mathbf{L}_j}} \tag{4.39}$$

One would expect this to be smaller in absolute value than r_{ij}, and so it is, since

$$\frac{\mathbf{L}_i^\top\mathbf{L}_i}{\sqrt{\mathbf{L}_i^\top \boldsymbol{\Sigma}_{ii}\mathbf{L}_i}} = \frac{\mathbf{w}_i^\top \boldsymbol{\Sigma}_{ii}^2\mathbf{w}_i}{\sqrt{\mathbf{w}_i^\top \boldsymbol{\Sigma}_{ii}^3\mathbf{w}_i}} \tag{4.40}$$

$$= \frac{\mathbf{w}_i^\top \boldsymbol{\Sigma}_{ii}^{1/2}\Sigma_{ii}^{3/2}\mathbf{w}_i}{\sqrt{\mathbf{w}_i^\top \Sigma_{ii}^3\mathbf{w}_i}} \leq \frac{\sqrt{\mathbf{w}_i^\top \boldsymbol{\Sigma}_{ii}\mathbf{w}_i}\sqrt{\mathbf{w}_i^\top \Sigma_{ii}^3\mathbf{w}_i}}{\sqrt{\mathbf{w}_i^\top \Sigma_{ii}^3\mathbf{w}_i}} = 1 \tag{4.41}$$

because of Cauchy–Schwarz. In general, mode A's composites, $\widetilde{\mathbf{c}}$, will not satisfy $\mathbf{B}\widetilde{\mathbf{c}}_{\text{endo}} = \mathbf{C}\widetilde{\mathbf{c}}_{\text{exo}} + \widetilde{\mathbf{z}}$ with $\widetilde{\mathbf{z}}$ uncorrelated with $\widetilde{\mathbf{c}}_{\text{exo}}$. Observe that we have $\widetilde{r}_{ij} = r_{ij}$ when and only when $\Sigma_{ii}\mathbf{w}_i \propto \mathbf{w}_i$ & $\boldsymbol{\Sigma}_{jj}\mathbf{w}_j \propto \mathbf{w}_j$, in which case each composite is *a principal component* of its corresponding block.

For the plim of the loadings, $\widetilde{\mathbf{L}}_i$, we note

$$\widetilde{\mathbf{L}}_i = \frac{1}{\sqrt{\mathbf{L}_i^\top \boldsymbol{\Sigma}_{ii}\mathbf{L}_i}}\boldsymbol{\Sigma}_{ii}\mathbf{L}_i \tag{4.42}$$

So mode A's loading vector is in the limit proportional to the true vector when and only when $\boldsymbol{\Sigma}_{ii}\mathbf{w}_i \propto \mathbf{w}_i$.

To summarize:

1. *Mode A will tend to underestimate the correlations in absolute value.*[14]
2. *The plims of the correlations between the composites for Mode A and Mode B will be equal when and only when each composite is a principal component of its corresponding block, in which case we have a perfect match between a model and two modes as far as the relationships between the composites are concerned.*
3. *The plims of the loading vectors for Mode A and Mode B will be proportional when and only when each composite is a principal component of its corresponding block.*

A final observation: we can "correct" mode A to yield the right results in the general situation via

$$\frac{\boldsymbol{\Sigma}_{ii}^{-1}\widetilde{\mathbf{w}}_i}{\sqrt{\widetilde{\mathbf{w}}_i^{\mathsf{T}}\boldsymbol{\Sigma}_{ii}^{-1}\widetilde{\mathbf{w}}_i}} = \mathbf{w}_i \tag{4.43}$$

and

$$\frac{\widetilde{\mathbf{w}}_i}{\sqrt{\widetilde{\mathbf{w}}_i^{\mathsf{T}}\boldsymbol{\Sigma}_{ii}^{-1}\widetilde{\mathbf{w}}_i}} = \mathbf{L}_i \tag{4.44}$$

4.3.3 Estimation of the Structural Equations

Given the estimate of $\mathbf{R}_c$ we now focus on the estimation of $\mathbf{B}\mathbf{c}_{\text{endo}} = \mathbf{C}\mathbf{c}_{\text{exo}} + \mathbf{z}$. We have exclusion constraints for the structural form matrices $\mathbf{B}$ and $\mathbf{C}$, i.e., certain coefficients are a priori known to be zero. There are no restrictions on $\text{cov}(\mathbf{z})$, or if there are, we will ignore them here (for convenience, not as a matter of principle). This seems to exclude Wold's recursive system where the elements of $\mathbf{B}$ on one side of the diagonal are zero, and the equation-residuals are *un*correlated. But we can always regress the first endogenous composite $c_{\text{endo},1}$ on $\mathbf{c}_{\text{exo}}$, and $c_{\text{endo},2}$ on $[c_{endo,1}; \mathbf{c}_{\text{exo}}]$, and $c_{\text{endo},3}$ on $[c_{\text{endo},1}; c_{\text{endo},2}; \mathbf{c}_{\text{exo}}]$ et cetera. The ensuing residuals are *by construction* uncorrelated with the explanatory variables in their corresponding equations, and by implication they are mutually uncorrelated. In a sense, there are no assumptions here, the purpose of the exercise (prediction of certain variables using a specific set of predictors) determines the regression to be performed; there is also no identifiability issue.[15]

[14] A working paper version of this paper said that the elements of the mode A loading vector would always be "larger" than the corresponding true values. I am obliged to Michel Tenenhaus for making me realize that the statement was not true.

[15] See Dijkstra (2014) for further discussion of Wold's approach to modeling. There is a subtle issue here. One could generate a sample from a system with $\mathbf{B}$ lower-triangular, a full matrix $\mathbf{C}$

Now consider $\mathbf{P}$, the regression matrix obtained from regressing the (estimated) endogenous composites on the (estimated) exogenous composites. It estimates $\boldsymbol{\Pi}$, the reduced form matrix $\mathbf{B}^{-1}\mathbf{C}$. We will use $\mathbf{P}$, and possible other functions of $\mathbf{R}_c$, to estimate the free elements of $\mathbf{B}$ and $\mathbf{C}$. There is no point in trying when $\boldsymbol{\Pi}$ is compatible with different values of the structural form matrices. So the crucial question is whether $\boldsymbol{\Pi} = \boldsymbol{B}^{-1}\mathbf{C}$, or equivalently $\mathbf{B}\Pi = \mathbf{C}$, can be solved uniquely for the free elements of $\mathbf{B}$ and $\mathbf{C}$. Take the ith equation[16]

$$\mathbf{B}_{i\cdot}\boldsymbol{\Pi} = \boldsymbol{C}_{i\cdot} \tag{4.45}$$

where the ith row of $\mathbf{B}$, $\mathbf{B}_{i\cdot}$, has 1 in the ith entry (normalization) and possibly some zeros elsewhere, and where the ith row of $\mathbf{C}$, $\mathbf{C}_{i\cdot}$, may also contain some zeros. The free elements in $\mathbf{C}_{i\cdot}$ are given when those in $\mathbf{B}_{i\cdot}$ are known, and the latter are to be determined by the zeros in $\mathbf{C}_{i\cdot}$. More precisely

$$\mathbf{B}_{(i,k:B_{ik}\text{ free or unit})} \cdot \boldsymbol{\Pi}_{\left(k:B_{ik}\text{ free or unit},\, j:C_{ij}=0\right)} = 0 \tag{4.46}$$

So we have a submatrix of $\boldsymbol{\Pi}$, the rows correspond with the free elements (and the unit) in the ith row of $\mathbf{B}$, and the columns with the zero elements in the ith row of $\mathbf{C}$. This equation determines $\mathbf{B}_{(i,k:B_{ik}\text{ free or unit})}$ uniquely, apart from an irrelevant nonzero multiple, *when and only when* the particular submatrix of $\boldsymbol{\Pi}$ has a rank equal to its number of rows minus one. This is just the number of elements to be estimated in the ith row of $\mathbf{B}$. To have this rank requires the submatrix to have at least as many columns. So a little thought will give that a necessary condition for unique solvability, *identifiability*, is that we must have as least as many exogenous composites in the system as coefficients to be estimated in any one equation. We emphasize that this *order condition* as it is traditionally called is indeed nothing more than necessary.[17] The *rank condition* is both necessary and sufficient.

---

and a full, non-diagonal covariance matrix for $\mathbf{z}$. Then no matter how large the sample size, we can never retrieve the coefficients (apart from those of the first equation which are just regression coefficients). The regressions for the other equations would yield values different from those we used to generate the observations, since the zero correlation between their equation-residuals would be incompatible with the non-diagonality of cov($\mathbf{z}$).

[16] What follows will be old hat for econometricians, but since non-recursive systems are relatively new for PLS-practitioners, some elaboration could be meaningful.

[17] As an example consider a square $\mathbf{B}$ with units on the diagonal but otherwise unrestricted, and a square $\mathbf{C}$ of the same dimensions, containing zeros only except the last row, where all entries are free. The order condition applies to all equations but the last, but *none* of the coefficients can be retrieved from $\boldsymbol{\Pi}$. This matrix is, however, severely restricted: it has rank one. How to deal with this and similar situations is handled by Bekker et al. (1994).

A very simple example, which we will use in a small Monte Carlo study in the next subsection is as follows. Let

$$\begin{bmatrix} 1 & b_{12} \\ b_{21} & 1 \end{bmatrix} \begin{bmatrix} c_{\text{endo},1} \\ c_{\text{endo},2} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & 0 & 0 \\ 0 & 0 & c_{23} & c_{24} \end{bmatrix} \begin{bmatrix} c_{\text{exo},1} \\ c_{\text{exo},2} \\ c_{\text{exo},3} \\ c_{\text{exo},4} \end{bmatrix} + \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \tag{4.47}$$

with $1 - b_{12}b_{21} \neq 0$. The order conditions are satisfied: each equation has three free coefficients and there are four exogenous composites.[18] Note that

$$\mathbf{\Pi} = \frac{1}{1 - b_{12}b_{21}} \begin{bmatrix} c_{11} & c_{12} & -b_{12}c_{23} & -b_{12}c_{24} \\ -b_{21}c_{11} & -b_{21}c_{12} & c_{23} & c_{24} \end{bmatrix} \tag{4.48}$$

The submatrix of $\mathbf{\Pi}$ relevant for an investigation into the validity of the rank condition for the first structural form equation is

$$\begin{bmatrix} \Pi_{13} & \Pi_{14} \\ \Pi_{23} & \Pi_{24} \end{bmatrix} = \frac{1}{1 - b_{12}b_{21}} \begin{bmatrix} -b_{12}c_{23} & -b_{12}c_{24} \\ c_{23} & c_{24} \end{bmatrix} \tag{4.49}$$

It should have rank one, and it does so in the generic case, since its first row is a multiple of its second row.[19] Note that we cannot have both c_{23} and c_{24} zero. Clearly, b_{12} can be obtained from $\mathbf{\Pi}$ via $-\Pi_{13}/\Pi_{23}$ or via -Π_{14}/Π_{24}. A similar analysis applies to the second structural form equation. We note that the model imposes two constraints on $\mathbf{\Pi}$: $\Pi_{11}\Pi_{22} - \Pi_{12}\Pi_{21} = 0$ and $\Pi_{13}\Pi_{24} - \Pi_{14}\Pi_{23} = 0$, in agreement with the fact that the 8 reduced form coefficients can be expressed in terms of 6 structural form parameters. For an extended analysis of the number and type of constraints that a structural form imposes on the reduced form see Bekker and Dijkstra (1990) and Bekker et al. (1994).

It will be clear that the estimate $\mathbf{P}$ of $\mathbf{\Pi}$ will not in general satisfy the rank conditions (although we do expect them to be close for sufficiently large samples), and using either $-P_{13}/P_{23}$ or $-P_{14}/P_{24}$ as an estimate for b_{12} will give different answers. Econometric methods construct explicitly or implicitly compromises between the possible estimates. 2SLS, as discussed above is one of them. See Dijkstra and Henseler (2015a,b) for a specification of the relevant formula (formula

[18]With 2SLS $c_{\text{endo},2}$ in the first equation is in the first stage replaced by its regression on the four exogenous variables. In the second stage we regress $c_{\text{endo},1}$ on the replacement for $c_{\text{endo},2}$ and two exogenous variables. So the regression matrix with three columns in this stage is spanned by four exogenous columns, and we should be fine in general. If there were four exogenous variables on the right-hand side, the regression matrix in the second stage would have five columns, spanned by only four exogenous columns, the matrix would not be invertible and 2SLS (and all other methods aiming for consistency) would break down.

[19]For more general models one could ask MATLAB, say, to calculate the rank of the matrices, evaluated for arbitrary values. A very pragmatic approach would be to just run 2SLS. If it breaks down and gives a singularity warning, one should analyze the situation. Otherwise you are fine.

(23)) for 2SLS that honors the motivation via two regressions. Here we will outline another approach based on Dijkstra (1989) that is close to the discussion about identifiability.

Consider a row vector[20] with ith subvector $\mathbf{B}_{i\cdot}\mathbf{P}-\mathbf{C}_{i\cdot}$. If $\mathbf{P}$ would equal $\mathbf{\Pi}$ we could get the free coefficients by making $\mathbf{B}_{i\cdot}\mathbf{P}-\mathbf{C}_{i\cdot}$ zero. But that will not be the case. So we could decide to choose values for the free coefficients that make each $\mathbf{B}_{i\cdot}\mathbf{P}-\mathbf{C}_{i\cdot}$ as "close to zero as possible." One way to implement that is to minimize a suitable quadratic form subject to the exclusion constraints and normalizations. We take

$$\left(\text{vec}\left[(\mathbf{BP}-\mathbf{C})^{\intercal}\right]\right)^{\intercal}\cdot\left(\mathbf{W}\otimes\widehat{\mathbf{R}}_{\text{exo}}\right)\cdot\text{vec}\left[(\mathbf{BP}-\mathbf{C})^{\intercal}\right] \tag{4.50}$$

Here $\otimes$ stands for Kronecker's matrix multiplication symbol, $\widehat{\mathbf{R}}_{\text{exo}}$ is the estimated p.d. correlation matrix of the estimated exogenous composites, $\mathbf{W}$ is a p.d. matrix with as many rows and columns as there are endogenous composites, and the operator "vec" stacks the columns of its matrix-argument one underneath the other, starting with the first. If we take a diagonal matrix $\mathbf{W}$ the quadratic form disintegrates into separate quadratic forms, one for each subvector, and minimization yields in fact 2SLS estimates. A non-diagonal $\mathbf{W}$ tries to exploit information about the covariances between the subvectors. For the classical econometric simultaneous equation model it is true that $\text{vec}\left[(\mathbf{BP}-\mathbf{C})^{\intercal}\right]$ is asymptotically normal with zero mean and covariance matrix $\text{cov}(\mathbf{z})\otimes\mathbf{R}_{\text{exo}}^{-1}$ divided by the sample size, adapting the notation somewhat freely. General estimation theory tells us to use the inverse of an estimate of this covariance matrix in order to get asymptotic efficiency. So $\mathbf{W}$ should be the inverse of an estimate for $\text{cov}(\mathbf{z})$. The latter is traditionally estimated by the obvious estimate based on 2SLS. Note that the covariances between the structural form residuals drive the extent to which the various optimizations are integrated. There is no or little gain when there is no or little correlation between the elements of $\mathbf{z}$. This more elaborate method is called 3SLS.

We close with some observations. Since the quadratic form in the parameters is minimized subject to zero constraints and normalizations only, there is an explicit solution, see Dijkstra (1989, section 5), for the formulae.[21] If the fact that the weights are estimated can be ignored, there is also an explicit expression for the asymptotic covariance matrix, both for 2SLS and 3SLS. But if the sampling variation in the weights does matter, this formula may not be accurate and 3SLS may not be more efficient than 2SLS. Both methods are essentially non-iterative and very fast, and therefore suitable candidates for bootstrapping. One potential advantage of 2SLS over 3SLS is that it may be more robust to model specification errors, because as opposed to its competitor, it estimates equation by equation, so that an error in one equation need not affect the estimation of the others.

[20] This is in fact, see below: $\left(\text{vec}\left[(\mathbf{BP}-\mathbf{C})^{\intercal}\right]\right)^{\intercal}$.

[21] For the standard approach and the classical formulae, see, e.g., Ruud (2000)

4.3.4 Some Monte Carlo Results

We use the setup from Dijkstra and Henseler (2015a,b) adapted to the present setting. We have

$$\begin{bmatrix} 1 & -0.25 \\ -0.50 & 1 \end{bmatrix} \begin{bmatrix} c_{\text{endo},1} \\ c_{\text{endo},2} \end{bmatrix} = \begin{bmatrix} -0.30 & 0.50 & 0 & 0 \\ 0 & 0 & 0.50 & 0.25 \end{bmatrix} \begin{bmatrix} c_{\text{exo},1} \\ c_{\text{exo},2} \\ c_{\text{exo},3} \\ c_{\text{exo},4} \end{bmatrix} + \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \tag{4.51}$$

All variables have zero mean, and we will take them jointly normal. Cov($\mathbf{c}_{\text{exo}}$) has ones on the diagonal and 0.50 everywhere else; the variances of the endogenous composites are also one and we take cov($c_{\text{endo},1}$,$c_{\text{endo},2}$) $= \sqrt{0.50}$. The values as specified imply for the covariance matrix for the structural form residuals $\mathbf{z}$:

$$\text{cov}(\mathbf{z}) = \begin{bmatrix} 0.5189 & -0.0295 \\ -0.0295 & 0.1054 \end{bmatrix} \tag{4.52}$$

Note that the correlation between z_1 and z_2 is rather small, -0.1261, so the setup has the somewhat unfortunate consequence to potentially favor 2SLS. The R-squared for the first reduced form equation is 0.3329 and for the second reduced form equation this is 0.7314.

Every composite is built up by three indicators, with a covariance matrix that has ones on the diagonal and 0.49 everywhere else. This is compatible with a one-factor model for each vector of indicators but we have no use nor need for that interpretation here.

The composites ($c_{\text{exo},1}$,$c_{\text{exo},2}$,$c_{\text{exo},3}$,$c_{\text{exo},4}$,$c_{\text{endo},1}$,$c_{\text{endo},2}$) need weights. For the first and fourth we take weights proportional to $[1, 1, 1]$. For the second and fifth the weights are proportional to $[1, 2, 3]$ and for the third and sixth they are proportional to $[1, 4, 9]$. There are no deep thoughts behind these choices.

We get the following weights (rounded to two decimals for readability): $[0.41, 0.41, 0.41]$ for blocks one and four, $[0.20, 0.40, 0.60]$ for blocks two and five, and $[0.08, 0.33, 0.74]$ for blocks three and six.

The loadings are now given as well: $[0.81, 0.81, 0.81]$ for blocks one and four, $[0.69, 0.80, 0.90]$ for blocks two and five, and $[0.61, 0.74, 0.95]$ for blocks three and six.

One can now calculate the 18 by 18 covariance/correlation matrix $\boldsymbol{\Sigma}$ and its unique p.d. matrix square root $\boldsymbol{\Sigma}^{1/2}$. We generate samples of size 300, which appears to be relatively modest given the number of parameters to estimate. A sample of size 300 is obtained via $\boldsymbol{\Sigma}^{1/2}\times\text{randn}(18, 300)$. We repeat this ten thousand times, each time estimating the weights via MAXVAR,[22] the loadings

[22]One might as well have used mode B of course, or any of the other canonical variables approaches. There is no *fundamental* reason to prefer one to the other. MAXVAR was available, and is essentially non-iterative.

via regressions and the correlations in the obvious way, and all structural form parameters via 2SLS and 3SLS using standardized indicators.[23]

The loadings and weights are on the average slightly underestimated, see Dijkstra (2015) for some of the tables: when rounded to two decimals the difference is at most 0.01. The standard deviations of the weights estimators for the endogenous composites are either the largest or the smallest: for the weights of $c_{\text{endo},1}$ we have resp. $[0.12, 0.12, 0.11]$ and for $c_{\text{endo},2}$ $[0.04, 0.04, 0.04]$; the standard deviations for the weights of the exogenous composites are, roughly, in between. And similarly for the standard deviations for the loadings estimators: for the loadings on $c_{\text{endo},1}$ we have resp. $[0.08, 0.07, 0.05]$ and for $c_{\text{endo},2}$ $[0.05, 0.04, 0.01]$; the standard deviations for the loadings on the exogenous composites are again, roughly, in between.

The following table gives the results for the coefficients in **B** and **C**, rounded to two decimals:

	Value	Mean 2SLS	Mean 3SLS	std 2SLS	std 3SLS
b_{12}	−0.25	−0.26	−0.26	0.08	0.08
b_{21}	−0.50	−0.50	−0.50	0.05	0.05
c_{11}	−0.30	−0.29	−0.28	0.05	0.05
c_{12}	+0.50	+0.49	+0.49	0.06	0.06
c_{23}	+0.50	+0.49	+0.49	0.03	0.03
c_{24}	+0.25	+0.25	+0.25	0.03	0.03

Clearly, for the model at hand 3SLS has nothing to distinguish itself positively from 2SLS[24] (its standard deviations are only smaller than those of 2SLS when we use three decimals). This might be different when the structural form residuals are materially correlated.

We also calculated, not shown, for each of the 10,000 samples of size 300 the theoretical (asymptotic) standard deviations for the 3SLS estimators. They are all on the average 0.01 smaller than the values in the table, they are relatively stable, with standard deviations ranging from 0.0065 for b_{12} to 0.0015 for c_{24}. They are not perfect but not really bad either.

It would be reckless to read too much into this small and isolated study, for one type of distribution. But the approach does appear to be feasible.

[23] The whole exercise takes about half a minute on a slow machine: 4CPU 2.40 Ghz; RAM 512 MB.

[24] It is remarkable that the accuracy of the 2SLS and 3SLS estimators is essentially as good, in three decimals, as those reported by Dijkstra and Henseler (2015a,b) for Full Information Maximum Likelihood (FIML) for the same model in terms of latent variables, i.e., FIML as applied to the *true* latent variable scores. See Table 2 on p. 18 there. When the latent variables are not observed directly but only via indicators, the performance of FIML clearly deteriorates (stds are doubled or worse).

4.4 Testing the Composites Model

In this section we sketch four more or less related approaches to test the appropriateness or usefulness of the model. In practice one might perhaps want to deploy all of them. Investigators will easily think of additional, "local" tests, like those concerning the signs or the order of magnitude of coefficients et cetera.

A thorny issue that should be mentioned here is *capitalization on chance*, which refers to the phenomenon that in practice one runs through cycles of model testing and adaptation until the current model tests signal that all is well according to popular rules-of-thumb.[25] This makes the model effectively stochastic, random. Taking a new sample and going through the cycles of testing and adjusting all over again may well lead to another model. But when we give estimates of the distribution functions of our estimators we imply that this helps to assess how the estimates will vary when other samples of the same size would be employed, while keeping the model *fixed*. It is tempting, but potentially very misleading, to ignore the fact that the sample (we/you, actually) favored a particular model after a (dedicated) model search, see Freedman et al. (1988), Dijkstra and Veldkamp (1988), Leeb and Pötscher (2006), and Freedman (2009)[26]. It is not clear at all how to properly validate the model on the very same data that gave it birth, while using *test* statistics as *design* criteria.[27] Treating the results conditional *on the sample at hand*, as purely descriptive (which in itself may be rather useful, Berk 2008), or testing the model on a fresh sample (e.g., a random subset of the data that was kept apart when the model was constructed), while bracing oneself for a possibly big disappointment, appear to be the best or most honest responses.

[25] "Capitalization on chance" is sometimes used when "small-sample-bias" is meant. That is quite something else.

[26] Freedman gives the following example. Let the 100×51 matrix $[\mathbf{y}, \mathbf{X}]$ consists of independent standard normals. So there is no (non-) linear relationship whatsoever. Still, a regression of $\mathbf{y}$ on $\mathbf{X}$ can be expected to yield an R-square of 0.50. On the average there will be 5 regression coefficients that are significant at 10%. If we keep the corresponding $\mathbf{X}$-columns in the spirit of "exploratory research" and discard the others, a regression could easily give a decent R-square and "dazzling t-statistics" (Freedman 2009, p.75). Note that here the "dedicated" model search consisted of merely two regression rounds. Just think of what one can accomplish with a bit more effort, see also, e.g., Dijkstra (1995).

[27] At one point I thought that "a way out" would be to condition on the set of samples that favor the chosen model using the same search procedure (Dijkstra and Veldkamp 1988): if the model search has led to the simplest true model, the conditional estimator distribution equals, asymptotically, the distribution that the practitioner reports. This conditioning would give substance to the retort given in practice that "we always condition on the given model." But the result referred to says essentially that we can ignore the search *if* we know it was *not* needed. So much for comfort. It is even a lot worse: Leeb and Pötscher (2006) show that convergence of the conditional distribution is only pointwise, not uniform, not even on compact subsets of the parameter space. The bootstrap cannot alleviate this problem, Leeb and Pötscher (2006), Dijkstra and Veldkamp (1988).

4.4.1 Testing Rank Restrictions on Submatrices

The covariance matrix of any subvector of $\mathbf{y}_i$ with any choice from the other indicators has rank one. So the corresponding regression matrix has rank one. To elaborate a bit, since $\mathrm{E}(c_1|c_2, c_3, \ldots, c_N)$ is a linear function of $\mathbf{y}$ the formula $\mathrm{E}(\mathbf{y}_1|\mathbf{y}_2, \mathbf{y}_3, \ldots, \mathbf{y}_N) = \mathbf{L}_1\mathrm{E}(c_1|c_2, c_3, \ldots, c_N)$ tells us that the regression matrix is a column times a row vector. Therefore its $p_1 \cdot (p - p_1)$ elements can be expressed in terms of just $(p-1)$ parameters (one row of $(p-p_1)$ elements plus $(p_1 - 1)$ proportionality factors). This number could be even smaller when the model imposes structural constraints on $\mathbf{R}_c$ as well. A partial check could be performed using any of the methods developed for restricted rank testing. A possible objection could be that the tests are likely to be sensitive to deviations from the Gaussian distribution, but jackknifing or bootstrapping might help to alleviate this. Another issue is the fact that we get many tests that are also correlated, so that simultaneous testing techniques based on Bonferroni or more modern approaches are required.[28]

4.4.2 Exploiting the Difference Between Different Estimators

We noted that a number of generalized canonical variable programs yield identical results when applied to a $\mathbf{\Sigma}$ satisfying the composites factor model. But we expect to get different results when this is not the case. So, when using the estimate for $\mathbf{\Sigma}$ one might want to check whether the differences between, say PLS mode B and MAXVAR (or any other couple of methods), are too big for comfort. The scale on which to measure this could be based on the probability (as estimated by the bootstrap) of obtaining a larger "difference" than actually observed.

4.4.3 Prediction Tests, via Cross-Validation

The path diagram might naturally indicate composites and indicators that are most relevant for prediction. So it would seem to make sense to test whether the model's rank restrictions can help improve predictions of certain selected composites or indicators. The result will not only reflect model adequacy but also the statistical phenomenon that the imposition of structure, even when strictly unwarranted, can help in prediction. It would therefore also reflect the sample size. The reference for an elaborate and fundamental discussion of prediction and cross-validation in a PLS-context is Shmueli et al. (2016).

[28] See, e.g., chapter 34 from DasGupta (2008).

4.4.4 Global Goodness-of-Fit Tests

In SEM we test the model by assessing the probability value of a distance measure between the sample covariance matrix $\mathbf{S}$ and an estimated matrix $\widehat{\boldsymbol{\Sigma}}$ that satisfies the model. Popular measures are

$$\frac{1}{2}\mathrm{tr}\left(\mathbf{S}^{-1}\left(\mathbf{S}-\widehat{\boldsymbol{\Sigma}}\right)\right)^2 \tag{4.53}$$

and

$$\mathrm{tr}\left(\mathbf{S}\widehat{\boldsymbol{\Sigma}}^{-1}\right)-\log\left(\det\left(\mathbf{S}\widehat{\boldsymbol{\Sigma}}^{-1}\right)\right)-p \tag{4.54}$$

They belong to a large class of distances, all expressible in terms of a suitable function f:

$$\sum_{k=1}^{p} f\left(\gamma_k\left(\mathbf{S}^{-1}\widehat{\Sigma}\right)\right). \tag{4.55}$$

Here $\gamma_k(\cdot)$ is the kth eigenvalue of its argument, and f is essentially a smooth real function defined on positive real numbers, with a unique global minimum of zero at the argument value 1.The functions are "normalized," $f''(1)=1$, entailing that the second-order Taylor expansions around 1 are identical.[29] For the examples referred to we have $f(\gamma)=\frac{1}{2}(1-\gamma)^2$ and $f(\gamma)=1/\gamma+\log(\gamma)-1$, respectively. Another example is $f(\gamma)=\frac{1}{2}(\log(\gamma))^2$, the so-called geodesic distance; its value is the same whether we work with $\mathbf{S}^{-1}\widehat{\boldsymbol{\Sigma}}$ or with $\mathbf{S}\widehat{\boldsymbol{\Sigma}}^{-1}$. The idea is that when the model fits perfectly, so $\mathbf{S}^{-1}\widehat{\boldsymbol{\Sigma}}$ is the identity matrix, then all its eigenvalues equal one, and conversely. This class of distances was first analyzed by Swain (1975).[30] Distance measures outside of this class are those induced by WLS with general fourth-order moments based weight matrices,[31] but also the simple ULS: $\mathrm{tr}\left(\mathbf{S}-\widehat{\boldsymbol{\Sigma}}\right)^2$. We can take any of these measures, calculate its value, and use the bootstrap to estimate the corresponding probability value. It is important to pre-multiply the observation vectors by $\widehat{\boldsymbol{\Sigma}}^{\frac{1}{2}}\mathbf{S}^{-\frac{1}{2}}$ before the bootstrap is implemented, in order to ensure that their empirical distribution has a covariance matrix that agrees with the assumed model.

[29]The estimators based on minimization of these distances are asymptotically equivalent. The value of the third derivative of f appears to affect the bias: high values tend to be associated with small residual variances. So the first example, "GLS," with $f'''(1)=0$, will tend to underestimate these variances more than the second example, "LISREL," with $f'''(1)=-4$. See Swain (1975).

[30]Swain (1975). See also Dijkstra (1990).

[31]The manual of EQS, Bentler (2006) is a treasure trove with information on goodness-of-fit testing with WLS, and Structural Equations Modeling generally. For related discussions, see Bentler and Dijkstra (1985) and Wansbeek and Meijer (2000).

For $\widehat{\boldsymbol{\Sigma}}$ one could take in an obvious notation $\widehat{\boldsymbol{\Sigma}}_{ii} := \mathbf{S}_{ii}$ and for $i \neq j$

$$\widehat{\boldsymbol{\Sigma}}_{ij} := \widehat{r}_{ij} \cdot \mathbf{S}_{ii}\widehat{\mathbf{w}}_i \cdot \widehat{\mathbf{w}}_j^{\top}\mathbf{S}_{jj}. \tag{4.56}$$

Here $\widehat{r}_{ij} = \widehat{\mathbf{w}}_i^{\top}\mathbf{S}_{ij}\widehat{\mathbf{w}}_j$ if there are no constraints on $\mathbf{R}_c$, otherwise it will be the *ij*th element of $\widehat{\mathbf{R}}_c$. If $\mathbf{S}$ is p.d., then $\widehat{\boldsymbol{\Sigma}}$ is p.d. (as follows from the appendix) and $\widehat{\boldsymbol{\Sigma}}^{\frac{1}{2}}\mathbf{S}^{-\frac{1}{2}}$ is well-defined.

4.5 Some Final Observations and Comments

In this chapter we outlined a model in terms of observables only while adhering to *the soft modeling principle* of Wold's PLS. Wold developed his methods against the backdrop of a particular *latent variables model*, the basic design. This introduces N additional *un*observable variables which by necessity cannot in general be expressed unequivocally in terms of the "manifest variables," the indicators. However, we can construct composites that satisfy the same structural equations as the latent variables, in an infinite number of ways in fact. Also, we can design composites such that the regression of the indicators on the composites yields the loadings. But in the regular case *we cannot have both*.

Suppose $\mathbf{y} = \boldsymbol{\Lambda}\mathbf{f} + \boldsymbol{\varepsilon}$ with $\mathrm{E}\mathbf{f}\boldsymbol{\varepsilon}^{\top} = 0$, $\boldsymbol{\Theta} := \mathrm{cov}(\boldsymbol{\varepsilon}) > 0$, and $\boldsymbol{\Lambda}$ has full column rank. The p.d. cov($\mathbf{f}$) will satisfy the constraints as implied by identifiable equations like $\mathbf{B}\mathbf{f}_{\text{endo}} = \mathbf{C}\mathbf{f}_{\text{exo}} + \boldsymbol{\zeta}$ with $\mathrm{E}\mathbf{f}_{\text{exo}}\boldsymbol{\zeta}^{\top} = 0$. All variables have zero mean. Let $\widehat{\mathbf{f}}$, of the same dimension as $\mathbf{f}$, equal $\mathbf{F}\mathbf{y}$ for a fixed matrix $\mathbf{F}$. If the regression of $\mathbf{y}$ on $\widehat{\mathbf{f}}$ yields $\boldsymbol{\Lambda}$ we must have $\mathbf{F}\boldsymbol{\Lambda} = \mathbf{I}$ because then

$$\boldsymbol{\Lambda} = \mathrm{E}\left[\mathbf{y}\,(\mathbf{F}\mathbf{y})^{\top}\right] \cdot [\mathrm{cov}\,(\mathbf{F}\mathbf{y})]^{-1} = \mathrm{cov}\,(\mathbf{y})\,\mathbf{F}^{\top}[\mathbf{F}\mathrm{cov}\,(\mathbf{y})\,\mathbf{F}^{\top}]^{-1} \tag{4.57}$$

Consequently

$$\widehat{\mathbf{f}} = \mathbf{F}\,(\boldsymbol{\Lambda} f + \boldsymbol{\varepsilon}) = \mathbf{f} + \mathbf{F}\boldsymbol{\varepsilon} \tag{4.58}$$

and $\widehat{\mathbf{f}}$ has a larger covariance matrix then $\mathbf{f}$ (the difference is p.s.d., usually p.d.). One example is[32] $\mathbf{F} = \left(\boldsymbol{\Lambda}^{\top}\boldsymbol{\Theta}^{-1}\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^{\top}\boldsymbol{\Theta}^{-1}$ with $\mathrm{cov}\left(\widehat{\mathbf{f}}\right) - \mathrm{cov}(\mathbf{f}) = \left(\boldsymbol{\Lambda}^{\top}\boldsymbol{\Theta}^{-1}\boldsymbol{\Lambda}\right)^{-1}$.

So, generally, if the regression of $\mathbf{y}$ on the composites yields $\boldsymbol{\Lambda}$, the covariance matrices cannot be the same, and the composites cannot satisfy the same equations

[32] One can verify directly that the regression yields $\boldsymbol{\Lambda}$. Also note that here $\mathbf{F}\boldsymbol{\Lambda} = \mathbf{I}$.

as the latent variables $\mathbf{f}$.[33] Conversely, if $\mathrm{cov}\left(\widehat{\mathbf{f}}\right) = \mathrm{cov}(\mathbf{f})$, then the regression of $\mathbf{y}$ on the composites cannot yield $\boldsymbol{\Lambda}$.

If we minimize $\mathrm{E}(\mathbf{y}-\boldsymbol{\Lambda}\mathbf{F}\mathbf{y})^{\mathsf{T}}\,\boldsymbol{\Theta}^{-1}\,(\mathbf{y}-\boldsymbol{\Lambda}\mathbf{F}\mathbf{y})$ subject to $\mathrm{cov}(\mathbf{F}\mathbf{y}) = \mathrm{cov}(\mathbf{f})$ we get the composites that LISREL reports. We can generate an infinite number of alternatives[34] by minimizing $\mathrm{E}(\mathbf{f}-\mathbf{F}\mathbf{y})^{\mathsf{T}}\,\mathbf{V}\,(\mathbf{f}-\mathbf{F}\mathbf{y})$ subject to $\mathrm{cov}(\mathbf{F}\mathbf{y}) = \mathrm{cov}(\mathbf{f})$ for any conformable p.d. $\mathbf{V}$. Note that each composite here typically uses *all* indicators. Wold takes composites that combine the indicators per block. Of course, they also cannot reproduce the measurement equations and the structural equations, but the parameters can be obtained (consistently estimated) using suitable corrections (PLSc.[35])

Two challenging research topics present themselves: *first*, the extension of the approach to more dimensions/layers, and *second*, the imposition of sign constraints on weights, loadings, and structural coefficients, while maintaining as far as possible the numerical efficiency of the approach.

Appendix

Here we will prove that $\boldsymbol{\Sigma}$ is positive definite when and only when the correlation matrix of the composites, $\mathbf{R}_c$, is positive definite. The "only when"-part is trivial. The proof that $\{\mathbf{R}_c$ is p.d.$\}$ implies $\{\boldsymbol{\Sigma}$ is p.d.$\}$ is a bit more involved. It is helpful to note *for that purpose* that we may assume that each $\boldsymbol{\Sigma}_{ii}$ is a unit matrix (pre-multiply and post-multiply by a block-diagonal matrix with $\boldsymbol{\Sigma}_{ii}^{-\frac{1}{2}}$ on the diagonal, and redefine $\mathbf{w}_i$ such that $\mathbf{w}_i^{\mathsf{T}}\mathbf{w}_i = 1$ for each i). So if we want to know whether the eigenvalues of $\boldsymbol{\Sigma}$ are positive it suffices to study the eigenvalue problem $\widetilde{\boldsymbol{\Sigma}}\mathbf{x} = \gamma\mathbf{x}$:

$$\begin{bmatrix} \mathbf{I}_{p_1} & r_{12}\mathbf{w}_1\mathbf{w}_2^{\mathsf{T}} & r_{13}\mathbf{w}_1\mathbf{w}_3^{\mathsf{T}} & \cdot & r_{1N}\mathbf{w}_1\mathbf{w}_N^{\mathsf{T}} \\ & \mathbf{I}_{p_2} & r_{23}\mathbf{w}_2\mathbf{w}_3^{\mathsf{T}} & \cdot & r_{2N}\mathbf{w}_2\mathbf{w}_N^{\mathsf{T}} \\ & & \cdot & \cdot & \cdot \\ & & & \mathbf{I}_{p_{N-1}} & r_{N-1,N}\mathbf{w}_{N-1}\mathbf{w}_N^{\mathsf{T}} \\ & & & & \mathbf{I}_{p_N} \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \cdot \\ \mathbf{x}_{N-1} \\ \mathbf{x}_N \end{bmatrix} = \gamma \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \cdot \\ \mathbf{x}_{N-1} \\ \mathbf{x}_N \end{bmatrix} \tag{4.59}$$

[33] One may wonder about the "best linear predictor" of $\mathbf{f}$ in terms of $\mathbf{y}$: $\mathrm{E}(\mathbf{f}|\,\mathbf{y})$. Since $\mathbf{f}$ equals $\mathrm{E}(\mathbf{f}|\,\mathbf{y})$ plus an uncorrelated error vector, $\mathrm{cov}(E\,(\mathbf{f}|\,\mathbf{y}))$ is not "larger" but "smaller" than $\mathrm{cov}(\mathbf{f})$. So $\mathrm{E}(\mathbf{f}|\,\mathbf{y})$ satisfies neither of the two desiderata.

[34] Dijkstra (2015).

[35] PLSc exploits the lack of correlation between some of the measurement errors *within* blocks. It is sometimes equated to a *particular* implementation (e.g., assuming all errors are uncorrelated, and a specific correction), but that is selling it short. See Dijkstra (2011, 2013a,b) and Dijkstra and Schermelleh-Engel (2014).

with obvious implied definitions. Observe that every nonzero solution of

$$\begin{bmatrix} \mathbf{w}_1^\mathsf{T} & \mathbf{0} & \cdot & \cdot & \mathbf{0} \\ \mathbf{0} & \mathbf{w}_2^\mathsf{T} & \mathbf{0} & \cdot & \mathbf{0} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \mathbf{0} & \mathbf{w}_{N-1}^\mathsf{T} & \mathbf{0} \\ \mathbf{0} & \cdot & \cdot & \mathbf{0} & \mathbf{w}_N^\mathsf{T} \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \cdot \\ \mathbf{x}_{N-1} \\ \mathbf{x}_N \end{bmatrix} = \mathbf{0} \tag{4.60}$$

corresponds with $\gamma = 1$, and there are $\sum_{i=1}^{N} p_i - N$ linearly independent solutions. The multiplicity of the root $\gamma = 1$ is therefore $\sum_{i=1}^{N} p_i - N$ and we need to find N more roots. By assumption $\mathbf{R}_c$ has N positive roots. Let $\mathbf{u}$ be an eigenvector with eigenvalue μ, so $\mathbf{R}_c\mathbf{u} = \mu \cdot \mathbf{u}$. We have

$$\widetilde{\boldsymbol{\Sigma}} \begin{bmatrix} u_1\mathbf{w}_1 \\ u_2\mathbf{w}_2 \\ \cdot \\ u_N\mathbf{w}_N \end{bmatrix} = \begin{bmatrix} (u_1 + r_{12}u_2 + \cdot + r_{1N}u_N)\,\mathbf{w}_1 \\ (r_{21}u_1 + u_2 + \cdot + r_{2N}u_N)\,\mathbf{w}_2 \\ \cdot \\ (r_{N1}u_1 + r_{N2}u_2 + \cdot + u_N)\,\mathbf{w}_N \end{bmatrix} = \mu \begin{bmatrix} u_1\mathbf{w}_1 \\ u_2\mathbf{w}_2 \\ \cdot \\ u_N\mathbf{w}_N \end{bmatrix} \tag{4.61}$$

In other words, the remaining eigenvalues are those of $\mathbf{R}_c$, and so all eigenvalues of $\widetilde{\boldsymbol{\Sigma}}$ are positive. Therefore $\boldsymbol{\Sigma}$ is p.d., as claimed.

Note for the determinant of $\boldsymbol{\Sigma}$ that

$$\det(\boldsymbol{\Sigma}) = \det(\mathbf{R}_c) \times \det(\boldsymbol{\Sigma}_{11}) \times \det(\boldsymbol{\Sigma}_{22}) \times \det(\boldsymbol{\Sigma}_{33}) \times \ldots \times \det(\boldsymbol{\Sigma}_{NN}) \tag{4.62}$$

and so the Kullback–Leibler' divergence between the Gaussian density for block-independence and the Gaussian density for the composites model is $-\frac{1}{2}\log(\det(\mathbf{R}_c))$. It is well known that $0 \leq \det(\mathbf{R}_c) \leq 1$, with 0 in case of a perfect linear relationship between the composites, so Kullback–Leibler divergence is infinitely large, and 1 in case of zero correlations between all composites, with zero divergence.

References

Bekker, P. A., & Dijkstra, T. K. (1990). On the nature and number of the constraints on the reduced form as implied by the structural form. *Econometrica, 58*(2), 507–514

Bekker, P. A., Merckens, A., & Wansbeek, T. J. (1994). *Identification, equivalent models and computer algebra*. Boston: Academic.

Bentler, P. M., & Dijkstra, T. K. (1985). Efficient estimation via linearization in structural models. In P. R. Krishnaiah (Ed.), *Multivariate analysis* (Chap 2, pp. 9–42). Amsterdam: North-Holland.

Bentler, P. M. (2006). EQS 6 structural equations program manual. Multivariate Software Inc.

Berk, R. A. (2008). *Statistical learning from a regression perspective.* New York: Springer.

Boardman, A., Hui, B., & Wold, H. (1981). The partial least-squares fix point method of estimating interdependent systems with latent variables. *Communications in Statistics-Theory and Methods, 10*(7), 613–639.

DasGupta, A. (2008). *Asymptotic theory of statistics and probability*. New York: Springer,
Dijkstra, T. K. (1983). Some comments on maximum likelihood and partial least squares methods. *Journal of Econometrics, 22*(1/2), 67–90 (Invited contribution to the special issue on the Interfaces between Econometrics and Psychometrics).
Dijkstra, T. K. (1981). *Latent variables in linear stochastic models* (PhD thesis, University of Groningen). Available on Research Gate.
Dijkstra, T. K. (1989). Reduced Form estimation, hedging against possible misspecification. *International Economic Review, 30*(2), 373–390.
Dijkstra, T. K. (1990). Some properties of estimated scale invariant covariance structures. *Psychometrika 55*(2), 327–336.
Dijkstra, T. K. (1995). Pyrrho's Lemma, or have it your way. *Metrika, 42*(1), 119–125.
Dijkstra, T. K. (2010). Latent variables and indices: Herman Wold's basic design and partial least squares. In V. E. Vinzi, W. W. Chin, J. Henseler & H. Wang (Eds.), *Handbook of partial least squares, concepts, methods and applications* (Chap, 1, pp. 23–46). Berlin: Springer.
Dijkstra, T. K. (2011). *Consistent partial least squares estimators for linear and polynomial factor models*. Technical Report. Research Gate. doi:10.13140/RG.2.1.3997.0405.
Dijkstra, T. K. (2013a). *A note on how to make PLS consistent*. Technical Report. Research Gate, doi:10.13140/RG.2.1.4547.5688.
Dijkstra, T. K. (2013b). *The simplest possible factor model estimator, and successful suggestions how to complicate it again.* Technical Report. Research Gate. doi:10.13140/RG.2.1.3605.6809.
Dijkstra, T. K. (2014). PLS' Janus face. *Long Range Planning, 47*(3), 146–153.
Dijkstra, T. K. (2015). *All-inclusive versus single block composites.* Technical Report. Research Gate. doi:10.13140/RG.2.1.2917.8082.
Dijkstra, T. K., & Henseler, J. (2015a). Consistent and asymptotically normal PLS estimators for linear structural equations. *Computational Statistics and Data Analysis, 81*, 10–23.
Dijkstra, T. K., & Henseler, J. (2015b). Consistent partial least squares path modeling. *MIS Quarterly, 39*(2), 297–316.
Dijkstra, T. K., & Schermelleh-Engel, K. (2014). Consistent partial least squares for nonlinear structural equation models. *Psychometrika, 79*(4), 585–604 [published online (2013)].
Dijkstra, T. K., & Veldkamp, J. H. (1988). Data-driven selection of regressors and the bootstrap. In T. K. Dijkstra (Ed.), *On model uncertainty and its statistical implications* (Chap. 2, pp. 17–38). Berlin: Springer.
Freedman, D. A. (2009). *Statistical models, theory and practice*. Cambridge: Cambridge University Press. Revised ed.
Freedman, D. A., Navidi, W., & Peters, S. C. (1988). On the impact of variable selection in fitting regression equations. In T. K. Dijkstra (Ed.), *On model uncertainty and its statistical implications* (Chap. 1, pp. 1–16). Berlin: Springer.
Haavelmo, T. (1944). The probability approach in econometrics. PhD-thesis *Econometrica 12*(Suppl.), 118pp. http://cowles.econ.yale.edu/
Leeb, H., & Pötscher, B. M. (2006). Can one estimate the conditional distribution of post-model-selection estimators? *The Annals of Statistics, 34*(5), 2554–2591.
Ruud, P. A. (2000). *Classical econometric theory*. New York: Oxford University Press.
Shmueli, G., Ray, S., Velasquez Estrada, J. M., & Chatla, S. (2016). The elephant in the room: Predictive performance of PLS models. *Journal of Business Research, 69*, 4552–4564.
Kettenring, J. R. (1971). Canonical analysis of several sets of variables. *Biometrika, 58*(3), 433–451.
Pearl, J. (2009). *Causality—models, reasoning and inference*. Cambridge: Cambridge University Press.
Swain, A. J. (1975). A class of factor analysis estimation procedures with common asymptotic sampling properties. *Psychometrika, 40*, 315–335.
Wansbeek, T. J. & Meijer, E. (2000). *Measurement error and latent variables in econometrics*. Amsterdam: North-Holland.
Wold, H. (1966). Nonlinear estimation by iterative least squares procedures. In F. N. David (Ed.), *Research papers in statistics. Festschrift for J. Neyman* (pp. 411–444). New York: Wiley.

Wold, H. (1975). Path models with latent variables: The NIPALS approach. In H. M. Blalock et al. (Eds.), *Quantitative sociology* (Chap. 11, pp. 307–358). New York: Academic.

Wold, H. (1982). Soft modeling: The basic design and some extensions. In K. G. Jöreskog & H. Wold (Eds.), *Systems under indirect observation, Part II* (Chap. 1, pp. 1–54). Amsterdam: North-Holland.

Chapter 5
Quantile Composite-Based Model: A Recent Advance in PLS-PM

A Preliminary Approach to Handle Heterogeneity in the Measurement of Equitable and Sustainable Well-Being

Cristina Davino, Pasquale Dolce, and Stefania Taralli

Abstract The aim of the present chapter is to discuss a recent contribution in the partial least squares path modeling framework: the quantile composite-based path modeling. We introduce this recent contribution from both a methodological and an applicative point of view. The objective is to provide an exploration of the whole dependence structure and to highlight whether and how the relationships among variables (both observed and unobserved) change across quantiles. We use a real data application, measuring the equitable and sustainable well-being of Italian provinces. Partial least squares path modeling is first applied to study the relationships among variables assuming homogeneity among observations. Afterwards, a multi-group analysis is performed, assuming that a specific factor (the geographic area) causes heterogeneity in the population. Finally, the quantile approach to composite-based path modeling provides a more in-depth analysis. Some relevant results are selected and described to show that the quantile composite-based path modeling can be very useful in this real data application, as it allows us to explore territorial disparities in depth.

C. Davino (✉)
University of Macerata, Macerata, Italy
e-mail: cristina.davino@unimc.it

P. Dolce
StatSC, Oniris, INRA, Nantes, France
e-mail: pasquale.dolce@unina.it

S. Taralli
Istat, Italian National Institute of Statistics, Marche Regional Office, Ancona, Italy
e-mail: taralli@istat.it

H. Latan, R. Noonan (eds.), *Partial Least Squares Path Modeling*,
DOI 10.1007/978-3-319-64069-3_5

5.1 Introduction

The aim of this chapter is to discuss a recent advance in the partial least squares path modeling (PLS-PM) framework: the quantile composite-based path modeling (QC-PM). QC-PM was recently introduced by Davino and Vinzi (2016) and further developments from the theoretical and applicative point of view have been described in few papers (Davino 2016; Davino et al. 2016a,b). QC-PM can be considered a complementary approach to PLS-PM, one of the most widespread methods used to analyse a network of relationships between unobserved and observed variables. QC-PM aims to broaden such analysis by going beyond the estimation of average effects in the network of relationships among variables. In particular, QC-PM aims to highlight whether and how the relationships among observed and unobserved variables as well as among the unobserved variables change according to the quantile of interest, thus providing an exploration of the whole dependence structure. To this purpose, quantile regression (QR) and quantile correlation (QC) are introduced in all the estimation phases of a PLS-PM algorithm.

In this chapter, QC-PM is examined from both a methodological and an applicative point of view. The former is based on the description of the main features and potentialities of QR. The latter is provided in the context of a highly debated issue: the measurement of Equitable and Sustainable Well-being (the so-called Bes, from the Italian acronym of Benessere Equo e Sostenibile). In 2011, the Italian National Institute of Statistics (ISTAT) launched two pilot projects to deepen the measurement of BES at the local level: UrBes and Provinces' BES. This chapter deals with the set of indicators that have been selected and implemented by the Provinces' BES Statistical Information System (SIS)[1] (Cuspi-Istat 2015a). A hierarchical construct model is implemented to measure the BES index and to study the relationships among the BES components, as BES is a multidimensional construct composed of several different domains and themes. The chapter provides a comparison of the results obtained using both the classical PLS-PM and the QC-PM approach.

The chapter also proposes a preliminary in-depth analysis of BES in the case of observed heterogeneity. It is a matter of fact that living conditions in Italy vary widely across the different geographic areas of the country. For this reason, it is advisable to support the global analysis with an evaluation of the differences among the Italian provinces.

This chapter is organized as follows. Section 5.2 provides a detailed description of the concept of equitable and sustainable well-being. A brief description of the data is provided in Sect. 5.2.2. Sections 5.3 and 5.4 illustrate the theoretical PLS-PM model and the main results obtained both on the whole set of Italian provinces and taking into account the geographic area. A detailed description of QR potentialities

[1] http://www.besdelleprovince.it/.

with respect to QC-PM is reported in Sect. 5.5 together with the BES results. Finally, some conclusions are given in Sect. 5.6.

5.2 Measuring Equitable and Sustainable Well-Being

5.2.1 The Reference Framework

Equitable and sustainable well-being is the official Italian statistical contribution to the commitment of the European Statistical System (ESS) to measure progress going beyond gross domestic product (GDP). This effort intends to develop high-quality statistics able to provide a shared vision of well-being and to support policy-making. The so-called BES Project, from the Italian acronym of Benessere Equo e Sostenibile, was jointly undertaken by the ISTAT and the Italian National Council for Economy and Labour (CNEL) in 2010 (Istat 2013).

The BES Project is an attempt to integrate the measurement of the well-being level with the assessment of its equitable distribution and future sustainability. Concerning the equity dimension, both the social and territorial cohesion are of great importance, but the BES indicators provide at best a regional breakdown (i.e. at the NUTS2 level.[2]) Thus, to deepen the measurement of BES at local level, in 2011 Istat launched two pilot projects: UrBes and Provinces' BES.

This work deals with the set of indicators that was selected and implemented by the Provinces' BES Statistical Information System (SIS),[3] which allows a breakdown at the NUTS3 level (Cuspi-Istat 2015a), thus meeting the need to highlight territorial disparities in order to assess territorial cohesion, which is of primary importance in the reference framework of local development policies.

BES is divided into 12 domains, most of which concern those items that have a direct impact on human and environmental well-being (Alkire 2012). These *outcome* domains are as follows: *health, education and training, work and life balance, economic well-being, subjective well-being, social relationships, security,*

[2]The acronym NUTS (from the French "Nomenclature des unitès territoriales statistiques") stands for Nomenclature of Territorial Units for Statistics, that is, the European Statistical System official classification for the territorial units. The NUTS is a partitioning of the EU territory for statistical purposes based on local administrative units. The NUTS codes for Italy have three hierarchical levels: NUTS1 (groups of regions), NUTS2 (regions) and NUTS3 (provinces). The current NUTS 2013 classification is valid from 1 January 2015, and for Italy at the NUTS3 level it includes 110 territorial units, coinciding with the 110 provinces that existed in Italy at the reference date. During 2016, following reforms by local authorities implemented by the Italian government, some provinces have become metropolitan cities, while other provinces have been suppressed due to regional laws (in particular the provinces of Sicily and Friuli-Venezia Giulia). As these changes have not yet been added to the statistical classification, in this chapter, the term provinces refers to the 110 units classified in NUTS3, so including the new metropolitan cities and the provinces that no longer exist.

[3]For more information see www.besdelleprovince.it.

landscape and cultural heritage and *environment*. The BES theoretical framework is completed by those domains that quantify the main elements underling the well-being itself. The *context* domains are as follows: *politics and institutions, research and innovation* and *quality of services*. The BES statistical construct consists of 134 basic indicators distributed across the above-mentioned domains.

Compared to BES, the Provinces' BES has some limitations: specifying BES at such detailed geographical level leads to a trade-off between information needs and data availability. At the first stage of the implementation (in 2012), the dataset contained just 29 indicators at the NUTS3 level that were equivalent to the national indicators. After further developments, the database was supplemented with additional "proxy" indicators, and other indicators were added to highlight gender or generational differences, thus reaching, in the 2014 release, a total of 87 basic indicators in 11 domains. Despite this, at the current stage, Provinces' BES SIS does not yet satisfactorily measures BES; thus, processing this data requires managing many constraints and distinctive features (Chelli et al. 2015; Taralli et al. 2015; Taralli and D'Andrea 2015).

First of all, subjective indicators at the NTUS3 level are missing in all domains and the equity and sustainability measures are poor. Moreover, the domains are not always suitably measured: as a whole, the most severe gaps affect the *Social relationships* and *Economic well-being* domains, besides *Subjective well-being* is completely missing. In order to increase the availability of indicators, several basic statistics were integrated; the input information is provided from 37 data sources managed by 13 different institutions. The quality of the indicators therefore varies because data have different reference periods and are affected by different kinds of bias depending on the features of the statistical source used (ESS 2001; Hall et al. 2010; Istat 2012).

The database contains data from primary surveys (social or administrative, total or sample), statistical compilations, information systems and administrative archives processed for statistical purposes. Finally, concerning the quality of indicators, and particularly the desirable properties of the indicators to be used for evaluation purposes, it should be noted that very detailed indicators are not always quite reliable, robust or relevant (Delvecchio 1995).

Given the above-mentioned issues, and considering that "what is badly defined is likely to be badly measured" OECD (2008), we carefully assessed and selected the Provinces' BES indicators to build the measurement model as consistently as possible with the BES theoretical framework and with our research goals.

Summarizing BES indicators at the local level requires emphasizing the differences among geographical areas; this raises the question of the weighted importance of each well-being indicator. The issue of each indicator's contribution to the composite index requires a trade-off between the need for synthesis and the composite responsiveness to territorial disparities: the question is whether the contribution of each determinant to the BES varies significantly or remains substantially the same according to the geographical area considered, the specific well-being structure and the global level of well-being.

The model was built using a "theory-driven" process. In fact, BES is an emerging construct, clearly based on a formative measurement model (Dolce and Lauro (2015)). Therefore, to measure BES, it is important to have a set of indicators suitable for capturing all the well-being components, including equity and sustainability. Furthermore, the elementary indicators should all be equally important from a theoretical point of view (Albers and Hildebrandt 2006; Diamantopoulos and Winklhofer 2001; Istat 2015).

Only the indicators of primary importance, that is, those that provide specific information that other indicators do not provide (so-called non-substitutable indicators), contribute to the synthesis of BES; thus, we included in each domain just those indicators that measure direct effects or impacts on well-being. Furthermore, as in the Provinces' BES SIS, the indicators in a given domain are grouped according to whether they pertain to the same subdimension (called as *theme*), we used this intermediate level among elementary indicators and domains to improve the specificity and sensitivity of the measurement model.

The Provinces' BES themes are the first-order constructs in the structural model described in Fig. 5.1, while domains are the second-order constructs, which are grouped into *outcome* and *context* domains at the higher level. The synthesis of *outcome* and *context* defines BES at the top of the model.

Within the limits described above (lower availability of relevant and reliable data at the NUTS3 level), the structural model explains for each Italian province or metropolitan city the impact of a single determinant (or group of determinants, i.e. *themes* or *domains*) on the overall well-being levels (*outcome*) and on the framework conditions (*context*), as well as the contribution of *outcome* and *context* to the BES of that area. This approach, in concept, is in line with the one followed by Istat to construct the BES domain composites at the regional level (Istat 2015). The main differences, compared to Istat's method, are in the aggregation mode applied and in the synthesis of the domain composites that, in this study, leads up to the construction of a global BES composite.

5.2.2 Data Description

The original Provinces' BES dataset covers 11 domains, structured into 31 themes, for a total of 87 indicators observed across the totality of the Italian Provinces and metropolitan cities (110 units).

Extensive data cleaning and data tuning was necessary to cover several issues that are crucial to start a multidimensional analysis and in particular carry out model-based path modeling (for example, missing data, multicollinearity, indicators' polarity).

After the data pre-processing, the data matrix contained 40 indicators, partitioned into 11 domains and 12 themes. A descriptive univariate analysis of each indicator is provided in Davino et al. (2016a, Sections 2.2 and 3).

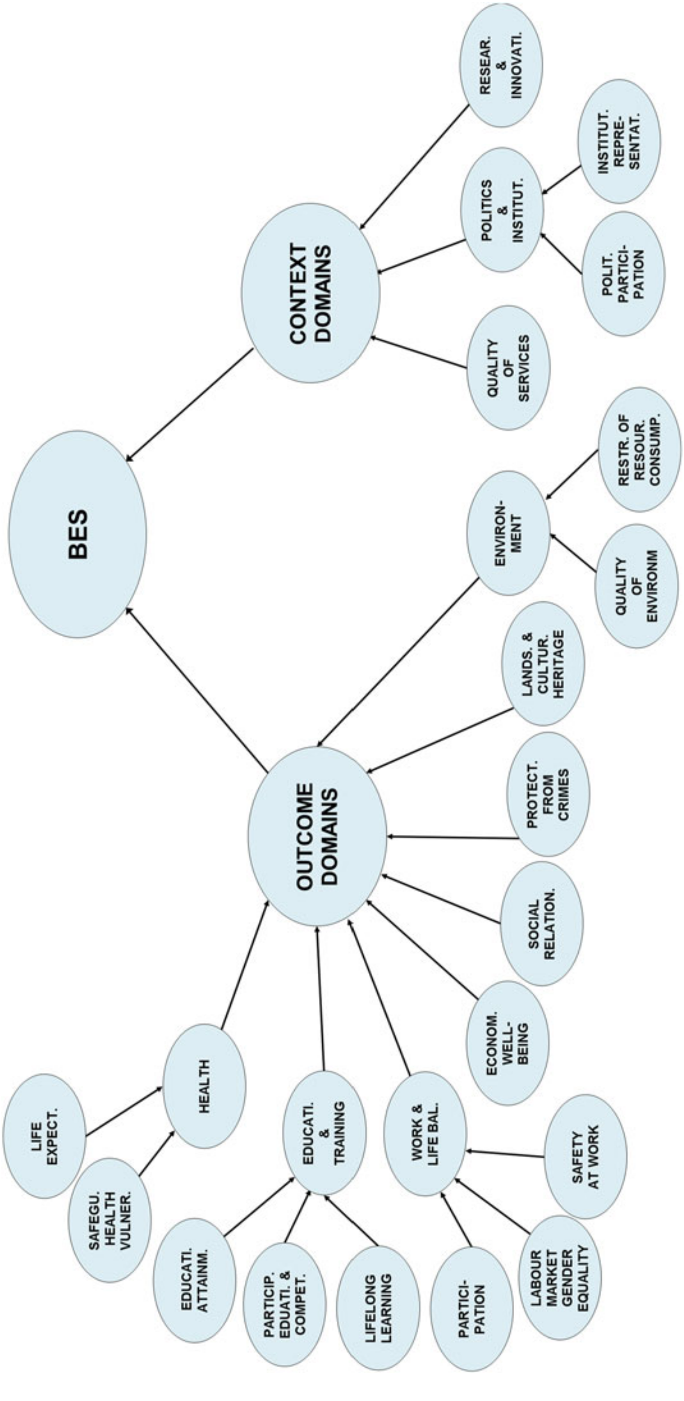


Fig. 5.1 Structural model describing *outcome* and *context* domains of the Provinces' BES

Tables 5.7 and 5.8 in the Appendix list the final domains, themes and indicators; labels in the last column are used in graphs and tables showing the results of the study. The indicators with reversed polarity are marked by the asterisk symbol in the tables.

5.3 A Hierarchical Composite Model for BES

As Fig. 5.1 shows, we consider a hierarchical structure for BES. The specification of this hierarchical structure derives from the available theoretical knowledge (Istat 2015). In particular, BES is considered as a multidimensional construct at the highest order. It is composed of several different domains, some of which in turn are composed of a number of themes (for example, the domain *health* is composed of two themes: *life expectancy* and *safeguard from health vulnerability*).

Figure 5.1 shows the specified hierarchical structure for BES and how domains and themes are connected to overall BES and to one another.

PLS-PM (Tenenhaus et al. 2005; Wold 1982) is an important method for assessing hierarchical models (Becker et al. 2012; Ringle et al. 2012; Wetzels et al. 2009; Wilson 2010) and particularly such complex models.

PLS-PM computes proxies of constructs as composites (i.e. weighted aggregates of the corresponding manifest variables, or MVs), assuming that all the information concerning the relationships among the blocks of MVs is conveyed by the constructs (Rigdon 2016; Sarstedt et al. 2016a). Consequently, hierarchical models in PLS-PM can be better defined as hierarchical composite models. Furthermore, proxies of constructs computed in hierarchical composite models are actually defined as composite indicators. In this specific model, the composite indicator for BES is the global score, while all the other composite indicators (for the domains and the themes) are partial scores (Guinot et al. 2001).

The weighting in PLS path modeling is aimed at maximizing variances in each block of variables and correlations between adjacent composites. If the MVs have different relationships with their own underlying construct, PLS-PM has important advantages compared to the traditional aggregation methods, such as principal component analysis or unit-weight composite model (i.e. a simple arithmetic mean of the MVs) (Henseler et al. 2014). In particular, depending on the chosen estimation mode for the calculation of the outer weights and schemes for the inner model, PLS-PM determines the weights of the MVs such that the more reliable MVs have larger weights; it also provides components that are as highly correlated as possible to each other while explaining the variances of their own set of variables.

We apply PLS-PM to estimate the parameters of the hierarchical composite model, using the most popular conceptualization of a hierarchical model: the so-called repeated indicators approach (Lohmöller 1989). As the name suggests, the repeated indicators approach is based on a repeated use of manifest variables. In particular, as constructs without associated MVs are not allowed in PLS-PM, higher-order constructs are defined considering all the MVs of the underlying lower-order constructs. For example, a second-order construct is directly measured by the MVs related to all the first-order constructs.

The repeated indicators method is not the only possible approach to assess hierarchical composite models (Becker et al. 2012; Wilson and Henseler 2007); indeed, the repeated indicator approach is said to have some limitations when the number of MVs is not balanced among blocks (Hair et al. (2014)). However, to the best of our knowledge, except for a study by Becker et al. (2012), there are hardly any studies in the literature that provide substantive reasons for this assumption or that analyse it in detail. Becker et al. (2012) analyse the effect of unbalanced blocks on the relationships among composites. The authors show that this effect mainly depends on the applied outer mode (i.e. *Mode B* and *Mode A*). When using *Mode B*, the repeated indicator approach does not seem to be affected by an unequal number of indicators per block.

We chose the repeated indicator approach because it considers simultaneously the entire path model in the estimation procedure while remaining in the well-established realm of the basic PLS-PM algorithm.

The path directions in the structural model are specified following the arrows in the path diagram in Fig. 5.1. In particular, the higher-order composites depend on the corresponding lower-order constructs.

In this study, the choice of the outer mode (the way to compute the outer weights) has nothing to do with the hypothesized measurement model. In the recent PLS-PM literature, researchers have started to clarify that the outer mode and the measurement model are separate in PLS-PM, and any association may be an illusion (Becker et al. 2013; Henseler et al. 2016; Rigdon 2016; Sarstedt et al. 2016a). *Mode A* and *Mode B* are just two different ways of computing the outer weights, and the choice between the two modes goes beyond the specified measurement model and requires a more thoughtful approach (Becker et al. 2013; Sarstedt et al. 2016a). Note that PLS-PM computes proxies for all constructs as weighted composites of the corresponding MVs, no matter which outer mode is used.

In order to maintain coherency with the path directions specified in the structural model (Dolce et al. 2016), we use *Mode A* for all the higher-order constructs and *Mode B* for the first-order constructs.

We apply PLS-PM to compute the BES global score and the partial scores and to examine the magnitude of effects of each domain and theme on overall BES, in order to search for the primary factors influencing BES.

First, we estimate a global model at the national level, considering all the provinces together. Afterwards, in Sect. 5.4.1, we introduce geographic area as a variable. In particular, we estimate a model for each of the following geographic areas of Italy: north-east, north-west, centre and south and islands.

5.3.1 Global Model PLS-PM Results

We present in this section the main results of application of PLS-PM to the global model. In particular, we show the estimated effects (i.e. the path coefficients) of the two drivers of BES (i.e. *outcome* and *context* domains), and the estimated effects of the different domains on the *outcome* and *context* drivers. Finally, as a measure of fit of the model, we consider the redundancy index. In hierarchical composite models the R^2 is very close to 1 for each structural equation, as higher-order constructs are almost fully explained by their lower-order constructs. Consequently, we do not consider the R^2 a measure of fit of the model.

The PLS-PM path coefficients for the global model are reported in Table 5.1. Both *outcome* and *context* have a significant effect on BES. As expected, *outcome* has a greater impact on BES than *context* because, as the number of variables for this domain is greater than for *context*. However, the effect of the *context* on BES is larger than expected.

In general, all the results are coherent with the hypothesized model, except for the *outcome* domain *environment*, which has a negative effect. However, in this case, the path coefficient is very small and not significant.

Except for the *environment*, all the domain effects on the outcome are significant. Table 5.1 gives an ordered list of the domains (the path coefficients are in decreasing order), in order to differentiate those factors that strongly influence BES from those domains with a smaller effect.

All the *context* domain path coefficients are significant, and the factor with the greatest influence for *context* is *quality of services*.

The redundancy index is a measure of the ability of the predictive composite to explain the variation in the dependent blocks. For example, the redundancy index of BES expresses how much of the variability in the MVs of BES is explained by the *outcome* and *context* composites. Redundancy index refers to each dependent MV.

Table 5.1 PLS-PM path coefficients of the global model (non-significant coefficients are shown in italics)

Composite		Path coefficient
BES	Outcome	0.674
	Context	0.353
Outcome	Work and life balance	0.270
	Economic well-being	0.236
	Health	0.234
	Social relationships	0.191
	Education	0.139
	Cultural heritage	0.110
	Protection from crimes	0.061
	Environment	*−0.007*
Context	Quality of services	0.542
	Politics and institutions	0.352
	Research and innovation	0.238

Table 5.2 Average redundancies for the global model

Composite	Redundancy
BES	0.315
Outcome	0.314
Context	0.367

Table 5.2 shows the redundancies for BES, *outcome* domain and *context* domain, computed as means of the redundancies of the corresponding blocks of variables.

Redundancy-based prediction is the relevant criterion for assessing the in-sample predictive ability of the structural model (Chin 2010a) as well as for comparing structural equation models. However, as noted by Lohmöller (1989), the average of all the redundancies can be considered as an index of goodness of fit of the global model.

Therefore, we compute the average redundancy index as the mean of all the block-redundancies weighted by the number of MVs of each block. The averaged redundancy index is equal to 0.357. Given the complexity of the model, and because redundancy values are generally small in PLS-PM, the fit of the model, in terms of redundancy measure, is judged as satisfactory and sufficient to justify interpretation of the results.

5.4 Observed and Unobserved Heterogeneity in PLS-PM

The global model assumes homogeneity across provinces. However, if heterogeneity is present, ignoring it may result in poor parameter estimates, and the model performance may seriously degrade.

In some special cases, heterogeneity in the models may be captured by observable variables that form homogeneous groups of statistical units. This special situation, usually referred to as observed heterogeneity, is actually very rare in practice. More frequently, unobserved heterogeneity (i.e. when the sources of heterogeneity are not known a priori) is the case in real applications (Rigdon et al. 2010). In general, regardless of the specific case, it is desirable to detect and treat heterogeneity in PLS-PM (Sarstedt et al. 2016b, 2011b).

A number of methods and approaches for identifying and treating unobserved heterogeneity (the so-called latent class techniques) have been proposed in the literature (Becker et al. 2013; Esposito Vinzi et al. 2008; Hahn et al. 2002; Lamberti et al. 2016; Ringle et al. 2010; Sarstedt 2008; Sarstedt et al. 2011a).

The desire to treat unobserved heterogeneity also led to the new method called QCPM, which is described in Sect. 5.5.

In Sect. 5.4.1, we treat observed heterogeneity, assuming that heterogeneity across provinces is due to the difference in geographic area. Thus, we apply a multi-group analysis in PLS-PM, creating four different groups of provinces generated from the four geographic areas. We test whether differences between pairs of path coefficients across areas are statistically significant, using the Chin and Dibberns (2010b) permutation test.

This first analysis focuses on the relationships in the structural model. The objective here is to determine which path coefficients are responsible for the differences between sub-models. An important limitation of this first study concerns the lack of measurement invariance assessment (Henseler et al. 2015). Assessing measurement invariance in PLS-PM is not a trivial task but should not be completely disregarded when comparing path coefficients. Ideally, we should have tested for measurement invariance to ensure that the differences between path coefficients do not result from a different meaning of the estimated composites.

5.4.1 PLS-PM Results Across Geographic Areas

A more in-depth analysis of the Province's BES must take into account that living conditions are quite different according to the geographic position of the province. Italian provinces are usually grouped into five areas: *north-west*, *north-east*, *centre*, *south* and *islands*.

In this study the last two areas were combined because of the small size of the group *islands*. Therefore, we obtain the following percentages of provinces: *centre* (20%), *north-east* (20%), *north-west* (23%) and *south and islands* (37%).

Estimating the structural model in Fig. 5.1 for each area, it is possible to highlight several differences among the areas from both a descriptive and an inferential point of view. Figures 5.2 and 5.3 show the distribution of the three main composite indicators of the model: BES, *outcome* and *context*. The graphs clearly show a different behaviour in the *south and islands* provinces. These provinces show not only lower median values but also higher variability and asymmetry. Moreover, in the South of Italy, the *context* domains are more critical than the *outcome* domains. Provinces located in the Centre also show a certain degree of heterogeneity, particularly in the lower tail of the construct distribution.

In Fig. 5.4, the 110 provinces are plotted according to BES, *context* and *outcome*. Different symbols are used to distinguish the geographic areas of location. Numbers in the lower diagonal are the correlation coefficients. There is a very high correlation between the BES and the *outcome*. The lower, but still strong, correlation between the *outcome* and the *context* indicates that the *context* domain is less discriminating between northern and southern provinces as compared to the *outcome* domain. In addition, the chart shows a marked divergence between the levels of *outcome* and *context* for a significant number of provinces of the north-east group (denoted by triangles in the figure).

In order to take into account observed heterogeneity among provinces, represented by the different geographic areas, path coefficients obtained in the multi-group analysis are compared with those derived from the global model. The impact of *context* and *outcome* on BES varies strongly moving from the north to the south of Italy. The bars in Fig. 5.5 represent the path coefficients obtained from separate analysis of each area while the horizontal lines represent the path coefficients estimated for the whole group of provinces. With respect to the *outcome*, the *south*

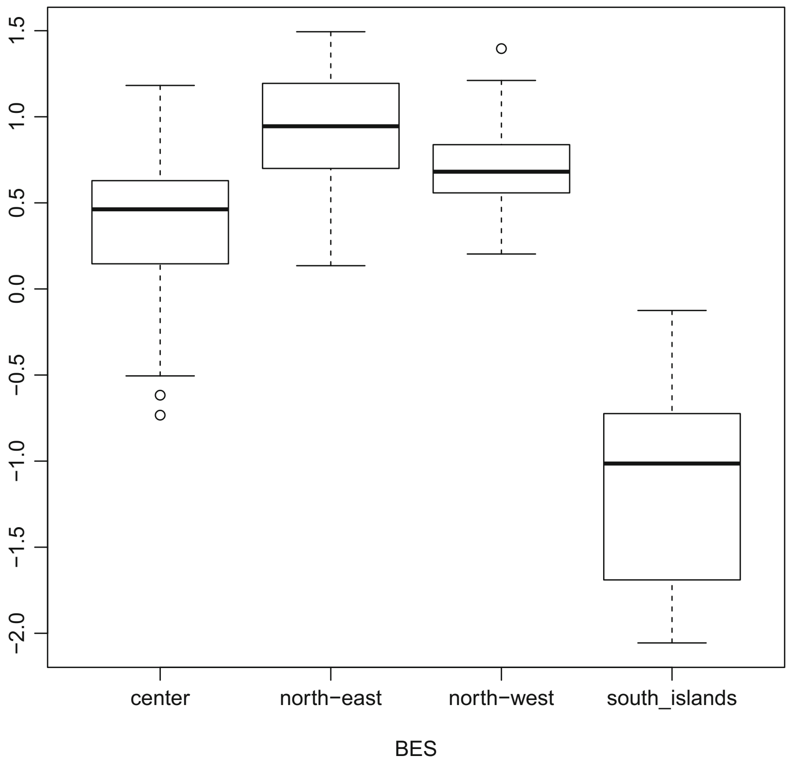


Fig. 5.2 *Boxplot* of the BES according to geographic area

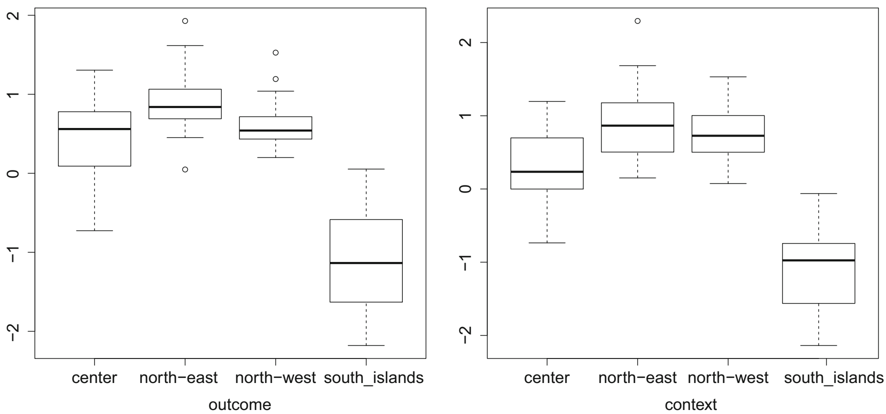


Fig. 5.3 *Boxplot* of the *outcome* and *context* according to geographic area

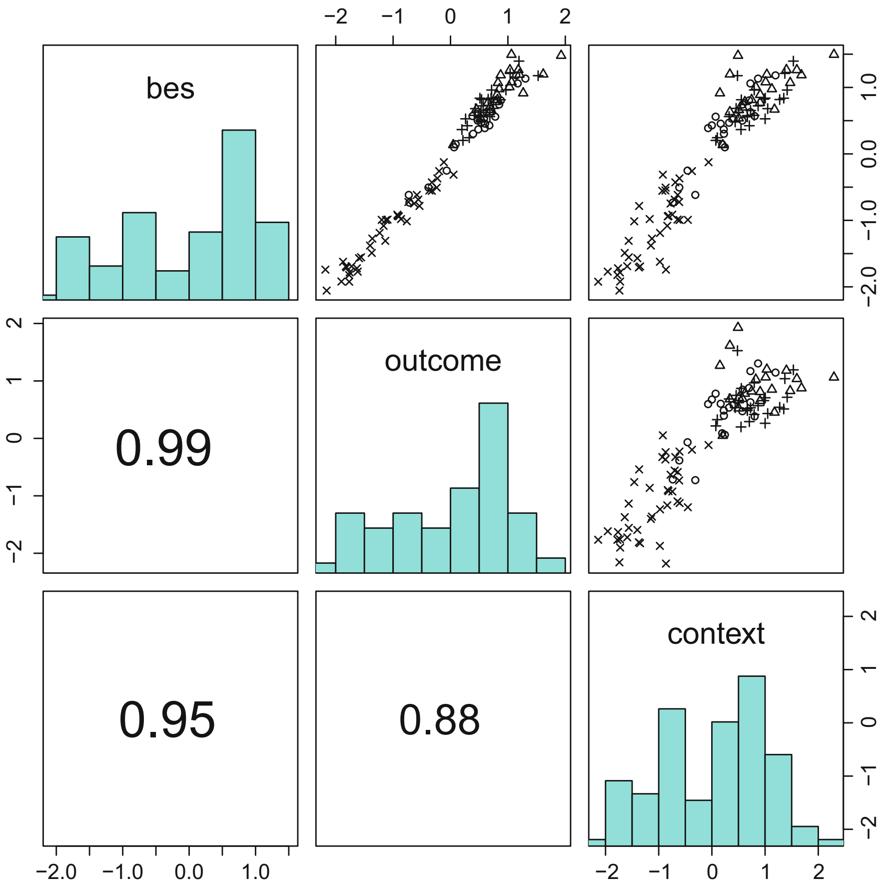


Fig. 5.4 The scatterplot matrix of BES against the *outcome* and the *context* according to geographic area

and islands area has a very high coefficient, even higher than the coefficient of the global analysis. The opposite happens for the *context* domain. Concerning north-east, the negative sign of the *context* is mainly due to those provinces (already highlighted in Fig. 5.4) where the *context* level differs significantly from that of the *outcome*.

Heterogeneity among geographic areas also occurs across the different domains of the *outcome* and the *context*. Figure 5.6 shows the path coefficients of the eight constructs influencing the *outcome* (left-hand side) and the three constructs influencing the *context* (right-hand side). The domains are represented in decreasing order according to the results of the global analysis (shown as black points). To simplify the left-hand graph, letters are used to denote each domain: A (*work/life*), B (*economic well-being*), C (*health*), D (*social relationships*), E (*education*), F (*cultural heritage*), G (*protection form crimes*) and H (*environment*). Figure 5.6

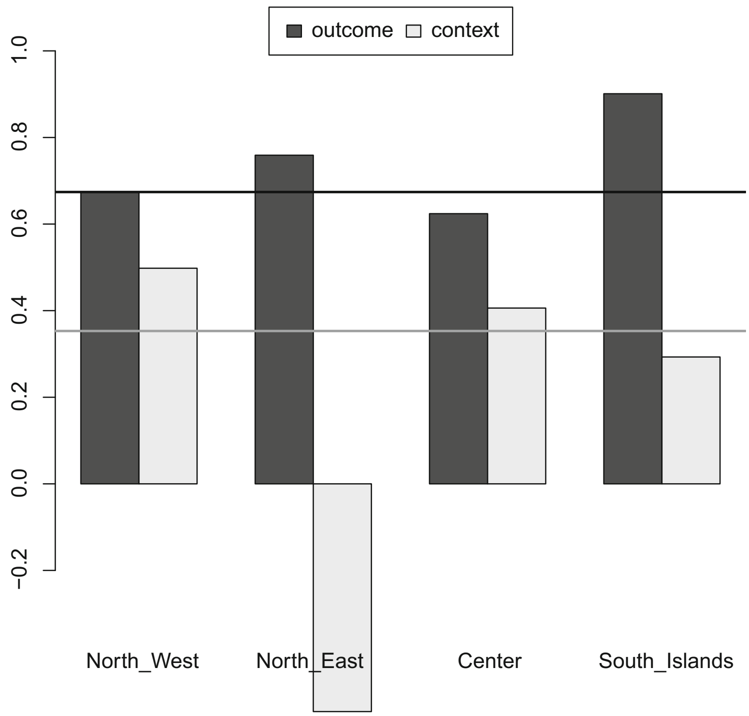


Fig. 5.5 BES path coefficients for the *outcome* and the *context* in the different geographical areas

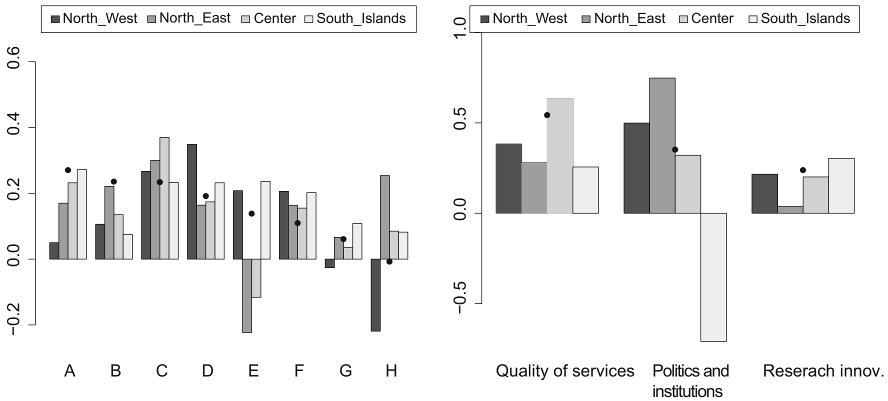


Fig. 5.6 Path coefficients for the *outcome* (*left-hand side*) and the *context* (*right-hand side*) domains in the different geographical areas

can be interpreted in several ways. For example, let us compare the results related to the *south and islands* group with the other areas and with the whole set of provinces. This area is characterized by a high impact on the *outcome* (even higher than the global path coefficient) with respect to *social relationships*, *education*, *cultural heritage* and *protection from crimes*. Among all the areas, *social relationships* have the highest impact in the north-west provinces. In the *context* framework, the negative role of *politics and institutions* is strongly indicated.

From an inferential point of view, it is possible to evaluate whether, for each pair of areas, differences between path coefficients are statistically significant. Tables 5.3 and 5.4 show p-values of the difference test between path coefficients. Low p-values indicate significantly different coefficients between areas.

The same approach can be followed for each theme of the *outcome* and *context* domains. To facilitate the interpretation of the results, Tables 5.5 and 5.6 show only the themes with a p-value lower than 0.1. The arrows indicate which path

Table 5.3 *P*-values of the difference test between the path coefficients of the *context* in the different geographical areas

	North-east	North-west	Centre	south and islands
North-east		0.03	0.02	0.01
North-west			0.36	0.01
Centre				0.03
South and islands				

Table 5.4 *P*-values of the difference test between the path coefficients of the *outcome* in the different geographical areas

	North-east	North-west	Centre	South and islands
North-east		0.94	0.49	0.01
North-west			0.47	0.01
Centre				0.01
South and islands				

Table 5.5 *Outcome* domains with a p-value lower than 0.10 in the different geographical areas

	North-east	North-west	Centre	South and islands
North-east				Education ↑
				Economic well-being ↓
				Cultural heritage↑
North-west			Social relationships ↑	Work and life balance ↑
				Economic well-being ↑
				Social relationships ↑
				Environment ↑
Centre				Education ↑
South and islands				

Table 5.6 *Context* domains with a *p*-value lower than 0.10 in the different geographical areas

	North-east	North-west	Centre	South and islands
North-east				Quality of services ↑
North-west			Quality of services ↓	Quality of services ↑
Centre				Quality of services ↓
South and islands				

coefficient is higher. For example, the *education* path coefficient in the *south and islands* provinces is significantly different from that obtained in the north-east group. Moreover, the impact of *education* on *outcome* increases in the north-east compared to the south of Italy (up arrow).

5.5 Quantile Composite-Based Path Modeling

5.5.1 Methodology

The QC-PM proposed by Davino and Esposito (2016) introduces both quantile regression (QR) (Koenker and Basset 1978) and quantile correlation (QC) (Li et al. 2015) in the traditional PLS-PM algorithm. While PLS-PM is based on simple and multiple ordinary least squares (OLS) regressions, in some particular cases, the estimates of coefficients may vary along the distribution of the dependent variable. In such cases, QC-PM can complement traditional PLS-PM, because it allows exploration of the entire dependence structure beyond the estimation of the average effects.

A brief introduction to QR is provided to better highlight the added value of QC-PM.

Quantile regression, as introduced by Koenker and Basset (1978), may be considered an extension of OLS regression because it is based on the estimation of a set of conditional quantiles of a response variable as a function of a set of covariates (Davino et al. 2013). The main features of QR can be summarized as follows:

- $Q_\theta\left(\hat{\mathbf{y}}|\mathbf{X}\right) = \mathbf{X}\hat{\beta}\left(\theta\right)$ represents the QR model for a given conditional quantile θ where $Q_\theta(.|.)$ denotes the conditional quantile function for the θth quantile. This chapter will refer to linear regression models.
- As $0 < \theta < 1$, it is potentially possible to estimate an infinite number of regression lines, but in practice a finite number is numerically distinct, which is known as the quantile process. For each quantile of interest, a regression line

is estimated and, consequently, a set of coefficients and a fitted response vector can be obtained.

- No parametric distribution assumptions are required for the error distribution.
- The estimation is realized by minimizing the weighted sum of absolute residuals which weights positive and negative residuals asymmetrically, respectively, with weights equal to $(1 - \theta)$ and θ.
- The parameter estimates have the same interpretation as those of any other linear model.
- The estimators are asymptotically normally distributed with different forms of the covariance matrix depending on the model assumptions (Koenker and Basset 1982a,b). Resampling methods can represent a valid alternative to the asymptotic inference (among many, see Kocherginsky et al. 2005) because they allow the estimation of parameter standard errors without requiring any assumption in relation to the error distribution.
- The assessment of goodness of fit exploits the general idea leading to the typical R^2 goodness of fit index in classical regression analysis. The most common goodness of fit index in the QR framework is called *pseudo*R^2 (Koenker and Machado 1999).

To appreciate QR features and results, a very simple example is provided using two indicators from the Provinces' BES dataset. A classical OLS regression and a QR for a set of selected quantiles (θ = [0.1, 0.25, 0.5, 0.75, 0.9]) are performed to explore the dependence of the *non-profit institutions* indicator from the *gross disposable income per household* indicator.

The graphical representation of the coefficients can aid in interpretation of the results. The left-hand side of Fig. 5.7 shows OLS (solid) and QR (dashed) lines imposed on the scatterplot of the dependent variable and the considered regressor.

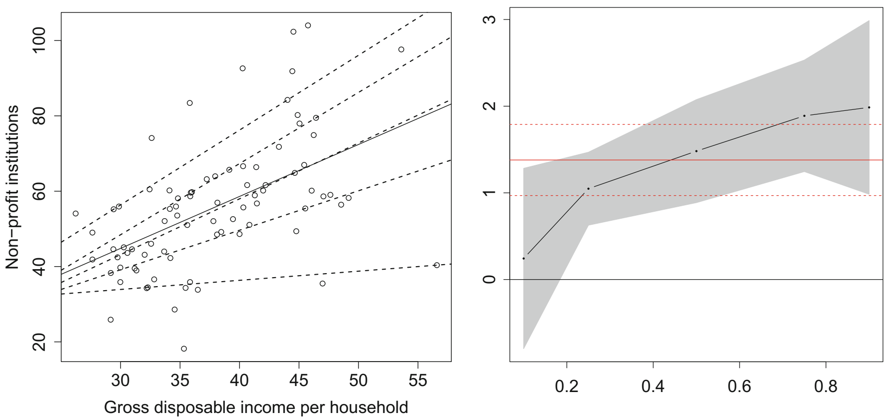


Fig. 5.7 Scatterplot with OLS and QR lines for five quantiles of interest (θ = [0.1, 0.25, 0.5, 0.75, 0.9])

Greater variability of the response variable for high *gross disposable income* values is evident from the scatterplot and from the fan shape of the QR lines. The effect of the regressor around the *non-profit institutions* variable remains stable around the average or median, but differs in size below and above the median. Moreover it is evident that in case of a generic quantile θ, $(100 \times (1 - \theta))\,\%$ of points lie above the quantile regression line and $(100 \times \theta)\,\%$ lie below the regression line (a number of points equal to the number of model parameters lies exactly on the line).

The right-hand side of Fig. 5.7 shows the QR coefficient trends. The horizontal axis displays the different quantiles, while the effect of the regressor is represented on the vertical axis. QR confidence bands (in grey) are obtained through the bootstrap method for $\alpha = 0.1$. The two solid lines parallel to the horizontal axis represent 0 and the OLS coefficients. The related confidence intervals are represented using dashed lines for $\alpha = 0.1$. This figure gives a graphic representation of the coefficient trend moving from lower to upper quantiles.

A quantile measure of correlation is introduced as well in the QC-PM in place of classical quantile correlation. Such a measure was introduced by Li et al. (2015) as a correlation measure between two random variables for a given quantile $\theta \in (0, 1)$. The quantile correlation index (QC) is constructed just like the Pearson correlation coefficient as the ratio between a covariance and the squared root of the product between the variances of the two variables. QC has the same properties as a correlation coefficient (it increases with the slope of the regression line and lies between -1 and 1) but it is not symmetric. Thus, it is necessary to identify which variable plays the role of dependent variable. With respect to the inferential aspects, a bootstrap approach is proposed, based on the complexity of the estimation of the variance–covariance matrix of the estimator (Davino and Esposito 2016).

QC-PM is in essence based on the introduction of a quantile approach in the estimation steps of the PLS-PM algorithm: the inner estimation, the outer estimation and the estimation of the path coefficients and loadings.

Several options for QC-PM can be implemented according to the choices adopted in the model estimation. For example, in the inner estimation (structural model), the LV estimation depends on the adopted weighting scheme. If the *path weighting* scheme is chosen, the inner weights linking the *j*th successor LV to its predecessors are estimated through a QR, while the inner weights among the *j*th LV and its successor LVs are determined using the QC. As in the quantile framework, the correlation is a non-symmetric measure and the use of QC distinguishes between predecessors and successors. QC is proposed as an alternative to the Pearson correlation coefficient if the *centroid* or the *factorial* scheme is adopted. As QC is an asymmetric index, the great benefit of this proposal is the possibility of taking into account the direction of the structural relationships. Therefore, by using QC, even the factorial and the centroid schemes take into account the role played by predecessor and successor LVs.

In the outer estimation (measurement model), simple (*Mode A*) or multiple (*Mode B*) quantile regressions allow computation of the composite scores for each quantile of interest. An innovative estimation mode (*Mode Q*) in the outer model is represented by the use of QC to compute the weights. *Mode Q* allows taking explicitly into account both outwards-directed and inwards-directed measurement models because, as stated before, unlike the Pearson correlation coefficient, QC is asymmetric. The advantage of *Mode Q* over *Mode A* and *Mode B* is mainly computational. QC is a descriptive bivariate measure while the choice of *Mode A* or *Mode B* in QC-PM requires the estimation of several QRs. Moreover, in case of inwards-directed relationships, the computational gain is greater because *Mode B* would require the estimation of multiple QRs.

Once convergence is reached and scores are computed, the path coefficients and the loadings are estimated by means of quantile regressions. For each quantile of interest, a vector of estimates including the intercept is obtained.

The proposed QC-PM method provides, for each quantile of interest θ, a set of weights, loadings, path coefficients and scores.

A final evaluation of the quality of the QC-PM results from both a descriptive (assessment) and inferential (validation) point of view has been proposed by Davino et al. (2016b). The goodness-of-fit measures typically used in PLS-PM have been extended to QC-PM and a non-parametric approach can be used to validate the significance of the estimates.

QC-PM is not meant to replace PLS-PM but rather to complement PLS-PM when the average effects are not sufficient to summarize the relationships among variables. In these cases, a more in-depth analysis of the whole distribution of the dependent variables is needed and can be obtained by exploiting its various quantiles.

5.5.2 Results

The estimation of a QC-PM provides a set of outer weights and path coefficients for each quantile of interest. The former represent the weights of each observed indicator on the corresponding theme or domain while the latter are an estimate of the impact of each theme on the corresponding domain as well as the impact of each domain on BES.

In the following, a selection of relevant results is described. A complete view of the PLS-PM and QC-PM results for all the themes is provided in Davino et al. (2016a).

The impacts of the different domains on the *outcome* and *context* drivers are shown in Fig. 5.8 where the PLS-PM and the QC-PM (for two quantiles of interest $\theta = 0.1$ and $\theta = 0.9$) path coefficients are represented. To facilitate the comparison, the eight constructs impacting on the *outcome* are shown in decreasing order, according to the PLS-PM results (black bars). If on average the most important

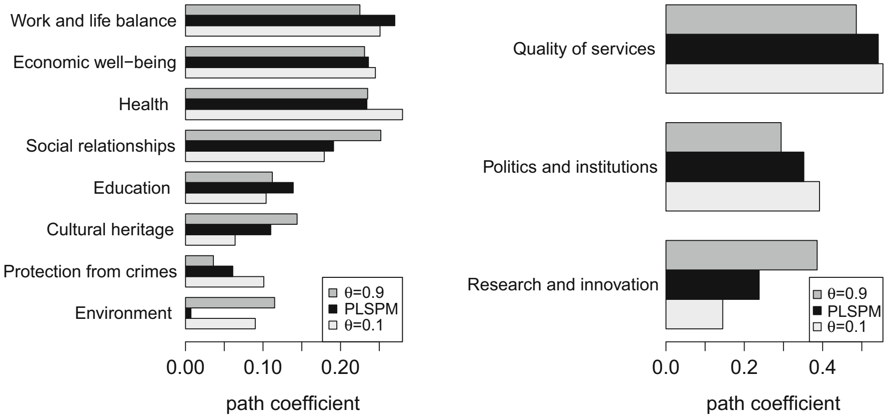


Fig. 5.8 The bar charts of the path coefficients for the *outcome* (*left-hand side*) and the *context* (*right-hand side*) domains from PLS-PM and from QC-PM for two quantiles of interest (θ = [0.1, 0.9])

domain for the *outcome* seems to be *work and life balance*, followed by *economic well-being*, in the 10% of the provinces with the lowest *outcome* values, *health* is the most important factor and *protection from crimes* and the *environment* come to light too. Results change in the upper quantile, where the *social relationships* domain exceeds all the others in terms of impact on the *outcome*. Intangible issues related to *cultural heritage* are also worth highlighting.

With respect to the *context* (right-hand side of Fig. 5.8), a central role is played by the *quality of services*, whose importance can be considered very high, while *research and innovation* assumes a different influence in the top (90%) or bottom (10%) of the *context* distribution.

PLS-PM and QC-PM also provide a weight for each indicator, representing the role played by the given indicator in constructing the corresponding *theme*. For illustrative purpose, Fig. 5.9 shows the outer weights of the *social relationships* theme for five selected quantiles (θ = [0.1, 0.25, 0.5, 0.75, 0.9]) and for the classical PLS-PM. For ease of interpretation, PLS-PM results are shown on the right-hand edge of the figure.

Limiting the analysis to the measurement of the average effects of the indicators V.7 (*non-profit institutions*) and V.8 (*volunteers in non-profit institutions (per 100 residents aged 14+)* on the level of *social relationships*, the results show that both have a positive weight. Such a result was expected as the two indicators are correlated. A more in-depth analysis is provided by the exploration of the results for the bottom and top part of the construct distribution. Provinces characterized by low *social relationships* are typically characterized also by a fragile institutional structure and consequently the number of volunteer-based associations increases

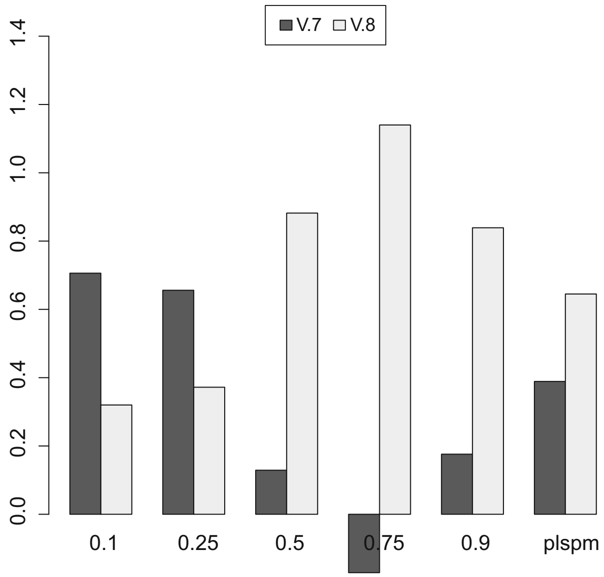


Fig. 5.9 PLS-PM and QC-PM outer weights for *social relationships* across a set of selected quantiles

even if the participation of volunteers is not adequate. On the other hand, when the social relationships are stable and strengthened, the weight of the participation becomes relevant even if the presence of the associative pattern is negligible.

A tentative approach to connect QC-PM results to the observed heterogeneity represented by the geographic area is provided in Fig. 5.10. This graph is a simultaneous representation of the first factorial plane of a principal component analysis performed on the data matrix of the QC-PM *outcome* path coefficients (on the rows the quantiles, in the columns the themes and in the cells the path coefficients). Path coefficients obtained from the multi-group PLS-PM are plotted as supplementary rows. The aim of the representation, purely descriptive, is twofold. First, it provides a simultaneous visualization of the obtained results for the five considered quantiles while Fig. 5.8 just compares the extreme quantiles. Second, it allows to explore the impact played by each domain in the different parts of the *outcome* distribution to the coefficients related to a given geographic area. For example, high values of *economic well-being* and *environment* characterize the top part of the *outcome* distribution, and similar values are obtained in the centre and north-east provinces.

Further research will be developed to take into account both observed and unobserved heterogeneity in QC-PM. In the quantile regression framework, different approaches have been proposed in the literature to analyse group effects in a dependence model (Geraci and Bottai 2014; Koenker 2004; Lamarche 2010),

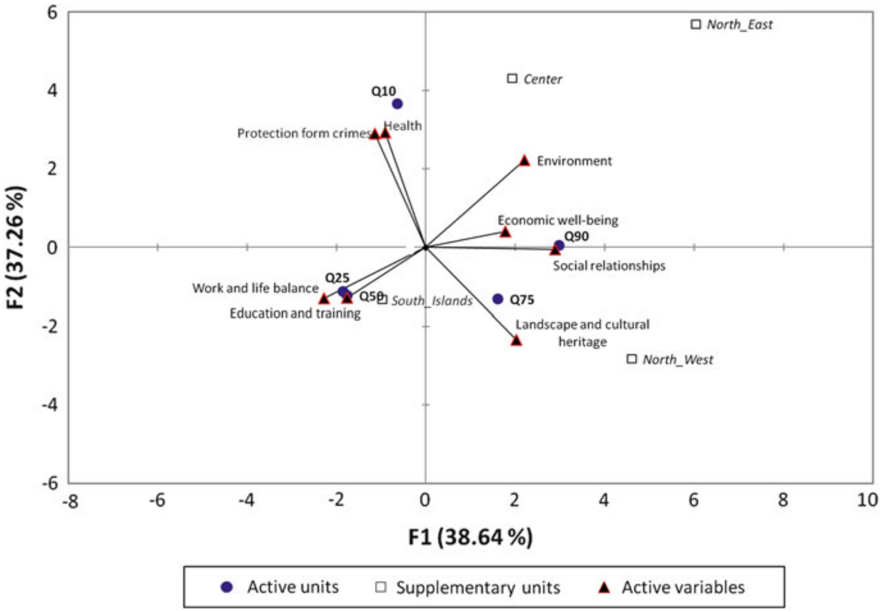


Fig. 5.10 PLS-PM and QC-PM outer weights for *social relationships* across a set of selected quantiles

although some of these studies deal with longitudinal data. With respect to the case described in this chapter, that is observed heterogeneity, we will explore the possibility of estimating different QC-PM models for each group or the introduction of dummy variables among the regressors to denote group membership. A more complex, but even more suitable approach, could be the adaptation in QC-PM of the procedure proposed by Davino and Vistocco (2007) and Davino and Vistocco (2015) in the quantile regression framework. The method provides a separate analysis of the dependence structure for each group but results are obtained on the whole sample, so to detect group effects and easily compare the coefficients estimated for each group

5.6 Concluding Remarks

The QC-PM method is complementary to PLS-PM and aims to broaden the analysis going beyond the estimation of average effects, exploring the network of relationships among variables.

The real data application presented in this chapter is related to the indicators of well-being released by the Italian National Institute of Statistics within the Provinces' Bes Project, which is one of the projects recently undertaken by Istat to study well-being at the local level. We apply PLS-PM to examine how and how much each driver contributes to well-being as a whole, by determining which factors primarily influence well-being both nationally and locally. For this last purpose, we introduce in the model the geographic area as a " stratification criterion", and then we estimate a global model at the national level and one model for each of the following four geographic areas of Italy: north-east, north-west, centre, south and islands. Consequently, we apply a multi-group analysis assuming that heterogeneity across provinces is due to their geographic location. The results of this analysis are statistically significant and very interesting from an analytical point of view, as they highlight different behaviours of the well-being drivers across geographic areas.

QC-PM provides a more in-depth analysis, introducing a quantile approach in the estimation steps of the PLS-PM algorithm. This method provides a set of outer weights and path coefficients for each quantile of interest, allowing estimation and comparison of the impact of each driver of well-being according to the distribution of the global well-being itself. This allows testing of whether and how much the importance of each component of well-being, as measured in average by PLS-PM, varies across the overall well-being distribution. Some relevant results are presented which show that QC-PM is very useful when analysing the well-being at the local level because it focuses on the distributional aspects and thus it reveals and measures in-depth the underlying structure of territorial disparities in the well-being distribution.

Appendix

See Tables 5.7 and 5.8.

Table 5.7 Dimensions, themes, indicators and labels

Health	Indicator	Label
Life expectancy	Life expectancy at birth (male)	I.1
	Life expectancy at birth (female)	I.2
Safeguard from specific health vulnerabilities	Infant mortality rate	I.3*
	Mortality rate for road accidents (15–34 years old)	I.4*
	Mortality rate for cancer (20–64 years old)	I.5*
	Mortality rate for dementia (65 years old and over)	I.6*
	Avoidable mortality rate (0–74 years old)	I.8*
Education and training	Indicator	Label
Educational attainment	Early leavers from education and training	II.1*
	People of working age with lower secondary education or less	II.2*
Participation and competencies	Participation in tertiary education (19–25 years old)	II.4
	Level of literacy and numeracy	II.6–7
Lifelong learning	Participation in lifelong learning (25–64 years old)	II.8
Work and life balance	Indicator	Label
Work participation	Employment rate	III.1–3
Labour market gender equality	Gender inequality in non-participation rate (M-F)	III.2*
	Gender inequality in employment rate (M-F)	III.4*
Safety at work	Incidence rate of occupational injuries, fatal or leading to permanent disability	III.9*
Economic well-being	Indicator	Label
Economic well-being	Gross disposable income per household	IV.1
	Households assets (total average amount)	IV.6
Social relationships	Indicator	Label
Social relationships	Non-profit institutions	V.7
	Volunteers in non-profit institutions (per 100 residents aged 14+)	V.8
Politics and institutions	Indicator	Label
Political participation	Electoral participation (European parliament elections)	VI.1
	Electoral participation (provincial elections)	VI.2
Institutional representation	Women and political representation in municipalities	VI.3
	Young people (<40 years old) and political representation in municipalities	VI.4

Table 5.8 Dimensions, themes, indicators and labels

Protection from crimes	Indicator	Label
Protection from homicides and violent crimes	Homicide rate	VII.1*
	Violent crimes reported	VII.3*
Landscape and cultural heritage	Indicator	Label
Landscape and cultural heritage	Conservation of historic urban fabric	VIII.1
	Presence of historic parks/gardens of significant public interest	VIII.2
	Museums and similar institutions	VIII.3
Environment	Indicator	Label
Quality of environment	Urban green areas	IX.1
	Low air pollution (PM10)	IX.3*
Restraint of resource consumption	Energy from renewable sources (electricity)	IX.7
	Restraint of landfill storage of waste	IX.8*
Research and innovation	Indicator	Label
Research and innovation	Propensity to patent	X.1
	Production industry specialization in knowledge-intensive sectors	X.7
Quality of services	Indicator	Label
Quality of services	Regular electricity supply	XI.1
	Separate collection of urban waste	XI.3
	No prisons exceeding capacity limits	XI.4
	Density of urban public transportation networks	XI.6
	Taking charge of users for early childhood services	XI.2

References

Albers, S., & Hildebrandt, L. (2006). Methodological problems in success factor studies - measurement error, formative versus reflective indicators and the choice of the structural equation model. *Zeitschrift für betriebswirtschaftliche Forschung, 58*(2), 2–33.

Alkire, S. (2012). Dimensions of human development. *World Development 30*(2), 181–205.

Becker, J. M., Klein, K., & Wetzels, M. (2012). Hierarchical latent variable models in PLS-SEM: guidelines for using reflective-formative type models. *Long Range Planning, 45*(5), 359–394

Becker, J. M., Rai, A., & Rigdon, E. E. (2013). Predictive validity and formative measurement in structural equation modeling: Embracing practical relevance. In *Proceedings of the International Conference on Information Systems (ICIS)*, Milan.

Becker, J. M., Rai, A., Ringle, C. M., & Volckner, F. (2013). Discovering unobserved heterogeneity in structural equation models to avert validity threats. *MIS Quarterly, 37*, 665–694.

Chin, W. W. (2010a). How to write up and report PLS analyses. In V. E. Vinzi, W. W. Chin, J. Henseler, & H. Wang (Eds.), *Handbook of partial least squares* (pp. 655–690). Berlin: Springer.

Chin, W. W., & Dibbern, J. (2010b). A permutation based procedure for multi-group PLS analysis: Results of tests of differences on simulated data and a cross cultural analysis of the sourcing of information system services between Germany and the USA. In V. E. Vinzi, W. W. Chin, J. Henseler & H. Wang (Eds.), *Handbook of partial least squares: Concepts, methods and applications*. Springer handbooks of computational statistics series (Vol. II, pp 171–193). Heidelberg: Springer.

Chelli, F. M., Ciommi, M., Emili, A., Gigliarano, C., & Taralli, S. (2015a). Comparing equitable and sustainable well-being (BES) across the Italian provinces. A factor analysis-based approach. *Rivista Italiana di Economia Demografia e Statistica 49*(3/4), 61–73. www.sieds.it.

Cuspi-Istat. (2015b). Il benessere Equo e Sostenibile delle province. Roma, Cuspi https://www.besdelleprovince.it/.

Davino, C. (2016). Combining PLS path modeling and quantile regression for the evaluation of customer satisfaction. *Statistica Applicata, 26*, 93–116.

Davino, C., Dolce, P., Taralli, S., & Vinzi, V. E. (2016a). A quantile composite-indicator approach for the measurement of equitable and sustainable well-being: A case study of the Italian Provinces. In *Social indicators research* (pp. 1–31). Dordrecht: Kluwer Academic Publishers.

Davino, C., Furno, M., & Vistocco, D. (2013). *Quantile regression: Theory and applications*. New York: Wiley.

Davino, C., & Vinzi, V. E. (2016). Quantile composite-based path modeling. *Advances in Data Analysis and Classification. Theory, Methods, and Applications in Data Science 10*(4), 491–520. doi:10.1007/s11634-015-0231-9.

Davino, C., Vinzi, V. E., & Dolce, P. (2016b). Assessment and validation in quantile composite-based path modeling. In H. Abdi, V. E. Vinzi, G. Russolillo, G. Saporta & L. Trinchera (Eds.), *The multiple facets of partial least squares methods*. Springer proceedings in mathematics & statistics (pp. 169–185). Berlin: Springer.

Davino, C., & Vistocco, D. (2007). The evaluation of University educational processes: A quantile regression approach. *Statistica, 3*, 267–278.

Davino, C., & Vistocco, D. (2015). Quantile regression for clustering and modeling data. In I. Morlini, T. Minerva & M. Vichi (Eds.), *Advances in statistical models for data analysis* (pp. 85–95). Cham: Springer.

Delvecchio, F. (1995). Scale di misura e indicatori sociali. Bari, Cacucci.

Diamantopoulos, A., & Winklhofer, H. M. (2001). Index construction with formative indicators: An alternative to scale development. *Journal of Marketing Research, 38*(2), 269–277.

Dolce, P., & Lauro, N. (2015). Comparing maximum likelihood and PLS estimates for structural equation modeling with formative blocks. 730 *Quality and Quantity, 49*, 891–902.

Dolce, P., Vinzi, V. E., & Lauro, C. (2016). Path directions incoherence in PLS path modeling: a prediction-oriented solution. In H. Abdi, V. E. Vinzi, G. Russolillo, G. Saporta & L. Trinchera (Eds.), *The multiple facets of partial least squares methods*. Springer proceedings in mathematics & statistics (pp. 59–70). Berlin: Springer.

ESS - European Statistical System. (2001). Summary Report from the Leadership Group (LEG) on Quality, Eurostat.

Geraci, M., & Bottai, M. (2014) Linear quantile mixed models. *Statistics and Computing, 24*, 461–479.

Guinot, C., Latreille, J., & Tenenhaus, M. (2001). PLS path modeling and multiple table analysis. Application to the cosmetic habits of women in Ile-de-France. *Chemometrics and Intelligent Laboratory Systems, 58*(2), 247–259.

Hahn, C., Johnson, M. D., Herrmann, A., & Huber, F. (2002). Capturing customer heterogeneity using a finite mixture PLS approach. *Schmalenbach Business Review, 54*, 243–269.

Hair, J. F., Hult, G. T. M., Ringle, C. M. & Sarstedt, M. (2014). *A primer on partial least squares structural equation modeling (PLS-SEM)*. Thousand Oaks, CA: Sage.

Hall, J., Giovannini, E., Morrone, A., & Ranuzzi, G. (2010). A framework to measure the progress of societies. OECD Statistics Working Papers, 2010/5, OECD Publishing (2010).

Henseler, J., Dijkstra, T. K., Sarstedt, M., Ringle, C. M., Diamantopoulos, A., Straub, D. W., Ketchen, D. J., Hair, J. F., Tomas, G., Hult, M., & Calantone, R. J. (2014). Common beliefs and reality about partial least squares: Comments on Rönkkö & Evermann (2013). *Organizational Research Methods, 17*(2), 182–209. doi:10.1177/1094428114526928.

Henseler, J., Hubona, G. S., & Ray, P. A. (2016). Using PLS path modeling in new technology research: Updated guidelines. *Industrial Management & Data Systems, 116*(1), 1–19.

Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science 43*(1), 115–135. doi:10.1007/s11747-014-0403-8.

Istat. (2012). Linee guida per la qualità dei processi statistici.

Istat. (2013). Bes 2013 - Il benessere equo e sostenibile in Italia, Roma, Istat.

Istat. (2015). Il Benessere Equo e sostenibile in Italia, Roma, Istat. http://www.misuredelbenessere.it/.

Kocherginsky, M., He, H., & Mu, Y. (2005). Practical confidence intervals for regression quantiles. *Journal of Computational and Graphical Statistics, 14*(1), 41–55.

Koenker, R. (2004). Quantile regression for longitudinal data. *Journal of Multivariate Data Analysis, 91*, 74–89.

Koenker, R., & Basset, G. (1978). Regression quantiles. *Econometrica, 46*, 33–50.

Koenker, R., & Basset, G. W. (1982a). Robust tests for heteroscedasticity based on regression quantiles. *Econometrica, 50*(1), 43–61.

Koenker, R., & Basset, G. W. (1982b). Tests for linear hypotheses and L1 estimation. *Econometrica, 46*, 33–50.

Koenker, R., & Machado, J. (1999). Goodness of fit and related inference processes for quantile regression. *Journal of the American Statistical Association, 94*(488), 1296–1310.

Lamarche, C. (2010). Robust penalized quantile regression estimation for panel data. *Journal of Econometrics, 157*, 396–498.

Lamberti, G., & Aluja, T. B. (2016). Sanchez G The Pathmox approach for PLS path modeling segmentation. *Applied Stochastic Models in Business and Industry, 32*(4), 453–468

Li, G., Li, Y., & Tsai, C. (2015). Quantile correlations and quantile autoregressive modeling. *Journal of the American Statistical Association, 110*, 246–261.

Lohmöller, J. (1989). *Latent variable path modeling with partial least squares*. Heidelberg: Physica-Verlag.

OECD. (2008). *Handbook on constructing composite indicators: Methodology and user guide*. Paris: OECD.

Rigdon, E. E. (2016). Choosing PLS path modeling as analytical method in European management research: A realist perspective. *European Management Journal, 34*(6), 598–605.

Rigdon, E. E., Ringle, C. M., & Sarstedt, M. (2010). Structural modeling of heterogeneous data with partial least squares. In N. K. Malhotra (Ed.), *Review of marketing research* (Vol. 7, pp. 255–296). Armonk: Sharpe.

Ringle, C. M., Sarstedt, M., & Mooi, E. A. (2010). Response-based segmentation using finite mixture partial least squares: Theoretical foundations and an application to American customer satisfaction index data. *Annals of Information Systems, 8*, 19–49.

Ringle, C. M, Sarstedt, M., & Straub, D. W. (2012). A critical look at the use of PLS-SEM in MIS quarterly. *MIS Quarterly, 36*(1), iii–xiv.

Sarstedt, M. (2008). A review of recent approaches for capturing heterogeneity in partial least squares path modelling. *Journal of Modelling in Management, 3*, 140–161.

Sarstedt, M., Becker, J.-M., Ringle, C. M., & Schwaiger, M. (2011a). Uncovering and treating unobserved heterogeneity with FIMIX-PLS: Which model selection criterion provides an appropriate number of segments? *Schmalenbach Business Review, 63*, 34–62.

Sarstedt, M., Hair, J. F., Ringle, C., Thiele, K., & Gudergan, S. (2016a). Estimation issues with PLS and CBSEM: where the bias lies! *Journal of Business Research, 69*(10), 3998–4010.

Sarstedt, M., Henseler, J., & Ringle, C. M. (2011b). Multi-group analysis in partial least squares (PLS) path modeling: Alternative methods and empirical results. In M. Sarstedt, M. Schwaiger & C. R. Taylor (Eds.), *Advances in international marketing* (Vol. 22, pp. 195–218). Bingley: Emerald.

Sarstedt, M., Ringle, C. M., & Gudergan, S. P. (2016b). Guidelines for treating unobserved heterogeneity in tourism research: A comment on Marques and Reis (2015). *Annals of Tourism Research, 57*, 279–284.

Taralli, S., Capogrossi, C., & Perri, G. (2015). Measuring equitable and sustainable well-being (BES) for policy-making at local level (NUTS3). Riv Ital di Economia Demografia e Statistica, 49(3/4), 95–107. www.sieds.it.

Taralli, S., & D'Andrea, P. (2015). BES indicators at local level: Application to policy cycle and evaluation. In Measuring progress at a local level, book of abstracts (pp. 13–16). Pisa: University Press.

Tenenhaus, M., Vinzi, V. E., Chatelin, M., & Lauro, C. (2005). PLS path modeling. *Computational Statistics and Data Analysis, 48*, 159–205.

Esposito Vinzi, V., Trinchera, L., Squillacciotti, S., & Tenenhaus, M. (2008). REBUS-PLS: A response-based procedure for detecting unit segments in PLS path modelling. *Applied Stochastic Models in Business and Industry, 24*, 439–458.

Wetzels, M., Odekerken-Schröoder, G., & van Oppen, C. (2009). Using PLS path modeling for assessing hierarchical construct models: Guidelines and empirical illustration. *MIS Quarterly, 33*(1), 177–195.

Wilson, B. (2010). Using PLS to investigate interaction effects between higher order branding constructs. In V. E. Vinzi, W. W. Chin, J. Henseler & H. Wang (Eds.), *Handbook of partial least squares: Concepts, methods and applications*. Springer handbooks of computational statistics series (Vol. II, pp. 521–652). Heidelberg: Springer.

Wilson, B., & Henseler, J. (2007). Modeling reflective higher-order constructs using three approaches with PLS path modeling: A Monte Carlo comparison. In *Australian and New Zealand Marketing Academy Conference*, Otago (pp. 791–800).

Wold, H. (1982). Soft modeling: The basic design and some extensions. In K. Jöreskog & H. Wold (Eds.), *Systems under indirect observation* (Vol. 2, pp. 1–54). Amsterdam: North-Holland.

Chapter 6
Ordinal Consistent Partial Least Squares

Florian Schuberth and Gabriele Cantaluppi

Abstract In this chapter, we present a new variance-based estimator called ordinal consistent partial least squares (OrdPLSc). It is a promising combination of consistent partial least squares (PLSc) and ordinal partial least squares (OrdPLS), respectively, which is capable to deal in structural equation models with common factors, composites, and ordinal categorical indicators. Besides providing the theoretical background of OrdPLSc, we present three approaches to obtain constructs scores from OrdPLS and OrdPLSc, which can be used, e.g., in importance-performance matrix analysis. Finally, we show its behavior on an empirical example and provide a practical guidance for the assessment of SEMs with ordinal categorical indicators in the context of OrdPLSc.

6.1 Introduction

Structural equation modeling (SEM) has become a customary method, in particular, in the fields of business and social sciences. Its capacity to model nomological networks, to take into account various forms of measurement error, and to test whole theories makes it a prime candidate for a variety of research issues.

The estimators for SEM can be generally divided into two groups: covariance- and variance-based estimators. Covariance-based parameter estimates are obtained by minimizing the discrepancy between the empirical and model-implied covariance matrix of the indicators. In contrast, variance-based estimators first build linear

---

Sections 6.1–6.4 are largely based on a published article by Schuberth et al. (2016).

F. Schuberth (✉)
Faculty of Business Management and Economics, University of Würzburg, Sanderring 2, 97070 Würzburg, Germany
e-mail: florian.schuberth@uni-wuerzburg.de

G. Cantaluppi
Faculty of Economics, Department of Statistical Science, Università Cattolica del Sacro Cuore, Largo A. Gemelli, 1, 20123 Milan, Italy
e-mail: gabriele.cantaluppi@unicatt.it

H. Latan, R. Noonan (eds.), *Partial Least Squares Path Modeling*,
DOI 10.1007/978-3-319-64069-3_6

combinations of the observed indicators as stand-ins for the theoretical constructs and subsequently estimate the parameters of interest using these stand-ins. Variance-based estimators are the proper estimators if the underlying model consists of constructs modeled as composites especially in an endogenous position, while covariance-based estimators are preferred if the constructs are modeled as common factors.

The most prominent estimator among the variance-based ones is probably partial least squares path modeling (PLS). Its application is prevalent in many fields, e.g., in management (Hair et al. 2012a) or business research (Gelhard and Von Delft 2016). Its capacity to estimate models with constructs modeled as common factors or as composites[1] makes it an attractive estimator for SEM.

Over the last decade, a lot of enhancements and extensions of PLS have been developed. The heterotrait–monotrait ratio of common factor correlations (Henseler et al. 2015b) for testing discriminant validity and a test for measurement invariance of composites (Henseler et al. 2016b) were introduced. Furthermore, the consistent PLS (PLSc) (Dijkstra and Henseler 2015b) and an exact test of overall-model fit (Dijkstra and Henseler 2015a) were developed which allow for the consistent estimation of SEMs with composites and common factors and their goodness-of-fit evaluation. Since all these approaches are based on the traditional PLS algorithm which uses ordinary least squares (OLS) regression analysis, it is implicitly assumed that all indicators are on a metric scale.

A lot of empirical studies are based on data collected by questionnaires, thus the indicators used are rarely on a metric scale: in many situations researchers are faced with data measured on ordinal categorical scales, e.g., in marketing research, in particular customer satisfaction surveys (Hair et al. 2012b; Coelho and Esteves 2007).

It is well known in the PLS literature, as well as in other fields, that treating categorical variables as continuous can lead to biased estimates and therefore to invalid inferences and erroneous conclusions. Lohmöller recognizes that the "[...] standard procedures cannot be used for the categorical and ordinal-scaled variables [...]" (Lohmöller 1989, chapter 4). Furthermore Hair et al. (2012b) mention that PLS is often used with categorical indicators but that their use in a procedure like PLS which is based on the OLS estimator can be problematic.

Several methods to address this issue in the context of PLS are provided in the literature, e.g., ordinal PLS (OrdPLS[2]) an innovative approach to deal with ordinal categorical indicators in a psychometric way (Cantaluppi 2012; Cantaluppi and Boari 2016). As OrdPLS is based on the traditional PLS algorithm, its use is limited to models where all constructs are modeled as composites. However, researchers are often faced with models containing constructs which are modeled as common

[1]For a comparison of constructs modeled as composites or common factors, see Rigdon (2012, 2016).

[2]In the original paper, ordinal PLS was abbreviated to OPLS. To avoid confusion with orthogonal partial least squares regression, in its latest version it is abbreviated to OrdPLS.

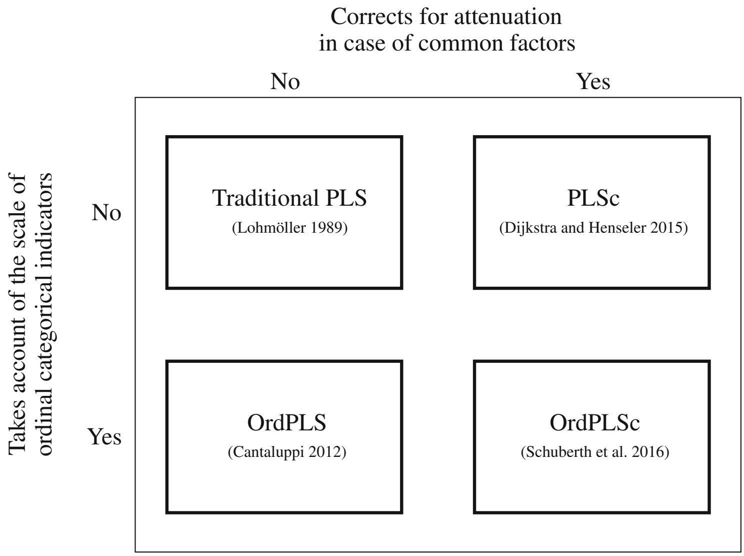


Fig. 6.1 A typology of PLS methods

factors instead of composites (Ringle et al. 2012; Hair et al. 2012b). So, there is a real need of methods like OrdPLSc which are able to deal with common factors, composites, and ordinal categorical indicators.

In this chapter, we present ordinal consistent partial least squares (OrdPLSc) (Schuberth et al. 2016) which is a combination of PLSc and OrdPLS, so providing the advantages of both. OrdPLSc is an estimator which is capable to consistently estimate SEMs including not only composites, but common factors and ordinal categorical indicators too. Figure 6.1 compares the properties of traditional PLS, PLSc, OrdPLS, and OrdPLSc with respect to dealing with common factors and taking into account the scale of ordinal categorical indicators.

As Schuberth et al. (2016) investigate the performance of OrdPLSc by a Monte Carlo simulation, we empirically examine the behavior of OrdPLSc. For this reason, we re-estimate the well-known European customer satisfaction model for the mobile phone industry (Tenenhaus et al. 2005) using OrdPLSc and compare the results with those from PLS, PLSc, and OrdPLS. We also provide commonly used instruments in PLS to assess construct validity and composite reliability taking into account the qualitative scale of ordinal categorical indicators. Additionally, we show how construct scores can be obtained from OrdPLSc.

The remainder of the chapter is organized as follows: the next section shows the development from PLS to PLSc and provides a reformulation of these two procedures in terms of indicators correlation matrices. In Sect. 6.3, we give a literature review of existing approaches dealing with categorical indicators in the framework of PLS, in particular we present the idea of the OrdPLS approach. Section 6.4 presents the ordinal consistent PLS (OrdPLSc). In the following

Sects. 6.5 and 6.6, we present methods to obtain construct scores and show ways to assess the results from OrdPLSc. In Sect. 6.7, we assess the results from OrdPLSc and compare them with those obtained from PLS, PLSc, and OrdPLS using an empirical example. The chapter closes with the conclusion in Sect. 6.8. Furthermore, we provide additional results in the Appendix.

6.2 The Development from PLS Path Modeling to Consistent PLS Path Modeling

PLS was developed by Wold (1975) for the analysis of high dimensional data in a low-structure environment and has undergone various extensions and modifications. It is an approach similar to generalized canonical correlation analysis (GCCA), and in addition able to emulate several of Kettenring's (1971) techniques for the canonical analysis of several sets of variables (Tenenhaus et al. 2005).

In its most modern appearance known as consistent PLS (PLSc) (Dijkstra and Henseler 2015a,b), it can be understood as a well-developed SEM method. It is capable to estimate recursive and non-recursive structural models with constructs modeled as composites and common factors. Figure 6.2 compares PLS with PLSc. Both obtain the outer weights and estimated stand-ins for the constructs by the classical PLS algorithm. While traditional PLS simply relies on OLS to estimate the model parameters, its extended version, PLSc, uses two-stage least squares (2SLS) to consistently estimate recursive path models. Furthermore, PLSc is able to handle both, constructs modeled as composites and common factors by using a post-correction for attenuation, for correlations affected by common factors.

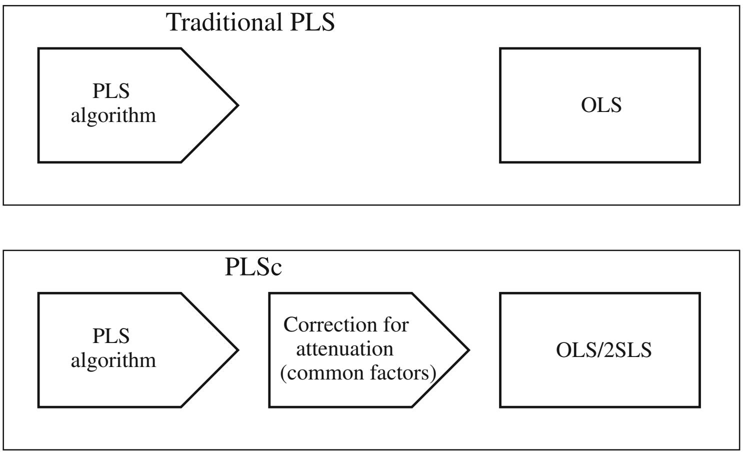
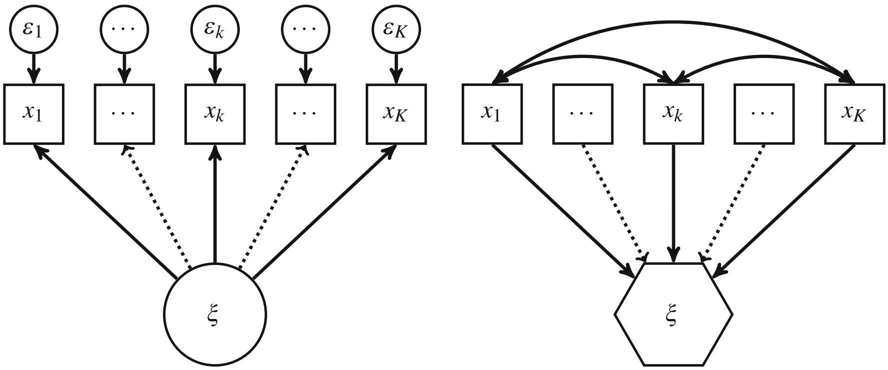


Fig. 6.2 Development of PLS to PLSc



(a) Common factor vs. (b) composite model

Fig. 6.3 Common factor versus composite

The classical common factor model assumes that the variance of a block of indicators $(x_1, \ldots, x_K)$ is completely explained by the underlying common factor (ξ in the large circle) and by their random errors $(\varepsilon_1, \ldots, \varepsilon_K)$, see Fig. 6.3a. This kind of model is commonly used in behavioral research. As Fig. 6.3b depicts, composites (ξ in the hexagon) are formed as linear combinations of their belonging indicators $(x_1, \ldots, x_K)$. Since the composite model does not put any restrictions on the covariances of the indicators belonging to one block, it relaxes the assumption that all of the covariation between the indicators has to be explained by the common factor.[3] Composites are often used as proxies for scientific concepts of interest (Ketterlinus et al. 1989; Maraun and Halpin 2008; Tenenhaus 2008; Rigdon 2012).

For the derivation of OrdPLS(c)[4] it is crucial to describe the well-known PLS algorithm (Wold 1975) and its extension to PLSc in terms of indicator covariances or correlations, respectively. Since in PLS no distinction between exogenous and endogenous constructs is made, we use the following notation: $\boldsymbol{\eta}$ is a $(J \times 1)$ vector containing all stand-ins for constructs which are connected by the structural model, whether they are modeled as common factors or as composites . The $(K \times 1)$ vector $\boldsymbol{x}$ contains the indicators which measure the common factor or build the composite, respectively.

6.2.1 Partial Least Squares

For a sample of size n, all observations of the K indicators are stacked in a data matrix $\boldsymbol{X}$ of dimension $(n \times K)$. For simplicity, the K_j indicators belonging to one

[3]See Bollen and Bauldry (2011) for a description of the different kinds of indicators.

[4]We use hereafter the notation OrdPLS(c) when we refer to OrdPLS and OrdPLSc.

common factor or one composite η_j are grouped to form block j with $j = 1, \ldots, J$. Observations of block j are stacked in the data matrix $\boldsymbol{X}_j$ of dimension $(n \times K_j)$ with $\sum_{j=1}^{J} K_j = K$. Without loss of generality, each indicator is standardized to have mean zero and a variance of one, such that the sample covariance matrix $\boldsymbol{S}$ equals the sample correlation matrix.

The PLS estimation procedure consists of three parts. In the first part, initial arbitrary outer weights $\hat{\boldsymbol{w}}_j^{(0)}$ $(K_j \times 1)$ for each block j are chosen which satisfy the following condition: $\hat{\boldsymbol{w}}_j^{(0)\prime} \boldsymbol{S}_{jj} \hat{\boldsymbol{w}}_j^{(0)} = 1$ where the $(K_j \times K_j)$ matrix $\boldsymbol{S}_{jj}$ contains the sample correlations of the indicators of block j. This condition holds for all outer weights in each iteration i and can be achieved by using a scaling factor $(\hat{\boldsymbol{w}}_j^{(i)\prime} \boldsymbol{S}_{jj} \hat{\boldsymbol{w}}_j^{(i)})^{-\frac{1}{2}}$ for the outer weights $\hat{\boldsymbol{w}}_j^{(i)}$ in each iteration.

In the second part, the iterative PLS algorithm starts with step one, the outer estimation of η_j:

$$\hat{\boldsymbol{\eta}}_j^{(i)} = \boldsymbol{X}_j \hat{\boldsymbol{w}}_j^{(i)} \quad \text{with} \quad \hat{\boldsymbol{w}}_j^{(i)\prime} \boldsymbol{S}_{jj} \hat{\boldsymbol{w}}_j^{(i)} = 1, \tag{6.1}$$

where $\hat{\boldsymbol{\eta}}_j^{(i)}$ is again a column vector of length n. Since outer weights are scaled, all outer estimates also have mean zero and unit variance.

In the second step, the inner estimate of η_j is calculated as a linear combination of inner weights and outer estimates of η_l:

$$\tilde{\boldsymbol{\eta}}_j^{(i)} = \sum_{l=1}^{J} e_{jl}^{(i)} \hat{\boldsymbol{\eta}}_l^{(i)}, \tag{6.2}$$

where $\tilde{\boldsymbol{\eta}}_j^{(i)}$ is a column vector of length n. The inner weight e_{jl} defines how the inner estimate $\tilde{\boldsymbol{\eta}}_j$ is built. Three different schemes for determination of e_{jl} exist: *centroid* (Wold 1982b), *factorial* (Lohmöller 1989), and *path weighting*. However, all schemes produce essentially the same results (Noonan and Wold 1982), hence, we only consider the *centroid* scheme.[5] The inner weights are chosen according to the signs of the correlations between the outer estimates

$$e_{jl}^{(i)} = \begin{cases} \operatorname{sign}(\hat{\boldsymbol{w}}_j^{(i)\prime} \boldsymbol{S}_{jl} \hat{\boldsymbol{w}}_l^{(i)}), \ \text{for } j \neq l & \text{if construct } j \text{ and } l \text{ are adjacent} \\ 0, & \text{otherwise,} \end{cases} \tag{6.3}$$

where adjacent refers to the constructs j and l being directly connected by the path model. All inner estimates $\tilde{\boldsymbol{\eta}}_j^{(i)}$ are again scaled to have unit variance.

In the third and last step of the iterative part, new outer weights are calculated. This can be done in three ways: *mode A*, *mode B*, and *mode C*. For *mode A*, estimated outer weights of block j equal the estimated coefficients of a multivariate regression

[5]For more details on the other schemes, see Tenenhaus et al. (2005).

from the indicators of block j on its related inner estimate. Due to standardization, new outer weights equal the correlations between the inner estimates of η_j and its related indicators:

$$\hat{\boldsymbol{w}}_j^{(i+1)} \propto \sum_{l=1}^{J} \boldsymbol{S}_{jl}\hat{\boldsymbol{w}}_l^{(i)} e_{jl}^{(i)} \quad \text{with} \quad \hat{\boldsymbol{w}}_j^{(i+1)\prime}\boldsymbol{S}_{jj}\hat{\boldsymbol{w}}_j^{(i+1)} = 1. \tag{6.4}$$

In contrast, for *mode B*, the new outer weights equal the estimated coefficients of a regression from the inner estimate on its connected indicators:

$$\hat{\boldsymbol{w}}_j^{(i+1)} \propto \boldsymbol{S}_{jj}^{-1} \sum_{l=1}^{J} \boldsymbol{S}_{jl}\hat{\boldsymbol{w}}_l^{(i)} e_{jl}^{(i)} \quad \text{with} \quad \hat{\boldsymbol{w}}_j^{(i+1)\prime}\boldsymbol{S}_{jj}\hat{\boldsymbol{w}}_j^{(i+1)} = 1. \tag{6.5}$$

Mode C, also known as *MIMIC mode*, is a mixture of mode A and B and is not considered here.

Since the traditional PLS algorithm has no single optimization criteria to be minimized, the new outer weights $\hat{\boldsymbol{w}}_j^{(i+1)}$ are checked for significant changes compared to the outer weights from the iteration step before $\hat{\boldsymbol{w}}_j^{(i)}$. If there is a significant change in the weights, the algorithm starts again at step one by building new outer estimates $\hat{\boldsymbol{\eta}}_j^{(i+1)}$ with the new outer weights $\hat{\boldsymbol{w}}_j^{(i+1)}$, otherwise it stops.

In the last part, at the algorithm convergence, the obtained stable outer weights $\hat{\boldsymbol{w}}_j$ are used to build the final composite stand-ins for both type of constructs:

$$\hat{\boldsymbol{\eta}}_j = \boldsymbol{X}_j\hat{\boldsymbol{w}}_j. \tag{6.6}$$

For constructs which are modeled as common factors, the factor loadings are estimated by OLS in accordance with the measurement model. In contrast, for constructs which are modeled as composites, the final weights equal the stable weights from the last iteration. Finally, path coefficients are estimated by OLS with respect to the structural model.

6.2.2 Consistent Partial Least Squares

As traditional PLS is based on composites, whose estimates are biased if constructs are modeled as common factors.[6] In general, a composite model has larger absolute inter composite correlations compared to the absolute inter common factor correlations of a model with the same structure but where all constructs are modeled as common factors. However, a transformation of the model-implied correlation

[6]Both, common factors and composites are possible ways of construct modeling, see Rigdon (2012).

matrix of a composite model into the model-implied correlation matrix of a common factor model can be achieved by a correction for attenuation (Cohen et al. 2013, chapter 2.10). Consistent PLS (PLSc) uses this correction to obtain consistent estimates for models where the underlying constructs are modeled by common factors (Dijkstra and Henseler 2015a,b).

The correction requires that each common factor is measured by at least two indicators and uses the linearity between population factor loadings and the population weights, $\boldsymbol{\lambda}_j = c_j \boldsymbol{w}_j$. The estimated correction factor for block j satisfies the following condition:

$$\text{plim}(\hat{c}_j) = \sqrt{\boldsymbol{\lambda}_j' \boldsymbol{\Sigma}_{jj} \boldsymbol{\lambda}_j}, \tag{6.7}$$

where $\boldsymbol{\lambda}_j$ is a column vector of length K_j containing the population loadings of common factor η_j and $\boldsymbol{\Sigma}_{jj}$ is the $(K_j \times K_j)$ population correlation matrix of the indicators of block j.[7] The correction factor $\hat{c}_j$ can be obtained by

$$\hat{c}_j^2 = \frac{\hat{\boldsymbol{w}}_j'(\boldsymbol{S}_{jj} - \text{diag}(\boldsymbol{S}_{jj}))\hat{\boldsymbol{w}}_j}{\hat{\boldsymbol{w}}_j'(\hat{\boldsymbol{w}}_j\hat{\boldsymbol{w}}_j' - \text{diag}(\hat{\boldsymbol{w}}_j\hat{\boldsymbol{w}}_j'))\hat{\boldsymbol{w}}_j}. \tag{6.8}$$

It is chosen such that the Euclidean distance between

$$\boldsymbol{S}_{jj} - \text{diag}(\boldsymbol{S}_{jj}) \quad \text{and} \quad (c_j\hat{\boldsymbol{w}}_j)(c_j\hat{\boldsymbol{w}}_j)' - \text{diag}((c_j\hat{\boldsymbol{w}}_j)(c_j\hat{\boldsymbol{w}}_j)') \tag{6.9}$$

is minimized (Dijkstra and Henseler 2015a). The factor loadings of block j are consistently estimated by

$$\hat{\boldsymbol{\lambda}}_j = \hat{c}_j \hat{\boldsymbol{w}}_j. \tag{6.10}$$

As PLSc is able to estimate recursive and non-recursive models, path coefficients are estimated by OLS or 2SLS according to the underlying structural model. Since all variables are standardized, the estimated path coefficients are based on the correlation between the columns of $\hat{\boldsymbol{\eta}}$. The correlation between the common factors j and l is consistently estimated by:

$$\widehat{\text{cor}}(\eta_j, \eta_l) = \frac{\hat{\boldsymbol{w}}_j' \boldsymbol{S}_{jl} \hat{\boldsymbol{w}}_l}{\sqrt{\hat{c}_j^2 \hat{\boldsymbol{w}}_j' \hat{\boldsymbol{w}}_j \hat{c}_l^2 \hat{\boldsymbol{w}}_l' \hat{\boldsymbol{w}}_l}}. \tag{6.11}$$

Using the corrected correlation of Eq. (6.11) for the estimation of the structural model, one obtains consistently estimated path coefficients between the common

[7] The use of *mode B* for common factors is not considered here. For a consistent version of PLS for mode B, see Dijkstra (2011).

factors.[8] For constructs which are modeled as composites no correction of the correlation is required because, by construction, they are not affected by attenuation. For example, if construct j is modeled as a common factor and construct l as a composite, the consistently estimated correlation is obtained by

$$\widehat{\text{cor}}(\eta_j, \eta_l) = \frac{\hat{\boldsymbol{w}}_j' \boldsymbol{S}_{jl} \hat{\boldsymbol{w}}_l}{\sqrt{\hat{c}_j^2 \hat{\boldsymbol{w}}_j' \hat{\boldsymbol{w}}_j}}. \tag{6.12}$$

6.3 The Development from PLS to Ordinal PLS

Since incorrectly handling ordinal categorical variables as continuous can lead to biased inferences and therefore to erroneous conclusions, the PLS literature provides approaches dealing with discrete indicators: dichotomize the ordinal categorical indicator, a mixture of PLS and correspondence analysis (CA), partial maximum likelihood PLS (PML-PLS), and non-metric PLS (NM-PLS).

Common practice in PLS is to replace a categorical indicator by a dummy matrix which is known as dichotomizing. Since the categorical indicator is replaced by $M-1$ dummy variables, where M is the number of observed categories, $M-1$ outer weights are obtained for the original variable. This contradicts the idea of treating an indicator as a whole.

Betzin and Henseler (2005) use correspondence analysis to quantify ex-ante categorical indicators. As the quantified indicators are obtained, PLS is used to estimate the model parameters. As a result, individual weights are obtained for each category of the categorical indicator. Again, this has the drawback that no single outer weight for a categorical indicator is calculated.

Partial maximum likelihood partial least squares (PML-PLS) (Jakobowicz and Derquenne 2007) is a modified version of the original PLS algorithm. It is a combination of PLS and generalized linear models, designed to deal with indicators of any scale. For categorical indicators, individual outer weights are computed for each category by ANOVA. Based on those, one "global" weight per categorical indicator is calculated. However, statistical properties like the proportionality of outer weights to factor loadings are unknown for the global weight and further investigation is needed. Moreover, the authors note that PML-PLS "is especially advantageous in the case of nominal or binary variables" (Jakobowicz and Derquenne 2007) but we focus on ordinal categorical indicators.

The last approach, non-metric partial least squares (NM-PLS) combines the PLS algorithm with optimal scaling to quantify qualitative indicators (Trinchera and

[8]For more details, e.g., the consistent estimation of non-recursive models and the correction for nonlinear structural equation models, see Dijkstra (1985, 1983, 2010, 2011); Dijkstra and Schermelleh-Engel (2014).

Russolillo 2010; Russolillo 2012). Optimal scaling is a procedure, which quantifies qualitative variables by preserving properties of the original measurement scales. In case of NM-PLS, the categorical indicator is quantified in a way that the correlation between the inner LV estimate and the quantified categorical indicator is maximized. As a result for each variable one outer weight is obtained as in traditional PLS for continuous indicators.

6.3.1 Ordinal Partial Least Squares

A further promising approach to deal with ordinal categorical variables is ordinal PLS (OrdPLS) (Cantaluppi 2012). It is a modified procedure for handling ordinal categorical variables in a classical psychometric way. In Sect. 6.2 we showed that all parameters can be obtained by the use of the correlation matrix $\boldsymbol{S}$. Traditional PLS uses the Bravais-Pearson (BP) correlation matrix, which requires all indicators to be continuous for consistency. The observation of an ordinal categorical variable is a qualitative measure, yet it is often coded as numeric and therefore mistakenly treated as quantitative by researchers. This routinely happens in applications with binary and ordinal categorical indicators, which results in biased BP correlation estimates (Quiroga 1992; O'Brien and Homer 1987; Wylie 1976; Carroll 1961). To fix this, OrdPLS uses a consistent correlation matrix as input to the traditional PLS algorithm. An advantage of OrdPLS over the approaches previously introduced is its transparent way of dealing with ordinal categorical variables. Furthermore, as Fig. 6.4 illustrates, the original PLS algorithm remains untouched and it is just provided by a consistent correlation matrix as input for the algorithm.

Since OrdPLS does not correct for attenuation, it shows the same drawbacks as PLS if common factors are included in the model. Nevertheless, we consider Ord-PLS as a powerful extension of PLS when applied under appropriate circumstances, i.e., for models with only composites. Furthermore, it is straightforward to extent by PLSc, to overcome its drawback, see Sect. 6.4. In the following subsection we present Pearson's considerations of ordinal categorical variables to provide a better understanding of the polychoric and polyserial correlation.

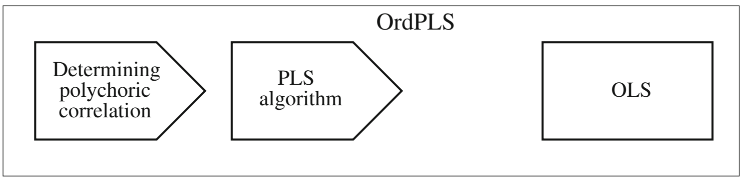


Fig. 6.4 Ordinal partial least squares

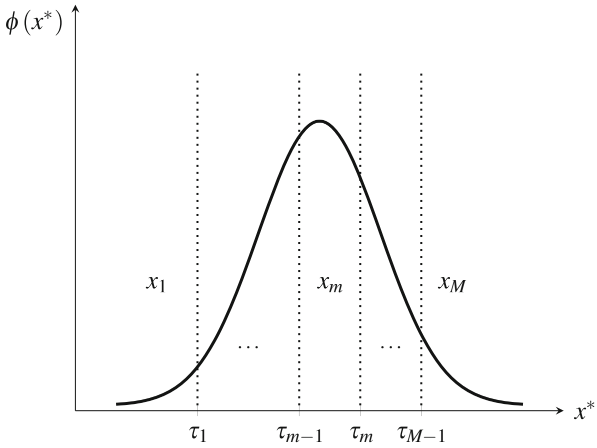


Fig. 6.5 Pearson's idea of an ordinal categorical variable

6.3.2 *Ordinal Categorical Variables According to Pearson*

Pearson (1900, 1913) considers an ordinal categorical variable as a crude measure of an underlying continuous random variable, while Yule (1900) assumes categorical variables being inherently discrete. In this chapter we follow the idea of Pearson: an observed ordinal categorical indicator x is the result of a polytomized standard normally distributed random variable x^*:

$$x = x_m \quad \text{if} \quad \tau_{m-1} \leq x^* < \tau_m \quad m = 1, \ldots, M \tag{6.13}$$

where the threshold parameters $\tau_0, \ldots, \tau_M$ determine the observed categories. The first and last threshold are fixed: $\tau_0 = -\infty$ and $\tau_M = \infty$. Moreover, the thresholds are assumed to be strictly increasing: $\tau_0 < \tau_1 < \ldots < \tau_M$.[9]

Figure 6.5 depicts the idea of an underlying continuous variable: for indicator x category x_m is observed if the realization of the underlying continuous variable x^* is in between the thresholds τ_{m-1} and τ_m.

6.3.3 *Polychoric and Polyserial Correlation*

Since an ordinal categorical variable is determined by an underlying continuous variable, it is more appropriate to consider the correlation between these underlying

[9]In empirical work it can happen that two consecutive threshold parameters are equal, $\tau_{m-1} = \tau_m$, if the corresponding category x_m is not observed.

quantitative continuous variables for evaluating the linear relationship of interest. This is achieved by using the polychoric or polyserial correlation (Drasgow 1986). To illustrate the principles of the polychoric correlation, we consider two ordinal categorical variables x_1 and x_2 with consecutive categories denoted by m_1 and m_2 for x_1 and x_2, respectively, with $m_1 = 1, \ldots, M_1$ and $m_2 = 1, \ldots, M_2$, see Eq. (6.13). The two underlying continuous variables are assumed to be jointly bivariate standard normally distributed with correlation ρ. The correlation between x_1^* and x_2^* can be consistently estimated by maximum likelihood using the following log-likelihood function:

$$\ln L = \ln(c) + \sum_{m_1=1}^{M_1} \sum_{m_2=1}^{M_2} n_{m_1 m_2} \ln(\pi_{m_1 m_2}), \tag{6.14}$$

where $\ln(c)$ is a constant term, $n_{m_1 m_2}$ denotes the observed joint absolute frequency of $x_1 = m_1$ and $x_2 = m_2$, and $\pi_{m_1 m_2}$ is the probability that category m_1 and m_2 are observed jointly. Due to the bivariate normality assumption, $\pi_{m_1 m_2}$ is obtained as:

$$\begin{aligned} &\pi_{m_1 m_2} = \qquad (6.15)\\ &\Phi_2(\tau_{m_1}, \tau_{m_2}, \rho) - \Phi_2(\tau_{m_1}, \tau_{m_2-1}, \rho) - \Phi_2(\tau_{m_1-1}, \tau_{m_2}, \rho_1) + \Phi_2(\tau_{m_1-1}, \tau_{m_2-1}, \rho), \end{aligned}$$

where Φ_2 is the cumulative distribution function of the bivariate standard normal distribution. The parameters τ_{m_1}, τ_{m_2}, and ρ are chosen to maximize the function $\ln L$. In order to reduce computational burden, a two-step procedure can be used (Olsson 1979). In the first step, estimated threshold parameters are separately calculated for both ordinal categorical indicators x_k, with $k = 1, 2$, as quantiles of cumulative marginal frequencies, $\hat{\tau}_{m_k} = \Phi^{-1}(F_{m_k})$ where F_{m_k} equals the cumulative marginal relative frequency up to category m_k. The function Φ^{-1} represents the quantile function of the univariate standard normal distribution. Second, given the estimated threshold parameters, Eq. (6.14) is maximized with respect to ρ. In case of a continuous and an ordinal categorical variable, the correlation between the two continuous variables is obtained by the polyserial correlation (Olsson et al. 1982). For more than two variables, a multivariate version is used to estimate the correlations (Poon and Lee 1987). Moreover, a less computational intensive two-step approach can be used for the multivariate version (Lee and Poon 1987). OrdPLS as well as OrdPLSc makes use of the polychoric and polyserial correlation when ordinal categorical indicators are part of the model.

6.4 Ordinal Consistent Partial Least Squares

Ordinal consistent partial least squares (OrdPLSc) is an approach which deals with ordinal categorical indicators in the same way as OrdPLS, but it additionally uses the correction for attenuation known from PLSc on the resulting model-implied

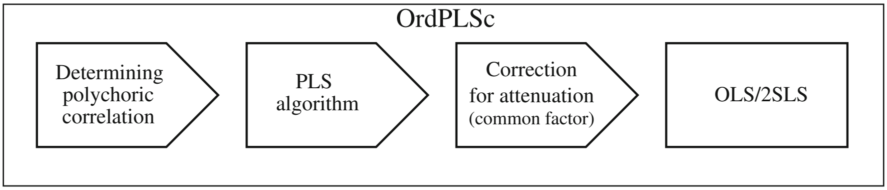


Fig. 6.6 Ordinal consistent partial least squares

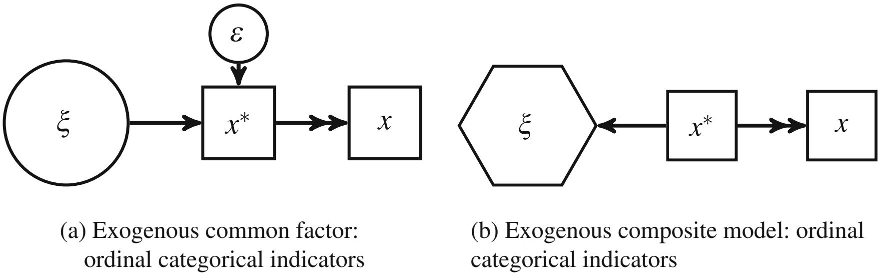


Fig. 6.7 Ordinal categorical indicators in common factor and composite models

correlation matrix if common factors are included in the model. Since, OrdPLS does not affect the original PLS algorithm, the statistical properties of the outer weights are maintained. This is important for the combination of PLSc and OrdPLS because the correction for attenuation of PLS estimates for common factors is heavily based on the proportional property of the outer weights. Figure 6.6 depicts the new approach and shows that OrdPLSc is PLSc provided with a consistent correlation matrix as input.

The role of an ordinal categorical indicator (x), more precisely its underlying continuous variable (x^*), is influenced by its position in the model. As Fig. 6.7a displays, when an ordinal categorical indicator belongs to a common factor, the outcome of the indicator variable is indirectly influenced by the underlying common factor ξ and a measurement error ε through the underlying continuous variable x^*. An ordinal categorical indicator that is part of a composite, see Fig. 6.7b, is simply a crude measure of an underlying continuous variable which actually builds the composite ξ along with other indicators belonging to this block (the double headed arrow represents the relationship between the underlying continuous and the observed ordinal categorical indicator).

Serious problems may appear when the nature of the ordinal categorical indicators is ignored. First, in common factor models the correlation between the indicator and its underlying factor is underestimated (Quiroga 1992; O'Brien and Homer 1987), which leads to biased estimates. Second, in the case of a composite, disregarding the scale of the ordinal categorical indicator leads to biased estimates, too. This is well known as the error-in-variables problem (Wooldridge 2012, chapter 15).

6.5 Evaluation of the Construct Scores in OrdPLS and OrdPLSc

A useful feature of PLS is that construct scores can be calculated directly enabling, e.g., importance-performance matrix analysis (IPMA).[10] In this section, we present three ways of calculating construct scores in the framework of OrdPLS and OrdPLSc[11] to obtain a category indication, which is expressed on the same common ordinal scale characterizing the manifest indicators[12] of the latent variable.[13] As the continuous variables $\boldsymbol{x}_j^*$ underlying each ordinal categorical indicator are not observable (see Fig. 6.7), unique construct scores $\hat{\eta}_j$ cannot be calculated directly. We are only able to determine the probability function of each composite $\hat{\eta}_j$ and an interval of possible values conditional on the threshold values pertaining to the latent variables x_{jk}^* that underlie each ordinal categorical indicator related to η_j. This information can be used for approximating construct scores.

Each underlying latent variable x_{jk}^* with $k = 1, \ldots, K_j$ is assumed to be standard normally distributed. Therefore, the composite $\hat{\eta}_j$, which is defined by the following linear combination (analogous to Eq. (6.6)):

$$\hat{\eta}_j = \sum_{k=1}^{K_j} \hat{w}_{jk} x_{jk}^*, \tag{6.16}$$

is also normally distributed and on a continuous scale. In order to assign a location value to the composite $\hat{\eta}_j$, a set of threshold parameters for the composite $\tau_m^{\hat{\eta}_j}$, $m = 1, \ldots, M-1$, can be derived from the individual indicators threshold parameters $\tau_m^{x_{jk}}$ referred to the underlying variable x_{jk}^*, $k = 1, \ldots, K_j$, where M denotes the number of categories for the indicator x_{jk}, as

$$\tau_m^{\hat{\eta}_j} = \sum_{k=1}^{K_j} \hat{w}_{jk} \tau_m^{x_{jk}}. \tag{6.17}$$

For practical reasons, threshold parameters equal to $\pm\infty$ are replaced by ± 4.

[10] IPMA is a technique aimed at finding which construct is better to act on, in order to improve the average level of a target construct. It is based on a scatterplot diagram, for each endogenous construct, representing summary location measures (performances) of its antecedent latent variables and their impacts (regression coefficients) on the analyzed endogenous construct.

[11] As the weights are unaffected by the correction for attenuation, the construct scores are the same for PLS and PLSc as well as for OrdPLS and OrdPLSc.

[12] Calculating construct scores in presence of indicators defined on Likert scales with different numbers of categories will be considered in future research.

[13] In PLS, a linear transformation of standardized scores is sufficient to assign location and scale to construct scores as shown, e.g., in Bayol et al. (2000).

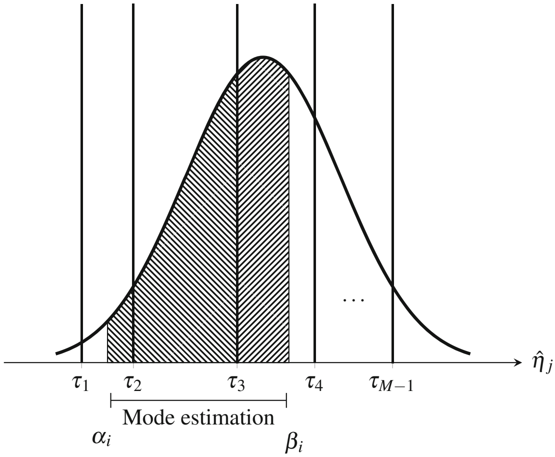


Fig. 6.8 Categorical construct scores, see Eq. (6.21)

In the following, we first define an interval of $\hat{\eta}_{ji}$, with $i = 1, \ldots, n$, which is the image of x_{jki}, $k = 1, \ldots, K_j$, where x_{jki} is the observation of subject i for indicator x_{jk} linked to the corresponding composite $\hat{\eta}_j$. Using the interval, we propose three ways to obtain construct scores on the common ordinal scale characterizing the indicators.

In case subject i chooses the same category m for all the indicators connected to $\hat{\eta}_j$, $x_{j1i} = \cdots = x_{jK_j i} = m$ with $m \in \{1, \ldots, M\}$, the image is of the type

$$A_m \equiv (\tau_{m-1}^{\hat{\eta}_j}, \tau_m^{\hat{\eta}_j}] \tag{6.18}$$

which we call *homogeneous thresholds*. Otherwise, as illustrated in Fig. 6.8, the set which is the image of all possible responses x_{jki}, will not correspond exactly to one subset A_m. Let us denote this set for subject i with

$$C_{ji} \equiv (\alpha_i^{\hat{\eta}_j}, \beta_i^{\hat{\eta}_j}], \tag{6.19}$$

where

$$\alpha_i^{\hat{\eta}_j} = \sum_{k=1}^{K_j} \hat{w}_{jk} \tau_{m-1}^{x_{jki}} \qquad \text{and} \qquad \beta_i^{\hat{\eta}_j} = \sum_{k=1}^{K_j} \hat{w}_{jk} \tau_m^{x_{jki}}. \tag{6.20}$$

The parameters $\tau_{m-1}^{x_{jki}}$ and $\tau_m^{x_{jki}}$ are the threshold parameters which determine the observed category for subject i of indicator x_{jk} and which can be used to define the values for the interval of x_{jk}^*. To assign a category to the i-th observation of the composite $\hat{\eta}_j$, we propose one of the following options:

1. **Mode estimation:** compute, see Fig. 6.8, the probabilities for C_{ji} to overlap each set A_m defined by the "homogeneous thresholds"

$$P\left(C_{ji} \cap A_m\right) \quad m = 1, \ldots, M \tag{6.21}$$

and select for observation i the set A_m with the maximum probability. To the set A_m corresponds the assignment of category m as a score estimate for the construct η_j. In Fig. 6.8 category 3 (interval from τ_2 to τ_3) is assigned.

2. **Median estimation:** compute the median for each observation i of the variable $\hat{\eta}_j$ over the interval C_{ji},

$$\text{median}(\hat{\eta}_{ji}|\hat{\eta}_{ji} \in C_{ji}) = \Phi^{-1}\left(\frac{1}{2}(\Phi(\alpha_i^{\hat{\eta}_j}) + \Phi(\beta_i^{\hat{\eta}_j})\right). \tag{6.22}$$

The category m pertaining the set A_m to which $\text{median}(\hat{\eta}_{ji}|\hat{\eta}_{ji} \in C_{ji})$ belongs is assigned to subject i.

3. **Mean estimation:** compute the mean of the variable $\hat{\eta}_{ji}$ over the interval C_{ji},

$$E(\hat{\eta}_{ji}|\hat{\eta}_{ji} \in C_{ji}) = \frac{\phi(\alpha_i^{\hat{\eta}_j}) - \phi(\beta_i^{\hat{\eta}_j})}{\Phi(\beta_i^{\hat{\eta}_j}) - \Phi(\alpha_i^{\hat{\eta}_j})}. \tag{6.23}$$

The category m pertaining the set A_m to which $E(\hat{\eta}_{ji}|\hat{\eta}_{ji} \in C_{ji})$ belongs is assigned to subject i.

6.6 Assessing the Results of OrdPLSc

The main focus of this section is the assessment of the OrdPLSc results. In the following, we give a brief overview of approaches commonly followed in PLS and PLSc which can be also used to assess the results from OrdPLSc. Furthermore, we present approaches for the model evaluation in the case of ordinal categorical indicators.

6.6.1 Overall Model Evaluation

Statistical tests for the evaluation of the overall-model fit were not available for a long time in PLS. Most recently, a bootstrap-based test was developed for PLSc (Dijkstra and Henseler 2015a). It is a combination of a bootstrap test about the model-implied covariance matrix (Beran and Srivastava 1985; Bollen and Stine 1992) and PLSc. In the context of OrdPLSc, this approach is under development and object of future research.[14] Nevertheless, the standardized root mean square residual (SRMR) (Hu and Bentler 1999) as an approximate model fit criteria can

[14]Applying simply the transformation proposed by Beran and Srivastava (1985) is not recommended, since the transformation of the qualitative categorical indicators is not clear.

be used. It captures the sum of the squared differences between the empirical and the model-implied correlation matrix. Hence, the smaller the SRMR, the better the model fit. Furthermore, as PLS was developed as prediction method (Wold 1982a), models estimated by PLS should be compared with regard to their predictive power (Shmueli et al. 2016).[15]

6.6.2 Measurement Model

In the following, we focus on criteria to assess convergent validity, discriminant validity, and internal consistency in the case of OrdPLSc.

6.6.2.1 Convergent Validity

Convergent validity refers to the extent to which the reflective indicators under each common factor are actually measuring the same common factor. A typically used measure for convergent validity is the average variance extracted (AVE) (Fornell and Larcker 1981; Fornell and Cha 1994; Farrell 2010), which can be appropriately used in the context of OrdPLSc.

6.6.2.2 Discriminant Validity

Discriminant validity refers to the extent to which a given common factor differs from the other common factors of a model. It can be examined for OrdPLSc in the same manner as for PLSc and is usually assessed by the Fornell–Larcker criterion (Fornell and Larcker 1981). In addition, in PLS(c) the heterotrait–monotrait ratio of common factor correlations (HTMT) (Henseler et al. 2015b) and the cross-loadings are used to investigate whether different common factors are also statistically different. Since all approaches mentioned are based on the factor loading estimates or the indicators correlation matrix, it is straightforward to use them in the context of OrdPLSc.

6.6.2.3 Internal Consistency

Internal consistency relates to the correlations among the indicators of one block and reflects the reliability of the measurement model. To evaluate internal consistency

[15]We investigate the predictive power of OrdPLSc and compare it to other approaches which are able to deal with ordinal categorical indicators in a future study.

for a block of indicators belonging to a common factor,[16] the use of the reliability measure ρ_A is recommended (Dijkstra and Henseler 2015b; Henseler et al. 2016a). Furthermore, measures of composite reliability like Dillon–Goldstein's ρ (also known as Jöreskog ρ_c) (Chin 1998) or Cronbach's α are usually considered. Since all these measures are based on the estimated factor loadings or the indicators correlations, they also can be used in the context of OrdPLSc.[17]

In general, it is noted that measures used in the context of OrdPLSc refer to the underlying continuous latent variables $\boldsymbol{x}_j^*$ instead to the ordinal categorical indicators $\boldsymbol{x}_j$ themselves.

6.6.3 Structural Model

Since OrdPLSc is also based on OLS, the coefficient of determination R^2 of the endogenous composites can be used to assess the structural model. It measures how well the explanatory composites explain an endogenous composite $\hat{\eta}_j$.

6.7 An Empirical Example: Customer Satisfaction

Traditional PLS has been successfully applied to models aiming at measuring customer satisfaction: first on a national level (Fornell et al. 1996; Fornell 1992) and later also in a business context (Johnson et al. 2001). We replicate the study from Bayol et al. (2000) and Tenenhaus et al. (2005) on customer satisfaction in the mobile phone industry to empirically investigate the performance of OrdPLSc in the presence of ordinal data collected on questionnaires with a large number of categories.[18] Furthermore, we compare the results of OrdPLSc to those of PLS, PLSc, and OrdPLS.

The assumed underlying customer satisfaction model refers to a version of the European Customer Satisfaction Index (ECSI) with one exogenous and 6 endogenous common factors, see Fig. 6.9 for a depiction of the structural model. The data set consists of 250 observations on 24 ordinal categorical indicators with ten categories each: 5 measures of *Image* (IMG), 3 measures of *Customers Expectations*

[16] Since composites are by assumption fully reliable, their assessment of internal consistency meaningless.

[17] The idea to calculate a Cronbach's α using the polychoric correlation is already known and denoted as ordinal alpha (Zumbo et al. 2007). Ordinal alpha avoids the negative bias of Cronbach's α in the case of ordinal categorical indicators.

[18] The data set is publicly available, e.g., from the R package *plspm* (Sanchez et al. 2015).

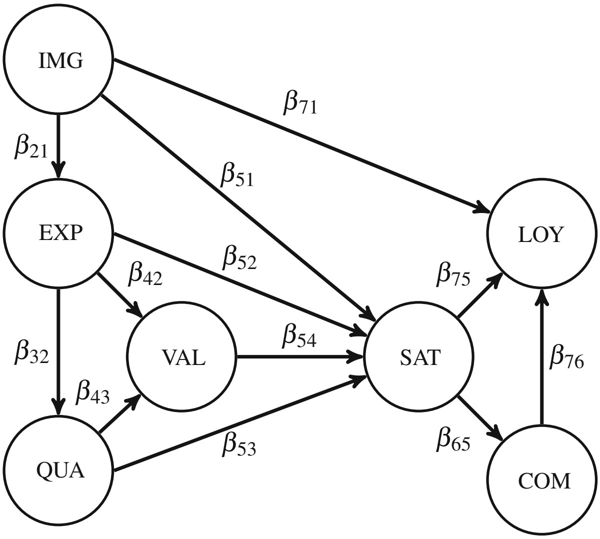


Fig. 6.9 Path diagram of the mobile phone industry customer satisfaction model

(EXP), 7 measures of *Perceived Quality* (QUA), 2 measures of *Perceived Value* (VAL), 3 measures of *Customer Satisfaction* (SAT), one measure of *Complaints* (COM), and 3 measures of *Loyalty* (LOY). For more details on the questionnaire, see Tenenhaus et al. (2005).

The estimation of the model is conducted in R (R Core Team 2016) using the package *matrixpls* (Rönkkö 2016).[19] For the estimation we use the *centroid scheme* for inner weighting, *mode A* for outer estimation, and the algorithm proposed by Wold (1982b) to obtain the final weights. To obtain the polychoric correlation for OrdPLS and OrdPLSc, we use a modified version[20] of the *polychoric* function of the *psych* package (Revelle 2016). Moreover, we use the NONE approach in the case of non-occupied cells, which is recommended for indicators with more than 2 categories (Savalei 2011). For the calculation of the construct scores we use user-written functions, which are provided upon request. In the bootstrap procedures, improper solutions are discarded. Figure 6.10 is built by the R package *ggplot2* (Wickham 2009).

Since the SRMR is below the recommended cut-off value of 0.08 (Hu and Bentler 1998) with regard to OrdPLSc, overall-model fit is established and we proceed with

[19] As *matrixpls* is still under development, we cross validated the results for PLS and PLSc with ADANCO (Henseler and Dijkstra 2015a).

[20] The original *polychoric* function does not allow to calculate the polychoric correlation between indicators with more than 8 categories.

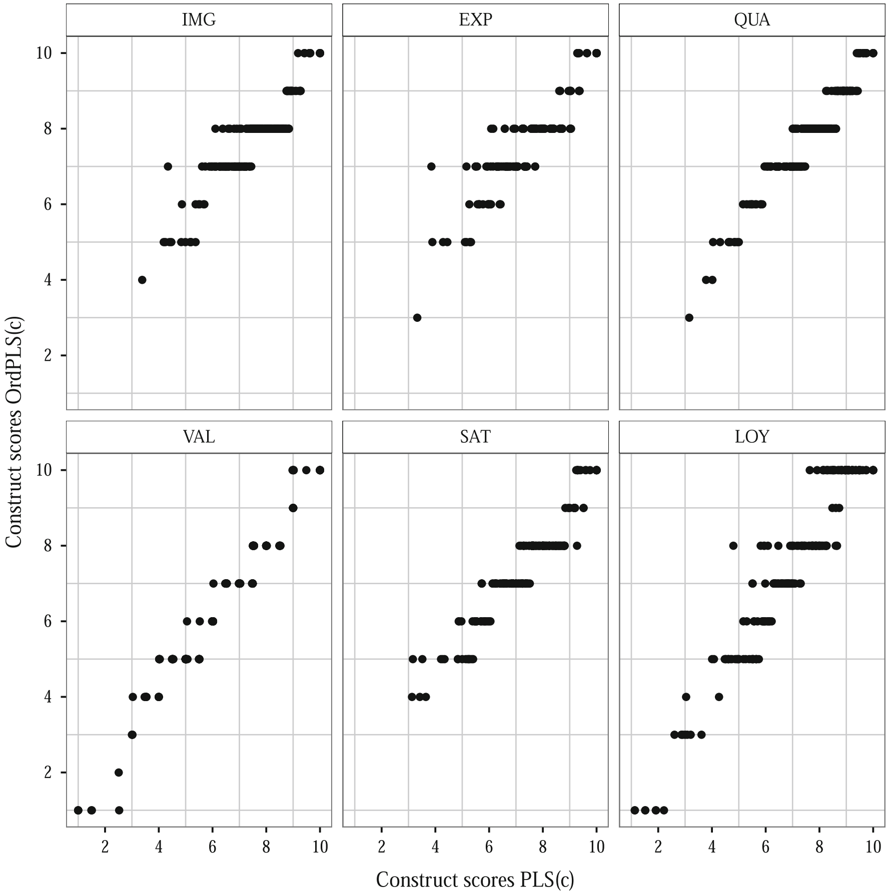


Fig. 6.10 Construct scores for PLS(c) and OrdPLS(c) (mode estimation)

considering the parameter estimates. In Table 6.1, we provide the path coefficient estimates of the mobile phone customer satisfaction model for PLS, PLSc, OrdPLS, and OrdPLSc.

The results show that OrdPLSc produces significant path coefficient estimates for β_{21}, β_{32}, β_{65}, and β_{75} while the other approaches produce a larger number of significant path coefficient estimates. Under the approaches considered, PLS and OrdPLS yield the most path coefficients which are significantly different from zero. Comparing the magnitude of the significant estimates, the consistent versions of PLS and OrdPLS lead to larger absolute path coefficient estimates. However, the path coefficient estimates for PLS and OrdPLS are known to be biased for common factor models (Schneeweiss 1993; Schuberth et al. 2016). Therefore, it is recommended to rely on the OrdPLSc estimates since they have been corrected for attenuation and take the scale of the ordinal indicators into account. Moreover,

Table 6.1 Path coefficient estimates of the mobile phone customer satisfaction model

	PLS	CI[a]		PLSc	CI[b]		OrdPLS	CI[a]		OrdPLSc	CI[c]	
β_{21}	**0.493**	0.379	0.610	**0.887**	0.646	0.940	**0.584**	0.474	0.669	**0.902**	0.740	0.963
β_{51}	**0.153**	0.060	0.262	0.109	−0.431	0.968	**0.199**	0.087	0.318	0.284	−0.130	1.127
β_{71}	**0.212**	0.068	0.362	0.044	−0.590	0.483	**0.261**	0.112	0.404	0.012	−0.680	0.695
β_{32}	**0.545**	0.441	0.653	**0.892**	0.675	0.950	**0.612**	0.507	0.697	**0.880**	0.736	0.959
β_{42}	0.066	−0.095	0.243	0.026	−0.644	0.795	0.037	−0.105	0.210	−0.115	−0.762	0.749
β_{52}	0.037	−0.050	0.139	−0.124	−0.573	0.505	0.035	−0.067	0.134	−0.120	−0.687	0.412
β_{43}	**0.540**	0.357	0.706	0.655	−0.093	1.236	**0.596**	0.416	0.746	0.801	−0.053	1.403
β_{53}	**0.544**	0.408	0.651	**0.833**	0.094	1.264	**0.517**	0.399	0.645	0.669	−0.023	1.042
β_{54}	**0.200**	0.081	0.317	0.191	−0.116	0.358	**0.198**	0.091	0.308	0.177	−0.028	0.306
β_{65}	**0.540**	0.424	0.643	**0.609**	0.483	0.703	**0.563**	0.454	0.657	**0.624**	0.509	0.722
β_{75}	**0.465**	0.290	0.617	**0.867**	0.346	1.472	**0.493**	0.331	0.651	**0.917**	0.226	1.540
β_{76}	0.050	−0.061	0.179	−0.068	−0.221	0.076	0.043	−0.075	0.156	−0.065	−0.178	0.104

Bold printed values are the coefficient estimates whose corresponding CI does not cover the zero
Percentile confidence intervals are calculated at a 95% confidence level
[a] based on 500 bootstrap samples
[b] based on 317 bootstrap samples
[c] based on 270 bootstrap samples

estimates obtained from PLS and OrdPLS as well as PLSc and OrdPLSc are quite similar for most path coefficients. Such a result is expected since the discrepancy between the polychoric and the BP correlation is reduced by the large number of categories per indicator (here: 10).

As all constructs are modeled as common factors, estimated factor loadings $\hat{\boldsymbol{\lambda}}_j$ instead of weights are of main interest in the measurement model. Considering the factor loadings, see Table 6.2, PLS and OrdPLS mostly produce larger estimates than their consistent counterparts. Moreover, all factor loading estimates are significantly different from zero, except the factor loading of indicator loy2 is not significant for OrdPLSc, OrdPLS, and PLSc. The factor loading estimates of PLS and OrdPLS as well as of PLSc and OrdPLSc are not comparable, since we report the standardized estimates of OrdPLS and OrdPLSc and the non-standardized factor loadings of PLS and PLSc.[21] The difference in the magnitude of PLS and PLSc as well as OrdPLS and OrdPLSc factor loading estimates is not surprising as PLS as well as OrdPLS estimates suffers from attenuation in the case of a common factor model. The consistent versions control for this bias leading in general to smaller factor loading estimates.

In the following, we investigate the convergent validity of the common factors. On the main diagonals in Table 6.3a–d, we provide the AVEs which are commonly used to investigate the convergent validity.

It is obvious that OrdPLS and PLS as well as OrdPLSc and PLSc lead to similar results, which is expected as the AVEs are based on the standardized factor loading estimates which only slightly differ because of the large number of categories per indicator, see Table 6.8. It is noteworthy that the AVEs obtained from PLS and OrdPLS estimates are not trustworthy if the common factor model holds, since the factor loading estimates are biased. Since only the AVEs for the common factors QUA, VAL, and SAT are larger than the recommended threshold of 0.5 (Fornell and Larcker 1981) using OrdPLSc, convergent validity cannot be established for the remaining common factors.

To assess discriminant validity, we first consider the standardized cross-factor loadings, see Tables 6.9, 6.10, 6.11, 6.12 in the Appendix. For PLS and OrdPLS, all factor loading estimates are larger than the estimated cross-loadings, which leads to believe that discriminant validity is achieved. In contrast, OrdPLSc as well as PLSc leads to a different conclusion. In almost every block is at least one indicator where a cross-loading exceeds the corresponding factor loading, except for the common factor VAL. This finding is supported by the Fornell–Larcker criterion. Using PLS and OrdPLS estimates, respectively, the Fornell–Larcker criterion indicates that discriminant validity is established for all common factors except for IMG and QUA. While the Fornell–Larcker criterion based on the PLSc and OrdPLSc estimates shows that only for the common factor VAL discriminant validity can be established. Following Henseler et al. (2015b), we recommend to interpret the

[21]Non-standardized factor loading estimates are easier to interpret. Standardized are referred to Table 6.8 in the Appendix.

Table 6.2 Factor loading and confidence interval estimates

	PLS	CI[a]		PLSc	CI[b]		OrdPLS	CI[a]		OrdPLSc	CI[c]	
ima1	**1.216**	0.931	1.462	**1.036**	0.814	1.251	**0.764**	0.690	0.818	**0.662**	0.555	0.760
ima2	**0.953**	0.665	1.209	**0.905**	0.649	1.169	**0.648**	0.525	0.734	**0.592**	0.471	0.691
ima3	**1.399**	1.015	1.698	**0.975**	0.653	1.256	**0.602**	0.463	0.695	**0.483**	0.318	0.610
ima4	**1.455**	1.203	1.669	**1.262**	0.989	1.483	**0.799**	0.729	0.854	**0.720**	0.604	0.818
ima5	**1.085**	0.854	1.359	**1.028**	0.821	1.232	**0.780**	0.724	0.826	**0.710**	0.623	0.791
exp1	**1.111**	0.808	1.367	**0.809**	0.562	1.106	**0.780**	0.654	0.848	**0.556**	0.426	0.686
exp2	**1.153**	0.724	1.547	**0.786**	0.518	1.198	**0.743**	0.589	0.832	**0.539**	0.391	0.707
exp3	**1.523**	1.022	1.931	**0.885**	0.585	1.242	**0.623**	0.476	0.738	**0.479**	0.360	0.620
qua1	**1.104**	0.916	1.281	**1.147**	0.969	1.337	**0.828**	0.779	0.875	**0.835**	0.761	0.895
qua2	**1.228**	0.929	1.529	**0.997**	0.751	1.306	**0.647**	0.531	0.737	**0.553**	0.447	0.683
qua3	**1.456**	1.199	1.685	**1.373**	1.131	1.564	**0.801**	0.747	0.852	**0.769**	0.679	0.849
qua4	**1.253**	1.051	1.428	**1.103**	0.894	1.294	**0.809**	0.744	0.871	**0.733**	0.656	0.819
qua5	**1.062**	0.874	1.237	**0.980**	0.777	1.142	**0.782**	0.718	0.837	**0.720**	0.616	0.821
qua6	**1.246**	1.069	1.402	**1.109**	0.940	1.294	**0.826**	0.752	0.889	**0.758**	0.684	0.849
qua7	**1.477**	1.252	1.738	**1.502**	1.276	1.732	**0.799**	0.735	0.849	**0.822**	0.758	0.891
val1	**2.032**	1.823	2.226	**1.646**	1.412	1.950	**0.914**	0.869	0.944	**0.771**	0.674	0.864
val2	**1.674**	1.413	1.907	**1.704**	1.504	1.965	**0.943**	0.925	0.958	**0.943**	0.881	0.996
sat1	**0.875**	0.715	1.031	**0.854**	0.694	0.981	**0.825**	0.768	0.870	**0.734**	0.644	0.813
sat2	**1.536**	1.329	1.751	**1.247**	1.035	1.445	**0.858**	0.814	0.893	**0.724**	0.656	0.802
sat3	**1.543**	1.319	1.774	**1.422**	1.246	1.653	**0.867**	0.830	0.901	**0.834**	0.780	0.910
loy1	**2.268**	1.936	2.538	**1.559**	1.257	2.002	**0.849**	0.766	0.903	**0.641**	0.524	0.739
loy2	**0.775**	0.037	1.396	0.513	−0.065	0.867	0.193	−0.056	0.393	0.163	−0.013	0.274
loy3	**1.923**	1.611	2.205	**1.893**	1.704	2.205	**0.924**	0.898	0.942	**0.903**	0.856	0.989

Bold printed values are the coefficient estimates where the corresponding CI does not cover the zero
Percentile confidence intervals are calculated at a 95% confidence level
[a] based on 500 bootstrap samples
[b] based on 317 bootstrap samples
[c] based on 270 bootstrap samples

Table 6.3 Average variance extracted and shared variance estimates

(a) PLS

	IMA	EXP	QUA	VAL	SAT	LOY
IMA	**0.476**	0.243	0.535	0.258	0.450	0.300
EXP	0.493	**0.471**	0.297	0.130	0.231	0.134
QUA	0.731	0.545	**0.574**	0.332	0.626	0.275
VAL	0.508	0.360	0.576	**0.850**	0.365	0.267
SAT	0.671	0.481	0.791	0.604	**0.683**	0.403
LOY	0.548	0.366	0.524	0.517	0.635	**0.520**

(b) PLSc

	IMA	EXP	QUA	VAL	SAT	LOY
IMA	**0.356**	0.786	0.835	0.433	0.786	0.600
EXP	0.887	**0.207**	0.796	0.373	0.693	0.460
QUA	0.914	0.892	**0.505**	0.460	0.906	0.455
VAL	0.658	0.611	0.679	**0.715**	0.567	0.475
SAT	0.886	0.833	0.952	0.753	**0.549**	0.747
LOY	0.775	0.678	0.674	0.689	0.864	**0.370**

(c) OrdPLS

	IMA	EXP	QUA	VAL	SAT	LOY
IMA	**0.522**	0.342	0.636	0.309	0.550	0.421
EXP	0.584	**0.517**	0.375	0.161	0.300	0.218
QUA	0.797	0.612	**0.619**	0.383	0.672	0.371
VAL	0.556	0.402	0.619	**0.862**	0.413	0.359
SAT	0.742	0.547	0.820	0.643	**0.723**	0.505
LOY	0.649	0.467	0.609	0.600	0.711	**0.537**

(d) OrdPLSc

	IMA	EXP	QUA	VAL	SAT	LOY
IMA	**0.409**	0.814	0.898	0.455	0.863	0.681
EXP	0.902	**0.276**	0.775	0.348	0.688	0.516
QUA	0.947	0.880	**0.557**	0.489	0.914	0.520
VAL	0.675	0.590	0.699	**0.742**	0.586	0.526
SAT	0.929	0.830	0.956	0.766	**0.586**	0.787
LOY	0.825	0.718	0.721	0.725	0.887	**0.417**

Correlations are below the diagonal, squared correlations are above the diagonal, and AVE estimates are presented on the diagonal (in boldface)

Table 6.4 HTMT results for PLS(c) and OrdPLS(c)

	IMA	EXP	QUA	VAL	SAT	LOY
IMA		0.917	0.949	**0.681**	0.929	0.943
EXP	**0.888**		**0.888**	**0.602**	**0.843**	**0.824**
QUA	0.929	**0.878**		**0.699**	0.958	**0.802**
VAL	**0.652**	**0.589**	**0.673**		**0.765**	**0.865**
SAT	0.910	**0.865**	0.954	**0.741**		1.001
LOY	**0.867**	**0.770**	**0.723**	**0.797**	0.957	

Fornell–Larcker criterion with caution when it is based on PLS or OrdPLS factor loading estimates, as these are known to be upward-biased.

Moreover, the HTMT further supports the suspicion that some indicators are incorrectly assigned. The lower triangular of Table 6.4 represents the results based on the BP correlations, while the upper triangular contains the results based on the polychoric correlations. Since the HTMT is solely based on indicators correlations, it leads to the same results for PLS and PLSc as well as for OrdPLS and OrdPLSc.

The HTMT based on the polychoric correlation indicates that the common factors IMA and EXP, IMA and QUA, IMA and SAT, IMA and LOY, QUAL and SAT, and SAT and LOY cannot be adequately distinguished since it is above the recommended threshold of 0.9 (Gold and Arvind Malhotra 2001).[22] The use of the HTMT based on the BP correlation in case of ordinal categorical indicators is not recommended because it does not take into account the qualitative character of the indicators. For a tutorial on how to proceed if the discriminant validity is not satisfied, see Farrell (2010) and Henseler et al. (2015b).

Furthermore, we assess internal consistency. In doing so, we consider Cronbach's alpha and ordinal alpha, respectively, which are lower-bound estimates for the reliability. Moreover, we refer to Dijkstra and Henseler's ρ_A and Dillon–Goldstein's ρ_c in order to examine the internal consistency. Table 6.5c provides the results of the three measures for the various PLS approaches.

The results for Dillon–Goldstein's ρ_c (Table 6.5c) for PLS and OrdPLS as well as their consistent version are again very similar, which is explained by the large number of indicators categories. Although the results for PLS and OrdPLS are more desirable, it is recommended to use them with caution as they are based on attenuated factor loading estimates (Zumbo et al. 2007). Comparing the values of Dillon–Goldstein's ρ_c for OrdPLSc with the recommended threshold of 0.7 (Henseler et al. 2016a), we conclude that construct scores are reliable except for the common factors EXP and LOY. This conclusion is fully supported by Cronbach's α and by ordinal alpha, see Table 6.5a. However, it is well known that the use of Cronbach's α should be done with circumspection as it can only be appropriately interpreted if the assumption of tau-equivalence holds. If this assumption is violated, Cronbach's α underestimates the reliability (Raykov 2004). Moreover, drawing

[22]Furthermore, the HTMT correlation can be used in a bootstrap procedure that allows the construction of confidence intervals (Henseler et al. 2015b), which is not done here.

Table 6.5 Internal consistency reliability

(a) Cronbach's α and ordinal alpha

	Cronbach's α	ordinal alpha
IMA	**0.723**	**0.768**
EXP	0.452	0.526
QUA	**0.877**	**0.896**
VAL	**0.824**	**0.842**
SAT	**0.779**	**0.809**
LOY	0.472	0.497

(b) Dijkstra–Henseler's ρ_A

	PLS(c)	OrdPLS(c)
IMA	**0.728**	**0.784**
EXP	0.425	0.536
QUA	**0.879**	**0.903**
VAL	**0.820**	**0.866**
SAT	**0.786**	**0.813**
LOY	0.687	**0.789**

(c) Dillon–Goldstein's ρ_c

	PLS	PLSc	OrdPLS	OrdPLSc
IMA	**0.818**	**0.731**	**0.844**	**0.772**
EXP	**0.727**	0.437	**0.761**	0.533
QUA	**0.904**	**0.876**	**0.919**	**0.897**
VAL	**0.919**	**0.832**	**0.926**	**0.850**
SAT	**0.865**	**0.784**	**0.887**	**0.809**
LOY	**0.735**	0.582	**0.736**	0.625

Results for COM are not reported, as it is measured by only one indicator

Table 6.6 R^2 of the endogenous composites

	PLS	PLSc	OrdPLS	OrdPLSc
IMA	0.000	0.000	0.000	0.000
EXP	0.243	0.786	0.342	0.814
QUA	0.297	0.796	0.374	0.773
VAL	0.335	0.461	0.384	0.490
SAT	0.672	0.931	0.719	0.936
COM	0.292	0.371	0.319	0.391
LOY	0.432	0.750	0.538	0.790

conclusion from Cronbach's α in the case of ordinal categorical indicators is not recommended (Zumbo et al. 2007) as it is based on the BP correlation.

Finally we consider Dijkstra–Henseler's ρ_A (Table 6.5b) which also largely supports our conclusion since most reliability estimates exceeds the threshold of 0.7, except for EXP. Even though Dijkstra–Henseler's ρ_A is quite similar for PLS and OrdPLS as well as for PLSc and OrdPLSc in our example, ρ_A based on the BP correlation should be used cautiously in the context of ordinal categorical indicators, as it does not take into account the qualitative scale of the indicators. Again, due to the large number of categories, ρ_A leads to quite similar results for PLS and OrdPLS as well as for PLSc and OrdPLSc.

Besides studying the significance levels of the path coefficients, it is common practice to consider the coefficient of determination R^2 of the endogenous composites (Table 6.6) to investigate the structural model.

Table 6.7 Coherency of construct scores between PLS(c) and OrdPLS(c)

Method	IMG	EXP	QUA	VAL	SAT	LOY
Mode estimation[a]	70.4	71.2	79.2	84.4	71.6	48.0
Median estimation[a]	75.2	74.8	78.0	88.0	70.4	51.6
Mean estimation[a]	73.2	76.8	75.6	86.8	71.6	49.2
Mode estimation[b]	98.8	98.0	100.0	99.6	99.6	90.8
Median estimation[b]	99.2	98.4	100.0	100.0	99.6	94.0
Mean estimation[b]	99.2	98.4	100.0	99.6	100.0	91.6

[a] Percentages of exact concordance after having rounded PLS scores to integer values
[b] Percentages of concordance with a difference between rounded values not larger than 1

Table 6.6 illustrates that the consistent PLS versions result in larger R^2s, which means that more variance of the endogenous composite is explained by the adjacent composites. The R^2s for OrdPLS and PLS as well as OrdPLSc and PLSc are again quite similar.

Finally, we investigate the construct scores for OrdPLSc. Figure 6.10 shows a comparison of the latent scores calculated with PLS(c) and OrdPLS(c).[23] Construct scores for the common factor COM are not reported since the variable is identical to its single indicator.

According to the approaches described in Sect. 6.5, we obtain ordinal construct scores for OrdPLS(c). For PLS(c) we use rounded scores for a better comparison. Table 6.7 shows the degree of coherency between the construct scores obtained from PLS(c) and the three procedures for OrdPLS(c). Using the rounded scores for PLS(c) the percentages of exact concordance are reported in the first three rows, while the remaining rows contain the percentages of concordance with a difference between rounded values not larger than 1. We have at least 70% exact concordance except for the common factor LOY. More than 90% of the cases for all common factors show a difference between rounded values lower than 1.

6.8 Conclusion

In our study, we present the development from PLS to OrdPLSc, a variance-based estimator which is capable to consistently estimate common factor models and to deal with ordinal categorical indicators. While Schuberth et al. (2016) examine the behavior and properties of OrdPLSc by a simulation study, we show how OrdPLSc is applied to a common factor model. Furthermore, we present the assessment of models with ordinal categorical indicators. For this purpose, we use a well-known empirical example from Tenenhaus et al. (2005). The results show that OrdPLS and

[23] As the weights are the same for PLS and PLSc as well as for OrdPLS and OrdPLSc, we only report the scores of PLS and OrdPLS.

OrdPLSc as well as PLS and PLSc produce quite similar estimates which is not surprising as the large number of categories per indicator reduces the discrepancy between the BP and the polychoric correlation. However, in contrast to the original study where traditional PLS was applied, OrdPLSc leads to substantially different results. This is mainly due to the correction for attenuation of the estimates and to the use of validity and reliability measures based on the polychoric correlation. Moreover, we present procedures which can be used to obtain construct scores on the indicators scale. In general, we recommend the use of OrdPLSc in case of ordinal categorical indicators, in particular when indicators with a small number of categories are included in the model.

Our study only considers situations where all indicators are measured on an ordinal categorical scale. Of course, in practice, continuous indicators are often part of the model. In such a context the polyserial correlation can be used to adequately address the issue of ordinal categorical indicators. Hence, future research should investigate the behavior of OrdPLSc for models containing a mixture of both ordinal categorical and continuous indicators. Furthermore, out-of-sample prediction plays an increasingly important role in PLS (Shmueli et al. 2016). Thus, the prediction power needs to be investigated for approaches dealing ordinal categorical indicators, e.g., using k-fold cross-validation. Moreover, we recommend the extension to the polychoric correlation for other variance-based estimators to deal with ordinal categorical indicators, e.g., generalized structural component analysis (Hwang and Takane 2014). Of particular interest is also their comparison to OrdPLS(c). Finally, tests for overall-model fit in case of ordinal categorical indicators need to be developed.

Acknowledgements Florian Schuberth thanks Prof. Jörg Henseler and Prof. Theo K. Dijkstra for their guidance and inspiring thoughts in the field of PLS. A special thanks goes to his supervisor Prof. Martin Kukuk for having introduced him to the polychoric correlation.
Gabriele Cantaluppi is grateful to Prof. Angelo Zanella for having introduced him to Structural Equation Models and Prof. Giuseppe Boari for his collaboration in fine tuning OrdPLS.

Appendix

Appendix 1: Standardized Loading and Cross-Loading Estimates

See Tables 6.8, 6.9, 6.10, 6.11, 6.12.

Table 6.8 Standardized factor loading estimates

	PLS	PLSc	OrdPLS	OrdPLSc
ima1	0.717	0.611	0.764	0.662
ima2	0.566	0.537	0.648	0.592
ima3	0.658	0.458	0.602	0.483
ima4	0.792	0.686	0.799	0.720
ima5	0.698	0.662	0.780	0.710
exp1	0.687	0.500	0.780	0.556
exp2	0.644	0.439	0.743	0.539
exp3	0.726	0.422	0.623	0.479
qua1	0.778	0.808	0.828	0.835
qua2	0.651	0.528	0.647	0.553
qua3	0.801	0.755	0.801	0.769
qua4	0.760	0.669	0.809	0.733
qua5	0.732	0.676	0.782	0.720
qua6	0.766	0.682	0.826	0.758
qua7	0.803	0.816	0.799	0.822
val1	0.933	0.756	0.914	0.771
val2	0.911	0.927	0.943	0.943
sat1	0.711	0.693	0.825	0.734
sat2	0.872	0.708	0.858	0.724
sat3	0.885	0.815	0.867	0.834
loy1	0.855	0.587	0.849	0.641
loy2	0.273	0.181	0.193	0.163
loy3	0.869	0.855	0.924	0.903

Table 6.9 Standardized cross-loading estimates of PLS

	IMA	EXP	QUA	VAL	SAT	COM	LOY
ima1	**0.717**	0.347	0.571	0.393	0.540	0.423	0.338
ima2	**0.566**	0.387	0.492	0.269	0.398	0.188	0.293
ima3	**0.658**	0.272	0.368	0.332	0.339	0.207	0.309
ima4	**0.792**	0.374	0.571	0.459	0.542	0.440	0.461
ima5	**0.698**	0.340	0.544	0.260	0.501	0.337	0.485
exp1	0.349	**0.687**	0.437	0.293	0.362	0.183	0.268
exp2	0.404	**0.644**	0.343	0.175	0.345	0.225	0.320
exp3	0.285	**0.726**	0.357	0.273	0.300	0.126	0.190
qua1	0.622	0.534	**0.778**	0.454	0.661	0.380	0.461
qua2	0.405	0.308	**0.651**	0.295	0.474	0.300	0.319
qua3	0.621	0.423	**0.801**	0.467	0.651	0.472	0.461
qua4	0.480	0.389	**0.760**	0.390	0.587	0.379	0.353
qua5	0.598	0.406	**0.732**	0.455	0.517	0.389	0.373
qua6	0.551	0.447	**0.766**	0.405	0.539	0.418	0.333
qua7	0.596	0.411	**0.803**	0.547	0.707	0.465	0.446
val1	0.405	0.314	0.477	**0.933**	0.495	0.287	0.435
val2	0.542	0.354	0.594	**0.911**	0.629	0.360	0.525
sat1	0.558	0.495	0.637	0.403	**0.711**	0.334	0.484
sat2	0.524	0.395	0.672	0.480	**0.872**	0.416	0.484
sat3	0.612	0.382	0.684	0.588	**0.885**	0.547	0.609
loy1	0.430	0.281	0.393	0.407	0.455	0.237	**0.855**
loy2	0.109	0.095	0.065	0.148	0.115	0.122	**0.273**
loy3	0.528	0.351	0.537	0.481	0.658	0.448	**0.869**

The bold printed value in each row shows on which construct the indicator loads most

Table 6.10 Standardized cross-loading estimates of PLSc

	IMA	EXP	QUA	VAL	SAT	COM	LOY
ima1	**0.611**	0.533	0.608	0.434	0.609	0.423	0.408
ima2	0.537	**0.594**	0.524	0.297	0.449	0.188	0.353
ima3	**0.458**	0.417	0.392	0.367	0.382	0.207	0.373
ima4	**0.686**	0.574	0.609	0.507	0.612	0.440	0.556
ima5	**0.662**	0.522	0.580	0.287	0.565	0.337	0.586
exp1	0.409	**0.500**	0.466	0.324	0.409	0.183	0.324
exp2	**0.473**	0.439	0.366	0.193	0.389	0.225	0.386
exp3	0.334	**0.422**	0.381	0.301	0.338	0.126	0.229
qua1	0.729	**0.819**	0.808	0.501	0.745	0.380	0.556
qua2	0.474	0.473	0.528	0.326	**0.535**	0.300	0.385
qua3	0.728	0.649	**0.755**	0.516	0.734	0.472	0.556
qua4	0.563	0.596	**0.669**	0.430	0.662	0.379	0.426
qua5	**0.701**	0.623	0.676	0.503	0.583	0.389	0.450
qua6	0.646	**0.686**	0.682	0.448	0.608	0.418	0.402
qua7	0.698	0.631	**0.816**	0.604	0.797	0.465	0.538
val1	0.475	0.482	0.509	**0.756**	0.558	0.287	0.525
val2	0.635	0.543	0.633	**0.927**	0.710	0.360	0.633
sat1	0.654	**0.760**	0.679	0.445	0.693	0.334	0.584
sat2	0.614	0.606	**0.717**	0.530	0.708	0.416	0.584
sat3	0.718	0.587	0.729	0.649	**0.815**	0.547	0.735
loy1	0.504	0.431	0.419	0.450	0.514	0.237	**0.587**
loy2	0.128	0.146	0.069	0.164	0.130	0.122	**0.181**
loy3	0.618	0.539	0.572	0.531	0.742	0.448	**0.855**

The bold printed value in each row shows on which construct the indicator loads most

Table 6.11 Cross-loading estimates of OrdPLS

	IMA	EXP	QUA	VAL	SAT	COM	LOY
ima1	**0.765**	0.423	0.620	0.453	0.614	0.458	0.440
ima2	**0.647**	0.445	0.559	0.324	0.474	0.253	0.399
ima3	**0.602**	0.325	0.408	0.362	0.378	0.247	0.376
ima4	**0.799**	0.453	0.652	0.517	0.609	0.491	0.545
ima5	**0.780**	0.456	0.610	0.346	0.570	0.389	0.559
exp1	0.415	**0.779**	0.475	0.334	0.407	0.245	0.336
exp2	0.490	**0.747**	0.428	0.240	0.421	0.273	0.400
exp3	0.348	**0.621**	0.415	0.293	0.346	0.164	0.263
qua1	0.688	0.559	**0.828**	0.522	0.719	0.416	0.563
qua2	0.460	0.351	**0.647**	0.329	0.511	0.314	0.398
qua3	0.671	0.474	**0.801**	0.509	0.678	0.502	0.531
qua4	0.593	0.466	**0.809**	0.457	0.658	0.435	0.435
qua5	0.664	0.495	**0.781**	0.495	0.564	0.435	0.453
qua6	0.642	0.534	**0.826**	0.482	0.620	0.458	0.445
qua7	0.641	0.466	**0.799**	0.577	0.731	0.495	0.505
val1	0.429	0.343	0.511	**0.912**	0.523	0.322	0.480
val2	0.589	0.397	0.627	**0.944**	0.661	0.385	0.620
sat1	0.633	0.526	0.691	0.473	**0.820**	0.379	0.575
sat2	0.572	0.440	0.696	0.524	**0.859**	0.451	0.549
sat3	0.681	0.435	0.705	0.633	**0.871**	0.592	0.677
loy1	0.524	0.383	0.461	0.492	0.513	0.275	**0.855**
loy2	0.095	0.066	0.076	0.138	0.114	0.123	**0.178**
loy3	0.625	0.445	0.610	0.563	0.726	0.496	**0.922**

The bold printed value in each row shows on which construct the indicator loads most

Table 6.12 Cross-loading estimates of OrdPLSc

	IMA	EXP	QUA	VAL	SAT	COM	LOY
ima1	0.666	0.578	0.653	0.486	**0.679**	0.458	0.497
ima2	0.587	**0.608**	0.588	0.348	0.525	0.253	0.451
ima3	**0.482**	0.443	0.429	0.388	0.419	0.247	0.425
ima4	**0.722**	0.619	0.686	0.554	0.674	0.491	0.615
ima5	**0.709**	0.623	0.641	0.371	0.630	0.389	0.631
exp1	0.469	**0.554**	0.500	0.358	0.450	0.245	0.379
exp2	**0.554**	0.545	0.450	0.257	0.466	0.273	0.452
exp3	0.393	**0.476**	0.436	0.314	0.383	0.164	0.297
qua1	0.777	0.764	**0.834**	0.559	0.795	0.416	0.635
qua2	0.519	0.480	0.557	0.353	**0.566**	0.314	0.449
qua3	0.758	0.648	**0.772**	0.546	0.750	0.502	0.599
qua4	0.670	0.637	**0.737**	0.490	0.729	0.435	0.491
qua5	**0.749**	0.676	0.713	0.531	0.624	0.435	0.512
qua6	0.725	0.729	**0.753**	0.516	0.685	0.458	0.502
qua7	0.724	0.637	**0.825**	0.619	0.808	0.495	0.570
val1	0.485	0.469	0.538	**0.764**	0.579	0.322	0.542
val2	0.665	0.543	0.660	**0.951**	0.731	0.385	0.700
sat1	0.714	0.719	**0.727**	0.507	0.713	0.379	0.649
sat2	0.646	0.601	**0.732**	0.561	0.725	0.451	0.619
sat3	0.769	0.595	0.742	0.678	**0.854**	0.592	0.764
loy1	0.591	0.524	0.485	0.527	0.567	0.275	**0.660**
loy2	0.108	0.090	0.080	**0.148**	0.127	0.123	0.142
loy3	0.705	0.609	0.641	0.603	0.804	0.496	**0.891**

The bold printed value in each row shows on which construct the indicator loads most

Appendix 2: Results for PLS and PLSc Based on the Indicators Correlation Matrix

See Tables 6.13, 6.14, 6.15, 6.16, 6.17, 6.18, 6.19.

Since the HTMT is based on the indicators correlations it is the same as in Table 6.4.

Table 6.13 Path coefficient estimates of the mobile phone customer satisfaction model based on the correlation matrix

	PLS	CI[a]		PLSc	CI[b]	
β_{21}	**0.505**	0.398	0.622	**0.864**	0.694	0.946
β_{51}	**0.179**	0.075	0.305	0.148	−0.275	0.841
β_{71}	**0.195**	0.051	0.345	−0.114	−0.721	0.594
β_{32}	**0.557**	0.459	0.663	**0.872**	0.707	0.951
β_{42}	0.051	−0.085	0.235	−0.051	−0.694	0.644
β_{52}	0.064	−0.040	0.157	0.036	−0.471	0.504
β_{43}	**0.557**	0.385	0.691	0.721	−0.012	1.291
β_{53}	**0.513**	0.376	0.626	0.667	−0.082	1.095
β_{54}	**0.192**	0.089	0.304	0.177	−0.029	0.311
β_{65}	**0.526**	0.405	0.632	**0.594**	0.472	0.690
β_{75}	**0.483**	0.314	0.633	**0.983**	0.208	1.517
β_{76}	0.071	−0.040	0.194	−0.036	−0.175	0.114

Bold printed values are the coefficient estimates whose corresponding CI does not cover the zero
Percentile confidence intervals are calculated at a 95% confidence level
[a] based on 500 bootstrap samples
[b] based on 483 bootstrap samples

Table 6.14 Factor loading and confidence interval estimates based on the correlation matrix

	PLS	CI[a]		PLSc	CI[b]	
ima1	**0.743**	0.636	0.807	**0.612**	0.439	0.746
ima2	**0.601**	0.474	0.696	**0.538**	0.375	0.657
ima3	**0.578**	0.438	0.686	**0.451**	0.305	0.581
ima4	**0.768**	0.665	0.838	**0.673**	0.515	0.797
ima5	**0.744**	0.675	0.802	**0.667**	0.576	0.776
exp1	**0.771**	0.663	0.845	**0.511**	0.365	0.673
exp2	**0.687**	0.454	0.818	**0.458**	0.303	0.644
exp3	**0.612**	0.450	0.746	**0.440**	0.317	0.572
qua1	**0.803**	0.752	0.847	**0.807**	0.748	0.879
qua2	**0.637**	0.517	0.734	**0.542**	0.419	0.651
qua3	**0.784**	0.714	0.833	**0.753**	0.669	0.831
qua4	**0.769**	0.674	0.849	**0.673**	0.570	0.779
qua5	**0.756**	0.674	0.821	**0.683**	0.560	0.783
qua6	**0.775**	0.629	0.868	**0.682**	0.551	0.811
qua7	**0.779**	0.705	0.836	**0.810**	0.740	0.877
val1	**0.904**	0.856	0.936	**0.754**	0.661	0.856
val2	**0.938**	0.922	0.951	**0.928**	0.860	0.986
sat1	**0.799**	0.744	0.849	**0.693**	0.616	0.778
sat2	**0.846**	0.798	0.885	**0.701**	0.630	0.773
sat3	**0.852**	0.813	0.886	**0.810**	0.754	0.880
loy1	**0.814**	0.701	0.877	**0.594**	0.486	0.721
loy2	**0.219**	0.010	0.413	0.173	−0.013	0.301
loy3	**0.917**	0.890	0.937	**0.869**	0.801	0.971

Bold printed values are the coefficient estimates where the corresponding CI does not cover the zero
Percentile confidence intervals are calculated at a 95% confidence level
[a] based on 500 bootstrap samples
[b] based on 483 bootstrap samples

Table 6.15 Average variance extracted and shared variance estimates based on the correlation matrix

(a) PLS

	IMA	EXP	QUA	VAL	SAT	LOY
IMAG	**0.478**	0.255	0.561	0.258	0.480	0.318
EXPE	0.505	**0.480**	0.311	0.131	0.260	0.144
QUAL	0.749	0.557	**0.577**	0.343	0.632	0.289
VAL	0.508	0.361	0.586	**0.849**	0.367	0.281
SAT	0.693	0.510	0.795	0.606	**0.693**	0.431
LOY	0.564	0.380	0.538	0.530	0.656	**0.517**

(b) PLSc

	IMA	EXP	QUA	VAL	SAT	LOY
IMAG	**0.353**	0.746	0.858	0.412	0.826	0.577
EXPE	0.864	**0.221**	0.761	0.333	0.716	0.419
QUAL	0.926	0.872	**0.507**	0.457	0.910	0.438
VAL	0.642	0.577	0.676	**0.715**	0.551	0.443
SAT	0.909	0.846	0.954	0.742	**0.542**	0.736
LOY	0.760	0.647	0.662	0.666	0.858	**0.379**

Correlations are below the diagonal, squared correlations are above the diagonal, and AVE estimates are presented on the diagonal (in boldface)

Table 6.16 Internal consistency reliability using the correlation matrix

(a) Dijkstra–Henseler's ρ_A		(b) Dillon–Goldstein's ρ_c		
	PLS(c)		PLS	PLSc
IMA	**0.740**	IMA	**0.819**	**0.720**
EXP	0.462	EXP	**0.733**	0.459
QUA	**0.884**	QUA	**0.905**	**0.877**
VAL	**0.849**	VAL	**0.918**	**0.833**
SAT	**0.785**	SAT	**0.871**	**0.780**
LOY	**0.746**	LOY	**0.724**	0.590

Values larger than 0.7 are printed in boldface. Results for COM are not reported, as it is measured by only one indicator

Table 6.17 R^2 of the endogenous composites

	PLS	PLSc
IMA	0.000	0.000
EXP	0.255	0.746
QUA	0.311	0.761
VAL	0.345	0.457
SAT	0.680	0.931
COM	0.277	0.353
LOY	0.457	0.739

Table 6.18 Cross-loadings PLS based on the correlation matrix

	IMAG	EXPE	QUAL	VAL	SAT	COMP	LOY
ima1	**0.743**	0.350	0.571	0.396	0.549	0.423	0.354
ima2	**0.601**	0.380	0.497	0.274	0.418	0.188	0.304
ima3	**0.578**	0.284	0.368	0.338	0.331	0.207	0.308
ima4	**0.768**	0.370	0.573	0.475	0.548	0.440	0.460
ima5	**0.744**	0.359	0.552	0.271	0.514	0.337	0.493
exp1	0.352	**0.771**	0.436	0.294	0.372	0.183	0.271
exp2	0.408	**0.687**	0.348	0.179	0.366	0.225	0.321
exp3	0.288	**0.612**	0.370	0.274	0.320	0.126	0.196
qua1	0.634	0.514	**0.803**	0.469	0.680	0.380	0.476
qua2	0.430	0.319	**0.637**	0.307	0.490	0.300	0.342
qua3	0.628	0.434	**0.784**	0.473	0.645	0.472	0.473
qua4	0.496	0.391	**0.769**	0.395	0.601	0.379	0.368
qua5	0.609	0.419	**0.756**	0.464	0.524	0.389	0.377
qua6	0.568	0.445	**0.775**	0.410	0.549	0.418	0.343
qua7	0.586	0.417	**0.779**	0.553	0.698	0.465	0.452
val1	0.395	0.312	0.474	**0.904**	0.486	0.287	0.431
val2	0.530	0.351	0.595	**0.938**	0.619	0.360	0.536
sat1	0.577	0.492	0.642	0.411	**0.799**	0.334	0.504
sat2	0.523	0.398	0.670	0.491	**0.846**	0.416	0.497
sat3	0.624	0.391	0.673	0.598	**0.852**	0.547	0.627
loy1	0.434	0.294	0.394	0.413	0.455	0.237	**0.814**
loy2	0.100	0.093	0.063	0.139	0.107	0.122	**0.219**
loy3	0.539	0.356	0.534	0.494	0.663	0.448	**0.917**

The largest loading of each indicator is printed in boldface

Table 6.19 Cross-loadings PLSc based on the correlation matrix

	IMAG	EXPE	QUAL	VAL	SAT	COMP	LOY
ima1	0.612	0.515	0.607	0.430	**0.619**	0.423	0.410
ima2	0.538	**0.560**	0.529	0.298	0.472	0.188	0.352
ima3	**0.451**	0.418	0.391	0.367	0.374	0.207	0.357
ima4	**0.673**	0.545	0.610	0.516	0.619	0.440	0.532
ima5	**0.667**	0.529	0.587	0.294	0.580	0.337	0.571
exp1	0.410	**0.511**	0.464	0.319	0.420	0.183	0.313
exp2	**0.475**	0.458	0.370	0.195	0.413	0.225	0.372
exp3	0.334	**0.440**	0.393	0.298	0.361	0.126	0.227
qua1	0.738	0.756	**0.807**	0.509	0.768	0.380	0.551
qua2	0.500	0.470	0.542	0.333	**0.553**	0.300	0.396
qua3	0.730	0.639	**0.753**	0.514	0.727	0.472	0.548
qua4	0.577	0.575	0.673	0.428	**0.678**	0.379	0.426
qua5	**0.708**	0.617	0.683	0.504	0.591	0.389	0.437
qua6	0.660	0.655	**0.682**	0.445	0.620	0.418	0.397
qua7	0.681	0.613	**0.810**	0.600	0.788	0.465	0.523
val1	0.459	0.458	0.504	**0.754**	0.549	0.287	0.499
val2	0.616	0.517	0.632	**0.928**	0.698	0.360	0.620
sat1	0.671	**0.724**	0.683	0.446	0.693	0.334	0.584
sat2	0.608	0.585	**0.712**	0.533	0.701	0.416	0.575
sat3	0.726	0.575	0.716	0.649	**0.810**	0.547	0.726
loy1	0.505	0.433	0.419	0.448	0.514	0.237	**0.594**
loy2	0.116	0.136	0.067	0.151	0.120	0.122	**0.173**
loy3	0.627	0.525	0.567	0.536	0.748	0.448	**0.869**

The bold printed value in each row shows on which construct the indicator loads most

References

Bayol, M.P., de la Foye, A., Tellier, C., & Tenenhaus, M. (2000). Use of PLS path modeling to estimate the European consumer satisfaction index (ECSI) model. *Statistica Applicata, 12*(3), 361–375.

Beran, R., & Srivastava, M. S. (1985). Bootstrap tests and confidence regions for functions of a covariance matrix. *The Annals of Statistics, 13*(1), 95–115.

Betzin, J., & Henseler, J. (2005). Looking at the antecedents of perceived switching costs. A PLS path modeling approach with categorical indicators. In *Proceedings of the PLS'05 International Symposium*, Barcelona.

Bollen, K. A., & Bauldry, S. (2011). Three Cs in measurement models: Causal indicators, composite indicators, and covariates. *Psychological Methods, 16*(3), 265.

Bollen, K. A., & Stine, R. A. (1992). Bootstrapping goodness-of-fit measures in structural equation models. *Sociological Methods & Research, 21*(2), 205–229. doi:10.1177/0049124192021002004.

Cantaluppi, G. (2012). A partial least squares algorithm handling ordinal variables also in presence of a small number of categories. arXiv preprint arXiv:12125049.

Cantaluppi, G., & Boari, G. (2016). A partial least squares algorithm handling ordinal variables. In H. Abdi, V. E. Vinzi, G. Russolillo, G. Saporta & L. Trinchera (Eds.), *The multiple facets* *of partial least squares and related methods: PLS*, Paris, 2014 (pp. 295–306). Cham: Springer International Publishing.

Carroll, J. B. (1961). The nature of the data, or how to choose a correlation coefficient. *Psychometrika, 26*(4), 347–372.

Chin, W. W. (1998). The partial least squares approach to structural equation modeling. In G. A. Marcoulides (Ed.), *Modern methods for business research* (pp. 295–336). London: Lawrence Erlbaum Associates.

Coelho, P. S., & Esteves, S. P. (2007). The choice between a five-point and a ten-point scale in the framework of customer satisfaction measurement. *International Journal of Market Research, 49*(3), 313–339.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2013). *Applied multiple regression/correlation analysis for the behavioral sciences*. New York: Routledge.

Dijkstra, T. K. (1983). Some comments on maximum likelihood and partial least squares methods. *Journal of Econometrics, 22*(1–2), 67–90.

Dijkstra, T. K. (1985). *Latent variables in linear stochastic models: Reflections on "maximum likelihood" and "partial least squares" methods* (Vol. 1). Amsterdam: Sociometric Research Foundation.

Dijkstra, T. K. (2010). Latent variables and indices: Herman Wold's basic design and partial least squares. In V. E. Vinzi, W. W. Chin, J. Henseler & H. Wang (Eds.), *Handbook of partial least squares* (pp. 23–46). Berlin: Springer.

Dijkstra, T. K. (2011). *Consistent partial least squares estimators for linear and polynomial factor models*. Technical Report. Research Gate. doi:10.13140/RG.2.1.3997.0405.

Dijkstra, T. K., & Henseler, J. (2015a). Consistent and asymptotically normal PLS estimators for linear structural equations. *Computational Statistics and Data Analysis, 81*(1), 10–23. doi:10.1016/j.csda.2014.07.008.

Dijkstra, T. K., & Henseler, J. (2015b). Consistent partial least squares path modeling. *MIS Quarterly, 39*(2), 297–316.

Dijkstra, T. K., & Schermelleh-Engel, K. (2014). Consistent partial least squares for nonlinear structural equation models. *Psychometrika, 79*(4), 585–604. doi:10.1007/S11336-013-9370-0.

Drasgow, F. (1986). Polychoric and polyserial correlations. In S. Kotz & N. Johnson (Eds.), *The encyclopedia of statistics* (Vol. 7, pp. 68–74). New York: Wiley.

Farrell, A. M. (2010). Insufficient discriminant validity: A comment on Bove, Pervan, Beatty, and Shiu (2009). *Journal of Business Research, 63*(3), 324–327.

Fornell, C. (1992). A national customer satisfaction barometer: The Swedish experience. *The Journal of Marketing, 56*(1), 6–21.

Fornell, C., & Cha, J. (1994). Partial least squares. In R. Bagozzi (Ed.), *Advanced methods of marketing research* (pp. 52–78). Oxford: Blackwell.

Fornell, C., Johnson, M. D., Anderson, E. W., Cha, J., & Bryant, B. E. (1996). The American customer satisfaction index: Nature, purpose, and findings. *The Journal of Marketing, 60*(4), 7–18.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*(1), 39–50. doi:10.2307/3151312.

Gelhard, C., & Von Delft, S. (2016). The role of organizational capabilities in achieving superior sustainability performance. *Journal of Business Research, 69*(10), 4632–4642 (Early Online Publication).

Gold, A. H., & Arvind Malhotra, A. H. S. (2001). Knowledge management: An organizational capabilities perspective. *Journal of Management Information Systems, 18*(1), 185–214.

Hair, J. F., Sarstedt, M., Pieper, T. M., & Ringle, C. M. (2012a). The use of partial least squares structural equation modeling in strategic management research: A review of past practices and recommendations for future applications. *Long Range Planning, 45*(5), 320–340. doi:10.1016/j.lrp.2012.09.008.

Hair, J. F., Sarstedt, M., Ringle, C. M., & Mena, J. A. (2012b). An assessment of the use of partial least squares structural equation modeling in marketing research. *Journal of the Academy of Marketing Science, 40*(3), 414–433. doi:10.1007/s11747-011-0261-6.

Henseler, J., & Dijkstra, T. K. (2015a). ADANCO 2.0. composite modeling GmbH, Kleve. www.composite-modeling.com.

Henseler, J., Hubona, G., & Ray, P. A. (2016a). Using PLS path modeling in new technology research: Updated guidelines. *Industrial Management & Data Systems, 116*(1), 2–20.

Henseler, J., Ringle, C. M., & Sarstedt, M. (2015b). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science, 43*(1), 115–135. doi:10.1007/s11747-014-0403-8.

Henseler, J., Ringle, C. M., & Sarstedt, M. (2016b). Testing measurement invariance of composites using partial least squares. *International Marketing Review, 33*(3), 405–431. doi:10.1108/IMR-09-2014-0304.

Hu, L., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods 3(4)*, 424–453. doi:10.1037/1082-989X.3.4.424.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1–55. doi:10.1080/10705519909540118.

Hwang, H., & Takane, Y. (2014). *Generalized structured component analysis: A component-based approach to structural equation modeling*. Boca Raton: CRC Press.

Jakobowicz, E., & Derquenne, C. (2007). A modified PLS path modeling algorithm handling reflective categorical variables and a new model building strategy. *Computational Statistics and Data Analysis, 51*(8), 3666–3678.

Johnson, M. D, Gustafsson, A., Andreassen, T. W, & Lervik, L., & Cha, J. (2001). The evolution and future of national customer satisfaction index models. *Journal of Economic Psychology, 22*(2), 217–245.

Kettenring, J. R. (1971). Canonical analysis of several sets of variables. *Biometrika, 58*(3), 433–451.

Ketterlinus, R. D., Bookstein, F. L., Sampson, P. D., & Lamb, M. E. (1989). Partial least squares analysis in developmental psychopathology. *Development and Psychopathology, 1*(4), 351–371.

Lee, S. Y., & Poon, W. Y. (1987). Two-step estimation of multivariate polychoric correlation. *Communications in Statistics - Theory and Methods, 16*(2), 307–320.

Lohmöller, J. B. (1989). *Latent variable path modeling with partial least squares*. Berlin: Springer.

Maraun, M. D., & Halpin, P. F. (2008). Manifest and latent variates. *Measurement: Interdisciplinary Research and Perspectives, 6*(1–2). 113–117.

Noonan, R., & Wold, H. (1982). PLS path modeling with indirectly observed variables: A comparison of alternative estimates for the latent variable. In K. G. Jöreskog & H. Wold (Eds.), *Systems under indirect observation: Causality, structure, prediction Part II* (Chap. 3, pp. 75–94). Amsterdam: North-Holland.

O'Brien, R. M., & Homer, P. (1987). Corrections for coarsely categorized measures: LISREL's polyserial and polychoric correlations. *Quality and Quantity, 21*(4), 349–360.

Olsson, U. (1979). Maximum likelihood estimation of the polychoric correlation coefficient. *Psychometrika, 44*(4), 443–460.

Olsson, U., Drasgow, F., & Dorans, N. J. (1982). The polyserial correlation coefficient. *Psychometrika, 47*(3), 337–347.

Pearson, K. (1900). Mathematical contributions to the theory of evolution. VII. on the correlation of characters not quantitatively measurable. *Philosophical Transactions of the Royal Society of London, Series A, 195*, 1–405 (Containing Papers of a Mathematical or Physical Character).

Pearson, K. (1913). On the measurement of the influence of "broad categories" on correlation. *Biometrika, 9*(1/2), 116–139.

Poon, W. Y., & Lee, S. Y. (1987). Maximum likelihood estimation of multivariate polyserial and polychoric correlation coefficients. *Psychometrika, 52*(3), 409–430.

Quiroga, A. M. (1992). *Studies of the polychoric correlation and other correlation measures for ordinal variables*. PhD thesis, Universitatis Upsaliensis.

R Core Team. (2016). R: A language and environment for statistical computing. R foundation for statistical computing, Vienna. https://www.R-project.org/.

Raykov, T. (2004). Behavioral scale reliability and measurement invariance evaluation using latent variable modeling. *Behavior Therapy, 35*(2), 299–331.

Revelle, W. (2016). Psych: Procedures for psychological, psychometric, and personality research. Northwestern University, Evanston, IL. http://CRAN.R-project.org/package=psych. R package version 1.6.4.

Rigdon, E. E. (2012). Rethinking partial least squares path modeling: In praise of simple methods. *Long Range Planning, 45*(5), 341–358. doi:10.1016/j.lrp.2012.09.010.

Rigdon, E. E. (2016). Choosing PLS path modeling as analytical method in European management research: A realist perspective. *European Management Journal, 34*(6), 598–605. doi:10.1016/j.emj.2016.05.006.

Ringle, C. M., Sarstedt, M., & Straub, D. (2012). A critical look at the use of PLS-SEM in MIS Quarterly. *MIS Quarterly, 36*(1), 3–14.

Rönkkö, M. (2016), matrixpls: Matrix-based partial least squares estimation. https://github.com/mronkko/matrixpls. R package version 1.0.1.

Russolillo, G. (2012). Non-metric partial least squares. *Electronic Journal of Statistics, 6*, 1641–1669.

Sanchez, G., Trinchera, L., & Russolillo, G. (2015), plspm: Tools for partial least squares path modeling (PLS-PM). https://CRAN.R-project.org/package=plspm. R package version 0.4.7.

Savalei, V. (2011). What to do about zero frequency cells when estimating polychoric correlations. *Structural Equation Modeling: A Multidisciplinary Journal, 18*(2), 253–273.

Schneeweiss, H. (1993). Consistency at large in models with latent variables. In K. Haagen, D. J. Bartholomew, M. Deistler (Eds.), *Statistical modelling and latent variables* (pp. 299–322). Amsterdam: Elsevier.

Schuberth, F., Henseler, J., & Dijkstra, T. K. (2016). Partial least squares path modeling using ordinal categorical indicators. *Quality and Quantity, 51*, 1–27. doi:10.1007/s11135-016-0401-7.

Shmueli, G., Ray, S., Estrada, J. M. V., & Chatla, S. B. (2016). The elephant in the room: Predictive performance of PLS models. *Journal of Business Research, 69*(10), 4552–4564.

Tenenhaus, M. (2008). Component-based structural equation modelling. *Total Quality Management, 19*(7–8), 871–886. doi:10.1080/14783360802159543.

Tenenhaus, M., Vinzi, V. E., Chatelin, Y. M., & Lauro, C. (2005). PLS path modeling. *Computational Statistics and Data Analysis, 48*(1), 159–205. doi:10.1016/j.csda.2004.03.005.

Trinchera, L., & Russolillo, G. (2010). *On the use of structural equation models and PLS path modeling to build composite indicators*. Working Papers 30–2010, Macerata University, Department of Studies on Economic Development (DiSSE).

Wickham, H. (2009). ggplot2: Elegant graphics for data analysis. New York: Springer. http://had.co.nz/ggplot2/book

Wold, H. (1975). Path models with latent variables: The NIPALS approach. In H. Blalock, A. Aganbegian, F. Borodkin, R. Boudon & V. Capecchi (Eds.), *Quantitative sociology, international perspectives on mathematical and statistical modeling* (Chap 11, pp. 307–357). New York: Academic.

Wold, H. (1982a). Models for knowledge. In J Gani (Ed.), *The making of statisticians* (pp 189–212). New York: Springer.

Wold, H. (1982b). Soft modeling: The basic design and some extensions. In K. G. Jöreskog & H. Wold (Eds.), *Systems under indirect observations, Part II* (pp. 1–54). Amsterdam: North-Holland.

Wooldridge, J. (2012). *Introductory econometrics: A modern approach*. Boston: Cengage Learning.

Wylie, P. B. (1976). Effects of coarse grouping and skewed marginal distributions on the Pearson product moment correlation coefficient. *Educational and Psychological Measurement, 36*(1), 1–7.

Yule, G. U. (1900). On the association of attributes in statistics: With illustrations from the material of the childhood society, &c. *Philosophical Transactions of the Royal Society of London, Series A 194*, 257–319 (Containing Papers of a Mathematical or Physical Character).

Zumbo, B. D., Gadermann, A. M., & Zeisser, C. (2007). Ordinal versions of coefficients alpha and theta for likert rating scales. *Journal of Modern Applied Statistical Methods, 6*(1), 21–29.

Part II
Methodological Issues

Chapter 7
Predictive Path Modeling Through PLS and Other Component-Based Approaches: Methodological Issues and Performance Evaluation

Pasquale Dolce, Vincenzo Esposito Vinzi, and Carlo Lauro

Abstract This chapter deals with the predictive use of PLS-PM and related component-based methods in an attempt to contribute to the recent debates on the suitability of PLS-PM for predictive purposes. Appropriate measures and evaluation criteria for the assessment of models in terms of predictive ability are more and more desirable in PLS-PM. The performance of the models can be improved by choosing the appropriate parameter estimation procedure among the different existing ones or by making developments and modifications of the latter. A recent example of this type of work is the non-symmetrical approach for component-based path modeling, which leads to a new method, called non-symmetrical composite-based path modeling. In the composites construction stage, this new method explicitly takes into account the directions of the relationships in the inner model. Results are promising for this new method, especially in terms of predictive relevance.

7.1 Introduction

Partial least square path modeling (PLS-PM) is certainly an important technique that deserves a prominent place in research applications when the aim of the analysis is prediction. In fact, the predictive ability of PLS-PM is more and more of interest (Becker et al. 2013; Evermann and Tate 2016; Shmueli et al. 2016; Ringle et al. 2012; Hair et al. 2012).

P. Dolce (✉)
Oniris, StatSC, Nantes, France
e-mail: pasquale.dolce@unina.it

V. Esposito Vinzi
ESSEC Business School, Cergy-Pontoise, France
e-mail: vinzi@essec.edu

C. Lauro
University of Naples "Federico II", Naples, Italy
e-mail: clauro@unina.it

H. Latan, R. Noonan (eds.), *Partial Least Squares Path Modeling*,
DOI 10.1007/978-3-319-64069-3_7

However, despite the predictive aim of many PLS-PM studies, most do not provide appropriate predictive ability metrics. The abilities of PLS-PM have not been fully explored and appreciated. As noted also by Rigdon (2012), 'researchers applying PLS path modeling often assert the "predictive" nature of their research, though researchers often seem to mean nothing more than aiming to maximise R^2 for dependent variables'. Moreover, in most cases, the main results produced by studies applying PLS-PM are statistical tests for evaluating the precision of PLS-PM estimates (mainly applying the bootstrap approach).

Improving the explanatory power of a model and reproducing model parameters do not imply good predictions about individual observations. Prediction modeling has a different goal in the analysis, one that cannot be achieved through explanatory measures.

A confusion between explanatory and predictive modeling can be the source of misunderstandings and misapplications of the component-based path modeling (Sarstedt et al. 2014). Only recently have researchers begun to discuss the suitability of PLS-PM for predictive purposes more accurately, clarifying what is really meant by prediction and what are the proper criteria for measuring PLS-PM predictive ability (Evermann and Tate 2016; Shmueli et al. 2016).

Predictive and explanatory modeling are two separate statistical concepts, each with different purposes and separate practical implications (Shmueli 2010; Shmueli and Koppius 2011). Explanatory modeling aims at describing a statistical model and testing theoretical hypotheses about the relationships between variables. Thus, in explanatory modeling, we want to evaluate the strength of the relationships between variables and the explanatory power of a model by using statistical tests on the parameters and explanatory measures (like the R^2) (Shmueli and Koppius 2011).

In the context of predictive modeling, the effect sizes between variables are not of primary importance. Instead, the goal is to make a good prediction of the manifest variable (MV) values of the dependent blocks from the MV values of predictor blocks. Consequently, the predictive performance of a model should be evaluated using appropriate measures, which refer to the ability of the model to make highly accurate predictions for dependent variables' values and new observations.

Evaluating the predictive performance of a model should be not of interest if the focus of the research is not prediction. In particular, there are several situations where researchers want to use component-based path modeling only for explanation and for reproducing model parameters, without evaluating out-of-sample predictive ability. In such a case, testing loadings, path coefficients and the explanatory power (using R^2, for example) can be can be largely sufficient.

However, it should be noted that even if explanation and prediction are two separate goals for the analysis, there are situations where the question of the research concerns both aspects; therefore, explanation and prediction can exist with each other.

As noted by Shmueli (2010), rather than considering explanation and prediction as extremes on a continuum, they can be seen as two different dimensions. A model possesses some level of explanatory power and some level predictive accuracy.

Consequently, the goal of the analysis will help in the choice of the appropriate method (Hair et al. 2017), but the performance of the model should be evaluated considering both explanatory and predictive qualities.

When prediction is not the goal, evaluating both the predictive performance of a model and the explanatory power may help when comparing different models. On the contrary, when prediction is the goal, considering explanation quality in the analysis helps in theory building.

Shmueli (2010) mentions some examples where the combination of explanatory and predictive modeling may be important, especially for bridging the gap between methodological development and practical application.

In component-based predictive path modeling, the quality of the measurement model and the structural model, in terms of how the variables' relationships are explained and how much variance is explained by the components, is of interest but should not be the sole objective if prediction is also if interest.

Certainly, parameter recovery is a very important aspect in PLS-PM and should be considered. Further studies in this direction are necessary to validate the component-based approach to structural equation modeling.[1]

However, because PLS-PM is also an oriented-prediction method, true predictive ability measures should be included in the evaluation criteria, and extensions and developments of further measures and evaluation criteria for the assessment of PLS-PM are desirable (Sarstedt et al. 2014). Starting from the criterion of interest, further extensions and modifications can be made on the basic PLS-PM algorithm to improve the predictive capabilities of the model.

The non-symmetrical approach for component-based path modeling proposed by Dolce (2015) is an example of work in this direction. The authors proposed a new algorithm that replaces the original PLS-PM iterative algorithm, with the desire to follow the directions of the relationships specified in the inner model and improve predictive ability. This new method, called non-symmetrical composite-based path modeling (NSCPM), seems to be very promising when one is interested in predictive power, as well as in understanding the relative importance of predictors on endogenous components and their own manifest variables.

In the next section, we will discuss the issue of prediction in component-based path modeling. In Sect. 7.3, we will look at the different ways for computing the outer weights in PLS-PM and how they affect the predictive ability of PLS-PM. Section 7.4 is dedicated to a discussion of the appropriate predictive ability metrics. Section 7.5 discusses the 'path directions incoherence' in PLS-PM, and a comparison between NSCPM with the classical PLS-PM is provided. Finally, Sect. 7.6 offers some general comments on predictive relevance in component-based path modeling and, in particular, in NSCPM.

[1]See, for example, the recent paper by Hair et al. (2017), which gives also some guidelines for choosing among composite-based approaches.

7.2 Predictive Ability in Component-Based Path Modeling

In the context of predictive-oriented PLS-PM, the investigator should focus on model performance in terms of prediction, without discarding the hypothesised relationships in the structural model, ones that also imply the predictive directions of the model. It is in this latter perspective that Lohmöller refers to 'predictor specification' in his thesis (Lohmöller 1989), stressing the relationship between a faithful representation of causal mechanisms underlying a phenomenon by the structural equations and the purposes of the prediction.

If PLS-PM were merely used either for testing causal mechanisms underlying a phenomenon or to develop a predictive model, then its use might be dubious. For the first case, if the theoretical model is correct and the standard assumptions underlying the factor-based approach are satisfied, covariance-based SEM (Jöreskog 1977) may do a better job. PLS-PM is considered also an alternative to covariance-based SEM because there are many situations where the assumptions of factor models are not fulfilled, even though this position has been harshly criticised in some academic literature, e.g., Rönkkö and Evermann (2013); Rönkkö et al. (2016); Rigdon (2012). One of the most common criticisms about PLS-PM is the lack of consistency. However, as recently noted by some authors (see, for example, Sarstedt et al. 2016; Rigdon 2016), PLS-PM estimators turned out to be inconsistent because simulations evaluate PLS-PM estimators using samples drawn from populations consistent with common factor models. The recent PLS-PM literature (e.g. Rigdon 2016; Sarstedt et al. 2016; Henseler et al. 2016; Becker et al. 2013; Henseler et al. 2014) provides an unambiguous distinction between the composite-based approach and common factor-based approach. The composite-based model is becoming more recognised as the reference model of the PLS-PM and has been recently used as the 'correct population' in simulation studies. When a correct population is used in simulation studies (i.e. composite-based populations), PLS-PM estimates are consistent. However, a number of works by Dijkstra and Henseler (see, for example, Dijkstra and Henseler 2015a,b) showed that the PLS-PM algorithm yields all the ingredients for obtaining CAN (consistent and asymptotically normal) estimations of loadings and latent variables squared correlations of the factor-based model.

As for the second justification behind using PLS-PM for developing a predictive model, it should be noted that modern prediction methods (Hastie et al. 2009) may perform better in terms of predictive capability. Undoubtedly, a drawback of data-driven methods is that the weighting criterion may not necessarily reflect the theoretical importance of all the components, and given that the weights are obtained from the data, they may not be constant over time and space.

For all these reasons, it is extremely important for the investigators to provide a clear statement of the position of their research and the objectives of the specific analysis.

Even though predictive-oriented PLS-PM is mainly a data analysis approach and a data-driven component method, the development of the models should not be a-theoretical if one wants to remain in the context of component-based path modeling (which is commonly used in real applications). It is worth bearing in mind that predictions in PLS-PM should be sensitive to the theory and that the theoretical model represented by the structural equations and prediction should not be separated (Lohmöller 1989). In developing a predictive-oriented component-based path modeling, the structural model specification should derive from theory while one can improve the predictive performance of the model by choosing the appropriate parameter estimation procedure among the different existing ones or by making developments and modifications of the latter. Dolce (2015); Shmueli et al. (2016) are examples of works in this direction.

7.3 Outer Weights Computation in PLS-PM

The literature on PLS-PM offers two main modes in the outer estimation step for computing the outer weights, which are known as *Mode A* and *Mode B* (Lohmöller 1989). By using *Mode A*, each MV is regressed on the corresponding instrumental component in the outer estimation step. *Mode B* applies multiple linear regression of the inner component on the corresponding MVs. Hence, *Mode B* takes into account both the MV-component correlation and the within-block intercorrelations. On the contrary, *Mode A* ignores correlations among MVs. However, because PLS-PM is a component-based method, components are computed as weighted aggregates of their MVs, whatever outer mode is applied.

Although *Mode B* and *Mode A* were originally presented as two different ways for computing outer weights (Wold 1982), researchers have traditionally associated the two modes with two different measurement model specifications, based essentially on the hypothesised relationships between the components and their own MVs. The PLS-PM literature has long indicated that the MV's weights in blocks defined as reflective (Fornell and Bookstein 1982), or outwards directed (Lohmöller 1989), are to be estimated using Mode A while MVs weights in blocks defined as inwards directed or formative are to be estimated using Mode B (e.g. Chin 1998; Hair et al. 2014; Esposito 2013; Dolce and Lauro 2015). Under conditions of low theoretical knowledge on the nature of the constructs, a rule of thumb in PLS-PM is to apply *Mode B* to the exogenous block and *Mode A* to the endogenous block (Wold 1980, 1982).

Consequently, there has been always a confusion between the two modes to compute outer weights (i.e. *Mode A* and *Mode B*) and the two theoretical measurement models (i.e. reflective and formative) (Rigdon 2016; Rönkkö et al. 2016).

Only recently, have researchers started to clarify this argument, recognising that the outer mode and the measurement model are separate in PLS-PM, and this association may be just an illusion (Rigdon 2016; Rönkkö et al. 2016; Sarstedt et al. 2016; Henseler et al. 2016; Becker et al. 2013). Actually, *Mode A* and *Mode B* are just two different ways for computing the outer weights, which are used as a vehicle for the computing components. The choice between the two modes goes beyond the specified measurement model and requires a more thoughtful approach (Sarstedt et al. 2016; Becker et al. 2013).

The important point here is that PLS-PM optimisation criteria change depending on the way the outer weights are calculated (Krämer 2007; Hanafi 2007; Tenenhaus and Tenenhaus 2011). When all weights are computed using *Mode B*, the solution can be characterised by a sum of absolute values of correlations between components. *Mode A* applied to all the blocks does not lead to a stationary equation related to the optimisation of a twice differentiable function (Krämer 2007). However, a modified version of Mode A, in which the outer weights rather than the components are normalised to unitary variance at each step of the algorithm, if applied to all blocks maximises the covariances between components (Krämer 2007; Tenenhaus and Tenenhaus 2011); this takes into account the component variances too. We believe that researchers applying PLS-PM should focus more on these latter findings than the distinction between reflective and formative measurement models.

Moreover, the choice between the two outer modes has certainly an effect in terms of the predictive ability of the model (Becker et al. 2013) and on how the predictive directions between components in the structural model are considered (Dolce et al. 2015). The only way to give an explanatory role to a block of manifest variables is to apply *Mode B*. Applying *Mode A* gives it a dependent role, whatever the path direction may be.

7.3.1 The Effects of the Different Outer Weighting Schemes on PLS-PM Predictive Ability

Mode A actually indicates that correlation weights are used to compute the outer weights. On the other hand, *Mode B* indicates that multiple regression weights are used. Therefore, the two different outer modes within PLS-PM methodology certainly lead models with different predictive.

Dana and Dawes (2004) demonstrated in the context of conventional regression that correlation weights (which ignore collinearity among the predictors) outperform multiple regression weights for out-of-sample predictions, unless the sample size is very large. For this reasons, the authors suggest avoiding using multiple regression weights for out-of-sample predictions. Dana and Dawes (2004) have also demonstrated that out-of-sample predictive ability depends on sample size.

As noted by Becker et al. (2013) and Rigdon (2012), Dana and Dawes's suggestions would translate into an advantage for the *Mode A* estimation of outer weights (which corresponds to the use of correlation weights) over *Mode B*. However, further studies are necessary to examine this issue further.

Within the literature on forecasting, it has been established that when the objective is to make as good a forecast as possible, then combinations of forecasts can yield improvements in terms of prediction compared to single forecasts (Bates and Granger 1969; Makridakis and Hibon 2000; Armstrong 2001) because each forecast nearly always contains some useful independent information. In this perspective, multiple-indicator approaches should have an advantage in prediction over single-indicator methods. Moreover, multiple linear regression adjusts for multicollinearity and gives less weight to predictors that are more redundant. Hence, *Mode B* in PLS-PM is certainly consistent for forecasting purposes (Becker et al. 2013).

In general, *Mode B* in PLS-PM produces higher R^2 in the structural model, providing more accurate multi-component prediction for individual endogenous component observations while *Mode A* produces higher loadings, leading to better multi-component individual observation predictions of MVs.

Becker et al. (2013) considered the sample size as an experimental condition; the authors conducted simulation studies aimed at analysing the out-of-sample prediction capability of PLS-PM. The results showed that if the criterion is out-of-sample predictive ability, PLS-PM performs poorly when the sample size is small. Sample sizes that would be adequate for the estimation of the parameters of the model may be highly inadequate for out-of-sample predictions.

7.3.2 PLS Regression for Outer Model Regularisation and Multi-component Estimation

PLS-PM has been recently criticised because it 'typically forms just one composite for each set of indicators, constraining the relationship between each set of predictors to be unidimensional' (Rigdon 2012, p. 354). The unidimensionality assumption may be very restrictive and limiting in several applications. Furthermore, unidimensional relationships may limit drastically the predictive ability of the model because most of the important information in the predictor blocks may be left out.

From a predictive point of view, the performance of PLS-PM can be certainly improved by extracting more than one component in each block of MVs. Certainly, this point of view severs all ties with the concept of one common factor, while putting emphasis on prediction ability. Consequently, multi-components PLS-PM conforms well when the goal is predicting, but it cannot be used to mimic the factor-based structural equation modeling.

Some proposals in this direction have already been introduced (Apel and Wold 1982; Höskuldsson 2009; Lohmöller 1989). An interesting approach was proposed by Esposito et al. (2010). The authors originally proposed this new approach as a solution to the issue of multicollinearity when applying *Mode B*. In fact, the problem of multicollinearity can be addressed by providing a PLS regression for estimating the outer weights, which would be an alternative to OLS regression. Moreover, by using this approach, we can extract more than one dimension for each block of MVs (the PLSR components) in alignment with a prediction purpose. The new approach, called Mode PLS, can be considered a fine-tuning between Mode A and Mode B, because it is based on the selection of a certain number of components of the PLS regression (Esposito et al. 2010) (Mode A corresponds to taking the first component from a PLS regression while Mode B corresponds to taking all the PLS regression components). This new mode is available in the PLS-PM module of the XLSTAT software.

7.4 Performance Metrics for Component-Based Predictive Path Modeling

When predictive relevance is of interest in component-based path modeling, one would be curious as to how the model performs in terms of prediction. It would be also interesting to compare different methods on their prediction performance for the same prediction problem; this would help show which method is the one to use for the specific real-data application. Nevertheless, which are the appropriate performance metrics to the specific prediction problem and real-data application?

The aim in predictive analysis is not to test whether the relationships among variables are significant, but instead to accurately predict values for individual cases. As noted by Shmueli et al. (2016), 'A statistically significant effect or relationship does not guarantee high predictive power, because the precision or magnitude of the causal effect might not be sufficient for obtaining levels of predictive accuracy that are practically meaningful'.

Prediction in composite-based path modeling could refer to the ability to predict the individual observations of components or individual observations for MVs of the endogenous blocks. Moreover, a distinction must be made between in-sample and out-of-sample predictions.

The R^2 in the structural model is an appropriate metric only for measuring the performance of the exogenous components in the endogenous components explanation. Instead, when prediction is to be made for individual observations of MVs of the endogenous blocks, redundancy-based measures are a more appropriate metric for assessing the in-sample predictive ability of the model.

However, metrics for the out-of-sample predictive power are needed to measure the performance of the model for prediction outside the dataset (on the data that has not been used during the model-building process). For this reason, a distinction between in-sample and out-of-sample predictions is necessary.

In this chapter, we focus on the predictive ability of the structural model because we want to take into account the predictive directions and the structural paths in the model and we want to maintain coherence with the theoretical structural model. To this purpose, redundancy-based prediction is the relevant criterion for assessing the predictive ability of the model (Chin 2010; Shmueli et al. 2016; Evermann and Tate 2016), as well as for comparing structural models. Communality-based measures are used if prediction of the individual cases is made through the components, but this is not the point of this study.

7.4.1 In-Sample Redundancy-Based measure

When investigators are interested in measuring the portion of the variance in the dependent blocks explained by their own predictors, they can use the redundancy index. Redundancy-based prediction is a relevant criterion for assessing the predictive ability of the structural model when a prediction is to be made for individual observations of MVs of the endogenous blocks (Chin 2010). In addition, the average of the redundancy indexes for each endogenous block can be also considered as an index of the goodness of fit of the global model (Lohmöller 1989).

Given two blocks of variables, $\boldsymbol{X}_1 = (\boldsymbol{x}_{11}, \ldots, \boldsymbol{x}_{P_1 1})$ and $\boldsymbol{X}_2 = (\boldsymbol{x}_{12}, \ldots, \boldsymbol{x}_{P_2 2})$, the redundancy index, as proposed by Stewart and Love (1968), measures the proportion of the variance in the dependent set $\boldsymbol{X}_2$ accounted for by the predictor set $\boldsymbol{X}_1$. The redundancy analysis model, proposed by Van den Wollenberg (1977), searches for the linear combination, $\hat{\xi}_1 = \boldsymbol{X}_1 \boldsymbol{w}_1$ (the first redundancy variate), which maximises the redundancy index, R_{X_2}, defined as

$$R_{X_2} = \sum_{p=1}^{P_2} \text{corr}(\hat{\boldsymbol{\xi}}_1, \boldsymbol{x}_{p2})^2 / P_2 \tag{7.1}$$

under the restriction that the variance of $\hat{\xi}_1 = 1$.

In the context of canonical correlation analysis (Hotelling 1935, 1936), the redundancy index [Eq. (7.1)] can be written as follows:

$$R_{X_2} = \rho^2 \sum_{p=1}^{P_2} \text{corr}(\hat{\boldsymbol{\xi}}_2, \boldsymbol{x}_{p2})^2 / P_2 \tag{7.2}$$

where ρ is the canonical correlation coefficient and $\boldsymbol{\xi}_2 = \boldsymbol{X}_2 \tilde{\boldsymbol{w}}_2$ is the first canonical component of $\boldsymbol{X}_2$ (Rencher 1998).

To define the redundancy index in the context of PLS-PM, assume that P variables are collected in a partitioned table of standardised data $\boldsymbol{X}$ in K blocks:

$$\boldsymbol{X} = [\boldsymbol{X}_1, \ldots, \boldsymbol{X}_J, \boldsymbol{X}_{J+1}, \ldots, \boldsymbol{X}_K],$$

where $\boldsymbol{X}_k$ $(k = 1, \ldots, J)$ are the exogenous blocks and $\boldsymbol{X}_k$ $(k = J+1, \ldots, K)$ are the endogenous blocks. We denote by $\boldsymbol{\xi}_k$ $(k = 1, \ldots, K)$ the corresponding components for each block of variables. A generic MV is denoted by $\boldsymbol{x}_{pk}$ $(p = 1, \ldots, P_k)$, $(k = 1, \ldots, K)$, where P_k is the number of MVs in the kth block.

For each endogenous block, in the PLS-PM context the redundancy index is computed as follows:

$$\text{Red}_k = \text{Com}_k \times R_k^2. \tag{7.3}$$

where Com_k is the average of the communalities in the kth block and R_k^2 is the multiple linear determination coefficient in the regression model of $\hat{\boldsymbol{\xi}}_q$ on its own predictor components.

Considering the way the redundancy index is obtained from the two different perspectives, it is clear that in PLS-PM the redundancy index is computed as in the context of CCA.

When the aim of the analysis is the maximisation of the explained variance of the MVs in the endogenous block from the MVs of the exogenous blocks (i.e. a redundancy-related criterion in a multi-block framework), the portion of the variability of each endogenous block's MV explained by its own predictors (represented by the explanatory components) can be defined as the following:

$$\mathrm{Red}_{\boldsymbol{x}_{pk}} = R^2(\boldsymbol{x}_{pk}, \{\hat{\boldsymbol{\xi}}'_{k'} s \text{ explaining } \hat{\boldsymbol{\xi}}_k\} \tag{7.4}$$

that is, as in the context of RA.

For a block k, the redundancy index is defined as follows:

$$\mathrm{Red}_k = \sum_{p=1}^{P_k} \mathrm{Red}_{\boldsymbol{x}_{pk}} / P_k \tag{7.5}$$

In this perspective, we can use the average of all the $\mathrm{Red}_{\boldsymbol{x}_{pk}}$ as a measure of global goodness of in-sample prediction. As a matter of fact, Lohmöller (1989) already considered the redundancy index as an index of goodness of fit of the global model and stated that the fit of the global model can be judged as satisfactory if the average of the redundancy indexes is sufficiently high. Because $\boldsymbol{X}_k$ ($k = J + 1, \ldots, K$) corresponds to the endogenous blocks, if we denote by $\tilde{P}$ the number of MVs of the endogenous blocks, the global goodness of prediction is defined as follows:

$$\overline{\mathrm{Red}} = \frac{1}{\tilde{P}} \sum_{k=J+1}^{K} P_k \times \mathrm{Red}_k \tag{7.6}$$

Just as with canonical correlations, no generally accepted guidelines have been established for the minimum acceptable redundancy index needed to judge a fit of the model as satisfactory. The researcher must judge the specific research problem being investigated to determine whether the redundancy index is enough to justify interpretation.

7.4.2 *Out-of-Sample Prediction*

To measure the performance of the model for prediction outside the dataset, out-of-sample prediction measures are needed.

7.4.2.1 The Stone-Geisser's Q^2 as a Measure of the Model's Predictive Relevance

To evaluate the model in terms of out-of-sample predictive ability, the so-called blindfolding procedure, which uses the Stone-Geisser's Q^2 (Geisser 1975; Stone 1974), is generally used in PLS-PM (Chin 1998). However, although the Q^2 is largely recommended in PLS-PM literature, it is not unanimously considered a true out-of-sample evaluation method (Shmueli et al. 2016) because there is no full observation omission (i.e. the omission of entire rows of the dataset). Therefore, there is no real 'test set' as in cross-validation, and this may lead to overestimation of the out-of sample predictive ability (Shmueli et al. 2016; Evermann and Tate 2016).

However, the blindfolding procedure is not even a purely in-sample evaluation method. In short, we may use it as an out-of-sample evaluation method, bearing in mind its limitations.

The PLS-PM adaptation of this approach follows a blindfolding procedure that, in brief, proceeds as follows: given a block of n cases and $\tilde{P}$ MVs (e.g. the MVs of the endogenous blocks), the procedure takes out a portion of the considered block during parameter estimations and then attempts to estimate the omitted part using the estimated parameters. The estimates for the omitted values are then compared to the observed values (Chin 1998). Beginning with the first data point (row 1, column 1) of this block, every mth observation is omitted, where m is the omission distance.

For an omission distance m, a proportion of $1/m$ of the sample will be discarded. Hence, a small omission distance will retain relatively less of the original sample for the parameter estimation than a large omission distance. Wold (1982, p. 33) suggests the use of an omission distance m equal to a number between the number of MVs per block and the sample size while (Chin 2010, p. 680) recommends a small value of around 5–10.

Blindfolding can be done on any set of variables. However, the predictive ability of the model typically involves the MVs of the endogenous blocks.

Different forms of Q^2 can be obtained based on different procedures for predicting observations from the model. In communality-based Q^2, prediction of observations is made by the computed composite and the estimated loadings and can be applied to all MVs (Chin 2010; Evermann and Tate 2012). Redundancy-based Q^2 is still based on the estimated loadings, but the composite are predicted from the structural model using the estimated path coefficients. It is applicable only to observations of the MVs of the endogenous blocks.

We use redundancy-based Q^2, which represents a measure of how well-observed values in endogenous blocks are reconstructed by the structural model and its parameter estimates (Chin 2010; Evermann and Tate 2012); therefore, it should be chosen when the predictive performance of the structural model is to be evaluated (Chin 2010; Evermann and Tate 2016). $Q^2 > 0$ implies that the model has predictive relevance, whereas $Q^2 < 0$ represents a lack of predictive relevance.

7.4.2.2 Cross-Validation Method to Evaluate Predictive Ability: The Mean Squared Prediction Errors

In some specific cases, which will be described in the last section, true out-of-sample evaluation methods (Hastie et al. 2009), such as the cross-validation method, can also be performed in the PLS-PM framework (Evermann and Tate 2016).

In the cross-validation method, entire rows of the dataset are omitted and represent the 'test set'. A sample of size n is split randomly into m sub-samples, each one of size equal to n/m. In the procedure, each sub-sample appears once as a test sample and $m-1$ times as training set. For each sub-sample used as a test set, the model parameters are estimated using the remaining $m-1$ sub-samples (the training sample). Then, the observation values of the dependent MVs for the test sub-sample are predicted using the observation values of predictor manifest variables of the test sub-sample and the estimated model parameters. This procedure is repeated m times.

We use the structural model parameter estimates for prediction so that we remain in the redundancy-based prediction context.

Following the description as in Evermann and Tate (2016), the definition of root-mean-squared error (RMSE), which measures the prediction error, is provided below.

Let $\hat{\boldsymbol{x}}_{h,i,p}$ be the predicted value for the subject i, $(i = 1, \ldots, n/m)$, of the sub-samples h, $(h = 1, \ldots, m)$, for the MVs of endogenous blocks, $(p = 1, \ldots, \hat{P})$, and $\boldsymbol{x}_{h,i,p}$ the corresponding true values. The prediction error is then defined as the RMSE averaged over all m sub-samples:

$$\overline{\text{RMSE}} = \frac{1}{m} \sum_{h=1\ldots m} \sqrt{\frac{m}{n\hat{P}} \sum_{i=1\ldots n/m} \sum_{p=1\ldots\hat{P}} (\hat{\boldsymbol{x}}_{h,i,p} - \boldsymbol{x}_{h,i,p})^2} \tag{7.7}$$

7.5 Predictive Directions of the Relationships Among Blocks

The PLS-PM algorithm proceeds in three stages (Wold 1982). The first stage computes the outer weights (the weights of the MVs) iteratively, which, in turn, serves to compute the weighted aggregates of their corresponding MVs (i.e. the weighted composites). Consequentially, this stage is of paramount importance because the quality of the overall results depends directly on it. The second stage estimates the multiplicative parameters of the model (path coefficients and loadings), and the third stage computes the location parameters. Only the second stage of the PLS-PM is concerned with the directionality of the structural model. The path directions are not taken into account in the first stage of the PLS-PM; therefore, the procedure fails to distinguish between dependent and explanatory composites in the weighted component construction.

This PLS-PM 'path directions incoherence' is mainly because of two factors. The first one concerns the way each component is defined in the inner step of the PLS-PM iterative algorithm. In particular, each component is defined as a linear combination of all the connected components. Two components are connected if a link exists between the two blocks: an arrow goes from one component to the other in the path diagram independently of the path direction. Consequently, the PLS-PM iterative algorithm does not distinguish between dependent and explanatory blocks. The directions of the links in the structural model do not play a role in the PLS-PM iterative algorithm, apart from the specific case of the so-called path-weighting scheme for the inner estimation (Tenenhaus et al. 2005). However, in the path-weighting scheme, the path direction is taken into account only in the way the inner weights are computed, but each component is still defined in the inner step of the algorithm as a function of all the connected components.

The second reason why the 'path directions incoherence' arises is because of the way the outer weights are computed. More specifically, the choice between the two outer modes has an effect on the predictive directions among blocks. The only way to give an explanatory role to a block is to apply *Mode B* while applying *Mode A* gives it a dependent role, whatever the path's direction is Dolce et al. (2015). Thus, the predictive direction in the structural model is given by the utilised outer mode.

If we wanted to respect the path directions, the use of Mode B to explanatory blocks and Mode A to dependent blocks may seem to be a natural choice. However, in the case of more than two blocks of variables, where some endogenous blocks may appear as both explanatory and dependent, this choice can be a much more complicated matter.

In short, the directions of the links in the structural model do not play a role in the PLS-PM iterative algorithm (the first stage of the PLS-PM). What actually has an effect on the predictive directions among blocks is the outer mode that is chosen.

7.5.1 A New Algorithm Coherent with the Predictive Directions

Because PLS-PM has been increasingly considered a predictive-oriented method, recent extensions and modifications have been made on the basic PLS-PM procedure to improve the predictive capabilities of the model (Dolce 2015; Shmueli et al. 2016).

The non-symmetrical approach for composite-based path modeling proposed by Dolce (2015) is an example of work in this direction. The authors proposed a new algorithm that replaces the original PLS-PM iterative algorithm to respect the directions of the relationships specified in the structural model and to improve the predictive ability of the model.

This new method (the NSCPM) explicitly takes into account the predictive direction among blocks and determines the MVs weights so that the corresponding composite is a good predictor of the dependent blocks (i.e. the composite fulfils its role in the inner model as well as possible).

The distinction between reflective and formative measurement models is disregarded while a great amount of emphasis is placed on the direction of relationships between components in the structural model. The main point is the distinction between explanatory blocks and dependent blocks in the model.

In particular, because some blocks play a double role in the model (i.e. they appear as both explanatory and dependent blocks in the model), in the NSC-PM method they are considered as explanatory when they play an explanatory role in the particular step of the algorithm and as dependent when they play a dependent role. When a block of variables plays an explanatory role in a specific step of the algorithm, we apply *Mode B* for computing the outer weights and apply *Mode A* when a block of variables plays a dependent role.

7.5.2 A Comparison Between Non-symmetrical Component-Based Path Modeling and PLS-PM

In general, because different component-based methods optimise different criteria, they must be judged by different metrics. However, in this study, we clearly state our position, as we are interested in evaluating the performance of the NSCPM and the classical PLS-PM in terms of predictive ability.

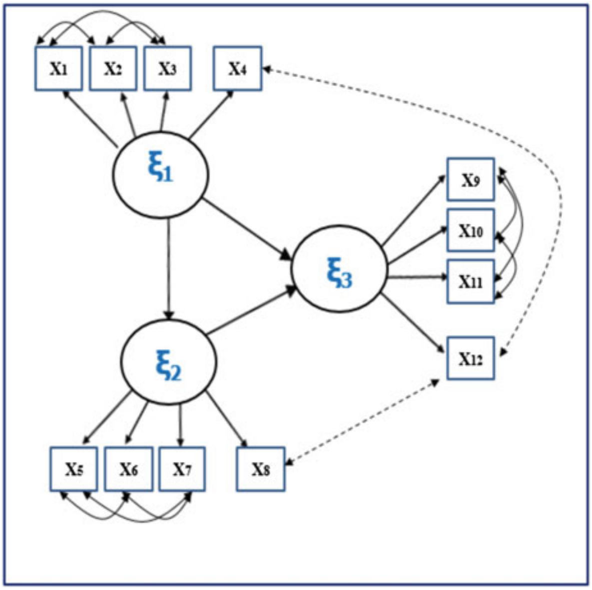
The comparison between the two methods is made using artificial data. Because in the case of a strong correlation within-blocks the results of most of the component-based methods are quite similar (because of the strength of the correlations), we generated a limit situation. In particular, we generated data from specific correlation structures, where the three blocks of variables in the model are defined as follows: three MVs, one in each set, are not characteristic of the whole set, but they are highly correlated among each other (see Fig. 7.1).[2] The variable $\boldsymbol{x}_4$ and the variable $\boldsymbol{x}_8$ are not highly correlated with the variables of their own blocks, but instead, they are both more correlated with the variable $\boldsymbol{x}_{12}$.

The mean of the correlations between the three related variables in each block is equal to 0.6. By including the fourth contaminating variable in each block, we get a mean correlation level within-blocks equal to 0.35. However, in each block, the Cronbach's α is about 0.7, and only the first eigenvalue is greater than one while the second one is slightly less than one. Therefore, this is a borderline situation in which the blocks of variables are generally considered consistent and unidimensional.

The data-generation process and the subsequent analysis are conducted in R (R Core Team 2014). We generated 500 Monte Carlo samples for six different levels of correlation averages between-blocks ($\bar{\rho} = 0.16, 0.19, 0.22, 0.25, 0.28, 0.31, 0.34,$ 0.37) to understand the effects of the different strengths of relationships between

[2]Note that we do not generate data using the standard factor-based structural equation modeling approach (i.e. using the implied covariance matrix). Instead, data are draw directly from a multivariate distribution with the pre-specified correlation matrix.

Fig. 7.1 Theoretical model for artificial data



blocks as well. However, because we did not find different results for different levels of correlation average between-blocks, we show only the results for the correlation averages between-blocks that are equal to 0.28, the conditions that represent a middle ground between the case of low correlation between blocks ($\bar{\rho} = 0.16$) and the case of high correlations between blocks ($\bar{\rho} = 0.37$).

We compared NSCPM with PLS-PM, applying *Mode B* for the exogenous block and *Mode A* for the endogenous blocks—we refer this model as PLS-PM(B,A,A)—and with PLS-PM applying *Mode B* for the first two blocks and *Mode A* for the last endogenous block—PLS-PM(B,B,A). We consider these two models, PLS-PM(B,A,A) and PLS-PM(B,B,A), because they are the most realistic for prediction use and the most coherent with the prediction directions among the path models. When either *Mode A* or *Mode B* is applied for all the blocks, predictive evaluation of the models may not be appropriate for the lack of predictive directoriality (Evermann and Tate 2016).

Consequently, the three statistics reported above (redundancy index, Q^2 and *RMSE*) can be applied to examine the predictive ability of the different methods considered here. As for the redundancy index, we use the index as it is computed in the context of PLS-PM to make the comparison possible.

As for the computation of Q^2, we chose an omission distance equal to 10, following the Wold (1982) and Chin (2010)'s recommendations. As for the computation of *RMSE*, following Evermann and Tate (2016), each sample is split randomly into ten sub-samples.

The computed statistics are not reported for each MVs and for each block of MVs because all three block have the same number of variables. Moreover, in the comparison between NSCPM and PLS-PM(B,B,A), the statistics are computed only for the last endogenous block.

In the following figures, the distributions of the three considered index values are reported for NSCPM, PLS-PM(B,A,A) and PLS-PM(B,B,A). These plots allow us to compare the performance of the two different PLS-PM estimation methods to one another and to NSCPM (Figs. 7.2, 7.3, and 7.4).

NSCPM performs slightly better than PLS-PM in terms of both in-sample and out-of-sample predictive relevance. In general, the worst model is the PLS-PM(B,B,A). Evidently, the differences are not large, but the results are in favour of NSCPM because prediction performance of the latter method is systematically

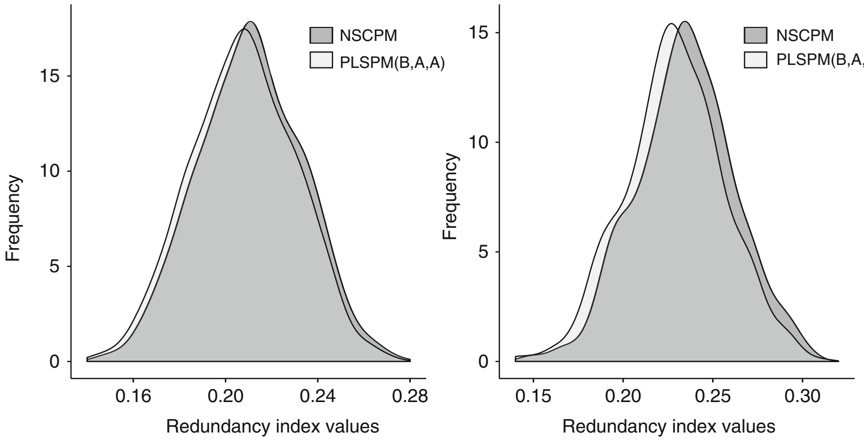


Fig. 7.2 Redundancy index density function for NSCPM and PLS-PM(B,A,A) (*left-hand side*) and NSCPM and PLS-PM(B,B,A) (*right-hand side*)

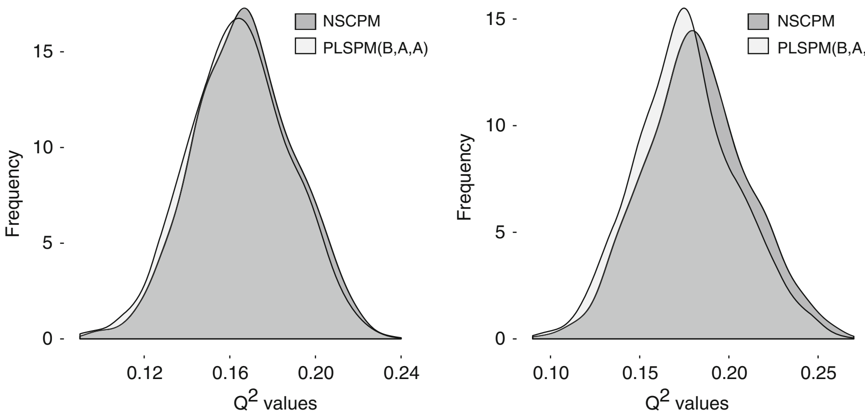


Fig. 7.3 Q^2 density function for NSCPM and PLS-PM(B,A,A) (*left-hand side*) and NSCPM and PLS-PM(B,B,A) (*right-hand side*)

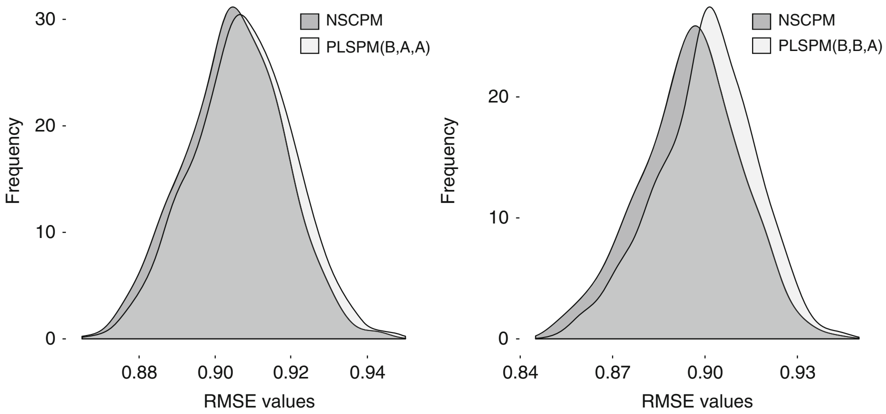


Fig. 7.4 *RMSE* density function for NSCPM and PLS-PM(B,A,A) (*left-hand side*) and NSCPM and PLS-PM(B,B,A) (*right-hand side*)

superior to the PLS-PM one. The largest advantage of NSCPM on PLS-PM is obtained when predictive power is assessed using RMSE, showing that better out-of-sample prediction can be achieved with NSCPM.

Even though we did not find very large differences between NSCPM and PLS-PM, we believe that this slight performance improvement may translate into real advantages in practice. Moreover, considering that the NSCPM algorithm is not more complicated than the PLS-PM algorithm and does not require more iterations to converge, we think that NSCPM can serve as an alternative procedure for the existing PLS-PM procedures and could be an important technique when the aim of the analysis is prediction. In addition, NSCPM is coherent with the direction of the relationship specified in the structural model.

7.6 Conclusions

PLS-PM is a statistical method that can certainly be considered an important technique in research applications, especially when the aim of the analysis is prediction and maintaining a strong interest in the explanation of the relationships between variables. The predictions of path models should be sensitive to the theory. In particular, the theoretical model represented by the structural equations and prediction should not be separated.

This chapter tries to help readers understand how to evaluate the predictive performance of component-based path models. Moreover, to improve the performance of the models in terms of predictive ability, researchers are recommended to apply further extensions and modifications to the basic PLS-PM algorithm.

A recent example of work in this direction, the NSCPM, was illustrated and compared to the classical PLS-PM. NSCPM seems to be very promising when researchers are interested in redundancy-based prediction. Certainly, this study is not exhaustive, but it is of help for appreciating the NSCPM performance in this specific example. To further evaluate the NSCPM performance in terms of predictive relevance, a Monte Carlo simulation study should be performed, one where several complementary conditions are considered. Accordingly, future research should aim to design a much broader simulation study for a more rigorous testing of the proposed method and to fully investigate the NSCPM properties.

Furthermore, future works should examine in detail Shmueli et al.'s PLSpredict procedure (Shmueli et al. 2016), for further investigating the NSCPM's predictive performance in the PLS-based method context.

References

Apel, H., & Wold, H. (1982). Soft modeling with latent variables in two or more dimensions: PLS estimation and testing for predictive relevance. In K. Jöreskog & H. Wold (Eds.), *Systems under indirect observations (Part II)* (pp. 209–247). Amsterdam: North-Holland.

Armstrong, J. S. (2001). Combining forecasts. In J. S. Armstrong (Ed.), *Principles of forecasting: A handbook for researchers and practitioners* (pp. 417–439). New York: Springer.

Bates, J. M., & Granger, C. W. J. (1969). The combination of forecasts. *Operations Research, 20*, 451–468. doi:10.2307/3008764.

Becker, J.-M., Rai, A., & Rigdon, E. E. (2013). Predictive validity and formative measurement in structural equation modeling: embracing practical relevance. In *Proceedings of the International Conference on Information Systems (ICIS)*, Milan.

Chin, W. W. (1998). The partial least squares approach for structural equation modeling. In G. Marcoulides (Ed.), *Modern methods for business research* (pp. 295–336). London: Lawrence Erlbaum Associates.

Chin, W. W. (2010). How to write up and report PLS analyses. In V. E. Vinzi, W. W. Chin, J. Henseler & H. Wang (Eds.), *Handbook of partial least squares: Concepts, methods and applications* (pp. 83–97, 655–690). New York: Springer.

Dana, J., & Dawes, R. M. (2004). The superiority of simple alternatives to regression for social science predictions. *Journal of Educational and Behavioral Statistics, 29*(3), 317–331.

Dijkstra, T. K., & Henseler, J. (2015a). Consistent and asymptotically normal PLS estimators for linear structural equations. *Computational Statistics and Data Analysis, 81*, 10–23. doi:10.1016/j.csda.2014.07.008.

Dijkstra, T. K., & Henseler, J. (2015b). Consistent partial least squares path modeling. *MIS Quarterly, 39*(2), 297–316.

Dolce, P. (2015). *Component-based path modeling: Open issues and methodological contributions*. Ph.D. Thesis, Università degli Studi di Napoli "Federico II".

Dolce, P., & Lauro, N. (2015). Comparing maximum likelihood and PLS estimates for structural equation modeling with formative blocks. *Quality and Quantity, 49*(3), 891–902.

Dolce, P., Vinzi, V. E., Lauro, C. (2015). Path directions incoherence in PLS path modeling: a prediction-oriented solution. In H. Abdi, et al. (Eds.), *The multiple facets of partial least squares methods and related methods*. Springer proceedings in mathematics & statistics (pp. 59–70).

Evermann, J., & Tate, M. (2012). Comparing the predictive ability of PLS and covariance models. In *Proceedings of the International Conference on Information Systems*, Orlando, FL (pp. 546–550).

Evermann, J., & Tate, M. (2016). Assessing the predictive performance of structural equation model estimators. *Journal of Business Research, 69*(10), 4565–4582.

Fornell, C., & Bookstein, F. L. (1982). Two structural equation models: LISREL and PLS applied to consumer exit-voice theory. *Journal of Marketing Research, 19*(4), 440–452.

Geisser, S. (1975). The predictive sample reuse method with applications. *Journal of the American Statistical Association, 70*(350), 320–328.

Hair, J. F., Hult, G. T. M, Ringle, C. M., & Sarstedt, M. (2014) A primer on partial least squares structural equation modeling (PLS-SEM). Thousand Oaks, CA: Sage.

Hair, J. F., Hult, G. T. M., Ringle, C. M., Sarstedt, M., & Thiele, K. O. (2017, in press). Mirror, mirror on the wall: A comparative evaluation of composite-based structural equation modeling methods. *Journal of the Academy of Marketing Science*. https://doi.org/10.1007/s11747-017-0517-x.

Hair, J. F., Sarstedt, M., Pieper, T., & Ringle, C. (2012). The use of partial least squares structural equation modeling in strategic management research: A review of past practices and recommendations for future applications. *Long Range Planning, 45*(5–6), 320–340. doi:10.1016/j.lrp.2012.09.008.

Hanafi, M. (2007). PLS path modeling: computation of latent variables with the estimation mode B. *Computational Statistics, 22*, 275–292.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction*. Berlin: Springer.

Henseler, J., Dijkstra, T. K., Sarstedt, M., Ringle, C. M., Diamantopoulos A, et al. (2014). Common beliefs and reality about partial least squares: Comments on Rönkkö & Evermann (2013). *Organizational Research Methods, 17*(2), 182–209. doi:10.1177/1094428114526928.

Henseler, J., Hubona, G. S., & Ray, P. A. (2016). Using PLS path modeling in new technology research: Updated guidelines. *Industrial Management & Data Systems, 116*(1), 2–20.

Höskuldsson, A. (2009). Modelling procedures for directed network of data blocks. *Chemometrics and Intelligent Laboratory Systems 97*(1), 3–10.

Hotelling, H. (1935). The most predictable criterion. *Journal of Education & Psychology, 26*(2), 139–142.

Hotelling, H. (1936). Relations between two sets of variates. *Biometrika, 28*(3–4), 321–377.

Jöreskog, K. (1977). Structural equation models in the social sciences: Specification, estimation and testing. In R. Krishnaiah (Ed.), *Applications of statistics* (pp. 265–287). Amsterdam: North-Holland.

Krämer, N. (2007). *Analysis of high-dimensional data with partial least squares and boosting*. PhD Thesis, Technische Universität Berlin.

Lohmöller, J. (1989). Latent variable path modeling with partial least squares. Berlin: Springer.

Makridakis, S., & Hibon, M. (2000), The M3 competition: Results, conclusions and recommendations. *International Journal of Forecasting, 16*(4), 451–476.

R Core Team. (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna.

Rencher, A. (1998). *Multivariate statistical inference and applications*. Wiley series in probability and statistics. New York: Wiley.

Rigdon, E. E. (2012). Rethinking partial least squares path modeling: In praise of simple methods. *Long Range Planning, 45*(5–6), 341–358. doi:10.1016/j.lrp.2012.09.010.

Rigdon, E. E. (2016). Choosing fPLSg path modeling as analytical method in European management research: A realist perspective. *European Management Journal, 34*(6), 598–605.

Ringle, C., Sarstedt, M., & Straub, D. (2012). A critical look at the use of PLS-SEM in MIS quarterly. *MIS Quarterly, 36*(1), iii–xiv.

Rönkkö, M., & Evermann, J. (2013). A critical examination of common beliefs about partial least squares path modeling. *Organizational Research Methods, 16*(3), 425–448.

Rönkkö, M., Antonakis, J., McIntosh, C., & Edwards, J. R. (2016). Partial least squares path modeling: Time for some serious second thoughts. *Journal of Operations Management, 47–48*, 9–27.
Sarstedt, M., Hair, J. F., Ringle, C., Thiele, K., & Gudergan, S. (2016). Estimation issues with PLS and CBSEM: Where the bias lies! *Journal of Business Research, 69*(10), 3998–4010. http://doi.org/10.1016/j.jbusres.2016.06.007.
Sarstedt, M., Ringle, C., Henseler, J., & Hair, J. (2014). On the emancipation of PLS-SEM: A commentary on Rigdon (2012). *Long Range Planning, 47*(3), 154–160. doi:10.1016/j.lrp.2014.02.007.
Shmueli, G. (2010). To explain or to predict? *Statistical Science, 25*(3), 289–310.
Shmueli, G., & Koppius, O. R. (2011) Predictive analytics in information systems research. *MIS Quarterly, 35*(3), 553–572.
Shmueli, G., Ray, S., Estrada, V. J. M., & Chatla, S. B. (2016). The elephant in the room: Predictive performance of PLS models. *Journal of Business Research, 69*(10), 4552–4564.
Stewart, D., & Love, W. (1968). A general canonical correlation index. *Psychological Bulletin, 70*(3), 160–163.
Stone, M. (1974). Cross-validatory choice and assessment of statistical predictions. *Journal of the Royal Statistical Society, 36*(2), 111–147.
Tenenhaus, A., & Tenenhaus, M. (2011). Regularized generalized canonical correlation analysis. *Psychometrika, 76*(2), 257–284. doi:10.1007/S11336-011-9206-8.
Tenenhaus, M., Vinzi, V. E., Chatelin, M., & Lauro, C. (2005). PLS path modeling. *Computational Statistics and Data Analysis, 48*(1), 159–205. doi:10.1016/j.csda.2004.03.005.
Van den Wollenberg, A. L. (1977). Redundancy analysis an alternative for canonical correlation analysis. *Psychometrika, 42*(2), 207–219.
Vinzi, V.E., & Russolillo, G. (2013). Partial least squares algorithms and methods. *WIREs Computational Statistics, 5*, 1–19
Vinzi, V.E., Trinchera, L., & Amato, S. (2010). PLS path modeling: Recent developments and open issues for model assessment and improvement. In V. E. Vinzi, W. Chin, J. Henseler & H. Wang (Eds.), *Handbook of partial least squares (PLS): Concepts, methods and applications* (pp. 47–82). Berlin: Springer.
Wold, H. (1980). Model construction and evaluation when theoretical knowledge is scarce: theory and application of partial least squares. In J. Kmenta & J. B. Ramsey (Eds.), *Evaluation of econometric models* (Chap. 3, pp. 47–74). Cambridge, MA: Academic.
Wold, H. (1982). Soft modeling: The basic design and some extensions. In K. Jöreskog & H. Wold (Eds.), *Systems under indirect observation: Part II* (pp. 1–54). Amsterdam: North-Holland.

Chapter 8
Mediation Analyses in Partial Least Squares Structural Equation Modeling: Guidelines and Empirical Examples

Gabriel Cepeda Carrión, Christian Nitzl, and José L. Roldán

Abstract Partial least squares structural equation modeling (PLS-SEM) is one of the options used to analyze mediation effects. Over the past few years, the methods for testing mediation have become more sophisticated. However, many researchers continue to use outdated methods to test mediation effects in PLS-SEM, which can lead to erroneous results in some cases. One reason for the use of outdated methods is that PLS-SEM tutorials do not draw on the newest statistical findings. This chapter illustrates how to perform modern procedures in PLS-SEM by challenging the conventional approach to mediation analysis and providing better alternatives.

These novel methods offer a wide range of testing options (e.g., multiple mediators) that go beyond simple mediation analysis alternatives, helping researchers to discuss their studies in a more accurate way. This chapter seeks to illustrate and help to operationalize the mediation in Nitzl et al.'s (Indus Manag Data Syst 116:1849–1864, 2016) paper about mediation in PLS, published in *Industrial Management & Data Systems*, with examples of two potential mediations: a multiple mediation with two mediators and a multistep multiple mediation.

8.1 Introduction

Partial least squares structural equation modeling (PLS-SEM) is a variance-based structural equation modeling technique that has been used to model latent variables, specifically composites, and the relationships between them (Henseler 2017). Therefore, it is a useful tool for testing hypotheses and answering research questions.

G.C. Carrión (✉) • J.L. Roldán
Faculty of Economics and Business, Department of Business Administration and Marketing, Universidad de Sevilla, Av. Ramón y Cajal 1, 41018 Seville, Spain
e-mail: gabi@us.es; jlroldan@us.es

C. Nitzl
Institute for Finance, Risk Management, and Management Accounting, University of the German Armed Forces Munich, Werner-Heisenberg-Weg 39, 85577 Neubiberg, Germany
e-mail: christian.nitzl@unibw.de

H. Latan, R. Noonan (eds.), *Partial Least Squares Path Modeling*,
DOI 10.1007/978-3-319-64069-3_8

One of these research questions investigates mediation. Mediation considers the presence of an intermediate variable or mechanism that transmits the effect of an antecedent variable to an outcome (Aguinis et al. 2017). For instance, mediation usually appears when the effect of reputation on customer loyalty is transmitted by customer satisfaction, such that reputation impact on customer satisfaction in turn influences customer loyalty (Hair et al. 2017). Hence, mediation refers to underlying effects that link antecedent and consequences variables. Despite the continuous use of mediation testing, studies in PLS-SEM often do not consider mediation effects in their hypotheses and therefore do not analyze the relevance in relevant structural models (Hair et al. 2017). In the worst case, researchers focus only on direct relationships and overlook mediation effects.

While there is a large body of literature on methods for testing mediation effects (Hayes and Scharkow 2013), the analytical tools that researchers have used in PLS-SEM studies to test mediation effects have generally been outdated compared to those for other statistical methods. Nitzl et al. (2016) have recently shown the misapplication of Baron and Kenny's procedure in the PLS-SEM field. Whereas researchers studying covariance-based structural equation modeling (CB-SEM) have often considered the latest findings when testing mediation (e.g., Iacobucci et al. 2007; Hair et al. 2010), most PLS-SEM researchers fail to do so, and in fact, they often avoid carrying out this kind of analysis. This is somewhat surprising because state-of-the-art applications for testing the significance of a mediator are also very suitable for PLS-SEM.

Therefore, we can state that one of the key reasons authors do not assess mediation effects in PLS path models is the lack of illustrative guidelines on conducting state-of-the-art mediation analysis with PLS-SEM. Furthermore, because these publications on PLS-SEM have been subjected to several recent changes (Henseler et al. 2016; Nitzl et al. 2016), an adequate illustration of these new guidelines related to mediation is badly needed.

Therefore, the objective of our chapter is to provide researchers with a nice illustration to implement mediation models in PLS-SEM. Thus, we offer complete examples and guidelines on how to conduct mediation analysis using PLS-SEM, inspired by Nitzl et al. (2016).

Our chapter is structured as follows: We first describe an advanced procedure for mediation analysis in PLS-SEM. We then list different types of mediation. Next, we offer illustrative examples of how to perform and discuss a mediation analysis with PLS. We also offer detailed guidelines for carrying out this type of analysis in PLS. Finally, we summarize our chapter and highlight potential avenues for future research.

8.2 Advanced Procedure for Mediation Analysis in PLS-SEM

8.2.1 The Mediation Effect

The core characteristic of a mediating effect (i.e., indirect effect or mediation) is that it involves a third variable that plays an intermediate role in the relationship between the independent and dependent variables. Technically speaking, the effect of the independent variable *X* on the dependent variable *Y* is mediated by a third variable, *M*, called the mediating variable or mediator (see Fig. 8.1). Thus, when we formulate mediation hypotheses, we focus on "how, or by what means, an independent variable (*X*) affects a dependent variable (*Y*) through one or more potential intervening variables, or mediators (*M*)" (Preacher and Hayes 2008). The researcher's objective in mediation analysis is mainly explanation (Henseler et al. 2016), although some scholars have also recently added the purpose of prediction (Shmueli et al. 2016).

Figure 8.1a shows the total effect *c* of the causal relationship between variables *X* and *Y*, and Fig. 8.1b shows a mediated effect in which *X* exerts an indirect effect $a \times b$ through *M* on *Y*.

Once we have defined the mediation effect, we briefly describe the procedure developed by Nitzl et al. (2016) to test mediation effects on PLS-SEM and also define the different types of mediation that researchers can find in their analysis. The procedure considers five important statements for testing mediating effects in PLS:

1. Testing the indirect effect $a \times b$ provides researchers with all the information they need to assess the significance of a mediation. Therefore, it is not necessary to conduct separate tests for paths *a* and *b* by applying PLS-SEM.

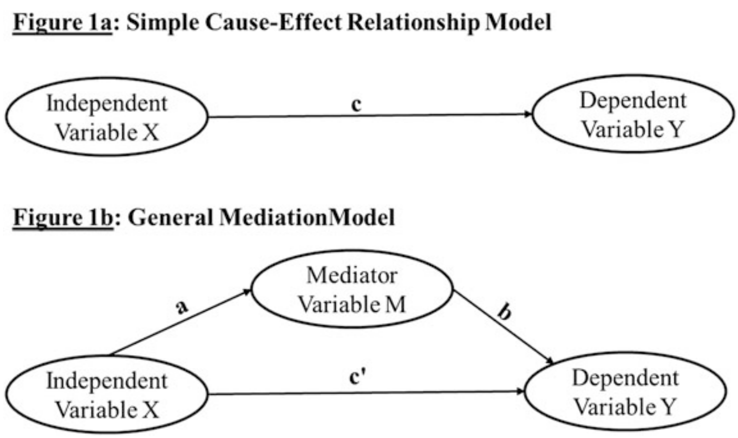


Fig. 8.1 (**a**) Simple cause-effect relationship and (**b**) general mediation model

2. The strength of the indirect effect $a \times b$ should determine the size of the mediation. Therefore, it is also not necessary to test the difference between c and c'.
3. A significant indirect effect $a \times b$ is the only prerequisite for establishing a mediation effect.
4. A bootstrap test should be used to test the significance of the indirect effect $a \times b$.
5. The significance of the direct effect (c') has to be tested in order to determine the type of effect and/or mediation.

These important statements are summarized in the procedure described by Nitzl et al. (2016). The procedure has two main steps (see Nitzl et al. (2016) for a more detailed description).

Step 1: Determining the significance of indirect effects and their magnitude
The indirect effect $a \times b$ must be significant in step 1 to establish a mediation effect. When testing mediation effects in PLS-SEM, researchers should bootstrap the sample of the indirect effects in order to obtain necessary information about the population distribution, in accordance with the nonparametric PLS-SEM method where bootstrap procedures are typically used for inference statistics, such as the calculation of the so-called pseudo t-value and confidence intervals (Henseler et al. 2009). The bootstrapping procedure is a nonparametric inferential technique that randomly withdraws several subsamples (e.g., 5000) with replacement from the original dataset. PLS-SEM uses each of the subsamples to estimate the underlying PLS path model.

The bootstrap routines of PLS-SEM software often provide results for at least direct effects (e.g., path a and path b). However, for a more detailed analysis of mediation, particularly in more complex model structures (e.g., multiple mediators), it is often necessary to compute the bootstrapping results for the indirect effects with the help of a spreadsheet application, such as Microsoft Excel or CALC in Open Office. For each bootstrapping subsample, the results of path a must be multiplied by path b to create the product term $a \times b$ of the indirect effect in a new column. For example, the computation of $k = 5000$ bootstrapping subsamples entails the generation of $k = 5000$ products $a \times b$ in a new column. The information about the characteristics of the distribution of mediation effects is obtained by calculating a ci% confidence interval for $a \times b$. For that, the subsamples (k) for $a \times b$ from the bootstrapping procedure must be arranged from smallest to largest (Hayes 2009). In the next step, a researcher has to select a specific alpha error; for example, for a probability of error of 5%, a 95% confidence interval must be determined with a 2.5% probability of error at each tail when conducting a two-sided test. The lower bound of $a \times b$ is in the $k \times (.5 - \text{ci}\%/2)$th ordinal position of the ordered list; for example, if one uses $k = 5000$ subsamples and a 95% confidence interval, the lower bound is the $5000 \times (.5 - 0.95/2) = 125$th ordinal position. Similarly, the $(1 + k \times (.5 + \text{ci}\%/2))$th ordinal determines the upper bound of the bootstrap confidence, which is the $1 + 5000 \times (.5 + 0.95/2) = 4876$th in the previous example. If *zero* is not included in the confidence interval, a researcher can assume that there is a significant indirect effect $a \times b$.

Another problem often occurs when the mean of the bootstrapped distribution for the indirect effect $a \times b$ is not equal to the estimated indirect effect $a \times b$ (Chernick 2011). As a result, researchers must correct for this bias in PLS-SEM. This can be accomplished by calculating the difference between the estimated indirect effect $a_{\mathrm{PM}} \times b_{\mathrm{PM}}$ from the path model (PM) and the mean value of the indirect effect $a_B \times b_B$ from the bootstrap sample (B). Consequently, the bias-corrected ci% confidence interval for an indirect effect $a \times b$ can be defined as

$$\begin{aligned}&[(k \times (.5 - \mathrm{ci}\%/2))\,\mathrm{th} + (a_{\mathrm{PM}} \times b_{\mathrm{PM}} - a_B \times b_B)\,;\\&(1 + k \times (.5 + \mathrm{ci}\%/2))\,\mathrm{th} + (a_{\mathrm{PM}} \times b_{\mathrm{PM}} - a_B \times b_B)]\,.\end{aligned} \tag{8.1}$$

Hayes and Scharkow (2013) show that the bias-corrected bootstrap confidence interval is the best approach for detecting mediation effects when a mediation effect is present (i.e., Type II error or power). Conversely, the percentile bootstrap confidence interval that is not bias corrected is a good compromise if a researcher is also concerned about Type I errors (Hayes and Scharkow 2013). Thus, the bias-corrected bootstrap confidence interval is the most reliable test if power is of the utmost concern, while the percentile bootstrap confidence interval provides a good compromise.

Step 2: Determining the type of effect and/or mediation
Step 2 involves defining the type of effect and/or mediation. A mediating effect always exists when the indirect effect $a \times b$ in step 1 is significant. The current mediation literature discusses two different types of mediation, full and partial mediation. Partial mediation can be subdivided into complementary and competitive partial mediation. We also discuss two effects that occur when the indirect effect is not significant, which means that only the direct effect is significant and no effect at all is significant. The latter cases do not represent a mediating effect in the narrow sense.

8.2.2 Full Mediation

A full mediation is indicated in the case where the direct effect c' is not significant, whereas the indirect effect $a \times b$ is significant. This means only the indirect effect via the mediator exists. In other words, full mediation means that the effect of the variable X on Y is completely transmitted with the help of another variable M. It also means the condition Y completely absorbs the positive or negative effect of X. In this way, it can completely pass an effect, or it can completely hinder the effect in terms of another effect. Technically speaking, the variable X extracts its influence only under a certain condition of M on Y. However, in the case of small samples, a researcher is to exercise some caution when talking about full mediation. As Rucker et al. (2011) showed, “the smaller the sample, the more likely mediation (when present) is to be labeled full as opposed to partial, because c' is more easily rendered non-significant" (p. 364). Hence, it is advisable to ensure that the sample size is great enough so that the necessary power of 0.8 for an alpha level of 0.05 for detecting effects in a PLS path model is obtained. For a simple mediation model, the necessary sample size can be quite low. Notwithstanding, a medium and small effect size would require a bigger sample. In contrast, in many cases, it can be observed that some small direct effects, c', remain even though the mediating effect is quite high in relation to the mediated direct effect. However, when this relation of the direct effect to the mediating effect becomes low but nevertheless stays significant, it can also be seen as full mediation. A researcher could indicate this with the help of the variance accounted for (VAF) value, which we will discuss in more detail below in our example. Conversely, when the absolute value of the indirect path $a \times b$ is larger than the absolute value of the total effect $(a \times b) + c'$, there is a suppressor effect (Cheung and Lau 2008); this situation could also be defined as full mediation (Hair et al. 2017).

8.2.3 *Partial Mediation*

All other situations under the condition that both the direct effect c' and the indirect effect $a \times b$ are significant represent partial mediation. Two types of partial mediation can be distinguished.

8.2.3.1 Complementary Partial Mediation

In a complementary partial mediation, the direct effect c' and indirect effect $a \times b$ point in the same (positive or negative) direction (Baron and Kenny 1986). It is an often observed result that $a \times b$ and c' are significant and $a \times b \times c'$ is positive, which indicates that a portion of the effect of X on Y is mediated through M, while X still explains a portion of Y that is independent of M. This complementary mediation hypothesis suggests that the intermediate variable explains, possibly confounds, or falsifies the relationships between the independent and dependent variables.

8.2.3.2 Competitive Partial Mediation

In a competitive partial mediation, the direct effect c' and indirect effect $a \times b$ point in a different direction. A negative $a \times b \times c'$ value indicates the presence of competitive mediation in step 2. As mentioned above, this indicates that a portion of the effect of X on Y is mediated through M, while X still explains a portion of Y that is independent of M. In the past, researchers often focused only on complementary mediation (Zhao et al. 2010). In the competitive partial mediation hypothesis, it is assumed that the intermediate variable will reduce the magnitude of the relationship between the independent and dependent variables. However, it is possible that the intermediate variable could increase the magnitude of the relationship between the independent and dependent variables. Competitive partial mediation has often been called a "negative confounding" or an "inconsistent" model. Thus, other types of mediation beyond complementary mediation should be considered in a PLS path model.

PLS researchers might also be interested in evaluating the strength (portion) in the case of a partial mediation. Mediation analyses regularly involve partial mediation, and therefore it can be helpful to have further information on the mediated portion. One approach for this is calculating the ratio of the indirect to total effect. This ratio is also known as the variance accounted for (VAF) value. VAF determines the extent to which the mediation process explains the dependent variable's variance. For a simple mediation, the proportion of mediation is defined as

$$\mathrm{VAF} = \frac{a \times b}{(a \times b) + c'}. \tag{8.2}$$

8.2.4 No Mediation

When the indirect effect is not significant, we can find another two situations. Although these cannot be considered mediation cases in a narrow sense, two types of effects can be distinguished.

8.2.4.1 Only Direct Effect

If the indirect effect $a \times b$ is not significant while the direct path *c'* is, the mediator variable has no impact; this indicates that a direct, non-mediating effect is present. In this case, the study was perhaps searching for a wrong mediation relationship. However, it is possible that an unrecognized mediation relationship still exists and another mediation variable is present that mediates an effect between X and Y (Shrout and Bolger 2002). Thus, a researcher should rethink his theoretical basis when he has not found the expected mediation relationship (cf. Zhao et al. 2010).

8.2.4.2 No Effect

There is no effect if neither the indirect effect $a \times b$ nor the direct effect c' is significant. The total effect can still be significant. First of all, in this case, a researcher should check if the sample size has enough power to show an effect when there is an effect. Putting the last two cases together—the indirect effect $a \times b$ is not significant and the direct path c' is or is not—frequently indicates a problematic or flawed theoretical framework (Zhao et al. 2010). In this case, a researcher has to thoroughly examine the hypothesized model. When, for example, the total effect c is significant, it can indicate that the mediation variable should be deleted because it brings no further degree of explanation. In the case where the mediation variable M has no real effect, it only dilutes the effect of the direct variable X and should be deleted.

8.2.5 Multiple Mediation

PLS is regularly applied in complex path models. There may be multiple relationships between one or more independent variables, one or more mediator variables, and one or more dependent variables. For instance, a complementary mediation variable (M_1) may mitigate the independent variable (X) to a dependent variable (Y), and at the same time, a competitive mediation variable (M_2) may also exist. From a naïve perspective, someone can assume that the independent variable is not relevant because there is no relevant total effect c. However, when one of the mediator variables has a strong influence in a certain situation, the independent variable also wins in terms of relevance. Such areas can become very challenging, for example, when analyzing which process improves or hinders the influence of the external pressure to work on the outcome in a PLS path model. However, when more than one mediating effect is present, the abovementioned differentiation between direct and indirect effects for detecting mediation relationships remains applicable, and the above recommendations are inalterable (Hayes 2009).

Figure 8.2 presents an example of a PLS path model with two mediators. The total effect is equal to the direct effect of X on Y, in addition to the sum of the indirect effects of M_1 and M_2. A given meditator's indirect effect is referred to as a specific indirect effect (e.g., through M_1). The sum of the two specific indirect effects is the complete indirect effect. Thus, the total effect is the sum of the direct effect and the complete indirect effects (i.e., the sum of the specific indirect effects includes the relationship between M_1 and M_2). For the example in Fig. 8.3, the calculation of the total effect is

$$c = c' + (a_1 \times b_1) + (a_2 \times b_2)\,. \tag{8.3}$$

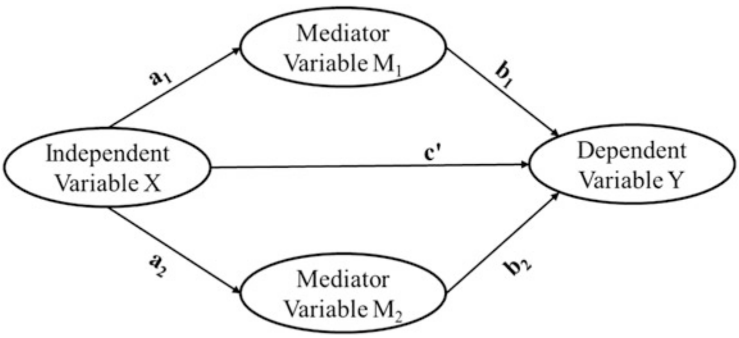


Fig. 8.2 Multiple mediator model

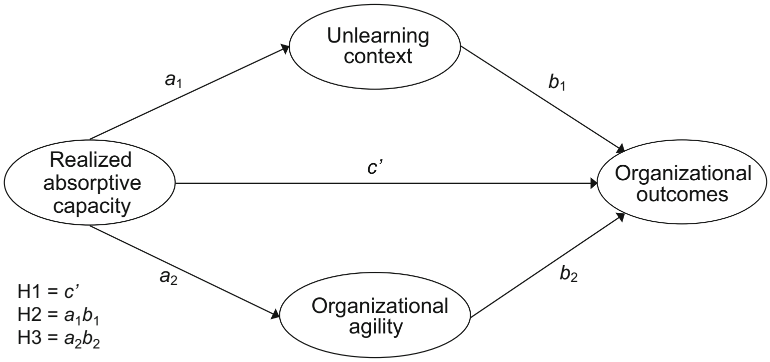


Fig. 8.3 An example of a multiple mediator model. Source: Roldán et al. (2014)

An interesting situation occurs when $a_1 \times b_1$ and $a_2 \times b_2$ in Eq. (8.2) have an opposite sign; this indicates that one effect functions as a complementary effect and the other functions as a competitive mediator effect. Such a model is called an inconsistent mediation model (MacKinnon et al. 2007). Consequently, even though significant specific indirect effects exist, the complete indirect effect [e.g., $(a_1 \times b_1) + (a_2 \times b_2)$] may not be significant.

Preacher and Hayes (2008) argue that the incorporation of multiple mediators and the comparison of their specific mediating effects are also useful for comparing different competing theories. Given this background, researchers are interested in comparing the strengths of specific mediating effects [e.g., $(a_1 \times b_1)$ and $(a_2 \times b_2)$] in complex models (Williams and MacKinnon 2008). For example, a researcher could test for two complementary mediator variables if mediator (M_1) has a stronger mediator effect than mediator (M_2). The previous explanation of how to compute bootstrap confidence intervals in PLS can be extended to test the significance of the difference between two specific mediating effects (Lau and Cheung 2012). For that purpose, a researcher must calculate the following equation:

$$D_{\mathrm{M}} = M_1 - M_2, \tag{8.4}$$

where M_1 and M_2 are the specific indirect effects and $\mathrm{D_M}$ is the difference between these two specific indirect effects. In this way, we test whether two specific indirect effects are equal or if they amount to zero. In the case examined in this study, the equation for Fig. 8.2 would be $D_{\mathrm{M}} = (a_1 \times b_1) - (a_2 \times b_2)$. Again, researchers can calculate the equation using a spreadsheet application to build a confidence interval with the help of the bootstrapping results of the PLS program (cf. Chin et al. 2013; Rodríguez-Entrena et al. 2016).

A frequently encountered case is that in which two mediators are connected to each other. This indicates an additional relationship between M_1 and M_2 in Fig. 8.2. Next, we provide examples of how to test such multiple mediation relationships in a PLS path model. In such a case, the total effect c can be calculated as follows: $c = c' + (a_1 \times b_1) + (a_2 \times b_2) + (a_1 \times a_3 \times b_2)$, where a_3 stands for the relation between M_1 and M_2. An interesting case in this situation is when a_2, b_2, and c' are not significantly different from zero, but the indirect effect $(a_1 \times a_3 \times b_2)$ is (e.g., when M_1 is the causal predecessor of M_2); this would mean that M_1 fully mediates the direct effect between X and M_2 and that M_2 fully mediates the direct effect between M_1 and Y, thus establishing a direct causal chain $X \rightarrow M_1 \rightarrow M_2 \rightarrow Y$ (Mathieu et al. 2008). Next, we illustrate this in our second example.

8.3 Illustrative Examples

8.3.1 A Case of a Multiple Mediator Model

In this first example, we take data from Roldán et al. (2014). This research examines the relationship between a key component of the absorptive capacity, the realized absorptive capacity (RACAP), and the organizational outcomes, this link being mediated by the unlearning context and the organizational agility (Fig. 8.3). These connections are examined through an empirical investigation of 112 large companies.

8.3.1.1 Data Collection and Measures

The population of this study consists of Spanish organizations that use Editran ™ and which have more than 100 employees. Editran™ is a software used to enhance communications over different platforms and is a de facto standard in the Spanish banking system. This population is suitable for our study, because these businesses are more familiar with knowledge and technology management. There were 464 companies identified from the SABI (Sistema de Análisis de Balances Ibéricos) database and invited to participate in the study, and 121 companies agreed. A total of 112 valid and completed questionnaires were collected.

We modeled RACAP and unlearning context as multidimensional constructs (composites). We measured RACAP by two first-order dimensions (composites): transformation and exploitation. The unlearning context variable was assessed using three first-order dimensions (composites): the examination of lens fitting (ELF), the framework for changing individual habits (CIH), and the framework for consolidation of emergent understandings (CEU). The example's constructs were estimated in Mode A, and the characteristics of the scales are the following:

(a) RACAP. Items were measured using a seven-point Likert scale from the study by Jansen et al. (2008). RACAP includes the transformation and exploitation of new external knowledge. The final cleansed scale consists of four items for the transformation dimension and three items for the exploitation dimension.

(b) Unlearning context. At the organizational level, it is viewed as memory elimination in general and as changing beliefs, norms, values, procedures, and routines in particular. As described above, the unlearning context has three dimensions: the consolidation of emergent understandings, the examination of lens fitting, and the framework for changing individual habits. The measures relating to consolidation of emergent understandings consisted of six items taken from a scale designed by Cegarra and Sanchez (2008). Five items were used to measure the examination of lens fitting. Finally, we measured the framework for changing individual habits using seven items.
(c) Organizational agility. A business-wide capability to deal with changes that often arise unexpectedly in business environments via rapid and innovative responses that exploit changes as opportunities to grow and prosper. The indicators of organizational agility are based on the measures of organizational agility used by Lu and Ramamurthy (2011). The scale was composed of six items.
(d) Organizational outcomes. It is understood as an assessment of the global performance of the business. The scale for organizational outcomes consisted of ten reflective items adapted from Quinn and Rohrbaugh (1983).

8.3.1.2 Hypotheses Development

Once the conceptual framework is shown, the next step is hypotheses development. The research model depicted in Fig. 8.3 includes one direct and two mediating hypotheses:

H1: RACAP is positively associated with organizational outcomes.
H2: The relationship between RACAP and organizational outcomes is positively mediated by the unlearning context.
H3: The relationship between RACAP and organizational outcomes is positively mediated by organizational agility.

8.3.1.3 PLS-SEM Practical Considerations

This chapter's goal is not to illustrate the complete data analysis with PLS, but to focus on the structural model, specifically on how to test this mediation model with PLS. Guidelines for a complete analysis with PLS can be found in Henseler et al. (2016) and Hair et al. (2017).

Significance of Direct and Indirect Effects

We assess the significance of one direct (c') and two indirect effects ($a_1 \times b_1$ and $a_2 \times b_2$). The critical issue is that if the significance of each indirect effect cannot be

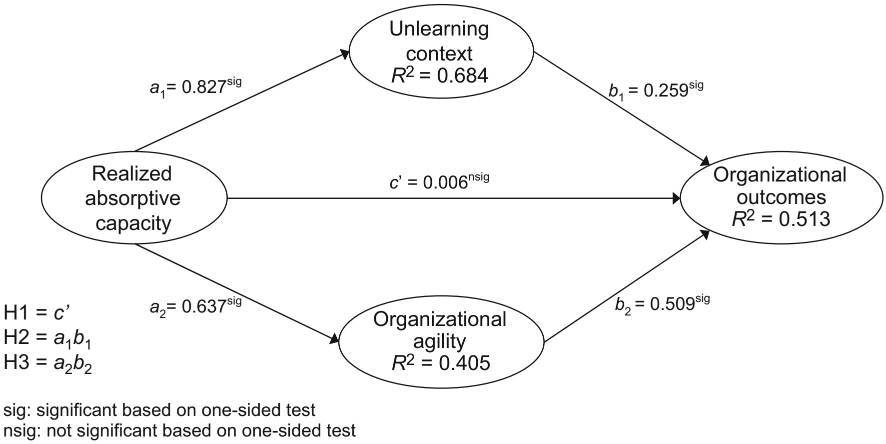


Fig. 8.4 An example of a multiple mediator model. Results. Source: Roldán et al. (2014)

established, there is no mediating effect. Consequently, having a significant indirect effect is the key to determining the type of mediation effect and its magnitude. Considering that our hypotheses have been formulated with direction (+), we will use a one-sided test. Accordingly, we will estimate 90% confidence intervals (CI).

Nitzl et al. (2016) suggested a procedure using a spreadsheet and multiplying the bootstrapping outputs (i.e., $a_1 \times b_1$ and $a_2 \times b_2$) to calculate the percentile and the bias-corrected confidence intervals. Therefore, once we run the model, we next perform the bootstrapping procedure with 5000 subsamples and no sign changes. In Fig. 8.4 we can see the estimates for direct effects.

In order to calculate the specific indirect effects and the different confidence intervals, we use a spreadsheet application (i.e., Excel or CALC) to obtain the significance of mediator effects ($a_1 \times b_1$ and $a_2 \times b_2$) in the relationship between RACAP and organizational outcomes. We suggest carrying out the following steps:

1. Take the 5000 sets of path coefficients from all direct effects created by the bootstrap procedure and copy and paste into a spreadsheet's columns (Fig. 8.5).
2. Create a new column for each indirect effect under assessment. In this case, we generate two new columns ($a_1 \times b_1$ and $a_2 \times b_2$) and explicitly calculate the product of the direct paths that form such indirect paths. In addition, we include another column for estimating the total indirect effect $(a_1 \times b_1) + (a_2 \times b_2)$ (Fig. 8.6).
3. Copy the original values (O) provided by PLS for the direct effects. Then calculate the product of the direct paths that form each indirect path. In the line below, calculate the mean (M) for each column of the paths obtained with the bootstrapping process (Fig. 8.7).
4. Insert a new line where you estimate the bias as original (O) − mean (M) for each column (Fig. 8.8).

	A	B	C	D	E	F
1						
2						
3						
4						
5						
6						
7						
8						
9		c'	a1	a2	b1	b2
10	Sample 0	-0,12	0,834	0,704	0,38	0,527
11	Sample 1	-0,136	0,885	0,552	0,348	0,589
12	Sample 2	-0,044	0,83	0,584	0,317	0,503
13	Sample 3	-0,065	0,823	0,605	0,461	0,349
14	Sample 4	0,037	0,863	0,71	0,173	0,559

Fig. 8.5 Example 1. Step 1

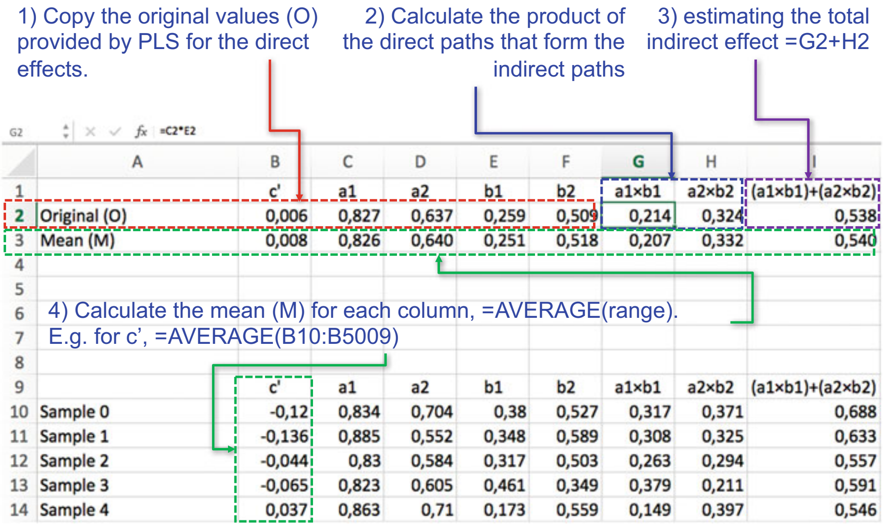
2) Insert a column for estimating the total indirect effect (a1 × b1)+(a2 × b2)

1) Create two new columns a1 × b1 and a2 × b2, and explicitly calculate the product of the direct paths that form the indirect paths under assessment

	A	B	C	D	E	F	G	H	I
1									
2									
3									
4									
5									
6									
7									
8									
9		c'	a1	a2	b1	b2	a1×b1	a2×b2	(a1×b1)+(a2×b2)
10	Sample 0	-0,12	0,834	0,704	0,38	0,527	0,317	0,371	0,688
11	Sample 1	-0,136	0,885	0,552	0,348	0,589	0,308	0,325	0,633
12	Sample 2	-0,044	0,83	0,584	0,317	0,503	0,263	0,294	0,557
13	Sample 3	-0,065	0,823	0,605	0,461	0,349	0,379	0,211	0,591
14	Sample 4	0,037	0,863	0,71	0,173	0,559	0,149	0,397	0,546

Fig. 8.6 Example 1. Step 2

5. Estimate the percentile bootstrap CI for each column using the function PERCENTILE (range, *k*), *k* being the percentile value between 0 and 1. In our case, given our hypotheses are postulated with direction (+), we will use one-sided test, and we will estimate 90% CI (Fig. 8.9).
6. Estimate the bias-corrected CI adding the bias to the previously calculated percentile CI (Fig. 8.10).



1) Copy the original values (O) provided by PLS for the direct effects.

2) Calculate the product of the direct paths that form the indirect paths

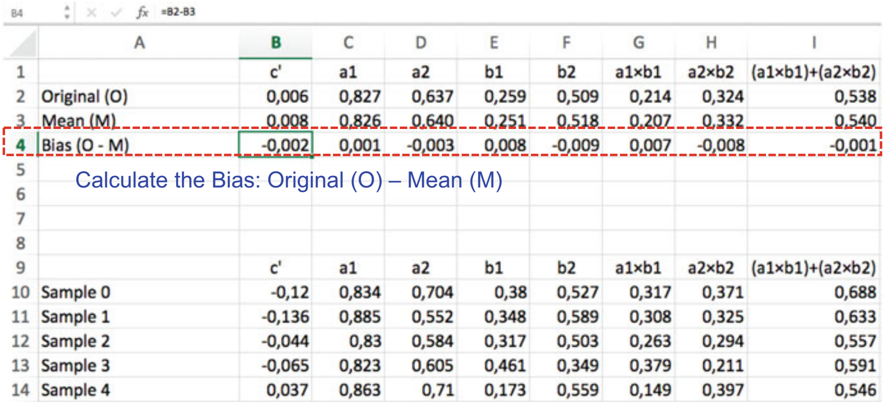
3) estimating the total indirect effect =G2+H2

G2 fx =C2*E2

	A	B	C	D	E	F	G	H	I
1		c'	a1	a2	b1	b2	a1×b1	a2×b2	(a1×b1)+(a2×b2)
2	Original (O)	0,006	0,827	0,637	0,259	0,509	0,214	0,324	0,538
3	Mean (M)	0,008	0,826	0,640	0,251	0,518	0,207	0,332	0,540
4									
5									
6									
7									
8									
9		c'	a1	a2	b1	b2	a1×b1	a2×b2	(a1×b1)+(a2×b2)
10	Sample 0	-0,12	0,834	0,704	0,38	0,527	0,317	0,371	0,688
11	Sample 1	-0,136	0,885	0,552	0,348	0,589	0,308	0,325	0,633
12	Sample 2	-0,044	0,83	0,584	0,317	0,503	0,263	0,294	0,557
13	Sample 3	-0,065	0,823	0,605	0,461	0,349	0,379	0,211	0,591
14	Sample 4	0,037	0,863	0,71	0,173	0,559	0,149	0,397	0,546

4) Calculate the mean (M) for each column, =AVERAGE(range).
E.g. for c', =AVERAGE(B10:B5009)

Fig. 8.7 Example 1. Step 3



B4 fx =B2-B3

	A	B	C	D	E	F	G	H	I
1		c'	a1	a2	b1	b2	a1×b1	a2×b2	(a1×b1)+(a2×b2)
2	Original (O)	0,006	0,827	0,637	0,259	0,509	0,214	0,324	0,538
3	Mean (M)	0,008	0,826	0,640	0,251	0,518	0,207	0,332	0,540
4	Bias (O - M)	-0,002	0,001	-0,003	0,008	-0,009	0,007	-0,008	-0,001
5									
6									
7									
8									
9		c'	a1	a2	b1	b2	a1×b1	a2×b2	(a1×b1)+(a2×b2)
10	Sample 0	-0,12	0,834	0,704	0,38	0,527	0,317	0,371	0,688
11	Sample 1	-0,136	0,885	0,552	0,348	0,589	0,308	0,325	0,633
12	Sample 2	-0,044	0,83	0,584	0,317	0,503	0,263	0,294	0,557
13	Sample 3	-0,065	0,823	0,605	0,461	0,349	0,379	0,211	0,591
14	Sample 4	0,037	0,863	0,71	0,173	0,559	0,149	0,397	0,546

Calculate the Bias: Original (O) – Mean (M)

Fig. 8.8 Example 1. Step 4

7. If the confidence interval (CI) for a mediation effect (products) does not include 0 value, it means the mediating effect is significantly different from 0. In our example, both indirect effects are significant. In addition, the total indirect effect is also significant (Table 8.1).

The key point to determine a mediation effect is the evaluation of the significance of the indirect effect (Table 8.1). In our example, both indirect effects are significant; therefore, H2 and H3 are supported. However, the direct effect is not significant; consequently H1 is not supported.

Percentile LOWER (5%) for *c'* =PERCENTILE(B10:B5009,0,05)
Percentile LOWER (95%) for *c'* =PERCENTILE(B10:B5009,0,95)

	A	B	C	D	E	F	G	H	I
1		c'	a1	a2	b1	b2	a1×b1	a2×b2	(a1×b1)+(a2×b2)
2	Original (O)	0,006	0,827	0,637	0,259	0,509	0,214	0,324	0,538
3	Mean (M)	0,008	0,826	0,640	0,251	0,518	0,207	0,332	0,540
4	Bias (O - M)	-0,002	0,001	-0,003	0,008	-0,009	0,007	-0,008	-0,001
5	Percentile LOWER (5%)	-0,189	0,757	0,509	0,022	0,365	0,018	0,217	0,373
6	Percentile UPPER (95%)	0,194	0,884	0,748	0,474	0,670	0,396	0,459	0,715
7									
8									
9		c'	a1	a2	b1	b2	a1×b1	a2×b2	(a1×b1)+(a2×b2)
10	Sample 0	-0,12	0,834	0,704	0,38	0,527	0,317	0,371	0,688
11	Sample 1	-0,136	0,885	0,552	0,348	0,589	0,308	0,325	0,633
12	Sample 2	-0,044	0,83	0,584	0,317	0,503	0,263	0,294	0,557
13	Sample 3	-0,065	0,823	0,605	0,461	0,349	0,379	0,211	0,591
14	Sample 4	0,037	0,863	0,71	0,173	0,559	0,149	0,397	0,546

Fig. 8.9 Example 1. Step 5

B7 fx =B5+B4

	A	B	C	D	E	F	G	H	I
1		c'	a1	a2	b1	b2	a1×b1	a2×b2	(a1×b1)+(a2×b2)
2	Original (O)	0,006	0,827	0,637	0,259	0,509	0,214	0,324	0,538
3	Mean (M)	0,008	0,826	0,640	0,251	0,518	0,207	0,332	0,540
4	Bias (O - M)	-0,002	0,001	-0,003	0,008	-0,009	0,007	-0,008	-0,001
5	Percentile LOWER (5%)	-0,189	0,757	0,509	0,022	0,365	0,018	0,217	0,373
6	Percentile UPPER (95%)	0,194	0,884	0,748	0,474	0,670	0,396	0,459	0,715
7	BC: P. LOWER (5%) + Bias	-0,191	0,758	0,506	0,030	0,356	0,025	0,208	0,372
8	BC: P. UPPER (95%) + Bias	0,192	0,885	0,745	0,482	0,661	0,403	0,451	0,714
9		c'	a1	a2	b1	b2	a1×b1	a2×b2	(a1×b1)+(a2×b2)
10	Sample 0	-0,12	0,834	0,704	0,38	0,527	0,317	0,371	0,688
11	Sample 1	-0,136	0,885	0,552	0,348	0,589	0,308	0,325	0,633
12	Sample 2	-0,044	0,83	0,584	0,317	0,503	0,263	0,294	0,557
13	Sample 3	-0,065	0,823	0,605	0,461	0,349	0,379	0,211	0,591
14	Sample 4	0,037	0,863	0,71	0,173	0,559	0,149	0,397	0,546

Fig. 8.10 Example 1. Step 6

Type of Mediation and Magnitude

Once we have determined the significance of the two mediation effects, we can go for the second step to determine the type of mediation and its magnitude. Table 8.1 shows the point estimate for the direct effect (c'), the indirect effects ($a_1 \times b_1$, $a_2 \times b_2$), and the total indirect effect [$(a_1 \times b_1) + (a_2 \times b_2)$]. Given that c' is not significant and both the indirect and the total indirect effects are significant, a full mediation can be defended. In addition, we can calculate VAF to assess the magnitude for each mediation. It can be said that almost 99% of the total effect is

Table 8.1 Example 1. Summary of mediating effects tests

Direct effects	Coefficient	Bootstrap 90% CI			
		Percentile		BC	
H1: c'	0.006^{nsig}	−0.189	0.194	−0.191	0.192
a_1	0.827^{sig}	0.757	0.884	0.758	0.885
a_2	0.637^{sig}	0.509	0.748	0.506	0.745
b_1	0.259^{sig}	0.022	0.474	0.030	0.482
b_2	0.509^{sig}	0.365	0.670	0.356	0.661

Indirect effects	Point estimate	Percentile		BC		VAF
H2: $a_1 \times b_1$	0.214^{sig}	0.018	0.396	0.025	0.403	39.3%
H3: $a_2 \times b_2$	0.324^{sig}	0.217	0.459	0.208	0.451	59.6%
Total indirect effect	0.538^{sig}	0.373	0.715	0.372	0.714	98.9%

Notes: *sig* significant, *nsig* not significant, *BC* bias corrected, *VAF* variance accounted for

due to two mediation effects jointly. Because the VAF exceeds 80%, this implies an additional argument for a full mediation.

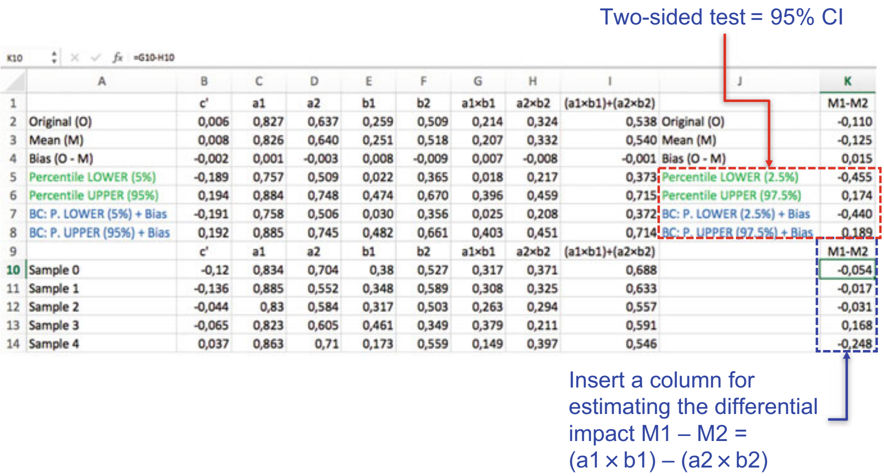
Comparison of Mediating Effects

When we evaluate a multiple mediator model, we can go further comparing the different mediating effects. In our example, we want to test whether the unlearning context (M_1) has a stronger mediator effect than the organizational agility (M_2) variable. With this aim in mind, we will assess the potential statistical difference between $a_1 \times b_1$ and $a_2 \times b_2$ following the guidelines provided by Chin et al. (2013) and Rodríguez-Entrena et al. (2016). Thus, we will include a new column where we estimate the difference between $a_1 \times b_1$ and $a_2 \times b_2$ and calculate percentile and bias-corrected CI. Because we have not postulated any hypothesis about the differential impact of both indirect effects, we will carry out a two-sided test (95% CI) (Fig. 8.11).

The test (Table 8.2) shows there is not a differential impact between M_1 and M_2 since both CIs contain the zero value. Accordingly, we cannot state that the unlearning context (M_1) has a stronger mediator effect than the organizational agility (M_2) variable and vice versa.

8.3.2 An Example of a Multistep Multiple Mediator Model

Our second example has been extracted from Roldán et al. (2017). This study examines post-adoption behaviors (i.e., frequency of use, routinization, and infusion) and their effects on the sense of community in the domain of social network


K10 fx =G10-H10

	A	B	C	D	E	F	G	H	I	J	K
1		c'	a1	a2	b1	b2	a1×b1	a2×b2	(a1×b1)+(a2×b2)		M1-M2
2	Original (O)	0,006	0,827	0,637	0,259	0,509	0,214	0,324	0,538	Original (O)	-0,110
3	Mean (M)	0,008	0,826	0,640	0,251	0,518	0,207	0,332	0,540	Mean (M)	-0,125
4	Bias (O - M)	-0,002	0,001	-0,003	0,008	-0,009	0,007	-0,008	-0,001	Bias (O - M)	0,015
5	Percentile LOWER (5%)	-0,189	0,757	0,509	0,022	0,365	0,018	0,217	0,373	Percentile LOWER (2.5%)	-0,455
6	Percentile UPPER (95%)	0,194	0,884	0,748	0,474	0,670	0,396	0,459	0,715	Percentile UPPER (97.5%)	0,174
7	BC: P. LOWER (5%) + Bias	-0,191	0,758	0,506	0,030	0,356	0,025	0,208	0,372	BC: P. LOWER (2.5%) + Bias	-0,440
8	BC: P. UPPER (95%) + Bias	0,192	0,885	0,745	0,482	0,661	0,403	0,451	0,714	BC: P. UPPER (97.5%) + Bias	0,189
9		c'	a1	a2	b1	b2	a1×b1	a2×b2	(a1×b1)+(a2×b2)		M1-M2
10	Sample 0	-0,12	0,834	0,704	0,38	0,527	0,317	0,371	0,688		-0,054
11	Sample 1	-0,136	0,885	0,552	0,348	0,589	0,308	0,325	0,633		-0,017
12	Sample 2	-0,044	0,83	0,584	0,317	0,503	0,263	0,294	0,557		-0,031
13	Sample 3	-0,065	0,823	0,605	0,461	0,349	0,379	0,211	0,591		0,168
14	Sample 4	0,037	0,863	0,71	0,173	0,559	0,149	0,397	0,546		-0,248

Fig. 8.11 Example 1. Comparison of mediating effects. Spreadsheet illustration

Table 8.2 Example 1. Comparison of mediating effects

Differential effect	Coefficient	Bootstrap 95% CI			
		Percentile		BC	
$M_1 - M_2 = (a_1 \times b_1) - (a_2 \times b_2)$	−0.110	−0.455	0.174	−0.440	0.189

sites. Specifically, this contribution formulates a multistep mediator model where frequency of use affects social integration via routinization and infusion (Fig. 8.12). The data was collected from 278 users of Tuenti, a popular social network site among the Spanish college student population during the period 2006–2012.

8.3.2.1 Data Collection and Measures

Undergraduate students, users of the Tuenti social network, were recruited from social studies at a public university in Southern Spain. A total of 278 questionnaires were collected from members who responded to an offline survey.

Frequency of use is defined as the number of times that an individual uses a social network site (SNS). It was operationalized by two self-reported measures. Routinization describes the state in which SNS use is no longer perceived as out of the ordinary but becomes institutionalized, being associated with habitual and standardized usage, that is, the integrating of the SNS into daily routines. We measure it by adapting a scale developed by Sundaram et al. (2007). Infusion is conceptualized as the extent to which a person uses an SNS to its highest level to maximize its potential, implying the notion of a deeper use. We use an adaptation of the measure developed by Jones et al. (2002). Finally, social integration measures

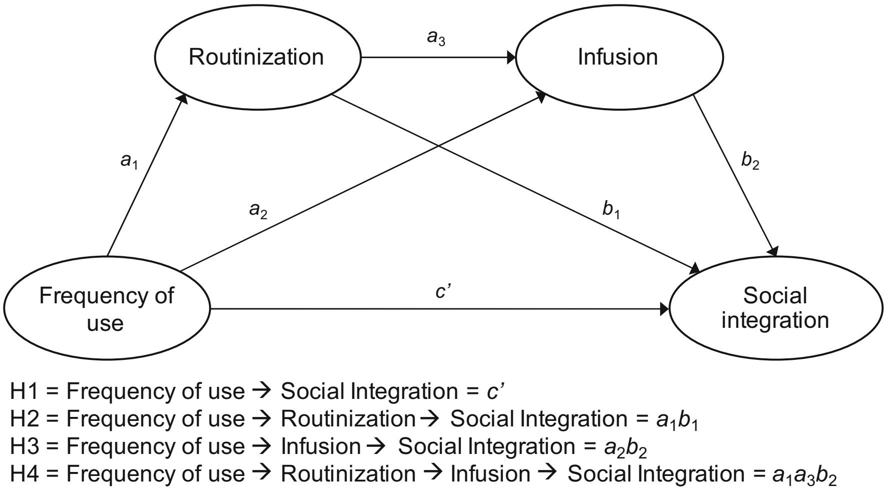


Fig. 8.12 An example of a model with a three-path mediated effect. Source: Roldán et al. (2017)

both the sense of belongingness to, and the identification with, the SNS and the social community's interactivity level. Consequently, social integration is modeled as a multidimensional construct composed of two dimensions: community participation and community integration. All variables have been estimated in Mode A.

8.3.2.2 Hypotheses Development

Considering the research model described in Fig. 8.12, we have postulated one direct and three mediating hypotheses, one of them proposing a three-path mediated effect:

H1: Frequency of use is positively related to social integration.
H2: The relationship between frequency of use and social integration is positively mediated by routinization.
H3: The relationship between frequency of use and social integration is positively mediated by infusion.
H4: The relationship between frequency of use and social integration is sequentially and positively mediated by routinization and infusion.

8.3.2.3 PLS-SEM Practical Considerations

We follow the guidelines described in the previous example. Therefore, we will show the final results of our analyses in order to avoid excessive redundancy in the explanation (Fig. 8.13).

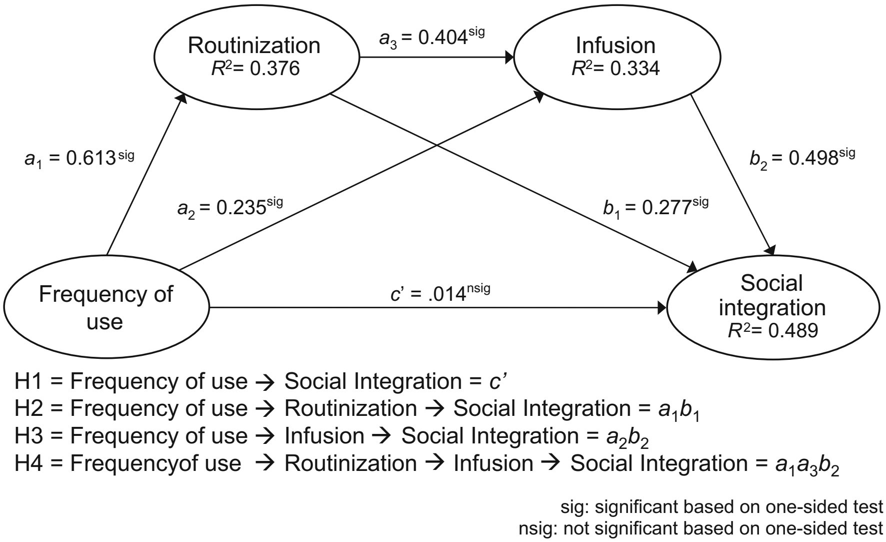


Fig. 8.13 An example of a model with a three-path mediated effect. Results. Source: Roldán et al. (2017)

	A	B	C	D	E	F	G	H	I	J	K
1		c'	a1	a2	a3	b1	b2	a1×b1	a2×b2	a1×a3×b2	(a1×b1)+(a2×b2)+(a1×a3×b2)
2	Original	0,014	0,613	0,235	0,404	0,277	0,498	0,170	0,117	0,123	0,410
3	Mean	0,015	0,614	0,235	0,404	0,277	0,498	0,170	0,117	0,124	0,411
4	Bias (O - M)	-0,001	-0,001	0,000	0,000	0,000	0,001	0,000	0,000	0,000	-0,001
5	Percentile LOWER (5%)	-0,073	0,548	0,132	0,286	0,165	0,408	0,101	0,063	0,080	0,343
6	Percentile UPPER (95%)	0,107	0,672	0,339	0,521	0,386	0,584	0,240	0,176	0,172	0,479
7	BC: P. LOWER (5%) + Bias	-0,074	0,547	0,131	0,286	0,166	0,408	0,101	0,063	0,079	0,343
8	BC: P. UPPER (95%) + Bias	0,106	0,671	0,339	0,521	0,386	0,584	0,240	0,176	0,171	0,478
9		c'	a1	a2	a3	b1	b2	a1×b1	a2×b2	a1×a3×b2	(a1×b1)+(a2×b2)+(a1×a3×b2)
10	Sample 0	0,049	0,531	0,228	0,515	0,290	0,480	0,154	0,109	0,131	0,395
11	Sample 1	0,011	0,589	0,210	0,441	0,287	0,434	0,169	0,091	0,113	0,373
12	Sample 2	-0,058	0,670	0,228	0,466	0,432	0,407	0,290	0,093	0,127	0,509
13	Sample 3	-0,003	0,565	0,407	0,295	0,252	0,504	0,143	0,205	0,084	0,431
14	Sample 4	0,089	0,622	0,261	0,372	0,264	0,488	0,164	0,127	0,113	0,404

Fig. 8.14 Example 2. Final spreadsheet with the estimation of indirect effects and confidence intervals

The evaluation of our research model involves estimating the significance of one direct (c') and the three indirect effects ($a_1 \times b_1$, $a_2 \times b_2$, and $a_1 \times a_3 \times b_2$). Considering that our hypotheses have been formulated with direction (+), we will use a one-sided test, calculating 90% confidence intervals (Fig. 8.14).

Significance of Direct and Indirect Effects

Frequency of use has no significant direct effect on social integration (H1: c') (Table 8.3). Therefore, H1 is not supported. On the other hand, all the indirect effects of

Table 8.3 Example 2. Summary of mediating effects tests

Direct effects	Coefficient	Bootstrap 90% CI			
		Percentile		BC	
H1: c'	0.014^{nsig}	−0.073	0.107	−0.074	0.106
a_1	0.613^{sig}	0.548	0.672	0.547	0.671
a_2	0.235^{sig}	0.132	0.339	0.131	0.339
a_3	0.404^{sig}	0.286	0.521	0.286	0.521
b_1	0.277^{sig}	0.165	0.386	0.166	0.386
b_2	0.498^{sig}	0.408	0.584	0.408	0.584

Indirect effects	Point estimate	Percentile		BC		VAF
H2: $a_1 \times b_1$	0.170	0.101	0.240	0.101	0.240	40.0%
H3: $a_2 \times b_2$	0.117	0.063	0.176	0.063	0.176	27.6%
H4: $a_1 \times a_3 \times b_2$	0.123	0.080	0.172	0.079	0.171	29.0%
Total indirect effect	0.410	0.343	0.479	0.343	0.478	96.6%

Notes: *sig* significant, *nsig* not significant, *BC* bias corrected, *VAF* variance accounted for

frequency of use on social integration are significant. This means that H2–H4 have been supported. Thus, routinization positively mediates the relationship between frequency of use and social integration (H2: $a_1 \times b_1$). Likewise, infusion mediates the path between frequency of use and social integration (H3: $a_2 \times b_2$). Finally, we find that frequency of use is positively associated with higher routinization and infusion, which relates to higher levels of social integration (H4: $a_1 \times a_3 \times b_2$).

Type of Mediation and Magnitude

Table 8.3 indicates that c' is not significant and all postulated indirect effects are significant. Consequently, this means that routinization and infusion fully and jointly mediate the influence of frequency of use on social integration. This is also supported by applying the variance accounted for (VAF) index. When the VAF has an outcome above 80%, a full mediation can be assumed. This occurs when we assess the total indirect effect of frequency of use on social integration (VAF = 96.6%).

Comparison of Mediating Effects

Finally, we will test whether routinization (M_1) has a stronger mediator effect than infusion (M_2). As we did in the previous example, we evaluate the statistical difference between $a_1 \times b_1$ and $a_2 \times b_2$ (Table 8.4). In this case, we do not observe a significant difference between both indirect effects.

Table 8.4 Example 2. Comparison of mediating effects

Differential effect	Coefficient	Bootstrap 95% CI			
		Percentile		BC	
$M_1 - M_2 = (a_1 \times b_1) - (a_2 \times b_2)$	0.053	−0.067	0.175	−0.067	0.175

8.4 Conclusion

PLS-SEM is a statistical procedure for structural equation modeling that social science researchers can consider when conducting research. This chapter helps readers to understand how PLS-SEM can be applied in mediation analysis through two illustrative examples. PLS-SEM seems not to be "a panacea for flaws in research design or execution" (Rigdon 2016: 604), but research must not ignore the proper model assessment prior to drawing a conclusion. This kind of advanced modeling (i.e., mediation analysis) can be performed by PLS-SEM as illustrated by this chapter. The adoption of these guidelines is advised for researchers who use PLS-SEM, particularly when they tackle multiple mediation models.

References

Aguinis, H., Edwards, J. R., & Bradley, K. J. (2017). Improving our understanding of moderation and mediation in strategic management research. *Organizational Research Methods* (forthcoming).

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic and statistical considerations. *Journal of Personality and Social Psychology, 51*(6), 1173–1182.

Cegarra, J. G., & Sánchez, M. T. (2008). Linking the individual forgetting context with customer capital from a seller's perspective. *The Journal of the Operational Research Society, 59*, 1614–1623.

Chernick, M. R. (2011). *Bootstrap methods: A guide for practitioners and researchers*. New York: Wiley.

Cheung, G. W., & Lau, R. S. (2008). Testing mediation and suppression effects of latent variables: Bootstrapping with structural equation models. *Organizational Research Methods, 11*(2), 296–325.

Chin, W. W., Kim, Y. J., & Lee, G. (2013). Testing the differential impact of structural paths in PLS analysis: A bootstrapping approach. In H. Abdi et al. (Eds.), *New perspectives in partial least squares and related methods*, *Springer proceedings in mathematics and statistics* (Vol. 56, pp. 221–229). New York: Springer.

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis*. Upper Saddle River, NJ: Pearson/Prentice Hall.

Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2017). *A primer on partial least squares structural equation modeling (PLS-SEM)* (2nd ed.). Thousand Oaks, CA: Sage.

Hayes, A. F. (2009). Beyond Baron and Kenny: Statistical mediation analysis in the new millennium. *Communication Monographs, 76*(4), 408–420.

Hayes, A. F., & Scharkow, M. (2013). The relative trustworthiness of inferential test of the indirect effect in statistical mediation analysis: Does method rally matter? *Psychological Science, 24*(10), 1918–1927.
Henseler, J. (2017). Bridging design and behavioral research with variance-based structural equation modeling. *Journal of Advertising, 46*(1), 178–192.
Henseler, J., Ringle, C. M., & Sinkovics, R. R. (2009). The use of partial least squares path modeling in international marketing. In R. R. Sinkovics & P. N. Ghauri (Eds.), *Advances in International Marketing* (Vol. 20, pp. 277–320). Bingley: Emerald.
Henseler, J., Hubona, G., & Ray, P. A. (2016). Using PLS path modeling in new technology research: Updated guidelines. *Industrial Management & Data Systems, 116*(1), 2–20.
Iacobucci, D., Saldanha, N., & Deng, X. (2007). A mediation on mediation: Evidence that structural equation models perform better than regression. *Journal of Consumer Psychology, 17*(2), 140–154.
Jansen, J. J. P., George, G., van den Bosch, F. A. J., & Volberda, H. W. (2008). Senior team attributes and organizational ambidexterity: The moderating role of transformational leadership. *Journal of Management Studies, 45*(5), 982–1007.
Jones, E., Sundaram, S., & Chin, W. W. (2002). Factors leading to sales force automation use: A longitudinal analysis. *Journal of Personal Selling & Sales Management, 22*, 145–156.
Lau, R. S., & Cheung, G. W. (2012). Estimating and comparing specific mediation effects in complex latent variable models. *Organizational Research Methods, 15*(1), 3–16.
Lu, Y., & Ramamurthy, K. (2011). Understanding the link between information technology capability and organizational agility: An empirical examination. *MIS Quarterly, 35*(4), 931–954.
MacKinnon, D. P., Fairchild, A. J., & Fritz, M. S. (2007). Mediation analysis. *Annual Review of Psychology, 58*, 593–614.
Mathieu, J. E., DeShon, R. P., & Bergh, D. D. (2008). Mediational inferences in organizational research: Then, now, and beyond. *Organizational Research Methods, 11*(2), 203–223.
Nitzl, C., Roldan, J. L., & Cepeda-Carrion, G. (2016). Mediation analysis in partial least squares path modeling: Helping researchers discuss more sophisticated models. *Industrial Management & Data Systems, 116*(9), 1849–1864.
Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods, 40*(3), 879–891.
Quinn, R. E., & Rohrbaugh, J. (1983). A spatial model of effectiveness criteria: Towards a competing values approach to organizational analysis. *Management Science, 29*(3), 363–377.
Rigdon, E. (2016). Choosing PLS path modeling as analytical method in European Management Research: A realist perspective. *European Management Journal, 34*(6), 598–605.
Rodríguez-Entrena, M., Schuberth, F., & Gelhard, C. (2016). Assessing statistical differences between parameters estimates in Partial Least Squares path modeling. *Quality and Quantity*, 1–13.
Roldán, J. L., Cegarra, J. G., & Cepeda, G. (2014). Building organisational agility through an unlearning Context. *European Conference on Knowledge Management, 2*, 834–842.
Roldán, J. L., Sanchez-Franco, M. J., & Real, J. C. (2017). From frequency of use to social integration: The mediation of routinization and infusion in Tuenti community. *European Research on Management and Business Economics, 23*(2), 63–69.
Rucker, D. D., Preacher, K. J., Tormala, Z. L., & Petty, R. E. (2011). Mediation analysis in social psychology: Current practices and new recommendations. *Social and Personality Psychology Compass, 5*(6), 359–371.
Shmueli, G., Ray, S., Estrada, J. M. V., & Chatla, S. B. (2016). The elephant in the room: Evaluating the predictive performance of PLS models. *Journal of Business Research, 69*(10), 4552–4564.
Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and non-experimental studies: New procedures and recommendations. *Psychological Methods, 7*(4), 422–445.

Sundaram, S., Schwarz, A., Jones, E., & Chin, W. W. (2007). Technology use on the front line: How technology enhances individual performance. *Journal of the Academy of Marketing Science, 35*, 101–112.

Williams, J., & MacKinnon, D. P. (2008). Resampling and distribution of the product methods for testing indirect effects in complex models. *Structural Equation Modeling, 15*(1), 23–52.

Zhao, X., Lynch, J. G., & Chen, Q. (2010). Reconsidering Baron and Kenny: Myths and truths about mediation analysis. *Journal of Consumer Research, 37*(2), 197–206. doi:10.1086/651257.

Chapter 9
Treating Unobserved Heterogeneity in PLS-SEM: A Multi-method Approach

Marko Sarstedt, Christian M. Ringle, and Joseph F. Hair

Abstract Accounting for unobserved heterogeneity has become a key concern to ensure the validity of results when applying partial least squares structural equation modeling (PLS-SEM). Recent methodological research in the field has brought forward a variety of latent class techniques that allow for identifying and treating unobserved heterogeneity. This chapter raises and discusses key aspects that are fundamental to a full and adequate understanding of how to apply these techniques in PLS-SEM. More precisely, in this chapter, we introduce a systematic procedure for identifying and treating unobserved heterogeneity in PLS path models using a combination of latent class techniques. The procedure builds on the FIMIX-PLS method to decide if unobserved heterogeneity has a critical impact on the results. Based on these outcomes, researchers should use more recently developed latent class methods, which have been shown to perform superior in recovering the segment-specific model estimates. After introducing these techniques, the chapter continues by discussing the means to identify explanatory variables that characterize the latent segments. Our discussion also broaches the issue of measurement invariance testing, which is a fundamental requirement for a subsequent comparison of parameters across groups by means of a multigroup analysis.

M. Sarstedt (✉)
Institute of Marketing, Otto-von-Guericke University Magdeburg, Magdeburg, Germany

Faculty of Business and Law, University of Newcastle, Callaghan, NSW, Australia
e-mail: marko.sarstedt@ovgu.de

C.M. Ringle
Institute of Human Resource Management and Organizations (HRMO), Hamburg University of Technology (TUHH), Hamburg, Germany

Faculty of Business and Law, University of Newcastle, Callaghan, NSW, Australia
e-mail: ringle@tuhh.de

J.F. Hair
University of South Alabama, Mobile, AL, USA
e-mail: joefhair@gmail.com

H. Latan, R. Noonan (eds.), *Partial Least Squares Path Modeling*,
DOI 10.1007/978-3-319-64069-3_9

9.1 Introduction

Researchers routinely create groups of data and analyze their group-specific results to account for heterogeneity (e.g., Kotler 1989; Völckner et al. 2010). But the sources of heterogeneity in the data are often difficult to know a priori. When researchers do not properly account for heterogeneity and the presence of substantial group-specific differences, results may be substantially biased. As a consequence, conclusions drawn are false and misleading (Sarstedt et al. 2009). In structural equation modeling (e.g., Byrne 2016; Diamantopoulos and Siguaw 2000; Rigdon 1998, 2005) in general and, more specifically, in partial least squares structural equation modeling (PLS-SEM) (e.g., Hair et al. 2017c; Lohmöller 1989; Wold 1975, 1982), uncovering unobserved heterogeneity is a key concern (Becker et al. 2013; Jedidi et al. 1997) researchers must address in their analysis and result evaluation (Hair et al. 2017b, 2012). Conversely, if researchers demonstrate that unobserved heterogeneity does not affect their results, they can analyze the data in a single model on the aggregate level (Hair et al. 2017a).

The application of standard data clustering procedures, such as k-means, is a common approach for dealing with unobserved heterogeneity. In the context of PLS-SEM, however, the traditional clustering approaches perform very poorly regarding the identification of group differences in the estimated path coefficients (Sarstedt and Ringle 2010). Hence, methodological PLS-SEM research has proposed different methods for identifying and treating unobserved heterogeneity, commonly referred to as latent class techniques (Sarstedt 2008).

Applying these latent class techniques is far from trivial, as evidenced in Sarstedt et al.'s (2016b) critical commentary on Marques and Reis's (2015) use of finite mixture PLS (FIMIX-PLS), which is the most commonly used latent class approach in PLS-SEM to date—Table 9.3 (in the Appendix) provides a summary of studies applying the technique. For this reason, this chapter raises and discusses key aspects that are fundamental to a full and adequate understanding of how to uncover and treat unobserved heterogeneity in PLS-SEM. More precisely, in this chapter, we introduce a systematic procedure for identifying and treating unobserved heterogeneity in PLS path models using a combination of latent class techniques. The procedure builds on the FIMIX-PLS method (Hahn et al. 2002; Sarstedt et al. 2011a) to decide if unobserved heterogeneity has a critical impact on the results. If so, researchers can use the method to identify the number of segments to retain from the data. Based on these results, researchers should use more recently developed latent class methods, which have been shown to perform superior in recovering the segment-specific model estimates. After introducing these techniques, the chapter continues by discussing the means to identify explanatory variables that characterize the latent segments, followed by an introduction of multigroup analysis that allows comparing parameters between two or more segments for significant differences. Our discussion also broaches the issue of measurement invariance testing, which is a fundamental requirement for comparing parameters across groups. Our aim is to offer further guidance to those researchers who work with PLS-SEM and related methods.

9.2 Guidelines for Uncovering and Treating Unobserved Heterogeneity in PLS Path Models

Treating unobserved heterogeneity in PLS path models is not a trivial endeavor. It requires careful application of latent class analysis in combination with measurement invariance assessment and multigroup analysis in a recurring process. Figure 9.1 illustrates the guidelines for uncovering and treating unobserved heterogeneity in PLS-SEM, which we will briefly introduce in this section and discuss in greater detail in the sections that follow.

Step 1 involves running FIMIX-PLS—by using the default settings (also see Ringle et al. 2010a) across a range of segments $\kappa = \{\kappa_1, \ldots, \kappa_{\max}\}$. Starting with a one-segment solution κ_1, the range depends on the relationships between the number of observations and the model complexity. Researchers need to compare the different segmentation results using a series of metrics that guide the decision of how many segments to retain. The assignment of the observations to separate segments follows. This assignment is based on the FIMIX-PLS membership probabilities. Each observation is therefore fully assigned to the segment where the FIMIX-PLS solutions revealed its highest membership probability. Again, in the final solution, the minimum number of observations per segment must be large enough to warrant a sufficient level of accuracy and statistical power in the group-specific model estimations. See Kock and Hadaya (2017) for two accurate and simple approaches for minimum sample size estimation in PLS-SEM and Hair et al. (2017d) for a recent evaluation of PLS-SEM's performance in small sample size constellations.

Step 2 uses the FIMIX-PLS solution from Step 1 as the starting partition for running a follow-up latent class analysis. This analysis draws on PLS prediction-oriented segmentation (PLS-POS) (Becker et al. 2013), PLS genetic algorithm segmentation (PLS-GAS) (Ringle et al. 2013, 2014), or PLS iterative reweighted regressions segmentation (PLS-IRRS) (Schlittgen et al. 2016) in order to further improve the FIMIX-PLS solution. The combined use of FIMIX-PLS and PLS-POS, as explained in Hair et al. (2017a), is easily accessible, since both methods are implemented in the SmartPLS 3 software (Ringle et al. 2015). Again, in the final solution, the minimum number of observations per segment of data must meet the PLS-SEM algorithm's minimum sample size requirements for producing accurate estimates in each group.

In Step 3, researchers have to run an ex post analysis to identify one or more explanatory variables that are a good match for the solution obtained in Step 2. A suitable (set of) explanatory variable(s) (e.g., age, gender, income) then facilitate(s) the forming of the final partition. Careful deliberation that examines whether the role of the explanatory variable(s) is theoretically and/or logically meaningful should

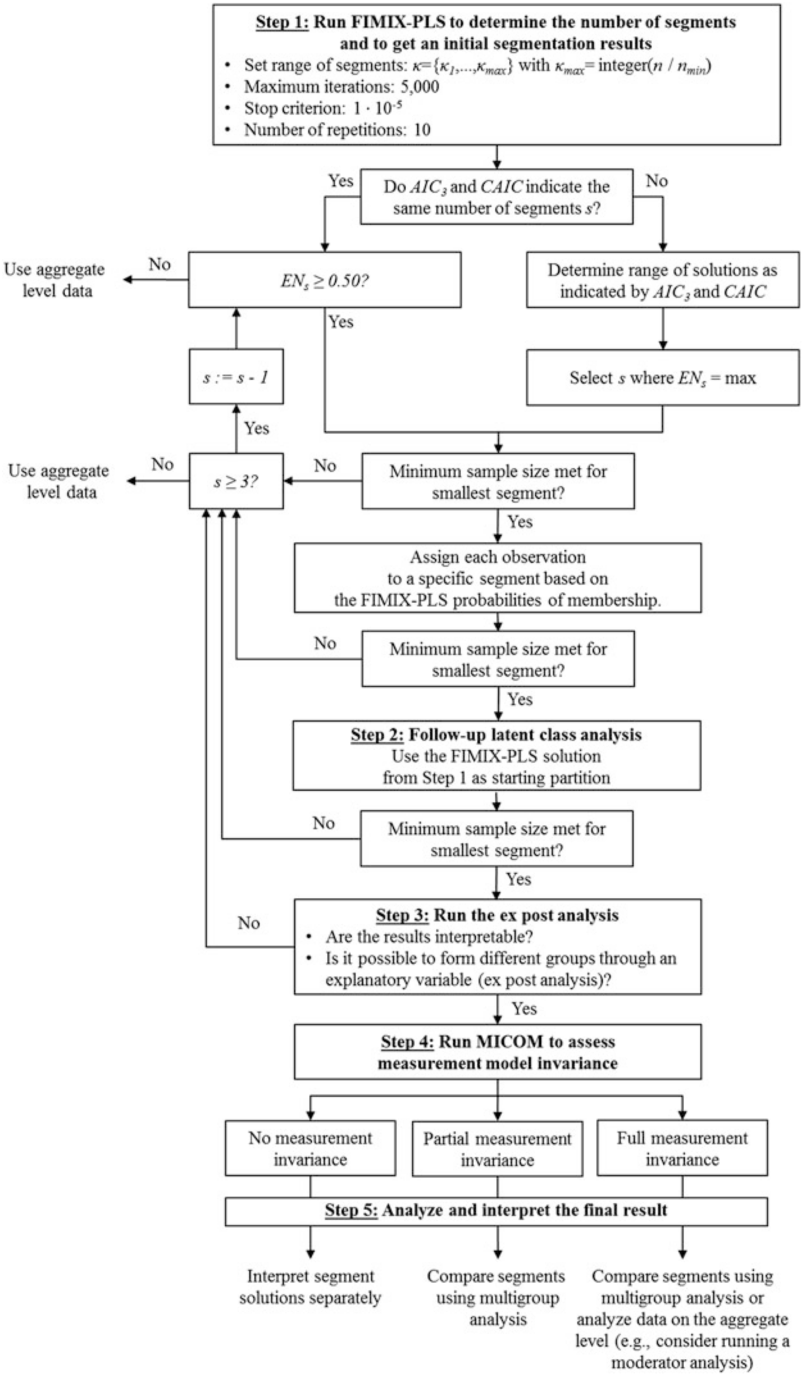


Fig. 9.1 Guidelines for uncovering and treating unobserved heterogeneity in PLS-SEM

supplement the statistical element comprising this step. Subsequently, to ensure that further assessments are based on actionable segments, researchers should create—based on the explanatory variables—observable segments and produce corresponding segment-specific PLS-SEM estimations.

Before comparing segment-specific PLS-SEM estimates, researchers must assess measurement invariance in Step 4 by means of the measurement invariance of composite models (MICOM) approach. Depending on the results of Step 4, researchers can run a multigroup analysis to compare the segment-specific results for significant differences.

9.3 Step 1: Run FIMIX-PLS

9.3.1 Basic Concept

As its name indicates, the FIMIX-PLS approach relies on the finite mixture model concept, which assumes that the overall population is a mixture of group-specific density functions. The aim of FIMIX-PLS is to disentangle the overall mixture distribution and estimate parameters (e.g., the path coefficients) of each group in a regression framework. Figure 9.2 shows an example of a mixture distribution that FIMIX-PLS aims to separate into segment-specific distributions.

FIMIX-PLS follows two steps. In the first step, the standard PLS-SEM algorithm (Lohmöller 1989) is run on the full set of data to obtain the scores of all the latent variables in the model. These latent variable scores then serve as input for a series of mixture regression analyses in the second step (Mclachlan and Peel 2000; Wedel and Kamakura 2000). The mixture regressions allow for the simultaneous probabilistic classification of observations into groups and the estimation of group-specific path

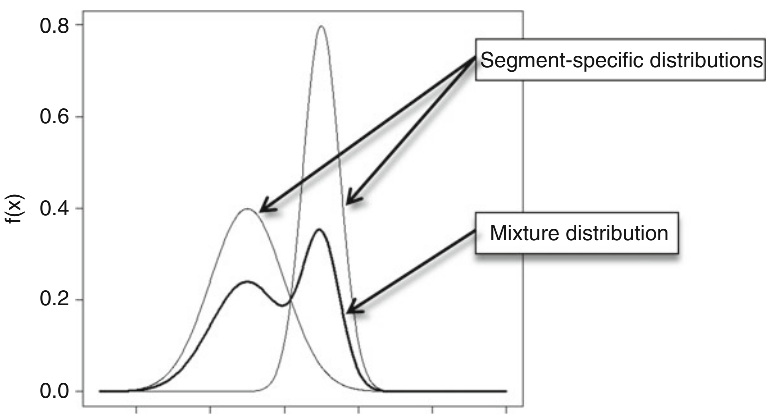


Fig. 9.2 Mixture distribution example (Hair et al. 2016, p. 66)

coefficients. While the researcher needs to explicitly define the number of segments, FIMIX-PLS offers several metrics that provide an indication of how many segments are present in the data (Hahn et al. 2002; Sarstedt et al. 2011a), which is a clear advantage of the method. Importantly, researchers can also use the FIMIX-PLS results to argue that unobserved heterogeneity is not a critical issue. In the latter case, one can interpret the aggregate data level as representing a single segment.

9.3.2 *Considerations When Using FIMIX-PLS*

9.3.2.1 Algorithm Settings

Running the FIMIX-PLS procedure requires the researcher to make several choices regarding the algorithm settings. In FIMIX-PLS, the model estimation process follows the likelihood principle, which asserts that all of the evidence in a sample that is relevant for the model parameters is contained in the likelihood function. This likelihood function is maximized by using the expectation-maximization (EM) algorithm (Dempster et al. 1977). The EM algorithm alternates between performing an expectation (E) step and a maximization (M) step. The E step creates a function for the expectation of the log-likelihood, which is evaluated by using the parameters' current estimate. The M step computes the parameters by maximizing the expected log-likelihood identified in the E step. The E and M steps are successively applied until the results stabilize. Stabilization is reached when there is no substantial improvement in the (log) likelihood value from one iteration to the next. A threshold value of 1×10^{-5} is recommended as a stop criterion to ensure that the algorithm converges at reasonably low levels of iterative changes in the log-likelihood values. When the stop criterion is set very low, the FIMIX-PLS algorithm may not converge within a reasonable time. Therefore, the researcher also needs to specify a maximum number of iterations, after which the algorithm will automatically terminate. Specifying a maximum number of 5,000 iterations ensures a sound balance between warranting acceptable computational running time and obtaining results that are precise enough.

Drawing on the EM algorithm for model estimation is attractive because it is very efficient and always converges. A downside of the EM algorithm, however, is its possible convergence in a local optimum (Steinley 2003). That is, the algorithm converges in a solution that is optimal within a neighboring set of candidate solutions but not among all possible solutions. To avoid convergence in a local optimum, researchers should run FIMIX-PLS several times for a predefined number of segments using random starting values. While simulation studies typically use 30 repetitions, empirical applications usually rely on ten repetitions. While more repetitions increase the likelihood of obtaining the global optimum FIMIX-PLS solution, they also take longer time to compute. Bearing this tradeoff in mind, we recommend using ten repetitions as this number typically yields stable results.[1] After comparing the results of multiple runs per given number of segments, the researcher chooses the solution with the highest log-likelihood value.

A further important consideration when running FIMIX-PLS involves the treatment of missing values. Kessel et al. (2010) have shown that as little as 5% missing values in one variable can cause severe problems in an FIMIX-PLS analysis when the missing values are replaced with the overall sample mean of that indicator's valid values (i.e., mean value replacement). In this case, the missing value treatment option creates a set of common scores, which FIMIX-PLS identifies as a distinct homogeneous segment. As a consequence, the number of segments will likely be over-specified, and observations that truly belong to other segments will be forced into this artificially generated one. Therefore, mean value replacement must not be used in a FIMIX-PLS context, even if there are only very few missing values in the data set. Research has put forward a wide range of alternative missing value imputation methods such as EM imputation or multiple imputation (Schafer 1997). However, as their efficacy has not been tested in a FIMIX-PLS context yet, researchers should remove all cases from the analysis that include missing values in any of the indicators used in the model (i.e., casewise deletion). While this approach also has its problems, particularly when values are missing at random (Sarstedt and Mooi 2014), it avoids the generation of an artificial segment, as is the case with mean value replacement. However, when missing values are at reasonable levels for running a PLS-SEM analysis [e.g., the number of missing values per observation and variable does not exceed 15%; (Hair et al. 2017c)], one may consider including these cases later after identifying a suitable explanatory variable in Step 3 of the procedure shown in Fig. 9.1.

Finally, the FIMIX-PLS algorithm needs to be run for alternating numbers of segments, starting with the one-segment solution. Since the number of segments is a priori unknown, researchers have to compare the solutions with the different segment numbers in terms of their statistical adequacy and interpretability. The range of possible segment numbers depends on the interplay between the sample size and the minimum sample size requirements to reliably estimate the given model (Kock and Hadaya 2017). Assume, for example, that the analysis of a certain model requires a minimum sample size of 50. When analyzing this model with 200 observations, it is not reasonable to run FIMIX-PLS with more than 4 segments. Specifically, when extracting more than 4 segments from a data set with 200 observations, at least 1 segment has less than 50 observations. If 50 observations is the minimum sample size to warrant sufficient accuracy and power given a certain degree of model complexity, extracting 5 or more segments does not produce solutions with the desired minimum sample size per extracted segment. It is therefore imperative to consider model-specific minimum segment sample size

[1] While ten repetitions achieve a sound balance between accuracy and computational time, future research should systematically explore the effect of model complexity and sample size on the number of repetitions to use.

requirements—as documented in, for example, Hair et al. (2017c)—before defining a range of segment solutions to consider in the FIMIX-PLS analysis. The largest integer gives the theoretical maximum number of segments to consider by dividing the sample size n by the minimum sample size $n_{\min}$: $\left\lfloor {n}/{n_{\min}} \right\rfloor$. However, since it is unlikely that the observations are evenly distributed across the segments, especially when the upper bound is high, considering a lower number of segments is generally preferred.

9.3.2.2 Determining the Number of Segments

One of the greatest challenges in the application of FIMIX-PLS relates to the determination of the number of segments to retain from the data (Sarstedt et al. 2011a). A misspecification can result in under- or over-segmentation and thus produce a flawed understanding of respondents' behaviors and of the ensuing managerial decisions, which is very likely to influence the specification of behaviors (Andrews and Currim 2003). Researchers can draw on a broad range of segment retention criteria to compare different segmentation solutions in terms of their model fit and avoid such misspecification. Well-known information criteria are, for example, Akaike's information criterion (AIC) (Akaike 1973), modified AIC with factor 3 (AIC_3) (Bozdogan 1994), consistent AIC (CAIC) (Bozdogan 1987), and Bayesian information criterion (BIC) (Schwarz 1978). The smaller these criteria's value, the better the segmentation solution. Hence, it is important to analyze and compare the segmentation solutions for different prespecified numbers of segments (e.g., 1–5) and to select the solution with the best information criterion outcome.

Sarstedt et al. (2011a) provide a formal representation of these criteria. Further, these authors have evaluated the efficacy of 18 different segment retention criteria in FIMIX-PLS across a broad range of data and model constellations. They demonstrate that researchers should jointly consider AIC_3 (Bozdogan 1994) and CAIC (Bozdogan 1987). More precisely, when both criteria indicate the same number of segments, this result is likely to be most appropriate. Alternative well-performing criteria are AIC with factor 4 (AIC_4) (Bozdogan 1994) and BIC. Other criteria show pronounced overestimation, such as AIC or underestimation tendencies such as minimum description length 5 [MDL_5; (Liang et al. 1992)]. Using these criteria, researchers can identify a reasonable range of segments. For example, if AIC indicates a five-segment solution, the research should consider a smaller number of segments. Conversely, if MDL_5 indicates a two-segment solution, the researcher should consider three or more segments. Table 9.1 highlights selected FIMIX-PLS information criteria and their performance.

Segment retention criteria are not a panacea for determining the most suitable number of segments in FIMIX-PLS. The relative differences in the segment retention criteria results are often marginal in terms of different numbers of segments. For example, in Rigdon et al. (2011), the average AIC_3 (CAIC) value in their industry sample 1 (2) is 31,055.31 (31,313.63), with a standard deviation of merely 303.06 (268.49)—for another example, see Navarro et al. (2011). In such a case,

Table 9.1 Selected information criteria and their performance in FIMIX-PLS (Hair et al. 2016, p. 70)

Abbreviation	Criterion	Performance in FIMIX-PLS
AIC	Akaike's information criterion	Weak performance Very strong tendency to overestimate the number of segments Can be used to determine the upper limit of reasonable segmentation solutions
AIC_3	Modified Akaike's information criterion with factor 3	Fair to good performance Tends to overestimate the number of segments Works well in combination with CAIC and BIC
AIC_4	Modified Akaike's information criterion with factor 4	Good performance Tends to over- and underestimate the number of segments
BIC	Bayesian information criterion	Good performance Tends to underestimate the number of segments Should be considered jointly with AIC_3
CAIC	Consistent Akaike's information criterion	Good performance Tends to underestimate the number of segments Should be considered jointly with AIC_3
MDL_5	Minimum description length with factor 5	Weak performance Very strong tendency to underestimate the number of segments Can be used to determine the lower limit of reasonable segmentation solutions

the criteria offer only limited means to differentiate between the segment solutions. More importantly, information criteria such as AIC_3 and CAIC do not give any indication of how well separated the segments are. For this reason, researchers are advised to consider the complementary use of the entropy-based measures, such as the entropy statistic (Ramaswamy et al. 1993):

$$EN_s = 1 - \frac{\left[\sum_{i=1}^{n}\sum_{j=1}^{s} -p_{ij}\cdot \ln p_{ij}\right]}{n\cdot \ln(s)}, \tag{9.1}$$

where p_{ij} is the probability that observation i ($i = 1, \ldots, n$) belongs to segment j ($j = 1, \ldots, s$). The EN ranges between 0 and 1; higher values indicate that more observations exhibit high probabilities of segment membership p_{ij} and thus uniquely belong to a certain segment. Sarstedt et al. (2011a, p. 52) note that "this criterion is critical to assessing whether the analysis produces well separated clusters, which is important for deriving management implications from any analysis." In line with prior research on this topic (e.g., Ringle et al. 2010a; Sarstedt and Ringle 2010), EN values of less than 0.50 indicate fuzzy segment memberships that prevent meaningful segmentation and limit the applied value of the solution.

Apart from the points above, the identified segments must meet certain standards, particularly in terms of their size. FIMIX-PLS relies on the EM algorithm, which always converges to the prespecified number of segments. While this characteristic is generally advantageous, especially compared with the problems that result from other optimization techniques, such as the Newton-Raphson method (Mclachlan and Peel 2000), it also entails two problems. First, the final solution strongly depends on the (random) starting values of the EM algorithm (Mclachlan 1988; Wedel and Kamakura 2000), which may converge in local optimum solutions. Therefore, it is of pivotal importance to run FIMIX-PLS analyses repeatedly (e.g., ten times) (Ringle et al. 2010a) for a prespecified number of segments and to select the best solution (i.e., the one with the highest likelihood value). This and similar settings should be reported to obtain insight into the specificities of the analyses and to have confidence in the conclusions. Second, the EM algorithm can "force" observations into an extraneous segment, even though they fit adequately into another segment. Such extraneous segments are usually very small, account for only a marginal portion of heterogeneity in the overall data set, and are unlikely to translate into meaningful market segmentation opportunities (Rigdon et al. 2010).

Finally, any similar data-driven approach to partitioning data should include practical considerations (e.g., Sarstedt et al. 2009), as the data can often only provide rough guidance to the number of segments to retain. Heuristics, such as information criteria or entropy measures, are fallible, as they are sensitive to data and model characteristics. For example, Becker et al. (2015) show that even low levels of collinearity in the structural model can have adverse consequences for the criteria's performance. Researchers might occasionally have a priori knowledge or a theory to rely on in making the choice. Researchers should thus ensure that the results are interpretable and meaningful. Not only must the number of segments be small enough to ensure parsimony and manageability, but each segment should also be large enough to warrant strategic attention (Kotler and Keller 2015). Finally, it is important to remember that unobserved heterogeneity can also exist within a priori formed segments (Rigdon et al. 2011).

9.4 Step 2: Follow-Up Latent Class Analysis

FIMIX-PLS offers important guidance in determining whether or not a significant level of unobserved heterogeneity is present. Furthermore, FIMIX-PLS provides a data-driven indication of the number of segments to retain from the data. Simulation studies show that FIMIX-PLS reliably reveals the existence of heterogeneity in PLS path models and correctly indicates the appropriate number of segments to retain from the data (Sarstedt et al. 2011a). At the same time, however, FIMIX-PLS is clearly limited in terms of correctly identifying the underlying segment structure that the group-specific path coefficients define (Ringle et al. 2013, 2014), especially when the path model includes formative measures (Becker et al. 2013).

Addressing these limitations, research has proposed a range of alternative latent class approaches. As a distance-based segmentation approach, Squillacciotti (2005,

2010) introduced the PLS typological path modeling procedure, which Esposito Vinzi et al. (2007, 2008) advanced by presenting the response-based procedure for detecting unit segments (REBUS-PLS). REBUS-PLS gradually reallocates observations from one segment to the other based on the objective of minimizing the model residuals.

Becker et al. (2013) further improve these latent class approaches by presenting the PLS-POS method. PLS-POS uses an improved distance measure and data reassignment procedure, which is applicable to all types of PLS path models, regardless of whether the latent variables draw on reflectively or formatively specified measurement models. Based on the distance measure, PLS-POS selects candidates for reassignment which improve the objective criterion. Becker et al. (2013) suggest using the sum of each group's sum of R^2 values as objective criterion. However, depending on the research goal, alternative objective criteria such as the sum of a specific target construct's R^2 value or the weighted sum of R^2 value(s) may be preferable. Becker et al.'s (2013) simulation study shows that PLS-POS performs well for segmentation purposes and provides favorable outcomes when compared with alternative segmentation techniques. As with FIMIX-PLS, researchers can apply PLS-POS by using this method's implementation in the software SmartPLS 3 (Ringle et al. 2015).

PLS-GAS (Ringle et al. 2013, 2014) is another versatile approach for uncovering and treating heterogeneity in measurement and structural models. This approach consists of two stages. The first stage applies a genetic algorithm that aims at finding the partition which minimizes the endogenous latent variables' unexplained variance. Implementing a genetic algorithm has the advantage that it can overcome local optimum solutions and covers a wide area of the potential search space before delivering a final best solution. In the second stage, a deterministic hill-climbing approach aims at delivering a solution that further improves the original one. The PLS-GAS method returns excellent results that usually outperform FIMIX-PLS and REBUS-PLS outcomes in particular. The downside, however, is that PLS-GAS is computationally very demanding.

For the latter reason, researchers have recently introduced PLS-IRRS (Schlittgen et al. 2016). The PLS-IRRS approach builds on Schlittgen's (2011) clusterwise robust regression. Robust regression reduces the weighting of observations with extreme values, which mitigates the influence of outliers in the data set. In the adaption of this concept for PLS-SEM-based segmentation, the weighting of observations with extreme values is not reduced. Instead, the observations are assigned to a separate segment. Hence, in PLS-IRRS groups of outliers represent distinct segments. At the same time, PLS-IRRS reduces the impact of inhomogeneous observations on the computation of segment-specific PLS-SEM solutions. Like PLS-POS and PLS-GAS, PLS-IRRS is generally applicable to all kinds of PLS path models. Moreover, initial simulation results show that PLS-IRRS performs well in terms of parameter recovery and predictive power (Schlittgen et al. 2015). The key advantage of PLS-IRRS, however, is its speed. PLS-IRRS is much faster than PLS-GAS but provides very similar results.

In light of their advantages, a combination of FIMIX-PLS with PLS-POS, PLS-GAS, or PLS-IRRS is particularly useful for treating unobserved heterogeneity as a multi-method approach in PLS-SEM. More specifically, researchers should first run FIMIX-PLS to ascertain whether or not heterogeneity is present in the data. If heterogeneity is present, the information criteria help to identify a suitable number of segments to retain. The final partition of observations that FIMIX-PLS produces serves as a starting partition for the other methods. Thereby, PLS-POS, PLS-GAS, or PLS-IRRS can further improve the FIMIX-PLS results. Combining FIMIX-PLS with one of the three other methods proves particularly useful because each method has its distinct strengths in the analysis process, for example, in terms of indicating the number of segments (FIMIX-PLS) or considering heterogeneity in the measurement models (PLS-GAS, PLS-POS, and PLS-IRRS). The PLS-POS method is a specifically suitable candidate for this kind of combination, because like FIMIX-PLS, it has been implemented in the SmartPLS 3 software package. Specifically, SmartPLS 3 allows running PLS-POS using the results of a prior FIMIX-PLS analysis in a single execution. In the following steps, we therefore refer to PLS-POS when discussing the further processing of results from Step 2 of the guidelines.

9.5 Step 3: Ex Post Analysis

The segments produced in Step 2 are—by definition—latent. Turning such a statistically derived insight into actionable understanding requires researchers to interpret the segments in terms of observable and managerially meaningful variables. They can do so by conducting an ex post analysis whose aim is to identify one or more explanatory variable(s) that match the method's partition in the best possible way (Hahn et al. 2002; Ringle et al. 2010a; Sarstedt and Ringle 2010).

Many researchers simply use available explanatory variables to characterize the latent segments (e.g., segment 1 includes 62% female customers, while segment 2 includes 24% female customers). However, PLS-POS-based ex post analysis goes far beyond a mere profiling of segments and is an integral part of any latent class analysis. In an ex post analysis, researchers need to partition the data by using an explanatory variable, or a combination of several explanatory variables, which yield a grouping of data that corresponds largely to the one produced by PLS-POS. Consider Table 9.2, which shows a sample partition of ten observations into two groups. The first six observations belong to the first latent segment, whereas the final four observations belong to the second latent segment. Partitioning the data based on the respondents' gender yields a grouping that fully corresponds to the one produced in the latent class analysis. The first latent segment comprises only females, whereas the second latent segment only comprises males. In this example, we are able to fully reproduce the latent segmentation result by means of an observable characteristic. In practice, however, reproducing latent segment structures by means of explanatory

Table 9.2 Ex post analysis example

Observation #	Latent class partition	Gender
1	1	Female
2	1	Female
3	1	Female
4	1	Female
5	1	Female
6	1	Female
7	2	Male
8	2	Male
9	2	Male
10	2	Male

variables is very challenging. As the observable characteristics do not usually match the latent segment structures well, an overlap of 60% between the PLS-POS partition and the one produced by the explanatory variable(s) can be considered satisfactory; also see Hair et al. (2017a).

To identify suitable explanatory variables, prior latent class analyses have relied on, among other methods, logistic regressions (Money et al. 2012; Wilden and Gudergan 2015) as well as on classification and regression trees (Ringle et al. 2010a; Sarstedt and Ringle 2010). More precisely, by using the latent class partition as the dependent variable and a set of observable characteristics as independent variables, these methods help to identify variables that enable the latent segments to be reproduced. Researchers may also use cross tabs (e.g., Sarstedt and Mooi 2014) to tabulate the PLS-POS partition against the partition produced by the explanatory variable with the aim of maximizing the overlap between the two (Matthews et al. 2016).

9.6 Step 4: Run the MICOM Procedure

Once the researcher has identified one or more suitable explanatory variable(s) that matches the latent class partition, the next step is to compare the segments by means of a multigroup analysis. Multigroup analysis allows researchers to test whether the numerical differences between segment-specific path coefficients are also significantly different (Hair et al. 2017a; Matthews 2018).

Prior to running a multigroup analysis, researchers need to test for measurement invariance to be confident; the group differences in model estimates do not result from either the distinctive content or meanings of the latent variables across groups or from the measurement scale. Research has proposed a variety of methods for measurement invariance testing whose applicability is, however, limited to factor-based SEM (Steenkamp and Baumgartner 1998; Vandenberg and Lance 2000). Henseler et al. (2015) introduced a procedure to assess the measurement invariance of composite models, which is consistent with PLS's nature as a composite-based approach to SEM (Sarstedt et al. 2016a). The MICOM approach involves three steps. Step 1 addresses establishing configural invariance (i.e., equal parameterization and way of estimation) to ensure that a composite has been equally specified for all the groups. A qualitative assessment of the composites' specification across all the groups must ensure the use of (1) identical indicators per measurement model, (2) identical data treatment, and (3) identical algorithm settings. Configural invariance is a precondition for the assessment of compositional invariance in Step 2 of the MICOM procedure. In the latter case, researchers must ensure that the differences in structural coefficients do not result from differences in the way the composite is formed. Compositional invariance assessment therefore involves testing whether the composite scores are created equally across groups, despite possible differences in the indicator weights. If the Step 2 results support invariance, the MICOM procedure continues with Step 3. This final step involves testing the equality of the composites' mean values and variances. In assessing whether the composite scores differ between two groups of data with regard to (1) the composition of the score vector (Step 2) and (2) the group means and variances (Steps 3a and 3b), MICOM follows a composite model logic analogous to the PLS-SEM approach (Rigdon et al. 2017). Hence, MICOM can be used for assessing the measurement invariance regardless of whether constructs have been specified reflectively or formatively.

Comparing path coefficients across groups in the course of a multigroup analysis requires establishing configural (Step 1) and compositional (Step 2) invariance, which is equivalent to partial measurement invariance. If, additionally, the composites have equal mean values and variances across the groups, full measurement invariance is supported. In this case, researchers can pool the data and interpret the results on an aggregate level. Henseler et al. (2015) provide full details of the MICOM procedure, including simulation study results and an empirical application (see also Schlägel and Sarstedt 2016). The establishment of partial measurement invariance allows comparing parameters between two or more segments for significant differences that result from the multigroup analysis. A group-specific importance-performance map analysis (IPMA) is a particularly useful tool comparing and interpreting the group-specific PLS-SEM results (Ringle and Sarstedt 2016; Schloderer et al. 2014).

9.7 Step 5: Assess the Group-Specific Solutions

Provided that Step 4 of the guidelines in Fig. 9.1 indicated at least partial measurement invariance, the researcher can test whether differences between group-specific path coefficients are statistically significant. For this purpose, one needs to conduct a multigroup analysis (Hair et al. 2017a). Technically, a multigroup analysis tests the null hypotheses that the path coefficients are not significantly different.

Standard approaches to multigroup analysis in PLS-SEM, such as the parametric approach, permutation test, and the PLS-MGA, enable testing differences between two segments. Among these approaches, Chin and Dibbern's (2010) permutation test is the most versatile approach and should be given preference (Hair et al. 2017b). When the latent class analysis involves comparing more than two segments, researchers should make use of Sarstedt et al.'s (2011b) omnibus test of group differences.[2] The omnibus test of group differences corresponds to an *F* test in that it tests for the equality of a parameter across multiple groups. The test applies a combination of bootstrapping and permutation to derive a probability value of the variance that the grouping variable explains. If this variance is significantly different from zero, researchers can conclude that at least one group-specific coefficient differs significantly from the others. To test for group-specific differences, researchers need to engage in pairwise comparisons, potentially correcting for an alpha inflation that occurs due to multiple testing.

Alternatively, if Step 4 (Fig. 9.1) did not indicate partial (or full) measurement invariance, researchers can only interpret the segment-specific PLS-SEM results in isolation without testing for significant differences. In this case, respondents may have ascribed different meanings to the construct measures, which implies that the constructs represent different concepts in the segments.

9.8 Summary

The impact of unobserved heterogeneity on PLS-SEM results can be considerable and, if not taken into account, may entail misleading interpretations (Becker et al. 2013). As a consequence, PLS-SEM analyses require the use of complementary latent class techniques, such as FIMIX-PLS, which allow testing for and treating unobserved heterogeneity (Hair et al. 2017c). Running such a latent class analysis is far from trivial, as it requires several choices that, if made incorrectly, can lead to incorrect findings, interpretations, and conclusions. For this reason, this chapter raises and discusses key aspects that are fundamental to a full and adequate understanding of how to uncover and treat unobserved heterogeneity in PLS-SEM. These include (1) the determination of the number of segments, (2) their specification by means of explanatory variables, (3) the comparison of path coefficients across segments by means of multigroup analysis, and (4) the requirements for doing so (i.e., the establishment of measurement invariance).

This chapter offers guidelines (Fig. 9.1) on how to systematically address the issue of unobserved heterogeneity following a multi-method approach, which draws

[2]The following link gives access to a Microsoft Excel file that supports computing the omnibus test of group differences for more than two segments: http://derwinchan.iwopop.com/MG-PLS.

on a combination of FIMIX-PLS and more advanced latent class techniques such as PLS-POS, PLS-GAS, or PLS-IRRS. These guidelines help researchers decide if unobserved heterogeneity affects their results and how to address these issues by partitioning their data into different segments. If unobserved heterogeneity is not a critical issue, researchers' FIMIX-PLS results substantiate that they can analyze the data on the aggregate level.

Acknowledgments This chapter builds on the articles published by Hair et al. (2016) and Matthews et al. (2016) in the *European Business Review* journal, the article by Sarstedt et al. (2016b) in *Annals of Tourism Research*, and the chapter on uncovering unobserved heterogeneity in the book on PLS-SEM advances by Hair et al. (2017a). This chapter refers to the use of the statistical software SmartPLS (http://www.smartpls.com). Ringle acknowledges a financial interest in SmartPLS.

Appendix

Table 9.3 Examples of FIMIX-PLS applications

Topic	Publication
Auction bidders' behavior	Mancha et al. (2014)
Behavior of supermarket shoppers	Teller and Gittenberger (2011)
Brand equity	Valette-Florence et al. (2011)
Brand loyalty	Loureiro (2012), Loureiro and Miranda (2011)
Brand value	Barnes and Mattson (2011),
Competitiveness in tourism destinations	Mazanec and Ring (2011)
Corporate reputation	Matthews et al. (2016), Sarstedt and Ringle (2010)
Customer satisfaction	Hahn et al. (2002), Human and Naudé (2014), Rigdon et al. (2011), Ringle et al. (2010a, b)
Dynamic capabilities	Wilden and Gudergan (2015)
Environmental orientation of firms	Mondéjar-Jiménez et al. (2015)
Environmental sustainable management	Ferrari et al. (2010)
Information systems user characteristics	Semina and Muris (2013)
SMEs' internet usage	Caniëls et al. (2015), Lenaerts and Gelderman (2015)
Knowledge sharing	Stewart Jr. et al. (2015)
Perceived value	Jiménez-Castillo et al. (2013)
Performance of organizations	Oyewobi et al. (2016)
Sport sponsorship	Alonso-Dos-Santos et al. (2016)
Stakeholder segmentation	Money et al. (2012)
Strategic marketing management	Navarro et al. (2011)
Switching costs in the ICT industry	Matzler et al. (2015)
Tourism management	Marques and Reis (2015)

References

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In B. N. Petrov & F. Csáki (Eds.), *Second international symposium on information theory* (pp. 267–281). Budapest: Académiai Kiadó.

Alonso-Dos-Santos, M., Vveinhardt, J., Calabuig-Moreno, F., & Montoro-Rios, F. J. (2016). Involvement and image transfer in sports sponsorship. *The Engineering Economist, 27*, 78–89.

Andrews, R. L., & Currim, I. S. (2003). Retention of latent segments in regression-based marketing models. *International Journal of Research in Marketing, 20*, 315–321.

Barnes, S. J., & Mattson, J. (2011). Segmenting brand value perceptions of consumers in virtual worlds: An empirical analysis using the FIMIX method. *International Journal of Online Marketing, 1*, 1–11.

Becker, J.-M., Rai, A., Ringle, C. M., & Völckner, F. (2013). Discovering unobserved heterogeneity in structural equation models to avert validity threats. *MIS Quarterly, 37*, 665–694.

Becker, J.-M., Ringle, C. M., Sarstedt, M., & Völckner, F. (2015). How collinearity affects mixture regression results. *Marketing Letters, 26*, 643–659.

Bozdogan, H. (1987). Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions. *Psychometrika, 52*, 345–370.

Bozdogan, H. (1994). Mixture-model cluster analysis using model selection criteria in a new information measure of complexity. In H. Bozdogan (Ed.), *Proceedings of the first US/Japan conference on frontiers of statistical modelling: An information approach* (pp. 69–113). Dordrecht: Kluwer Academic.

Byrne, B. M. (2016). *Structural equation modeling with AMOS: Basic concepts, applications, and programming (multivariate applications)*. New York: Routledge.

Caniëls, M. C. J., Lenaerts, H. K. L., & Gelderman, C. J. (2015). Explaining the internet usage of SMEs: The impact of market orientation, behavioural norms, motivation and technology acceptance. *Internet Research, 25*, 358–377.

Chin, W. W., & Dibbern, J. (2010). A permutation based procedure for multi-group PLS analysis: Results of tests of differences on simulated data and a cross cultural analysis of the sourcing of information system services between Germany and the USA. In V. Esposito Vinzi, W. W. Chin, J. Henseler, & H. Wang (Eds.), *Handbook of partial least squares: Concepts, methods and applications*, *Springer handbooks of computational statistics series* (Vol. II, pp. 171–193). Heidelberg: Springer.

Dempster, A. P., Laird, N., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society B, 39*, 1–38.

Diamantopoulos, A., & Siguaw, J. A. (2000). *Introducing LISREL*. Thousand Oaks, CA: Sage.

Esposito Vinzi, V., Ringle, C. M., Squillacciotti, S., & Trinchera, L. (2007). *Capturing and treating unobserved heterogeneity by response based segmentation in PLS path modeling: A comparison of alternative methods by computational experiments*. ESSEC Research Center, Working Paper No. 07019, Cergy Pontoise Cedex.

Esposito Vinzi, V., Trinchera, L., Squillacciotti, S., & Tenenhaus, M. (2008). REBUS-PLS: A response-based procedure for detecting unit segments in PLS path modelling. *Applied Stochastic Models in Business and Industry, 24*(5), 439–458.

Ferrari, G., Mondéjar-Jiménez, J., & Vargas-Vargas, M. (2010). Environmental sustainable management of small rural tourist enterprises. *International Journal of Environmental Research, 4*, 407–414.

Hahn, C., Johnson, M. D., Herrmann, A., & Huber, F. (2002). Capturing customer heterogeneity using a finite mixture PLS approach. *Schmalenbach Business Review, 54*, 243–269.

Hair, J. F., Sarstedt, M., Ringle, C. M., & Mena, J. A. (2012). An assessment of the use of partial least squares structural equation modeling in marketing research. *Journal of the Academy of Marketing Science, 40*(3), 414–433. doi:10.1007/s11747-011-0261-6.

Hair, J. F., Sarstedt, M., Matthews, L., & Ringle, C. M. (2016). Identifying and treating unobserved heterogeneity with FIMIX-PLS: Part I – Method. *European Business Review, 28*(1), 63–76.

Hair, J. F., Sarstedt, M., Ringle, C. M., & Gudergan, S. P. (2017a). *Advanced issues in partial least squares structural equation modeling (PLS-SEM)*. Thousand Oaks, CA: Sage.

Hair, J. F., Hollingsworth, C. L., Randolph, A. B., & Chong, A. Y. L. (2017b). An updated and expanded assessment of PLS-SEM in information systems research. *Industrial Management & Data Systems, 117*, 442–458.

Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2017c). *A primer on partial least squares structural equation modeling (PLS-SEM)* (2nd ed.). Thousand Oaks, CA: Sage.

Hair, J. F., Hult, G. T. M., Ringle, C. M., Sarstedt, M., & Thiele, K. O. (2017d). Mirror, mirror on the wall: A comparative evaluation of composite-based structural equation modeling methods. *Journal of the Academy of Marketing Science* (in press).

Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science, 43*(1), 115–135. doi:10.1007/s11747-014-0403-8.

Human, G., & Naudé, P. (2014). Heterogeneity in the quality–satisfaction–loyalty framework. *Industrial Marketing Management, 43*, 920–928.

Jedidi, K., Jagpal, H. S., & Desarbo, W. S. (1997). Finite-mixture structural equation models for response-based segmentation and unobserved heterogeneity. *Marketing Science, 16*(1), 39–59.

Jiménez-Castillo, D., Sánchez-Fernández, R., & Iniesta-Bonillo, M. Á. (2013). Segmenting university graduates on the basis of perceived value, image and identification. *International Review on Public and Nonprofit Marketing, 10*, 235–252.

Kessel, F., Ringle, C. M., & Sarstedt, M. (2010). *On the impact of missing values on model selection in FIMIX-PLS*. Proceedings of the 2010 INFORMS Marketing Science Conference, Cologne.

Kock, N., & Hadaya, P. (2017). Minimum sample size estimation in PLS-SEM: The inverse square root and gamma-exponential methods. *Information Systems Journal* (in press).

Kotler, P. (1989). From mass marketing to mass customization. *Planning Review, 17*, 10–47.

Kotler, P., & Keller, K. L. (2015). *Marketing management*. Upper Saddle River, NJ: Pearson/Prentice-Hall.

Lenaerts, H. K. L., & Gelderman, C. J. (2015). Explaining the internet usage of SMEs. *Internet Research: Electronic Networking Applications and Policy, 25*, 358–377.

Liang, Z., Jaszak, R. J., & Coleman, R. E. (1992). Parameter estimation of finite mixtures using the EM algorithm and information criteria with applications to medical image processing. *IEEE Transactions on Nuclear Science, 39*, 1126–1133.

Lohmöller, J.-B. (1989). *Latent variable path modeling with partial least squares*. Berlin: Springer.

Loureiro, S. (2012). Love and loyalty in car brands: Segmentation using finite mixture partial least squares. In A. W. Gaul, A. Geyer-Schulz, L. Schmidt-Thieme, & J. Kunze (Eds.), *Challenges at the interface of data analysis, computer science, and optimization: Proceedings of the 34th Annual Conference of the Gesellschaft für Klassifikation, Karlsruhe, July 21–23, 2010* (pp. 503–510). Berlin: Springer.

Loureiro, S. M. C., & Miranda, F. J. (2011). Brand equity and brand loyalty in the internet banking context: FIMIX-PLS market segmentation. *Journal of Service Science and Management, 4*, 476–485.

Mancha, R., Leung, M. T., Clark, J., & Sun, M. (2014). Finite mixture partial least squares for segmentation and behavioral characterization of auction bidders. *Decision Support Systems, 57*, 200–211.

Marques, C., & Reis, E. (2015). How to deal with heterogeneity among tourism constructs? *Annals of Tourism Research, 52*, 172–174.

Matthews, L. (2018). Applying multi-group analysis in PLS-SEM: A step-by-step process. In H. Latan & R. Noonan (Eds.), *Partial least squares structural equation modeling: Basic concepts, methodological issues and applications*. New York: Springer.

Matthews, L., Sarstedt, M., Hair, J. F., & Ringle, C. M. (2016). Identifying and treating unobserved heterogeneity with FIMIX-PLS: Part II – A case study. *European Business Review, 28*(2), 208–224. doi:10.1108/EBR-09-2015-0095.

Matzler, K., Strobl, A., Thurner, N., & Füller, J. (2015). Switching experience, customer satisfaction, and switching costs in the ICT industry. *Journal of Service Management, 26*, 117–136.

Mazanec, J. A., & Ring, A. (2011). Tourism destination competitiveness: Second thoughts on the world economic forum reports. *Tourism Economics, 17*, 725–751.

Mclachlan, G. J. (1988). On the choice of starting values for the EM algorithm in fitting mixture models. *Journal of the Royal Statistical Society Series D (The Statistician), 37*, 417–425.

Mclachlan, G. J., & Peel, D. (2000). *Finite mixture models*. New York: Wiley.

Mondéjar-Jiménez, J., Segarra-Oña, M., Peiró-Signes, Á., Payá-Martínez, A. M., & Sáez-Martínez, F. J. (2015). Segmentation of the Spanish automotive industry with respect to the environmental orientation of firms: Towards an ad-hoc vertical policy to promote eco-innovation. *Journal of Cleaner Production, 86*, 238–244.

Money, K. G., Hillenbrand, C., Henseler, J., & Da Camara, N. (2012). Exploring unanticipated consequences of strategy amongst stakeholder segments: The case of a European revenue service. *Long Range Planning, 45*, 395–423.

Navarro, A., Acedo, F. J., Losada, F., & Ruzo, E. (2011). Integrated model of export activity: Analysis of heterogeneity in managers' orientations and perceptions on strategic marketing management in foreign markets. *Journal of Marketing Theory and Practice, 19*, 187–204.

Oyewobi, L. O., Windapo, A. O., & Rotimi, J. O. B. (2016). Environment, competitive strategy, and organizational characteristics: A path analytic model of construction organizations' performance in South Africa. *Canadian Journal of Administrative Sciences/Revue Canadienne des Sciences de l'Administration, 33*, 213–226.

Ramaswamy, V., Desarbo, W. S., Reibstein, D. J., & Robinson, W. T. (1993). An empirical pooling approach for estimating marketing mix elasticities with PIMS data. *Marketing Science, 12*, 103–124.

Rigdon, E. E. (1998). Structural equation modeling. In G. A. Marcoulides (Ed.), *Modern methods for business research* (pp. 251–294). Mahwah, NJ: Erlbaum.

Rigdon, E. E. (2005). Structural equation modeling: Nontraditional alternatives. In B. Everitt & D. Howell (Eds.), *Encyclopedia of statistics in behavioral science* (pp. 1934–1941). New York: Wiley.

Rigdon, E. E., Ringle, C. M., & Sarstedt, M. (2010). Structural modeling of heterogeneous data with partial least squares. In N. K. Malhotra (Ed.), *Review of marketing research* (pp. 255–296). Armonk, NY: Sharpe.

Rigdon, E. E., Ringle, C. M., Sarstedt, M., & Gudergan, S. P. (2011). Assessing heterogeneity in customer satisfaction studies: Across industry similarities and within industry differences. *Advances in International Marketing, 22*(1), 169–194.

Rigdon, E. E., Sarstedt, M., & Ringle, C. M. (2017). On comparing results from CB-SEM and PLS-SEM: Five perspectives and five recommendations. *Marketing ZFP - Journal of Research and Management* (in press).

Ringle, C. M., & Sarstedt, M. (2016). Gain more insight from your PLS-SEM results: The importance-performance map analysis. *Industrial Management & Data Systems, 116*(9), 1865–1886.

Ringle, C. M., Sarstedt, M., & Mooi, E. A. (2010a). Response-based segmentation using finite mixture partial least squares: Theoretical foundations and an application to American Customer Satisfaction Index data. *Annals of Information Systems, 8*, 19–49.

Ringle, C. M., Wende, S., & Will, A. (2010b). Finite mixture partial least squares analysis: Methodology and numerical examples. In V. Esposito Vinzi, W. W. Chin, J. Henseler, & H. Wang (Eds.), *Handbook of partial least squares: Concepts, methods and applications*, *Springer handbooks of computational statistics series* (Vol. II, pp. 195–218). New York: Springer.

Ringle, C. M., Sarstedt, M., Schlittgen, R., & Taylor, C. R. (2013). PLS path modeling and evolutionary segmentation. *Journal of Business Research, 66*, 1318–1324.

Ringle, C. M., Sarstedt, M., & Schlittgen, R. (2014). Genetic algorithm segmentation in partial least squares structural equation modeling. *OR Spectrum, 36*(1), 251–276.
Ringle, C. M., Wende, S., & Becker, J.-M. (2015). *SmartPLS 3*. Bönningstedt: SmartPLS GmbH. Retrieved from http://www.smartpls.com.
Sarstedt, M. (2008). A review of recent approaches for capturing heterogeneity in partial least squares path modelling. *Journal of Modelling in Management, 3*, 140–161.
Sarstedt, M., & Mooi, E. A. (2014). *A concise guide to market research: The process, data, and methods using IBM SPSS statistics*. Berlin: Springer.
Sarstedt, M., & Ringle, C. M. (2010). Treating unobserved heterogeneity in PLS path modelling: A comparison of FIMIX-PLS with different data analysis strategies. *Journal of Applied Statistics, 37*, 1299–1318.
Sarstedt, M., Schwaiger, M., & Ringle, C. M. (2009). Do we fully understand the critical success factors of customer satisfaction with industrial goods? – Extending Festge and Schwaiger's model to account for unobserved heterogeneity. *Journal of Business Market Management, 3*(3), 185–206.
Sarstedt, M., Becker, J.-M., Ringle, C. M., & Schwaiger, M. (2011a). Uncovering and treating unobserved heterogeneity with FIMIX-PLS: Which model selection criterion provides an appropriate number of segments? *Schmalenbach Business Review, 63*, 34–62.
Sarstedt, M., Henseler, J., & Ringle, C. M. (2011b). Multi-group analysis in partial least squares (PLS) path modeling: Alternative methods and empirical results. In M. Sarstedt, M. Schwaiger, & C. R. Taylor (Eds.), *Advances in international marketing* (Vol. 22, pp. 195–218). Bingley: Emerald. doi:10.1108/S1474-7979(2011)0000022012.
Sarstedt, M., Hair, J. F., Ringle, C. M., Thiele, K. O., & Gudergan, S. P. (2016a). Estimation issues with PLS and CBSEM: Where the bias lies! *Journal of Business Research, 69*(10), 3998–4010. doi:10.1016/j.jbusres.2016.06.007.
Sarstedt, M., Ringle, C. M., & Gudergan, S. P. (2016b). Guidelines for treating unobserved heterogeneity in tourism research: A comment on Marques and Reis (2015). *Annals of Tourism Research, 57*, 279–284. doi:10.1016/j.annals.2015.10.006.
Schafer, J. L. (1997). *Analysis of incomplete multivariate data*. London: Chapman and Hall.
Schlägel, C., & Sarstedt, M. (2016). Assessing the measurement invariance of the four-dimensional cultural intelligence scale across countries: A composite model approach. *European Management Journal, 34*(6), 633–649.
Schlittgen, R. (2011). A weighted least-squares approach to clusterwise regression. *Advances in Statistical Analysis, 95*, 205–217.
Schlittgen, R., Ringle, C. M., Sarstedt, M., & Becker, J.-M. (2015). Segmentation of PLS path models by iterative reweighted regressions. In J. Henseler, C. M. Ringle, J. L. Roldán, & G. Cepeda (Eds.), Second international symposium on partial least squares path modeling: The conference for PLS users, Seville.
Schlittgen, R., Ringle, C. M., Sarstedt, M., & Becker, J.-M. (2016). Segmentation of PLS path models by iterative reweighted regressions. *Journal of Business Research, 69*(10), 4583–4592.
Schloderer, M. P., Sarstedt, M., & Ringle, C. M. (2014). The relevance of reputation in the nonprofit sector: The moderating effect of socio-demographic characteristics. *International Journal of Nonprofit and Voluntary Sector Marketing, 19*, 110–126.
Schwarz, G. (1978). Estimating the dimensions of a model. *The Annals of Statistics, 6*, 461–464.
Semina, H., & Muris, C. (2013). Segmentation of information systems users: The finite mixture partial least squares method. *Journal of Organizational and End User Computing, 25*, 1–26.
Squillacciotti, S. (2005). Prediction oriented classification in PLS path modeling. In T. Aluja, J. Casanovas, V. Esposito Vinzi, & M. Tenenhaus (Eds.), *PLS & marketing: Proceedings of the 4th international symposium on PLS and related methods* (pp. 499–506). Paris: DECISIA.
Squillacciotti, S. (2010). Prediction-oriented classification in PLS path modeling. In V. Esposito Vinzi, W. W. Chin, J. Henseler, & H. Wang (Eds.), *Handbook of partial least squares: Concepts, methods and applications* (pp. 219–233). New York: Springer.
Steenkamp, J.-B. E. M., & Baumgartner, H. (1998). Assessing measurement invariance in cross-national consumer research. *Journal of Consumer Research, 25*, 78–107.

Steinley, D. (2003). Local optima in K-means clustering: What you don't know may hurt you. *Psychological Methods, 8*, 294–304.

Stewart, W. H., Jr., May, R. C., & Ledgerwood, D. E. (2015). Do you know what I know? Intent to share knowledge in the US and Ukraine. *Management International Review, 55*, 737–773.

Teller, C., & Gittenberger, E. (2011). Patronage behaviour of elderly supermarket shoppers: Antecedents and unobserved heterogeneity. *The International Review of Retail, Distribution and Consumer Research, 21*, 483–499.

Valette-Florence, P., Guizanib, H., & Merunka, D. (2011). The impact of brand personality and sales promotions on brand equity. *Journal of Business Research, 64*, 24–28.

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods, 3*, 4–70.

Völckner, F., Sattler, H., Hennig-Thurau, T., & Ringle, C. M. (2010). The role of parent brand quality for service brand extension success. *Journal of Service Research, 13*(4), 379–396.

Wedel, M., & Kamakura, W. A. (2000). *Market segmentation: Conceptual and methodological foundations*. Boston, MA: Kluwer.

Wilden, R., & Gudergan, S. P. (2015). The impact of dynamic capabilities on operational marketing and technological capabilities: Investigating the role of environmental turbulence. *Journal of the Academy of Marketing Science, 43*, 181–199.

Wold, H. (1975). Path models with latent variables: The NIPALS approach. In H. M. Blalock et al. (Eds.), *Quantitative sociology: International perspectives on mathematical and statistical modeling* (pp. 307–357). New York: Academic Press.

Wold, H. (1982). Soft modeling: The basic design and some extensions. In K. G. Jöreskog & H. Wold (Eds.), *Systems under indirect observations: Part II* (pp. 1–54). Amsterdam: North-Holland.

Chapter 10
Applying Multigroup Analysis in PLS-SEM: A Step-by-Step Process

Lucy Matthews

Abstract This book chapter identifies the importance and different uses for multigroup analysis, such as research interests in cross-cultural or gender differences. Multigroup analysis via partial least squares structural equations modeling, which tests a single structural relationship at a time, is an effective way to evaluate moderation across multiple relationships versus standard moderation. Step-by-step instructions and guidelines using SmartPLS 3.0 are provided using a sample dataset. The instructions include an assessment of the measurement characteristics of the constructs by including the MICOM procedure, which adds an additional level of accuracy. Examples of both positive and negative outcomes as well as potential solutions to problems are provided in order to help users understand how to apply multigroup analysis to their own dataset. By using multigroup analysis, researchers are able to uncover differences of subsamples within the total population that is not evident when examined as a whole. Researchers having the ability to run multigroup analysis considerably improve the likelihood of identifying significant and meaningful differences in various relationships across group-specific results.

10.1 Introduction

Many research studies report their findings based on an analysis of a single population. Unfortunately, studies that pool data as a single population fail to assess whether there are significant differences across two or more subgroups of data (Chin and Dibbern 2010). As a result, interpreting results from a single population can be misleading (Sarstedt et al. 2016a). But if categorical moderating variables are available in the dataset, group-specific path coefficient estimates that are significantly different can be identified efficiently, thereby accounting for observed heterogeneity (Sarstedt et al. 2011) and minimizing the potential for misrepresentation of the results (Sarstedt et al. 2009).

L. Matthews (✉)
Middle Tennessee State University, Murfreesboro, TN, USA
e-mail: lucy.matthews@mtsu.edu

H. Latan, R. Noonan (eds.), *Partial Least Squares Path Modeling*,
DOI 10.1007/978-3-319-64069-3_10

Multigroup analysis (MGA) or between-group analysis as applied using partial least squares structural equations modeling (PLS-SEM) is a means of testing predefined data groups to determine if there are significant differences in group-specific parameter estimates (e.g., outer weights, outer loadings, and path coefficients) (Hair et al. 2014a; Henseler and Chin 2010). By applying MGA, researchers are therefore able to test for differences between two identical models for different groups. The ability to identify the presence or absence of multigroup differences can be based on either a bootstrapping or permutation result for every group. Partial least squares structural equation modeling multigroup analysis (PLS-MGA) can be instrumental, therefore, in identifying differences among a priori-specified groups within the dataset (e.g., Hair et al. 2014a; Horn and McArdle 1992; Keil et al. 2000).

MGA is particularly useful for globally focused research, such as cross-cultural studies. For example, the method has been used to compare antecedents of market orientation across three countries (Brettel et al. 2008), to test the determinants and outcomes of cultural intelligence (Schlagel and Sarstedt 2016), and to examine company stakeholder orientation in five European countries (Patel et al. 2016). Alternatively, MGA has been used to understand the differences between consumers with high vs. low tendency toward loyalty (Picon-Berjoyo et al. 2016). This type of analysis enabled researchers to uncover differences between groups such as the low impact of switching costs for individuals exhibiting high loyalty (Picon-Berjoyo et al. 2016). Additionally, by gaining insight into group differences, a more accurate assessment is conducted, and strategy implementation based on the outcomes can be more specific for the heterogeneous groups in the data. Finally, the differences can highlight the error associated with incorrectly treating these subpopulations as a single homogeneous group (Schlagel and Sarstedt 2016).

MGA via PLS-SEM is an efficient way to assess moderation across multiple relationships as opposed to standard moderation, which examines a single structural relationship at a time (Hair et al. 2010, 2011, 2012c). According to Hair et al. (2014a, p. 246) "...this approach offers a more complete picture of the moderator's influence on the analysis results as the focus shifts from examining its impact on one specific model relationship to examining its impact on all model relationships."

Continuous moderators are relatively easy to examine in PLS-SEM, but deserve special attention. They are often measured with multiple items, which will increase the predictive validity compared to single-item measurements (Diamantopoulous et al. 2012; Sarstedt et al. 2016b). Within the context of moderation, this can be particularly problematic as moderation is usually associated with rather limited effect sizes (Aguinis et al. 2005). As a result, any lack of predictive power makes it more difficult to identify significant relationships. Moreover, when modeling moderating effects, the measurement model construct is contained in the model twice. The construct is the moderator variable itself in addition to being in the interaction term. The result amplifies the limitations of single-item measurement when used to execute moderation.

10.2 Overview of Steps for Running MGA in PLS-SEM

Comparison of group-specific outcomes can be undertaken in three simple to apply steps, thus increasing the rigor of the data analysis and reducing misleading results. An overview of the steps (Fig. 10.1) is provided as an introduction to the process that draws upon information from Hair et al. (2014a). The overview is followed by an in-depth discussion of each step, including examples for an easy application to your own dataset. To facilitate the discussion, the SmartPLS software is used to describe the process (Ringle et al. 2015).

Step 1 involves generating data groups that are based on the categorical variable of interest [e.g., gender (Rutherford et al. 2011), country of origin (Brettel et al. 2008), urban vs. rural (Rasoolimanesh et al. 2016)]. Data groups are generated in SmartPLS by double-clicking on the data for the model of interest and selecting "Generate Data Groups." A name can be specified for the group being generated, and then, the group is established by selecting the categorical variable of interest from your dataset. For example, if your theory or judgment suggests that males and females produce different results, then your analysis would be set up to examine gender. For more extensive analyses, more than one categorical variable can be selected (e.g., gender and marital status), which would create multiple outcome groups (single female, single male, married females, married males, etc.). For the example in this chapter, however, a single categorical variable will be used.

After specifying the variable of interest, the data groups are generated. Output is provided on a separate tab labeled "data groups." The groups are listed based on the coding of your data. Additionally, the number of records for each data group is provided. Each line item can be edited to rename the subgroup a more identifiable name [e.g., female, rather than gender (1.0)]. Once the data is subdivided, it is important to confirm that the new subgroups are large enough and comparable in size so as not to introduce error (Becker et al. 2013; Hair et al. 2014a). The minimum sample size recommendation in PLS-SEM has differing views by researchers. One view is that the number of responses for each subgroup should equal (or be comparable to) the sample size recommendations for a statistical power of 80%, as recommended by Cohen (1992) and Hair et al. (2014a). Groups with fewer observations than that recommended for a statistical power of 80% in most situations should not be used (Table 10.1).

Alternatively, Kock and Hadaya (2016) analyze the gamma-exponential method and the inverse square root method. They demonstrate that while the gamma-exponential method is much more complex of an application, for PLS-SEM users who are not methodological researchers, the inverse square root method may be a simpler equation for minimum sample size estimations at the early stage of the research design (Table 10.2) (Kock and Hadaya 2016). Although the method leads to a small overestimation, the slight imperfection allows for a safe minimum sample size (Kock and Hadaya 2016).

Step 2 involves using the three-step procedure to analyze the measurement invariance of composite models (MICOM) (Henseler et al. 2016). Measurement

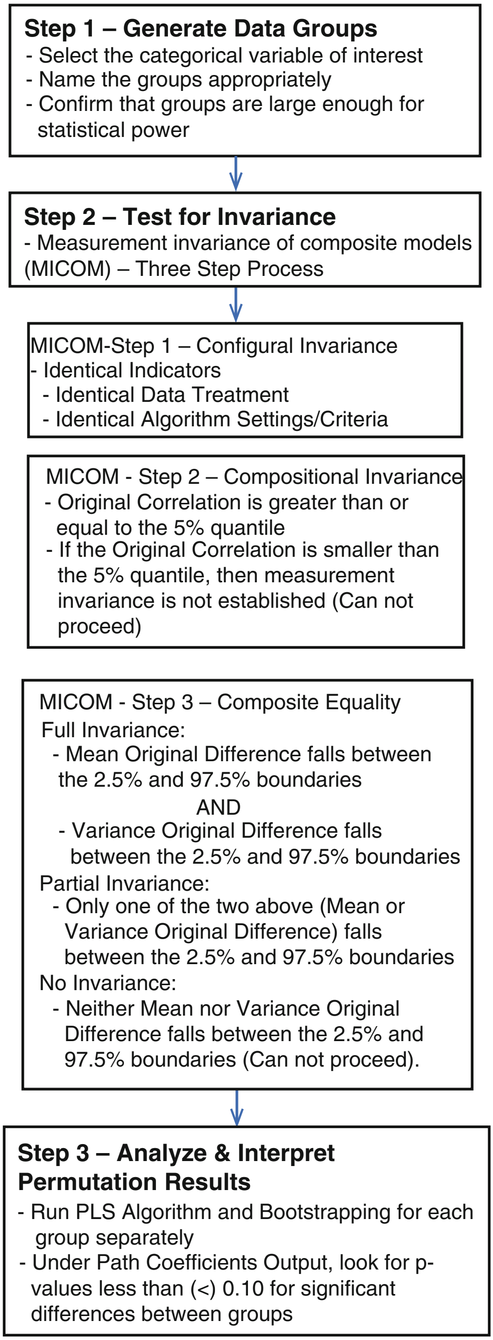


Fig. 10.1 Guidelines for running MGA in PLS-SEM

Table 10.1 Sample size recommendation in PLS-SEM for a statistical power of 80% (Cohen 1992; Hair et al. 2014a)

Maximum number of arrows pointing at a construct	Significance level											
	1%				5%				10%			
	Minimum R^2											
	0.10	0.25	0.50	0.75	0.10	0.25	0.50	0.75	0.10	0.25	0.50	0.75
2	158	75	47	38	110	52	33	26	88	41	26	21
3	176	84	53	42	124	59	38	30	100	48	30	25
4	191	91	58	46	137	65	42	33	111	53	34	27
5	205	98	62	50	147	70	45	36	120	58	37	30
6	217	103	66	53	157	75	48	39	128	62	40	32
7	228	109	69	56	166	80	51	41	136	66	42	35
8	238	114	73	59	174	84	54	44	143	69	45	37
9	247	119	76	62	181	88	57	46	150	73	47	39
10	256	123	79	64	189	91	59	48	156	76	49	41

Table 10.2 Alternative sample size recommendation in PLS-SEM using inverse square root method (Kock and Hadaya 2016)

Maximum number of arrows pointing at a construct	Minimum R^2 in the model			
	0.10	0.25	0.50	0.75
2	110	52	33	26
3	124	59	38	30
4	137	65	42	33
5	147	70	45	36
6	157	75	48	39
7	166	80	51	41
8	174	84	54	44
9	181	88	57	46
10	189	91	59	48

invariance (also referred to as equivalence) is a means of determining if the measurement models specify measures of the same attribute under different conditions (Henseler et al. 2015, 2016). This is a critical issue that must be addressed in MGA. Testing for measurement invariance determines "whether or not, under different conditions of observing and studying phenomena, measurement models yield measures of the same attribute" (Henseler et al. 2015, p. 117). When measurement invariance is established, researchers can conclude that different model estimation parameters are not the result of the distinct content or meaning of the latent variables/constructs that comprise the measurement model of any one group (Chin and Dibbern 2010; Henseler et al. 2016). Additionally, by not establishing invariance in the measurement model constructs, measurement error may be introduced leading to biased results (Hult et al. 2008). Therefore, when analyzing differences between groups, type II errors are minimized (Hult et al. 2008), and the resulting differences are the result of actual group-specific differences in the parameters and not measurement invariance (Henseler et al. 2016). The MICOM procedure compares group parameters and identifies if there is no measurement invariance, partial measurement invariance, or full measurement invariance (Henseler et al. 2016).

In *Step 3*, results of the statistical tests for multigroup comparisons are assessed. A number of approaches can be used to compare the path coefficients of the group SEMs. Three tests are included in the SmartPLS MGA (PLS-MGA) option—Henseler et al.'s (2009) PLS-MGA procedure, parametric, and Welch-Satterthwaite. A fourth approach to making group comparisons—permutation (Hair et al. 2014a; Henseler et al. 2016)—is a separate option in the SmartPLS software. The parametric test results tend to be liberal and subject to type I errors and are also limited because they are based on distribution assumptions that are not consistent with the nonparametric PLS-SEM method (Hair et al. 2017b; Sarstedt et al. 2011). The Welch-Satterthwaite test is a variant of the parametric test, but does not assume equal variances when comparing the means of two groups. The Henseler et al.'s PLS-MGA procedure (Henseler et al. 2009) and the permutation test are both nonparametric approaches. The Henseler et al.'s PLS-MGA procedure (Henseler et al. 2009) approach is included in the regular multigroup option. The Henseler et al.'s PLS-MGA procedure (Henseler et al. 2009) derives a probability value for a one-tailed test by comparing each bootstrap estimate of one group to all the bootstrap estimates of the same parameter in the other group (Hair et al. 2011). While considered to be an appropriate test, the results may be a bit challenging to interpret due to the one-tailed test. Moreover, since bootstrap distributions are not necessarily symmetrical, the Henseler et al.'s PLS-MGA procedure (Henseler et al. 2009) cannot be used to test two-tailed hypotheses. In contrast, the permutation test is a separate option and is run during Step 2 of the analysis as part of the test for measurement invariance. The output of the path coefficients from the measurement invariance option is another means of comparing the path coefficients of the subgroups. The permutation test is more conservative than the parametric test and controls well for type I error. Moreover, most researchers recommend the permutation test (Hair et al. 2017b); therefore, that approach is examined in this chapter.

10.3 Example Application of PLS-MGA

As a means of providing clarity for the execution of these steps, examples are provided for each step. The examples include output and explanation for the results that are the primary focus of the analysis. The next section revisits each step in the PLS-MGA process and provides specific details and interpretation.

10.3.1 Step 1: Generate Data Groups

Before executing an MGA, you must generate groups in your data. When you do this, the analysis is able to statistically assess the differences between the group-specific parameters, most often path coefficients resulting from different subpopulations (Brettel et al. 2008; Grewal et al. 2008). This procedure enables researchers to evaluate observed heterogeneity in model relationships (Lohmöller 1989).

MGA is similar to moderator analysis where the moderator is a categorical variable, and it is anticipated that the moderator will affect at least one and perhaps all of the model relationships (Sarstedt et al. 2011). The models in Fig. 10.2 examine the differences between female and male sales representatives. The question to be investigated is "Do the sales roles of females and males differ with regard to autonomy, skill discrepancy, and cognitive engagement?"

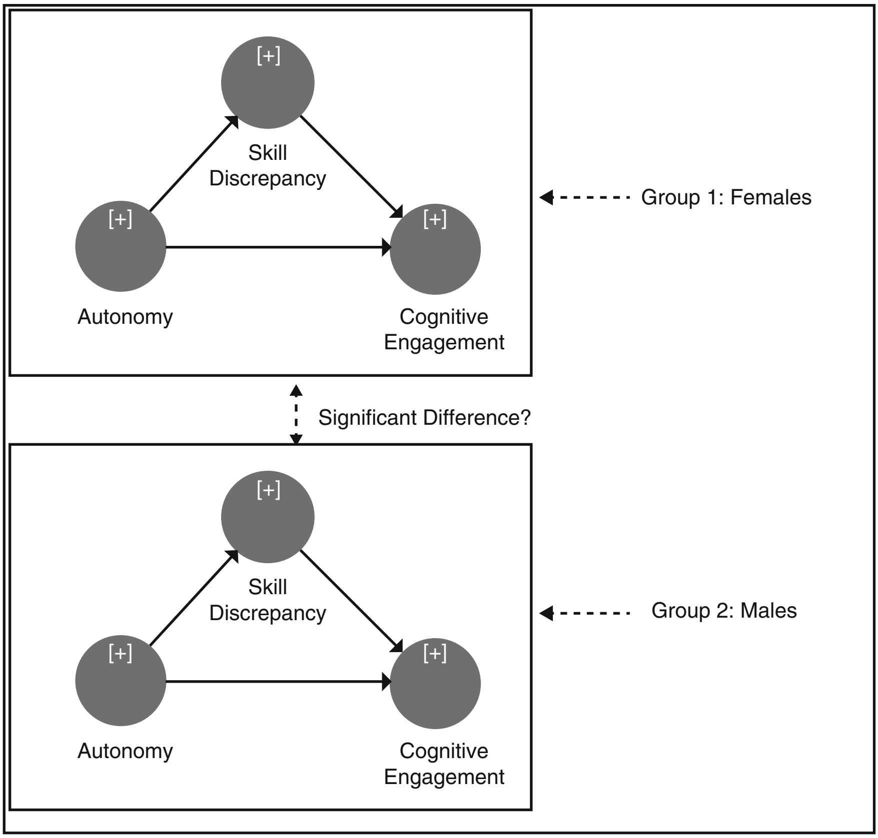


Fig. 10.2 Categorical variable PLS-MGA example

The data used for the examples in this chapter was collected via a Qualtrics online panel of business-to-business salespeople. After removal of one outlier and two straight-line responses, the final sample size is 235. The outlier identified himself or herself as a business-to-business salesperson, but when responding to the number of customers, these respondents stated 10,000 customers while the remaining respondents identified a number in the range of 1–400. Since a portion of the research is related to levels of customer service, this participant was identified as being more closely related to retail rather than business to business and was therefore removed. The questionnaire employed established scales, when available. Modifications were made to adapt the scales to the context of sales.

Recall that the sample sizes of the subpopulations must be large enough to meet statistical power guidelines. Therefore, groups that do not meet sample size recommendations should not be utilized. You may consider combining one or more smaller groups with another group if the groups exhibit similar characteristics.

For this example, the sample size for the female subgroup is 101 and for the male subgroup is 134. Each of these subpopulations exceeds the minimum for the theoretical model in Fig. 10.2 that has two arrows pointing at a construct—i.e., cognitive engagement (10% with a minimum R^2 of 0.10 = 88). In order to exceed the minimum R^2 of 0.10 at a 5% significance level, both the male and female subgroups would need to exceed 110 (Hair et al. 2014a). Ultimately, subsamples of 158 for both males and females would provide a significance level of 1%.

The two subpopulations (male and female) meet the minimum sample size criteria, but are not the same size. While two subpopulations do not have to be exactly the same size, they do need to be comparable in size. The guideline to consider regarding group sample size differences is when one group is more than 50% larger than the other, the difference is likely to bias the results of the statistical test of differences (Hair et al. 2016a). The recommended procedure when confronted with groups that differ substantially in sizes, therefore, is to randomly withdraw respondents from the larger subgroup (males) to make the groups comparable in size, with each subpopulation totaling 101 (Hair et al. 2016a). Another option, when possible, is to collect more data for the subgroup with the smaller sample size. When the sample sizes of the groups are considered comparable, it is appropriate to move to Step 2. Note that the groups are not required to be the exact same sample size but should be comparable to avoid producing biased results (Hair et al. 2017b).

The theoretical model for the example (Fig. 10.2) includes three constructs: autonomy, skill discrepancy, and cognitive engagement. Autonomy measures the extent to which salespeople have the freedom to determine which customers are pursued or not, how resources should be distributed among the firm's customers, as well as which customer relationships to continue and which to end. The autonomy construct had 13 items and was measured using 7-point Likert-type scales (rho_A = 0.954 male sample, 0.973 female sample; AVE = 0.600 male sample, 0.608 female sample). The skill discrepancy construct had four items and was measured using 11-point (0–10) Likert-type scales (rho_A = 0.915 male sample, 0.860 female sample; AVE = 0.785 male sample, 0.69 female sample). Finally, the cognitive engagement construct had five items measured using 11-point Likert-type scales (rho_A = 0.940 male sample, 0.942 female sample; AVE = 0.773 male sample, 0.797 female sample) (Table 10.3). In addition to meeting recommended guidelines for reliability and convergent validity, the heterotrait-monotrait ratio (HTMT) was used to assess discriminant validity. All measures were well below the 0.90 thresholds, thus indicating discriminant validity (Hair et al. 2014a) (Table 10.4).

To generate the subpopulations within the dataset in SmartPLS, go to the Project Explorer window and identify the dataset your model is using. Next, double-click on the data icon for your model. The “Generate Data Groups” icon will appear at the top of your SmartPLS screen. You will need to assign an initial name for your groups. The name can be edited to be more specific once the subpopulations have been established. Next, the categorical variable of interest is selected in the group column section (see Fig. 10.3). Note that more than one group variable can be included in the analysis (e.g., gender and age), but only one group variable will be discussed in the example for this chapter. The “Prune groups” option can remain at the default of 10 for the initial processing. Groups can be manually pruned (eliminated) after they are generated if the subpopulation is not large enough to meet statistical power guidelines (Cohen 1992; Hair et al. 2014a).

As noted earlier, the categorical variable we are using in this analysis is gender. After specifying the overall group name of gender, the output displayed in the data groups tab indicates two groups were generated (see Fig. 10.4). The first group is specified as Group_Gender_Q31_Gender(1.0) and shows the number of records (responses) associated with that group. The second group is specified as Group_Gender_Q31_Gender(2.0) and shows the number of respondents associated with this particular group. To rename these initial group labels, highlight (click on) the row for the first group, and two buttons will appear on the right side of the row (Delete and Edit). By selecting the edit button, the group name can be edited for this subpopulation. In this case, the coding for gender 1.0 represents the male portion of the population. Therefore, the group name is revised to read male (see Fig. 10.5). Similarly, Group_Gender_Q31_Gender(2.0) is renamed female.

10.3.2 Step 2: Test for Invariance

The next step in the process is to test for measurement invariance. To test for measurement invariance in PLS-SEM, the MICOM procedure is executed (Henseler et al. 2016). This procedure requires three steps to test for configural and compositional invariance, as well as equality of composite mean values and variances (Henseler et al. 2016).

The first step in the MICOM procedure involves examining configural invariance (Henseler et al. 2016). The assessment of configural invariance consists of an evaluation of the measurement models for all groups to determine if the same basic factor structure exists in all the groups (same number of constructs as well as items for those constructs). Establishing configural invariance involves the fulfillment of

Table 10.3 Outer loadings, rho_A, and AVE for example application

	Group			Males			Females		
	Outer loadings	rho_A	AVE	Outer loadings	rho_A	AVE	Outer loadings	rho_A	AVE
Autonomy (SD = 1, SD = 7)		0.954	0.603		0.954	0.600		0.973	0.608
I can choose not to pursue a prospective customer	0.723			0.691			0.791		
I can decide on my own whether or not to pursue a prospective customer	0.737			0.664			0.818		
I have control over which prospects pursue	0.796			0.764			0.852		
I am empowered to determine which customers to pursue or not pursue	0.803			0.805			0.813		
I have control over which of my customers are designated as *most* important by my firm	0.838			0.843			0.816		
I have control over which of my customers are designated as *least* important by my firm	0.725			0.678			0.758		
I have significant autonomy in determining which of my customers should receive preferential treatment from the firm	0.808			0.839			0.737		
I can decide on my own whether or not one of my customers should receive a high priority status	0.747			0.837			0.605		
I am empowered to lower my customers' priority status within my firm	0.721			0.780			0.624		
I have significant autonomy in determining which customer relationships to end	0.824			0.826			0.815		
I can choose not to end a relationship with a particular customer	0.783			0.785			0.805		
I can decide on my own whether or not should terminate the relationship with one of my customers	0.739			0.706			0.777		
I have control over which of my customer relationships to continue	0.839			0.817			0.877		

Skill discrepancy (SD = 0, SA = 10)		0.889	0.744		0.915	0.785		0.860	0.690
My job requires me to learn new things	0.870			0.885			0.842		
My job requires me to be creative	0.860			0.880			0.814		
My job requires a high level of skill	0.844			0.861			0.854		
I get to do a variety of different things on my job	0.876			0.917			0.812		
Cognitive engagement (SD = 0, SA = 10)		0.936	0.785		0.940	0.773		0.942	0.797
At work, my mind is focused on my job	0.922			0.915			0.931		
At work, I focus a great deal of my attention on my job	0.886			0.856			0.910		
At work, I am absorbed by my job	0.782			0.840			0.733		
At work, I pay a lot of attention to my job	0.909			0.889			0.925		
At work, I devote a lot of attention to my job	0.924			0.894			0.946		

Table 10.4 Heterotrait-monotrait ratio (HTMT) for example application

	Group			Males			Females		
	Autonomy	Cognitive engagement	Skill discrepancy	Autonomy	Cognitive engagement	Skill discrepancy	Autonomy	Cognitive engagement	Skill discrepancy
Autonomy									
Cognitive engagement	0.229			0.270			0.195		
Skill discrepancy	0.341	0.403		0.493	0.414		0.182	0.423	

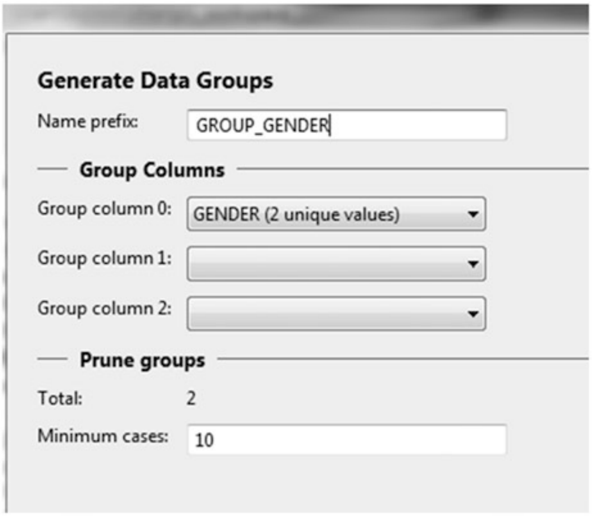


Fig. 10.3 Assign initial name to group and selection of categorical variable

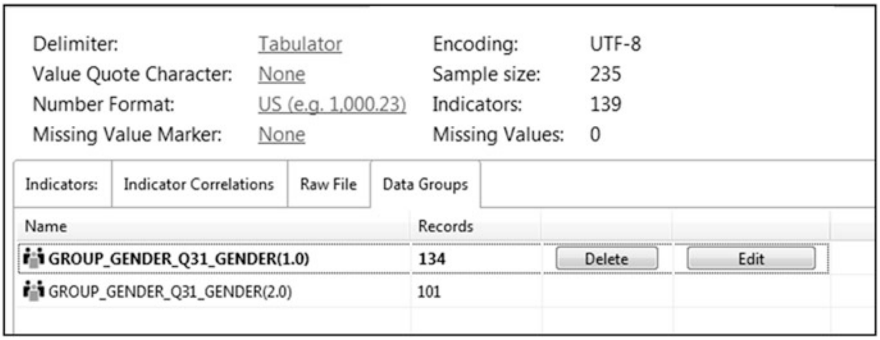


Fig. 10.4 Subpopulations generated

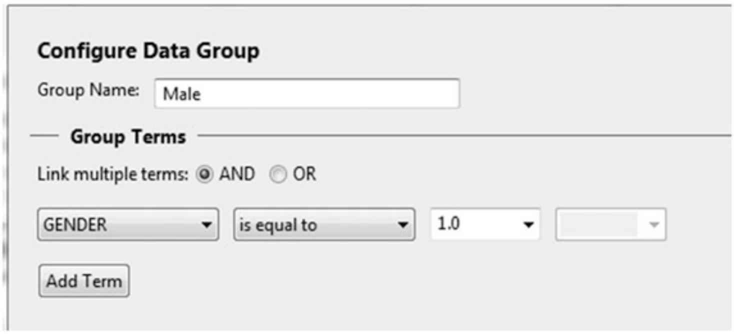


Fig. 10.5 Subpopulation Group_GENDER_Q31_GENDER(1.0) is renamed to male

the following criteria: (a) identical indicators per measurement model, (b) identical data treatment, and (c) identical algorithm settings or optimization criteria (Henseler et al. 2016). All measurement indicators must be included in the constructs across all groups.

The execution of the MICOM procedure includes reviewing the process that was followed during the survey development. Specifically, this evaluation involves a further qualitative assessment of the items to ensure the constructs were initially designed as equivalent. For example, with international studies, translation and back translation practices should have been employed. Additionally, all the data must have been treated identically (e.g., dummy coding, reverse coding, or other forms of recoding, standardization, or missing value treatment). Outliers should also be identified and treated in a similar manner. The items loading on each construct must be invariant across groups as well. Finally, algorithm settings must be identical, and optimization criteria should be applied (Henseler et al. 2016). All of these guidelines are observed to ensure that a composite is a unidimensional entity with the same nomological net across all the groups (Henseler et al. 2016). If all of the above criteria have been met, which is the case for the present example, then configural invariance is established.

The objective of the second step of the MICOM procedure is to examine compositional invariance, which occurs when composite scores are created equally across groups (Dijkstra and Henseler 2011). Permutation tests are also conducted to statistically assess whether compositional invariance is present. Permutation tests are nonparametric (Henseler et al. 2016). For each permutation run, the correlations between the composite scores using the weights obtained from the first group are computed against the composite scores using the weights obtained from the second group (Henseler et al. 2016).

First, select calculate, then Permutation. Under Setup (see Fig. 10.6), specify the desired subpopulations for Group A and Group B. Permutations should be set at 5000. The test type option generally can remain at the two-tailed default as can the significance level of 0.05. But if the sample sizes are smaller and directional hypotheses are involved, a one-tailed test can be applied. Parallel Processing can also remain as the default. Under the Partial Least Squares tab, the Maximum Iterations should be set to 5000, and the Weighting Scheme should remain on Path. The Stop Criterion defaults to 7, but can be adjusted to another small number, such as 5. With regard to missing values, mean replacement is recommended when there are less than 5% of the values missing per indicator (Hair et al. 2014a). Casewise or listwise deletion removes all cases from the analysis that include missing values in any of the indicators used. This may result in a much lower dataset due to missing values. Pairwise deletion uses all observations that contain complete responses for the calculations within the model parameters. Therefore, based on the volume of missing data, select the appropriate setting. The default is set to Mean Replacement. All other defaults are appropriate (e.g., No Weighting Vector).

After the calculation, the output report will default to the Path Coefficients. Under the Quality Criteria options at the bottom of the screen, the desired output report is MICOM. Tabs are available in the MICOM results for the second and third steps.

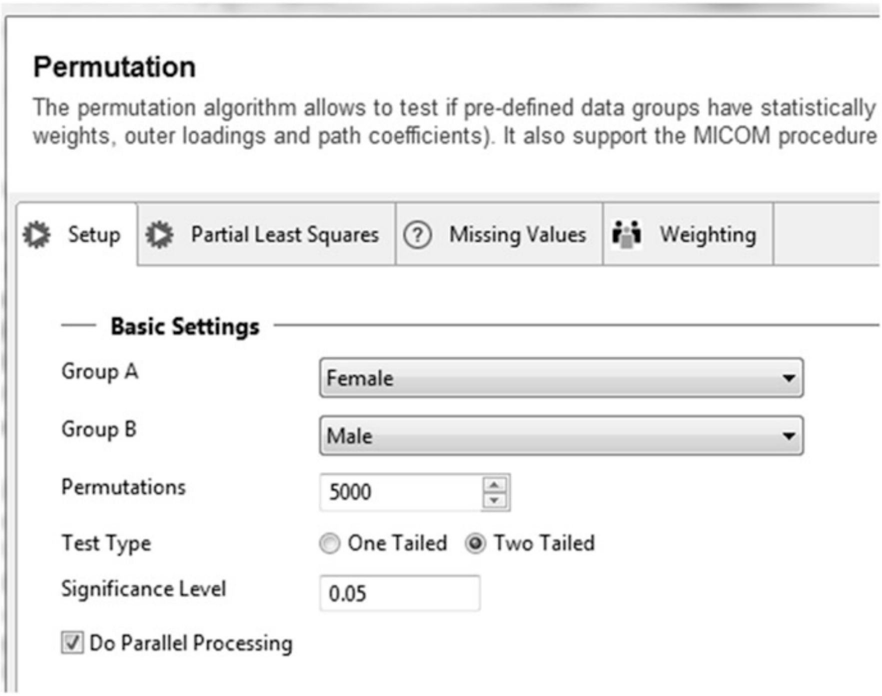


Fig. 10.6 Setup for permutation

The MICOM permutation results report also includes the subsequent Step 3 portion of the MICOM procedure (Henseler et al. 2016).

We continue our example with the three-construct theoretical model that examines gender and salesperson roles. As shown in Table 10.5, the MICOM results report for the second step which indicates that compositional invariance has been demonstrated for all the constructs. This is evident based on the original correlations being equal to or greater than the 5.00% quantile correlations (shown in the 5% column).

A permutation test compares the composite scores of the first and second group to determine if the correlation c is significantly different from the empirical distribution of c_u (represented by the 5.00% quantile) (Henseler et al. 2016). If the results indicate that compositional invariance is a problem for one or more of the constructs, items can be deleted from the constructs in an effort to achieve invariance. Another much less desirable option is to remove entire constructs from the group-specific comparisons, provided that doing so is supported by theory (Henseler et al. 2016).

The next step is to evaluate the results tab for the third step of the MICOM procedure. Table 10.6 shows the first portion of the results. In this step, we assess the composites' (constructs) equality of mean values and variances across the groups. For invariance to be established, the first column (mean original difference) must

Table 10.5 MICOM Step 2 results report

	Original correlation	Correlation permutation mean	5.00%	Permutation *p*-values
Autonomy	0.99	0.995	0.987	0.093
Cognitive engagement	0.999	0.998	0.995	0.485
Skill discrepancy	0.999	0.998	0.994	0.722

Table 10.6 MICOM Step 3 results report—part 1

	Mean original difference (males − females)	Mean permutation mean difference (males − females)	2.50%	97.50%	Permutation *p*-values
Autonomy	0.098	−0.005	−0.268	0.245	0.442
Cognitive engagement	0.117	−0.006	−0.253	0.26	0.4
Skill discrepancy	−0.217	0.001	−0.245	0.269	0.088

Table 10.7 MICOM Step 3 results report—part 2

	Variance original difference (males − females)	Variance permutation mean difference (males − females)	2.50%	97.50%	Permutation *p*-values
Autonomy	−0.207	0.011	−0.378	0.398	0.242
Cognitive engagement	**−0.544**	0.014	**−0.518**	**0.495**	0.03
Skill discrepancy	0.187	0.002	−0.581	0.586	0.569

be a number that falls within the 95% confidence interval. This is assessed by comparing the mean original difference to the lower (2.5%) and upper (97.5%) boundaries shown in columns three and four. If the mean original difference is a number that falls within the range of the lower and upper boundaries, then the first part of step three has been met, thus providing initial evidence of invariance. The constructs in Table 10.6 all pass this portion of the test for invariance.

The second portion of the results for the MICOM step three is shown in Table 10.7. Within SmartPLS, these results will appear to the right of the output presented in Table 10.6. Additionally, for illustration purposes, the construct titles have been displayed again with the output for this second assessment. Similar to the assessment conducted using Table 10.6, the data in column one (variance original difference) must be a number that falls within the 95% confidence interval. Therefore, the first column is again compared to the lower (2.5%) and upper (97.5%) confidence interval. In order to conclude full measurement invariance for the composites (Henseler et al. 2016), all the constructs must fall within the 95% confidence interval. However, in Table 10.7, note that the variance original difference value for the construct cognitive engagement does not fall within the

95% confidence interval. The first portion of Step 3 (Table 10.6) indicated partial invariance for cognitive engagement. But the construct did not meet the guidelines in this step for establishing full invariance. Therefore, only partial invariance is confirmed for this construct. Partial invariance is present when a construct passes only one of the two confidence interval tests, as illustrated with the example shown in Tables 10.6 and 10.7. The permutation *p*-values greater than 0.05 in Table 10.6 provide additional support for the cognitive engagement construct passing the measurement invariance test.

By establishing full measurement invariance, the composites (measurement models) of the two groups can be analyzed using the pooled data. However, using such pooled data without first establishing full measurement invariance could be misleading if there are differences in the structural model that have not been accounted for (Henseler et al. 2016).

If a construct does not pass the third MICOM step (e.g., had cognitive engagement failed both tests in Tables 10.6 and 10.7), and there is a significant difference in the composites' equality of mean values and variances across groups, then that construct should be removed from the analysis. Another possibility, however, is the group differences in the structural model can be accounted for by using the non-invariant construct as a moderator (e.g., cognitive engagement could be the moderator) (Henseler et al. 2016). This would be similar to using gender as the moderator. That is, since we know there are differences in the measurement for the construct, it could be used as the categorical variable of interest. For example, respondents with high cognitive engagement could be compared to those with low cognitive engagement. Using a mean or median split to divide the respondents into high and low groups is not a good approach since the division into groups is arbitrary and non-theoretical. Rather, a better methodology is to apply a cluster analysis to the variable/construct to identify high and low groups (Hair et al. 2016a).

10.3.3 Step 3: Analyze and Interpret Permutation Results

Once invariance is established, the focus is to determine if the path coefficients of the theoretical models for the two groups are significantly different. We will first begin by analyzing the group separately prior to determining if there are group-specific differences. In order to run each group separately, a data file containing only the male participants and another containing only the female participants is needed. For this example, those files were generated in SPSS, converted to .csv files, and imported into my current project. Therefore, this project in SmartPLS contains one model and three data files. Using the guidelines set out for evaluation of a measurement model (Hair et al. 2014a), run the model for each group separately. As noted in Table 10.8, the relationship between autonomy and skill discrepancy is significant for males (p-value $= 0.00$) and is not for females (p-value $= 0.222$). The other relationships, autonomy and cognitive engagement as well as skill discrepancy and cognitive engagement, do not indicate a major difference between males and

Table 10.8 Bootstrapping results for males and females separately

	Males					Females				
	Original sample (O)	Sample mean (M)	Standard deviation (STDEV)	T statistics (|O/STDEV|)	*p*-values	Original sample (O)	Sample mean (M)	Standard deviation (STDEV)	T statistics (|O/STDEV|)	*p*-values
Autonomy → cognitive engagement	0.116	0.118	0.099	1.172	0.241	0.151	0.144	0.132	1.144	0.253
Autonomy → skill discrepancy	0.471	0.485	0.072	6.548	0	0.171	0.209	0.14	1.221	0.222
Skill discrepancy → cognitive engagement	0.334	0.339	0.098	3.396	0.001	0.36	0.393	0.143	2.514	0.012

females. Next, it is necessary to determine if the difference between the two groups is significant. This can be accomplished by returning to the output for the permutation test.

The permutation test results were obtained as a part of testing for measurement invariance. As you may recall, when running the permutation test, the output reports were set to default on the path coefficient output. If you again review the permutation output reports, you will note that the initial column under the path coefficient results (see Table 10.9) displays the hypothesized structural relationships. The next two columns show the original path coefficients for the groups of interest, in this example, males and females. The following two columns indicate the path coefficient original differences as well as the permutation mean differences, which are followed by the lower and upper boundaries for the 95% confidence interval. The final column contains the permutation *p*-Value. A permutation *p*-value of less than or equal to 0.10 designates a significant difference between the two groups of interest. The relationship between autonomy and skill discrepancy indicates a significant difference between males and females. This is evident by the permutation *p*-value of 0.04 in Table 10.9.

Using the information from the group-specific bootstrapping as well as the above permutation test, we can now indicate that there is a significant difference between male and female salespeople as it relates to autonomy and skill discrepancy. This is important for sales managers to understand. Specifically, the findings reveal that male and female salespeople operate differently in their roles as salespersons. For example, there is a significant relationship between autonomy and skill discrepancy for male salespersons, as well as a significant relationship between skill discrepancy and cognitive engagement. Therefore, skill discrepancy partially mediates the relationship between autonomy and cognitive engagement. Since sales managers want their sales force to be cognitively engaged in the sales process, these findings indicate that for males, autonomy is indirectly related to sales engagement, as well as being directly related. In contrast, the link between autonomy and skill discrepancy is not significant for females, indicating that skill discrepancy does not mediate the relationship between autonomy and cognitive engagement. Therefore, additional research is needed to identify alternative antecedents for female salespersons that will lead to stronger cognitive engagement in their sales roles.

MGA allows researchers to determine significant differences among observed characteristics such as age, gender, or country of origin. While the path coefficients for the partitioned groups will almost always indicate numerical differences, understanding when those differences are significant is the role that MGA plays. These differences may not be evident in aggregate data since significant positive and negative group-specific results may offset one another.

Table 10.9 Permutation test path coefficient results

	Path coefficients original (males)	Path coefficients original (females)	Path coefficients original difference (males − females)	Path coefficients permutation mean difference (males − females)	2.50%	97.50%	Permutation *p*-values
Autonomy → cognitive engagement	0.116	0.151	−0.035	0.004	−0.315	0.339	0.825
Autonomy → skill discrepancy	0.471	0.171	0.3	−0.009	−0.291	0.283	0.04
Skill discrepancy → cognitive engagement	0.334	0.36	−0.026	−0.007	−0.347	0.347	0.961

10.4 Summary

Virtually, all previous social sciences research focused on understanding a single sample of data. Researchers examined the total sample, without considering subsamples, primarily because they assumed this approach provided an accurate understanding of the findings. They also analyzed the total sample findings because there were limited scientific, rigorous analytical procedures to divide the sample into meaningful subgroups. The most widely applied approach was to use theoretical a priori-defined simple approaches, such as size of firms, age, gender, or income. While helpful, subgroup analyses based on a priori-defined groups often did not assess measurement characteristics and were typically limited to attempting to understand a single relationship or model parameter. What was lacking was a simple, straightforward, efficient method for examining multiple relationships/parameters simultaneously and efficiently. The PLS-MGA, particularly in combination with assessing invariance, represents considerable progress in filling this void for researchers that are using PLS-SEM.

PLS-MGA substantially improves the ability of researchers to identify meaningful and significant differences in multiple relationships across group-specific results (Picon-Berjoyo et al. 2016; Sarstedt et al. 2014; Schlagel and Sarstedt 2016). Specifically, multiple model parameters can be examined simultaneously, and if statistically significant differences are present in the theoretical model, they can be efficiently identified. Moreover, in combination with the MICOM procedure, researcher can add an additional level of accuracy to their findings by including an assessment of the measurement characteristics of their constructs. Since research that does not examine group-specific differences often leads to misinterpretation of the results (Hult et al. 2008), it is important for researchers to apply this procedure when meaningful subgroups are present in the data or when they are subsequently identified using methods for assessing unobserved heterogeneity in sample data.

A primary concern of social science researchers, when comparing path coefficients among groups, should be to confirm that the construct measures are equivalent across the groups. Therefore, testing for measurement invariance is necessary to avoid introducing bias into research findings (Hair et al. 2014a; Henseler et al. 2016). MGA can be easily executed by following the approach provided in this chapter (Fig. 10.1): Step 1, Generate Data Groups; Step 2, Test for Invariance; and Step 3, Analyze and Interpret Results. By observing and following the guidelines, predefined data groups can be examined using PLS path modeling, and if meaningful and significant differences are present in the data, they can be reported and explained (Hair et al. 2014a, 2017b; Lohmöller 1989), therefore improving the rigor of research publication practices (Hair et al. 2012a, b, 2013, 2014b; Sarstedt and Mooi 2014) and improving our understanding of previously misunderstood theoretical relationships.

This study focused on differences between two groups; however, there are times when more than two groups are involved. Future research should provide step-by-step instructions on comparing more than two groups. Additionally, this study explained the permutation procedure for conducting MGA. Since there are several means of conducting MGA, future research is needed that compares the various methods of MGA to better understand the differences in the results of each. Similarly, since the introduction of consistent PLS (PLSc) (Dijkstra and Henseler 2015), MGA has yet to be combined with that process to determine the proper use of traditional bootstrapping and consistent bootstrapping. Finally, the guidelines provided for conducting a MGA with PLS-SEM include the MICOM test for invariance; future research should explore the possibility of simplifying the process allowing for simultaneous assessment of configural and compositional invariance. Finally, due to the controversy over minimum sample size, the establishment of a procedure to test for a lack of statistical power due to sample size is encouraged.

Beyond the analysis conducted in this chapter based on a priori knowledge, researchers should also conduct an analysis to examine potential differences that may not have been identified via a priori-defined categorical variables. That is, unobserved heterogeneity should be examined (Hair et al. 2016b; Jedidi et al. 1997; Matthews et al. 2016; Sarstedt et al. 2018). Both MGA and tests to uncover unobserved heterogeneity can be used to identify differences among subpopulations within a larger dataset. MGA, however, uses categorical variables that have been identified a priori and collected in the dataset. However, sources of such differences can be difficult to identify; therefore, recent methods such as PLS-GAS, PLS-POS, and PLS-IRRS can be instrumental for uncovering other areas that partition data into groups (Hair et al. 2014a). In an effort to improve the validity of PLS-SEM results, which continue to be enhanced (Hair et al. 2017a), researchers are encouraging the routine application of such techniques (Hair et al. 2011, 2012c, 2013; Wilson et al. 2014).

Acknowledgments This chapter refers to the use of the statistical software SmartPLS (http://www.smartpls.com) (Ringle et al. 2015).

References

Aguinis, H., Beaty, J. C., & Boik, R. K. (2005). Effect size and power in assessing moderating effects of categorical variables using multiple regression: A 30-year review. *The Journal of Applied Psychology, 90*, 94–107.

Becker, J.-M., Rai, A., Ringle, C. M., & Völckner, F. (2013). Discovering unobserved heterogeneity in structural equation models to avert validity threats. *MIS Quarterly, 37*, 665–694.

Brettel, M., Engelen, A., Heinemann, F., & Vadhanasindhu, P. (2008). Antecedents of market orientation: A cross-cultural comparison. *Journal of International Marketing, 16*, 84–119.

Chin, W. W., & Dibbern, J. (2010). A permutation based procedure for multi-group PLS analysis: Results of tests of differences on simulated data and a cross cultural analysis of the sourcing of information system services between Germany and the USA. In V. Esposito Vinzi, W. W. Chin, J. Henseler, & H. Wang (Eds.), *Handbook of partial least squares: Concepts, methods and applications*, *Springer Handbooks of Computational Statistics Series* (Vol. II, pp. 171–193). Heidelberg: Springer.

Cohen, J. A. (1992). A power primer. *Psychological Bulletin, 112*, 155–519.

Diamantopoulos, A., Sarstedt, M., Fuchs, C., Wilczynski, P., & Kaiser, S. (2012). Guidelines for choosing between multi-item and single-item scales for construct measurement: A predictive validity perspective. *Journal of Academy of Market Science, 40*(3), 434–449. doi:10.1007/s11747-011-0300-3.

Dijkstra, T. K., & Henseler, J. (2011). Linear indices in nonlinear structural equation models: Best fitting proper indices and other composites. *Quality and Quantity, 45*, 1505–1518. doi:10.1007/s11135-010-9359-z.

Dijkstra, T. K., & Henseler, J. (2015). Consistent partial least squares path modeling. *MIS Quarterly, 39*, 297–316.

Grewal, R., Chakravarty, A., Ding, M., & Liechty, J. (2008). Counting chickens before the eggs hatch: Associating new product development portfolios with shareholder expectations in the pharmaceutical sector. *International Journal of Research in Marketing, 25*, 261–272.

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. (2010). *Multivariate data analysis*. Upper Saddle River, NJ: Pearson/Prentice Hall.

Hair, J. F., Ringle, C. M., & Sarstedt, M. (2011). PLS-SEM: Indeed a silver bullet. *Journal of Marketing Theory and Practice, 19*(2), 139–152.

Hair, J. F., Ringle, C. M., & Sarstedt, M. (2012a). Partial least squares: The better approach to structural equations modeling? *Journal of Long Range Planning, 45*(5–6), 312–319. doi:10.1016/j.lrp.2012.09.011.

Hair, J. F., Sarstedt, M., Pieper, T. M., & Ringle, C. M. (2012b). The use of partial least squares structural equation modeling in strategic management research: A review of past practices and recommendations for future applications. *Long Range Planning, 45*(5–6), 320–340. doi:10.1016/j.lrp.2012.09.008.

Hair, J. F., Sarstedt, M., Ringle, C. M., & Mena, J. A. (2012c). An assessment of the use of partial least squares structural equation modeling in marketing research. *Journal of the Academy of Marketing Science, 40*(3), 414–433. doi:10.1007/s11747-011-0261-6.

Hair, J. F., Ringle, C. M., & Sarstedt, M. (2013). Partial least squares structural equation modeling: Rigorous applications, better results and higher acceptance. *Journal of Long Range Planning, 46*(1–2), 1–12.

Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2014a). *A primer on partial least squares structural equation modeling (PLS-SEM)*. Thousand Oaks, CA: Sage.

Hair, J. F., Sarstedt, M., Hopkins, L., & Kuppelwieser, V. (2014b). Partial least squares structural equation modeling (PLS-SEM): An emerging tool in business research. *European Business Review, 26*(2), 106–121. doi:10.1108/EBR-10-2013-0128.

Hair, J. F., Celsi, M., Money, A., Samouel, P., & Page, M. (2016a). *Essentials of business research methods*. New York: Routledge.

Hair, J. F., Sarstedt, M., Matthews, L., & Ringle, C. (2016b). Identifying and treating unobserved heterogeneity with FIMIX-PLS: Part I – Method. *European Business Review, 28*(1), 63–76.

Hair, J. F., Hult, G. T. M., Ringle, C. M., Sarstedt, M., & Thiele, K. O. (2017a). Mirror, mirror on the wall: A comparative evaluation of composite-based structural equation modeling methods. *Journal of Academy of Marketing Science* (in press)

Hair, J. F., Sarstedt, M., Ringle, C. M., & Gudergan, S. (2017b). *Advanced issues in partial least squares structural equation modeling (PLS-SEM)*. Thousand Oaks, CA: Sage.

Henseler, J., & Chin, W. W. (2010). A comparison of approaches for the analysis of interaction effects between latent variables using partial least squares path modeling. *Structural Equation Modeling, 17*(1), 82–109. doi:10.1080/10705510903439003.

Henseler, J., Ringle, C. M., & Sinkovics, R. R. (2009). The use of partial least squares path modeling in international marketing. In R. R. Sinkovics & P. N. Ghauri (Eds.), *Advances in International Marketing* (Vol. 20, pp. 277–320). Bingley: Emerald.

Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science, 34*, 115–135.

Henseler, J., Ringle, C. M., & Sarstedt, M. (2016). Testing measurement invariance of composites using partial least squares. *International Marketing Review, 33*(3), 405–431. doi:10.1108/IMR-09-2014-0304.

Horn, J. L., & McArdle, J. J. (1992). A practical and theoretical guide to measurement invariance in aging research. *Experimental Aging Research, 18*(3–4), 117–144.

Hult, G. T. M., Ketchen, D. J., Griffith, D. A., et al. (2008). Data equivalence in cross-cultural international business research: Assessment and guidelines. *Journal of International Business Studies, 39*(6), 1027–1044. doi:10.1057/palgrave.jibs.8400396.

Jedidi, K., Jagpal, H. S., & Desarbo, W. S. (1997). Finite-mixture structural equation models for response-based segmentation and unobserved heterogeneity. *Marketing Science, 16*(1), 39–59.

Keil, M., Tan, B. C. Y., Wei, K.-K., et al. (2000). A cross-cultural study on escalation of commitment behavior in software projects. *MIS Quarterly, 24*(2), 299–325.

Kock, N., & Hadaya, P. (2016). Minimum sample size estimation in PLS-SEM: The inverse square root and gamma-exponential methods. *Information Systems Journal* (in press). doi:10.1111/isj.12131.

Lohmöller, J.-B. (1989). *Latent variable path modeling with partial least squares*. Berlin: Springer.

Matthews, L., Sarstedt, M., Hair, J. F., & Ringle, C. (2016). Identifying and treating unobserved heterogeneity with FIMIX-PLS: Part II – A case study. *European Business Review, 28*(2), 208–224. doi:10.1108/EBR-09-2015-0095.

Patel, V. K., Manley, S., Ferrell, O. C., Hair, J. F., & Pieper, T. M. (2016). Is stakeholder theory relevant for European firms? *European Management Journal, 34*(6), 650–660.

Picon-Berjoyo, A., Ruiz-Moreno, C., & Castro, I. (2016). A mediating and multi-group analysis of customer loyalty. *European Management Journal, 34*(6), 701–713.

Rasoolimanesh, S. M., Roldan, J. L., Jaafar, M., & Ramayah, T. (2016). Factors influencing residents' perceptions toward tourism development; differences across rural and urban world heritage sites. *Journal of Travel Research*. doi:10.1177/0047287516662354.

Ringle, C., Wende, S., & Becker, J. (2015). *SmartPLS 3*. Bönningstedt: SmartPLS GmbH. Retrieved from http://www.smartpls.com.

Rutherford, B. N., Park, J., & Han, S. L. (2011). Increasing job performance and decreasing salesperson propensity to leave: An examination of an Asian sales force. *Journal of Personal Selling & Sales Management, 31*(2), 171–183.

Sarstedt, M., & Mooi, E. A. (2014). *A concise guide to market research: The process, data, and methods using IBM SPSS statistics*. Berlin: Springer.

Sarstedt, M., Schwaiger, M., & Ringle, C. M. (2009). Do we fully understand the critical success factors of customer satisfaction with industrial goods? – Extending Festge and Schwaiger's model to account for unobserved heterogeneity. *Journal of Business Market Management, 3*(3), 185–206.

Sarstedt, M., Henseler, J., & Ringle, C. M. (2011). Multi-group analysis in partial least squares (PLS) path modeling: Alternative methods and empirical results. In M. Sarstedt, M. Schwaiger, & C. R. Taylor (Eds.), *Advances in international marketing* (Vol. 22, pp. 195–218). Bingley: Emerald. doi:10.1108/S1474-7979(2011)0000022012.

Sarstedt, M., Ringle, C. M., & Hair, J. F. (2014). PLS-SEM: Looking back and moving forward. *Long Range Planning, 47*(3), 132–137. doi:10.1016/j.lrp.2014.02.008.

Sarstedt, M., Ringle, C. M., & Gudergan, S. P. (2016a). Guidelines for treating unobserved heterogeneity in tourism research: A comment on Marques and Reis (2015). *Annals of Tourism Research, 57*, 279–284. doi:10.1016/j.annals.2015.10.006.

Sarstedt, M., Hair, J. F., Ringle, C. M., Thiele, K. O., & Gudergan, S. P. (2016b). Estimation issues with PLS and CBSEM: Where the bias lies! *Journal of Business Research, 69*(10), 3998–4010. doi:10.1016/j.jbusres.2016.06.007.

Sarstedt, M., Ringle, C. M., & Hair, J. F. (2018). Treating unobserved heterogeneity in PLS-SEM: A multi-method approach. In H. Latan & R. Noonan (Eds.), *Partial least squares structural equation modeling: Basic concepts, methodological issues and applications*. Heidelberg: Springer.

Schlagel, C., & Sarstedt, M. (2016). Assessing the measurement invariance of the four-dimensional cultural intelligence scale across countries: A composite model approach. *European Management Journal, 34*(6), 633–649.

Wilson, S. R., Whitmoyer, J. G., Pieper, T. M., et al. (2014). Method trends and method needs: Examining methods needed for accelerating the field. *Innovative and Established Research Methods in Family Business, 5*(1), 4–14. doi:10.1016/j.jfbs.2014.01.011.

Chapter 11
Common Method Bias: A Full Collinearity Assessment Method for PLS-SEM

Ned Kock

Abstract In the context of structural equation modeling employing the partial least squares (PLS-SEM) method, common method bias is a phenomenon caused by common variation induced by the measurement method used and not by the network of causes and effects in the model being studied. Two datasets were created through a Monte Carlo simulation to illustrate our discussion of this phenomenon: one contaminated by common method bias and the other not contaminated. A practical approach is presented for the identification of common method bias based on variance inflation factors generated via a full collinearity test. Our discussion builds on an illustrative model in the field of e-collaboration, with outputs generated by the software WarpPLS. We demonstrate that the full collinearity test is successful in the identification of common method bias with a model that nevertheless passes standard convergent and discriminant validity assessment criteria based on a confirmation factor analysis.

11.1 Introduction

The foundation on which structural equation modeling (SEM) rests owes much of its existence to one of the greatest evolutionary biologists in history: Sewall Wright. A key element of that foundation is the method of path analysis, which has been developed by Wright (1934, 1960) to study causal assumptions in the field of evolutionary biology (Kock 2011). Both path analysis and SEM rely on the creation of models expressing causal relationships through links among variables. Two main types of SEM find widespread use today: covariance-based and PLS-based SEM. While the former relies on the minimization of differences between covariance matrices, the latter employs the partial least squares (PLS) method developed by Herman Wold (1980). PLS-based SEM is often referred to simply as PLS-SEM and is widely used in the field of e-collaboration and many other fields. A third SEM

N. Kock (✉)
Division of International Business and Technology Studies, Texas A&M International University, 5201 University Boulevard, Laredo, TX 78041, USA
e-mail: nedkock@gmail.com

H. Latan, R. Noonan (eds.), *Partial Least Squares Path Modeling*,
DOI 10.1007/978-3-319-64069-3_11

type has recently been gaining increasing attention, factor-based PLS-SEM, which combines elements of both covariance based and PLS-SEM (Kock 2014, 2015c).

Regardless of SEM flavor, models expressing causal assumptions include latent variables. These latent variables are measured indirectly through other variables generally known as indicators (Maruyama 1998; Mueller 1996). Indicator values are usually obtained from questionnaires where answers are provided on numeric scales, of which the most commonly used are Likert-type scales (Cohen et al. 2003). Using questionnaires answered on Likert-type scales constitutes an integral part of an SEM study's measurement method. Common method bias is a phenomenon that is caused by the measurement method used in an SEM study and not by the network of causes and effects among latent variables in the model being studied.

We provide a discussion of common method bias in PLS-SEM and of a method for its identification based on full collinearity tests (Kock and Lynn 2012). Our discussion builds on an illustrative model in the field of e-collaboration, with outputs from the software WarpPLS, version 5.0 (Kock 2015a). This software provides the most extensive set of outputs of any PLS-SEM software and thus is a good choice for our illustrative discussion.

The algorithm used to generate latent variable scores based on indicators was PLS Mode A, employing the path weighting scheme. While this is the algorithm-scheme combination most commonly used in PLS-SEM, it is by no means the only combination available. The recent emergence of factor-based PLS-SEM algorithms further broadened the space of existing combinations (Kock 2014, 2015c).

We created two datasets based on a Monte Carlo simulation (Robert and Casella 2005; Paxton et al. 2001). One of the two datasets was contaminated by common method bias; the other was not. We demonstrate that the full collinearity test is successful in the identification of common method bias with a model that nevertheless passes standard validity assessment criteria based on a confirmation factor analysis.

In our discussion, all variables are assumed to be standardized, i.e., scaled to have a mean of zero and standard deviation of one. This has no impact on the generality of the discussion. Standardization of any variable is accomplished by subtraction of its mean and division by its standard deviation. A standardized variable can be rescaled back to its original scale by reversing these operations.

11.2 What Is Common Method Bias?

Common method bias, in the context of PLS-SEM, is a phenomenon that is caused by the measurement method used in an SEM study and not by the network of causes and effects in the model being studied. For example, the instructions at the top of a questionnaire may influence the answers provided by different respondents in the same general direction, causing the indicators to share a certain amount of common variation. Another possible cause of common method bias is the implicit social desirability associated with answering questions in a questionnaire in a particularly way, again causing the indicators to share a certain amount of common variation.

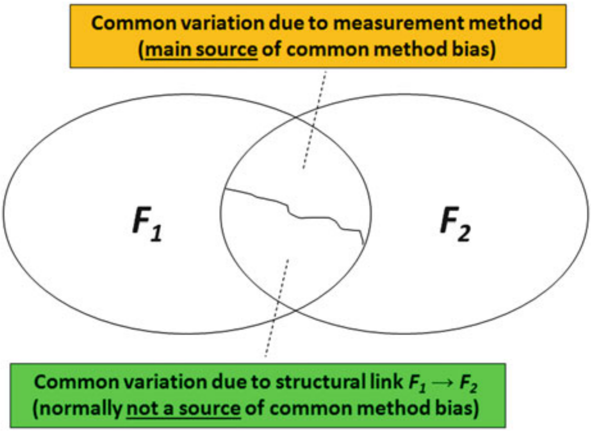


Fig. 11.1 Common method bias in a simple model with two factors

To better illustrate this point, let us consider a simple model with two latent variables represented by two factors F_1 and F_2. In this model, we hypothesize that F_1 causes F_2. Because of this causal link, we expect the two factors to share common variation but not the type of common variation that is associated with common method bias. In other words, while the two factors are expected to be correlated, the common variation that induces this correlation is not normally a cause of common method bias.

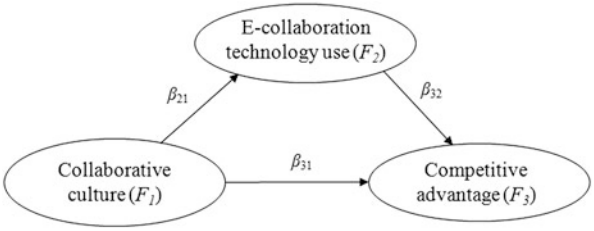
This is illustrated in Fig. 11.1 where the areas within the ovals provide a schematic representation of variation in the factors, with the shared variation being in the overlap area, in the middle of the figure. The overlap representing shared variation is divided into two parts: one for common variation due to measurement method and the other for common variation due to the structural link going from F_1 to F_2.

Since common method bias is caused by common variation that emerges in the measurement model (factor-indicator links) and not in the structural model (factor-factor links), it is particularly difficult to detect and isolate. The difficulty comes from the fact that non-pathological common variation ("good" common variation) nearly always exists in correctly specified SEM models. Otherwise, all factors would be uncorrelated—which would defeat the purposes of the SEM analyses in most cases—to uncover actual associations among latent variables that may have a causal basis.

11.3 Illustrative Model

A mathematical understanding of common method bias is likely to be useful in clarifying some aspects of the nature of the phenomenon. The adoption of an illustrative model can in turn help reduce the level of abstraction of a mathematical exposition. Therefore, our discussion is based on the illustrative model depicted

Fig. 11.2 Illustrative model



in Fig. 11.2, which is inspired by an actual empirical study in the field of e-collaboration (Kock 2005, 2008; Kock and Lynn 2012). The illustrative model incorporates three latent variables, each measured through six indicators. It assumes that the unit of analysis is the firm.

The latent variables are *collaborative culture* (F_1), the perceived degree to which a firm's culture promotes continuous collaboration among its members to improve the firm's productivity and the quality of the firm's products; e-*collaboration technology use* (F_2), the perceived degree of use of e-collaboration technologies by the members of a firm; and c*ompetitive advantage* (F_3), the perceived degree of competitive advantage that a firm possesses when compared with firms that compete with it. Mathematically, if our model were not contaminated with common method bias, each of the six indicators x_{ij} would be derived from its latent variable F_i (of which there are three in the model) according to Eq. (11.1), where λ_{ij} is the loading of indicator x_{ij} on F_i, θ_{ij} is the standardized indicator error term, and $\omega_{\theta j}$ is the weight of θ_{ij} with respect to x_{ij}:

$$x_{ij} = \lambda_{ij} F_i + \omega_{\theta j} \theta_{ij}, \quad i = 1 \ldots 3, \; j = 1 \ldots 6. \tag{11.1}$$

Since θ_{ij} and F_i are assumed to be uncorrelated, the value of $\omega_{\theta j}$ in this scenario can be easily obtained as:

$$\omega_{\theta j} = \sqrt{1 - {\lambda_{ij}}^2}.$$

If our model *were* contaminated with common method bias, each of the six indicators x_{ij} would be derived from its latent variable F_i according to Eq. (11.2), where M is a standardized variable that represents common method variation and ω_M is the common method weight (aka common method loading or the positive square root of the common method variance):

$$x_{ij} = \lambda_{ij} F_i + \omega_M M + \omega_{\theta j} \theta_{ij}, \quad i = 1 \ldots 3, \; j = 1 \ldots 6. \tag{11.2}$$

In this scenario, the value of $\omega_{\theta j}$ can be obtained as:

$$\omega_{\theta j} = \sqrt{1 - {\lambda_{ij}}^2 - {\omega_M}^2}.$$

In Eq. (11.2), we assume that the common method weight ω_M is the same for all indicators. An alternative perspective assumes that the common method weight ω_M is not the same for all indicators, varying based on a number of factors. Two terms are used to refer to these different perspectives, namely, congeneric and non-congeneric, although there is some confusion in the literature as to which term refers to what perspective.

Note that the term $\omega_M M$ introduces common variation that is shared by all indicators in the model. Since latent variables aggregate indicators in PLS-SEM, this shared variation has the effect of artificially increasing the level of collinearity among latent variables. As we will see later, this also has the predictable effect of artificially increasing path coefficients.

Nevertheless, it is important to point out that in some cases, path coefficients may be suppressed by common method variation, although this is much less common. Latent variables are linked in pairs in SEM models, each pair including one predictor and one criterion latent variable. If the common variation is in one direction for one of the latent variables of a linked pair and in the other direction for the other latent variable, then it is likely that the corresponding path coefficient will underestimate the true value. This is illustrated in Fig. 11.3.

The scenario illustrated in the figure is one in which the use of a technology is forced upon employees in an organization, leading to a hostile work environment. This occurs as part of an action research study of the impact that the technology has on employees' performance at work. When answering questions on a questionnaire administered as part of the study, employees consistently overestimate their use of

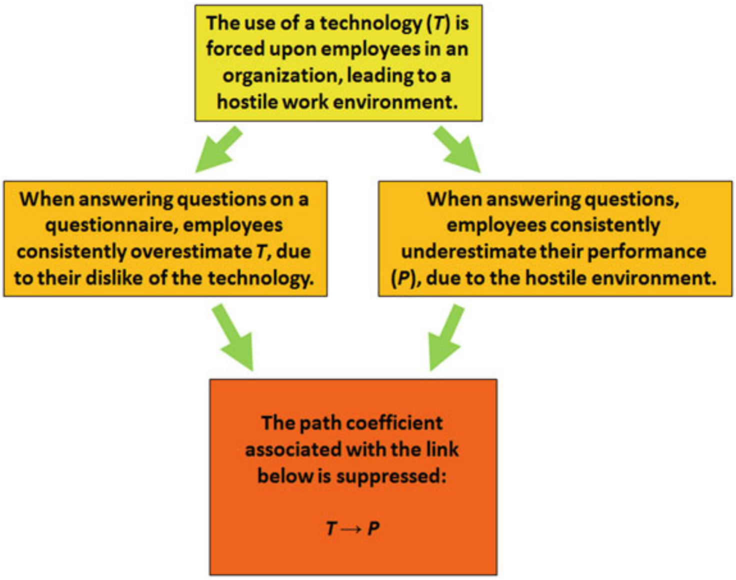


Fig. 11.3 Path coefficient suppression example

the technology, due to their dislike of the technology. That is, the employees dislike the technology so much that any use is perceived as too much use.

On the other hand, when answering questions about their performance, employees consistently underestimated their performance, due to the hostile work environment. That is, being in a hostile work environment gives the employees the impression that they are performing below their usual level. This common perception shared by employees, together with the common perception that they use the technology too much, leads to a suppression of the path coefficient associated with the causal link going from technology use to performance at work. The common source of bias here is not the questionnaire used for data collection but the hostile environment associated with the forced technology use.

11.4 Data Used in the Analysis

We created two datasets of 300 rows of data, equivalent to 300 returned questionnaires, with answers provided on Likert-type scales going from 1 to 7. This was done based on a Monte Carlo simulation (Robert and Casella 2005; Paxton et al. 2001). The data was created for the 3 latent variables and the 18 indicators (6 per latent variable) in our illustrative model.

Using this method, we departed from a "true" model, which is a model for which we know the nature and magnitude of all of the relationships among variables beforehand. One of the two datasets was contaminated by common method bias; the other was not. In both datasets, path coefficients and loadings were set as follows:

$$\beta_{21} = \beta_{31} = \beta_{32} = .45$$

$$\lambda_{ij} = .7, \quad i = 1 \ldots 3, j = 1 \ldots 6.$$

That is, all path coefficients were set as .45 and all indicator loadings as .7. In the dataset contaminated by common method bias, the common method weight was set to values lightly lower than the indicator loadings:

$$\omega_M = .6.$$

In Monte Carlo simulations where samples of finite size are created, true sample coefficients vary. Usually, true sample coefficients vary according to a normal distribution centered on the true population value. Given this, and since we created a single sample of simulated data, our true sample coefficients differed from the true population coefficients. Nevertheless, when we compared certain coefficients obtained via a PLS-SEM analysis for the two datasets, with and without contamination, the effects of common method bias became visible. This is particularly true for path coefficients, which tend to be inflated by common method bias. As noted earlier, path coefficient inflation is a predictable outcome of shared variation among latent variables.

11.5 Path Coefficient Inflation

Table 11.1 shows the path coefficients for the models not contaminated by common method bias (No CMB) and contaminated (CMB). As we can see, all three path coefficients were greater in the model contaminated by common method bias. The differences among path coefficients ranged from a little over 20 to nearly 40%.

This path coefficient inflation effect is one of the key reasons why researchers are concerned about common method bias, as it may cause type I errors (false positives). Nevertheless, as illustrated earlier, common method bias may also be associated with path coefficient deflation, potentially leading to type II errors (false negatives).

As we can see, the inflation effect can lead to marked differences in path coefficients. In the case of the path coefficient β_{21}, the difference is of approximately 39.82%. As noted earlier, path coefficient inflation occurs because common variation is introduced, being shared by all indicators in the model. As latent variables aggregate indicators, they also incorporate the common variation, leading to an increase in the level of collinearity among latent variables. Greater collinearity levels in turn lead to inflated path coefficients. One of the goals of a confirmatory factor analysis is to assess two main types of validity in a model: convergent and discriminant validity. Acceptable convergent validity occurs when indicators load strongly on their corresponding latent variables. Acceptable discriminant validity occurs when the correlations among a latent variable and other latent variables in a model are lower than a measure of communality among the latent variable indicators.

Given these expectations underlying acceptable convergent and discriminant validity, one could expect that a confirmatory factor analysis would allow for the identification of common method bias. In fact, many researchers in the past have proposed the use of confirmatory factor analysis as a more desirable alternative to Harman's single-factor test—a widely used common method bias test that relies on exploratory factor analysis. Unfortunately, as we will see in the next section, conducting a confirmatory factor analysis is not a very effective way of identifying common method bias. Models may pass criteria for acceptable convergent and discriminant validity and still be contaminated by common method bias.

Table 11.1 Path coefficients

	β_{21}	β_{31}	β_{32}
No CMB	.447	.409	.357
CMB	.625	.512	.435

Note: *CMB* common method bias

11.6 Confirmatory Factor Analysis

Table 11.2 is a combined display showing loadings and cross-loadings. Loadings, shown in shaded cells, are unrotated. Cross-loadings are oblique-rotated. Acceptable convergent validity would normally be assumed if the loadings were all above a certain threshold, typically .5. As we can see, all loadings pass this test. This is the case for both models, with and without common method bias contamination. That is, both models present acceptable convergent validity.

These results highlight one interesting aspect of the common method bias phenomenon in the context of PLS-SEM. There appears to be a marked inflation in loadings, similarly to what was observed for path coefficients. Since convergent validity relies on the comparison of loadings against a fixed threshold, then it follows that common method bias would tend to artificially increase the level of convergent validity of a model.

Table 11.3 shows correlations among latent variables and square roots of average variances extracted (AVEs). The latter are shown in shaded cells, along diagonals. Acceptable discriminant validity would typically be assumed if the number in the

Table 11.2 Assessing convergent validity

	No CMB			CMB		
	F_1	F_2	F_3	F_1	F_2	F_3
x_{11}	.742	.010	-.095	.902	.072	-.075
x_{12}	.730	.029	.010	.912	.060	-.100
x_{13}	.772	.051	-.043	.900	-.075	.054
x_{14}	.771	-.061	.109	.891	.004	-.064
x_{15}	.766	.004	.042	.913	-.085	.176
x_{16}	.729	-.033	-.044	.890	.026	.001
x_{21}	.022	.690	-.102	.011	.900	.031
x_{22}	-.060	.709	-.027	-.003	.892	-.063
x_{23}	.049	.701	.005	.080	.893	-.113
x_{24}	.018	.766	.031	-.068	.921	.077
x_{25}	-.106	.731	.040	.020	.905	.002
x_{26}	.055	.766	.033	-.036	.924	.057
x_{31}	.022	-.003	.721	.020	-.005	.911
x_{32}	-.039	.029	.712	.052	-.013	.908
x_{33}	-.029	-.063	.693	-.003	-.012	.913
x_{34}	-.018	-.008	.724	-.037	.035	.909
x_{35}	.013	-.060	.754	-.065	-.072	.920
x_{36}	.041	.088	.762	.030	.065	.903

Notes: loadings are unrotated and cross-loadings are oblique-rotated; loadings shown in shaded cells

CMB common method bias

Table 11.3 Assessing discriminant validity

	No CMB			CMB		
	F_1	F_2	F_3	F_1	F_2	F_3
F_1	.752	.447	.568	.901	.625	.785
F_2	.447	.728	.540	.625	.906	.756
F_3	.568	.540	.728	.785	.756	.911

Notes: Square roots of average variances extracted (AVEs) shown on shaded diagonal

diagonal cell for each column is greater than any of the other numbers in the same column.

That is, if the square root of the AVE for a given latent variable is greater than any correlation involving that latent variable, and this applies to all latent variables in a model, then the model presents acceptable discriminant validity. As we can see, this is the case for both of our models, with and without common method bias contamination. Both models can thus be assumed to display acceptable discriminant validity.

Here, we see another interesting aspect of the common method bias phenomenon in the context of PLS-SEM. While correlations among latent variables increase, the same happens with the AVEs. This simultaneous increase in correlations and AVEs is what undermines the potential of a discriminant validity check in the identification of common method bias.

In summary, two key elements of a traditional confirmatory factor analysis are a convergent validity test and a discriminant validity test. According to our analysis, neither test seems to be very effective in the identification of common method bias. An analogous analysis was conducted by Kock and Lynn (2012), which prompted them to offer the full collinearity test as an effective alternative for the identification of common method bias.

11.7 The Full Collinearity Test

Collinearity has classically been defined as a predictor-predictor phenomenon in multiple regression models. In this traditional perspective, when two or more predictors measure the same underlying construct, or a facet of such construct, they are said to be collinear. This definition is restricted to classic, or vertical, collinearity. Lateral collinearity is defined as a predictor-criterion phenomenon, whereby a predictor variable measures the same underlying construct, or a facet of such construct, as a variable to which it points in a model. The latter is the criterion variable in the predictor-criterion relationship of interest.

Kock and Lynn (2012) proposed the full collinearity test as a comprehensive procedure for the simultaneous assessment of both vertical and lateral collinearity (see, also, Kock and Gaskins 2014). Through this procedure, which is fully automated by the software WarpPLS, variance inflation factors (VIFs) are generated for all latent variables in a model. The occurrence of a VIF greater than 3.3 is proposed as an indication of pathological collinearity and also as an indication that a model may be contaminated by common method bias. Therefore, if all VIFs resulting from a full collinearity test are equal to or lower than 3.3, the model can be considered free of common method bias.

Table 11.4 shows the VIFs obtained for all the latent variables in both of our models, based on a full collinearity test. As we can see, the model contaminated with common method bias includes a latent variable with VIF greater than 3.3, which is shown in a shaded cell. That is, the common method bias test proposed by Kock and

Table 11.4 Full collinearity VIFs

	F_1	F_2	F_3
No CMB	1.541	1.472	1.739
CMB	2.619	2.347	3.720

Note: *CMB* common method bias

Lynn (2012), based on the full collinearity test procedure, seems to succeed in the identification of common method bias.

While it is noteworthy that the full collinearity test was successful in the identification of common method bias in a situation where a confirmation factor analysis was not, this success is not entirely surprising given our previous discussion based on the mathematics underlying common method bias. That discussion clearly points at an increase in the overall level of collinearity in a model, corresponding to an increase in the full collinearity VIFs for the latent variables in the model, as a clear outcome of common method bias.

11.8 Discussion and Conclusion

Figure 11.4 summarizes key themes emerging from the results presented here. Common method variation leads to variation that is shared by all factors (i.e., latent variables) in an SEM model. This leads, more often than not, to inflation in factor-indicator loadings and linked factor-factor path coefficients. This joint inflation makes it difficult to detect possible common method bias using classic factor analysis and related tests. Our results demonstrate the need for a test that unveils excessive and pathological common variation. We offer a discussion of one such test, the full collinearity test, with promising outcomes.

There is disagreement among methodological researchers about the nature of common method bias, how it should be addressed, and even whether it should be addressed at all. Richardson et al. (2009) discuss various perspectives about common method bias, including the perspective put forth by Spector (1987) that common method bias is an "urban legend." Assuming that the problem is real, what can we do to avoid common method bias in the first place? A seminal source in this respect is Podsakoff et al. (2003), who provide a number of suggestions on how to avoid the introduction of common method bias during data collection.

Our discussion focuses on the identification of common method bias based on full collinearity assessment, whereby a model is checked for the existence of both vertical and lateral collinearity (Kock and Gaskins 2014; Kock and Lynn 2012). If we find evidence of common method bias, is there anything we can do to eliminate or at least reduce it? The answer is arguably "yes," and, given the focus of our discussion, the steps discussed by Kock and Lynn (2012) for dealing with collinearity are an obvious choice: indicator removal, indicator reassignment, latent variable removal, latent variable aggregation, and hierarchical analysis. Readers are referred to that publication for details on how and when to implement these steps.

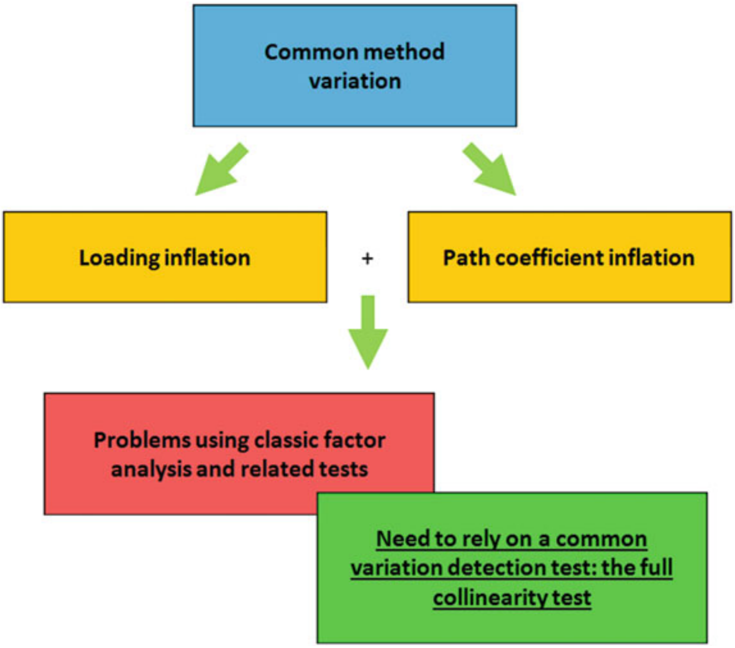


Fig. 11.4 The need for the full collinearity test

Full collinearity VIFs tend to increase with model complexity, in terms of number of latent variables in the model, because (a) the likelihood that questions associated with different indicators will overlap in perceived meaning goes up as the size of a questionnaire increases, which should happen as the number of constructs covered grows and (b) the likelihood that latent variables will overlap in terms of the facets of the constructs to which they refer goes up as more latent variables are added to a model.

Models found in empirical research studies in the field of e-collaboration typically contain more than three latent variables. This applies to many other fields where path analysis and SEM are employed. Therefore, we can reasonably conclude that our illustration of the full collinearity test of common method bias discussed here is conservative in its demonstration of the likely effectiveness of the test in actual empirical studies.

Kock and Lynn (2012) pointed out that classic PLS-SEM algorithms are particularly effective at reducing model-wide collinearity, because those algorithms maximize the variance explained in latent variables by their indicators. Such maximization is due in part to classic PLS-SEM algorithms not modeling measurement error, essentially assuming that it is zero. As such, the indicators associated with a latent variable always explain 100% of the variance in the latent variable. Nevertheless, one of the key downsides of classic PLS-SEM algorithms is that path coefficients tend to be attenuated (Kock 2015b). In a sense, they reduce collinearity levels “too much.” The recently proposed factor-based PLS-SEM algorithms (Kock 2014) address this problem. Given this, one should expect the use of factor-based PLS-SEM algorithms to yield slightly higher full collinearity VIFs than classic PLS-SEM algorithms, with those slightly higher VIFs being a better reflection of the true values. Consequently, the VIF threshold used in common method bias tests should arguably be somewhat higher than 3.3 when factor-based PLS-SEM algorithms are used. In their discussion of possible thresholds, Kock and Lynn (2012) note that a VIF of 5 could be employed when algorithms that incorporate measurement error are used. Even though they made this remark in reference to covariance-based SEM algorithms, the remark also applies to factor-based PLS-SEM algorithms, as both types of algorithms incorporate measurement error.

Our goal here is to help empirical researchers who need practical and straightforward methodological solutions to assess the overall quality of their measurement frameworks. To that end, we discussed and demonstrated a practical approach whereby researchers can conduct common method bias assessment based on a full collinearity test of a model. Our discussion was illustrated with outputs of the software WarpPLS (Kock 2015a), in the context of e-collaboration research. Nevertheless, our discussion arguably applies to any field where path analysis and SEM can be used.

Acknowledgments This book chapter is based on a previous version published as an article in the International Journal of e-collaboration. The author is the developer of the software WarpPLS, which has over 7000 users in more than 33 different countries at the time of this writing, and moderator of the PLS-SEM e-mail distribution list. He is grateful to those users and to the members of the PLS-SEM e-mail distribution list, for questions, comments, and discussions on topics related to SEM and to the use of WarpPLS.

References

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences*. Mahwah, NJ: L. Erlbaum.

Kock, N. (2005). What is e-collaboration. *International Journal of E-collaboration, 1*(1), 1–7.

Kock, N. (2008). E-collaboration and e-commerce in virtual worlds: The potential of second life and world of warcraft. *International Journal of E-collaboration, 4*(3), 1–13.

Kock, N. (2011). A mathematical analysis of the evolution of human mate choice traits: Implications for evolutionary psychologists. *Journal of Evolutionary Psychology, 9*(3), 219–247.

Kock, N. (2014). *A note on how to conduct a factor-based PLS-SEM analysis*. Laredo, TX: ScriptWarp Systems.

Kock, N. (2015a). *WarpPLS 5.0 user manual*. Laredo, TX: ScriptWarp Systems.

Kock, N. (2015b). One-tailed or two-tailed P values in PLS-SEM? *International Journal of E-collaboration, 11*(2), 1–7.

Kock, N. (2015c). A note on how to conduct a factor-based PLS-SEM analysis. *International Journal of E-collaboration, 11*(3), 1–9.

Kock, N., & Gaskins, L. (2014). The mediating role of voice and accountability in the relationship between Internet diffusion and government corruption in Latin America and Sub-Saharan Africa. *Information Technology for Development, 20*(1), 23–43.

Kock, N., & Lynn, G. S. (2012). Lateral collinearity and misleading results in variance-based SEM: An illustration and recommendations. *Journal of the Association for Information Systems, 13*(7), 546–580.

Maruyama, G. M. (1998). *Basics of structural equation modeling*. Thousand Oaks, CA: Sage.

Mueller, R. O. (1996). *Basic principles of structural equation modeling*. New York: Springer.

Paxton, P., Curran, P. J., Bollen, K. A., Kirby, J., & Chen, F. (2001). Monte Carlo experiments: Design and implementation. *Structural Equation Modeling, 8*(2), 287–312.

Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *The Journal of Applied Psychology, 88*(5), 879–903.

Richardson, H. A., Simmering, M. J., & Sturman, M. C. (2009). A tale of three perspectives: Examining post hoc statistical techniques for detection and correction of common method variance. *Organizational Research Methods, 12*(4), 762–800.

Robert, C. P., & Casella, G. (2005). *Monte Carlo statistical methods*. New York: Springer.

Spector, P. E. (1987). Method variance as an artifact in self-reported affect and perceptions at work: Myth or significant problem? *The Journal of Applied Psychology, 72*(3), 438–443.

Wold, H. (1980). Model construction and evaluation when theoretical knowledge is scarce: Theory and application of partial least squares, Chap. 3. In J. Kmenta & J. B. Ramsey (Eds.), *Evaluation of econometric models* (pp. 47–74). Cambridge, MA: Academic Press.

Wright, S. (1934). The method of path coefficients. *Annals of Mathematical Statistics, 5*(3), 161–215.

Wright, S. (1960). Path coefficients and path regressions: Alternative or complementary concepts? *Biometrics, 16*(2), 189–202.

Chapter 12
Integrating Non-metric Data in Partial Least Squares Path Models: Methods and Application

Francesca Petrarca, Giorgio Russolillo, and Laura Trinchera

Abstract In this chapter we discuss how to include non-metric variables (i.e., ordinal and/or nominal) in a PLS path model. We present the Non-Metric PLS approach for handling these type of variables, and we integrate the logistic regression into the PLS path model for predicting binary outcomes. We discuss features and properties of these PLS Path Modeling enhancements via an application on real data. We use data collected by merging the archives of Sapienza University of Rome and the Italian Ministry of Labor and Social Policy. The analysis of this data measured quantitatively, for the first time in Italy, the impact of graduates' Educational Performance on the first 3 years of their job career.

12.1 Introduction

In recent years, Italian universities have been urged to improve their understanding of labor market demand for their graduates. Due to the ongoing crisis in Italy, assessing young people's employment conditions, especially those with higher education qualifications, has become crucial.

In March 2008, the Italian government initiated administrative archives for the labor market. Since then, all employers (individuals, companies and public entities) are required to fill out several forms regarding the start and end dates and any extension or alteration of their employees' contracts. These communications are collected in the *Compulsory Communication (CO)* administrative archive. In 2011,

F. Petrarca
Department of Methods and Models for Economics, Territory and Finance, Sapienza University of Rome, Via del Castro Laurenziano 9, 00161 Rome, Italy
e-mail: francesca.petrarca@uniroma1.it

G. Russolillo (✉)
CNAM - Laboratoire Cédric, 292 Rue Saint-Martin, 75141 Paris Cedex 03, France
e-mail: giorgio.russolillo@cnam.fr

L. Trinchera
NEOMA Business School, 1 Rue du Maréchal Juin, BP 215, 76825 Cedex Mont Saint Aignan, France
e-mail: laura.trinchera@neoma-bs.fr

H. Latan, R. Noonan (eds.), *Partial Least Squares Path Modeling*,
DOI 10.1007/978-3-319-64069-3_12

for the first time in Italy, the Italian Ministry of Labor archive was merged with that of the Italian University "Sapienza," the largest European university by enrollments and the oldest of Rome's four state universities. This amalgamation has generated a new archive, UNI.CO, which we were able to use to investigate quantitatively, for the first time in Italy, the employability of Sapienza alumni in the Italian labor market, and how their careers develop during the first three years after graduation.

We modeled data from the UNI.CO database by means of Partial Least Squares Path Modeling (PLS-PM). To take into account the scale of ordinal and nominal manifest variables correctly in the outer model we introduce the Non-Metric PLS-PM approach (NM-PLSPM) (Russolillo 2012, 2014). Moreover, we use logistic regression instead of ordinary least squares (OLS) regression to estimate structural coefficients for predicting binary outcomes within the inner model.

The chapter is structured as follows. In Sect. 12.2, we briefly describe PLS-PM and NM-PLSPM. We also present an extension of PLS-PM for endogenous latent variables with a binary single indicator. We present the UNI.CO archive in Sect. 12.3. In Sects. 12.4 and 12.5, we present and analyze two different models for analyzing Sapienza alumni careers. The first model describes alumni who had not yet obtained employment when they graduated. The second model refers to alumni who had already obtained a job when they graduated. Finally, in Sect. 12.6 we draw our conclusions.

12.2 Methods

12.2.1 Partial Least Squares Path Modeling

The aim of Partial Least Square Path Modeling (PLS-PM) is to analyze path models with unobserved variables. The key idea behind PLS-PM is that each unobserved variable [commonly referred to as *construct* or *latent variable (LV)*] in the path model can be approximated by a linear combination (that is a *composite*, according to Lohmöller 1989 and Bollen and Bauldry 2011) of related observed variables [commonly referred to as *manifest variable (MV)*]. The system of weights to be applied to the manifest variables to compute LV proxies is obtained as the result of an iterative algorithm, i.e., the PLS algorithm. The PLS algorithm comprises two different steps, the inner and outer estimation steps. In the inner estimation step, composites are obtained as weighted aggregates of connected composites, while in the outer estimation step, composites are calculated as weighted aggregates of their corresponding MVs. Different options exist for calculating the weights in the two steps. In practice, the option chosen for the inner weights does not substantially affect the results of the algorithm (Noonan and Wold 1982). Here, we focus on two options, called the *centroid scheme* and the *factorial scheme*, for which theoretical convergence proprieties have been proved. As for the outer weights, two main modes can be used for a set of MVs related to a LV: Mode A and Mode B, which are graphically represented in the path diagram by outward and inward directed arrows, respectively. Mode A uses correlation weights, while Mode B uses OLS regression weights. Since correlation weights do not account for shared variability among MVs, their stability is not affected by multicollinearity. Moreover, correlation weights tend to yield higher out-of-sample R^2 when sample size and true predictability are moderate. On the other hand, OLS regression weights are recommended for very high sample sizes and low multicollinearity (Becker et al. 2013). Once the algorithm converges, the composites are obtained as weighted sums of corresponding MVs with the system of weights given by the PLS algorithm. Finally, path coefficients and loadings are estimated as simple or multiple regression coefficients according to the system of interdependent equations represented by the path diagram. A complete description of the PLS algorithm and of the statistical properties of the PLS-PM can be found in Esposito et al. (2010).

The PLS algorithm does not maximize a unique criterion: the solution depends on the chosen inner weighting scheme and on the selected mode for each block in the model. Although Dijkstra (1981) first pointed out a relationship between Mode B and Generalized Canonical Correlation Analysis as early as 1981, for a long while no other significant results on the properties of the PLS-PM algorithm were published. In 2007, 26 years after Dijkstra's first findings, Hanafi (2007) proved that when Mode B is selected for all blocks in the model, the PLS algorithm, as proposed by Wold (1985), monotonically converges to at least a local maximum of the function:

$$f(\mathbf{w}_1, \mathbf{w}_2, \dots \mathbf{w}_Q) = \sum_{q \neq q'} c_{qq'} g(\text{cov}(\mathbf{X}_q \mathbf{w}_q, \mathbf{X}_{q'} \mathbf{w}_{q'})) \tag{12.1}$$
$$\text{s.t.} \quad \text{var}(\mathbf{X}_q \mathbf{w}_q) = 1, \quad q = \{1, \dots, Q\}$$

where q is an index that generically refers to one of the Q LVs in the model; $\mathbf{X}_q$ refers to the block of MVs related to the qth LV; $\mathbf{w}_q$ refers to the system of weights to be applied to $\mathbf{X}_q$ to compute the LV proxy; $c_{qq'} = 1$ if two LVs are connected in the path diagram, and $c_{qq'} = 0$ otherwise; $g()$ is a function depending on the inner weight scheme, i.e., $g() = \text{square}()$ if the factorial scheme is used, and $g() = \text{abs}()$ if the centroid scheme is used.

In the same year, Krämer (2007) showed that the "full Mode A" PLS-PM algorithm is not based on a stationary equation related to the optimization of a twice differentiable function. She proposed a slightly adjusted Mode A (known in the literature as New Mode A), in which a normalization constraint is put on outer weights rather than composites, to obtain a stationary point of the optimization problem linked to the maximization of the following criterion:

$$f(\mathbf{w}_1, \mathbf{w}_2, \dots \mathbf{w}_Q) = \sum_{q \neq q'} c_{qq'} g(\text{cov}(\mathbf{X}_q \mathbf{w}_q, \mathbf{X}_{q'} \mathbf{w}_{q'})) \tag{12.2}$$
$$\text{s.t.} \quad \|\mathbf{w}_q\|^2 = 1, \quad q = \{1, \dots, Q\}$$

Tenenhaus and Tenenhaus (2011) extended Hanafi's results and proposed a more flexible criterion, called Regularized Generalized Canonical Correlation Analysis (RGCCA), in which "New Mode A" and "Mode B" are mixed by means a vector of regularization parameters τ_q $(q = 1, \ldots, Q)$; when τ_q is either zero or 1, RGCCA becomes PLS-PM when "New Mode A" is used for some blocks and "Mode B" for others. In this case, PLS-PM procedure monotonically converges towards the maximization of the criterion

$$f(\mathbf{w}_1, \mathbf{w}_2, \ldots \mathbf{w}_Q) = \sum_{q \neq q'} c_{qq'} g(\mathrm{cov}(\mathbf{X}_q \mathbf{w}_q, \mathbf{X}_{q'} \mathbf{w}_{q'})) \tag{12.3}$$

$$\text{s.t.} \quad \tau_q ||\mathbf{w}_q||^2 + (1 - \tau_q)\mathrm{var}(\mathbf{X}_q \mathbf{w}_q) = 1, \quad q = \{1, \ldots, Q\}, \quad \tau_q = \{0, 1\}$$

Moreover, Tenenhaus and Tenenhaus (2011) rewrote criterion (12.3) as a function of the inner estimates of the latent variables $\mathbf{z}_q$:

$$f(\mathbf{w}_1, \mathbf{w}_2, \ldots \mathbf{w}_Q) = \sum_{q} \mathrm{cov}(\mathbf{X}_q \mathbf{w}_q, \mathbf{z}_q) \tag{12.4}$$

From criterion (12.3) it is clear that, if Mode B is chosen, the PLS algorithm provides a system of weights that maximizes the sum of the correlations between connected composites. If the New Mode A is used instead, the algorithm searches for weights providing composites that are correlated as much as possible to both the connected composites and to their own MVs. Even if Mode A cannot be proven to converge to criterion (12.2), in practice it yields solutions very close to those of new Mode A.

12.2.2 PLS-PM with Binary Single-Indicator Endogenous Latent Variables

As discussed in Sect. 12.2.1, once the PLS algorithm has converged, the path coefficients are obtained as regression coefficients of linear models among composite, $\widehat{\boldsymbol{\xi}}_q$. Hence, each endogenous composite $\widehat{\boldsymbol{\xi}}_q$ is modeled as:

$$\widehat{\boldsymbol{\xi}}_q = \sum_{q' : \widehat{\boldsymbol{\xi}}_{q'} \in \Omega_q} \beta_{q'q} \widehat{\boldsymbol{\xi}}_{q'} + \boldsymbol{\zeta}_q \tag{12.5}$$

where Ω_q is the set of composites explaining $\widehat{\boldsymbol{\xi}}_q$, $\beta_{q'q}$ is the path coefficient linking the q'th composite to the qth endogenous composite, and $\boldsymbol{\zeta}_q$ is a residual vector.

In the case of single-indicator latent variables (i.e., LVs associated with a single MV), the computed composite and the associated MV are perfectly collinear. If the single indicator is a binary variable, the connected composite also has a binary structure. In such particular cases, the linear model in Eq. (12.5) is no longer appropriate. Therefore, we propose to model the composite as a Bernoulli random variable of parameter π, and to use a generalized linear model to predict an endogenous single-indicator binary composite:

$$\pi_q = E(\widehat{\boldsymbol{\xi}}_q = 1|\widehat{\boldsymbol{\xi}}_{q' \in \Omega_q}) = \frac{\exp(\sum_{q':\widehat{\boldsymbol{\xi}}_{q'} \in \Omega_q} \beta_{q'q}\widehat{\boldsymbol{\xi}}_{q'})}{1 + \exp(\sum_{q':\widehat{\boldsymbol{\xi}}_{q'} \in \Omega_q} \beta_{q'q}\widehat{\boldsymbol{\xi}}_{q'})} \tag{12.6}$$

where π_q is the conditional probability of $\widehat{\boldsymbol{\xi}}_q = 1$.

In such models, it is possible to interpret each path coefficient as the marginal effect on the (logarithmic) odds ratio and to estimate the probability that each single-indicator binary endogenous composite will be equal to 1. This yields new insights into the interpretation of the structural model involving binary single-indicator endogenous LVs.

12.2.3 Non-metric Partial Least Squares Path Modeling

PLS-PM assumes that

- each manifest variable is measured on an interval (or ratio) scale;
- relations between variables and latent constructs are linear and, consequently, monotone.

Therefore, traditional PLS-PM can only handle data measured on a scale with metric properties, but it cannot properly handle ordinal and nominal variables, which are by definition (Stevens 1946) sets in which the notion of distance (metric) between elements in the set is not defined. Most software suggests performing PLS analyses by replacing nominal MVs with the corresponding indicator matrix. However, this solution is not recommended because, as pinpointed by Russolillo (2012):

- complete disjunctive coding conflicts with the idea of the variable as a whole: it considers categories as if they were variables in themselves;
- binary coding inflates the dimensionality of the data matrix;
- the weight of a dummy variable representing a category mainly associated with central values of the corresponding LV score distribution is systematically underestimated.

Ordinal MVs, frequently measured on Likert scales, are usually handled as if they were numerical variables; this pragmatic approach leads to a (non-optimal) scaling procedure which assigns a numerical value to each level of the ordinal variable under the assumption that consecutive levels are equally spaced.

Three main approaches have been proposed in the literature for handling specifically nominal and ordinal MVs in PLS-PM. Jakobowicz and Derquenne

(2007) proposed a modified PLS-PM algorithm which interprets the categories of a nominal variable as distinct variables, and computes a weight for each of them. The second approach (Cantaluppi 2012; Cantaluppi and Boari 2016) applies the standard PLS-PM algorithm to a polychoric correlation matrix rather than the traditional Bravais-Pearson correlation matrix, to handle ordinal categorical MVs. The use of the polychoric correlation matrix has recently been combined with Consistent PLS (Dijkstra and Henseler 2015) to obtain a consistent estimator in the framework of traditional common factor models using ordinal MVs (Schuberth et al. 2016). The third approach, recently proposed by Russolillo (2012, 2014), is Non-Metric Partial Least Squares (NM-PLS).

NM-PLS refers to a new class of PLS type algorithms that allow the iterative PLS algorithm to act as an Optimal Scaling (OS) algorithm. To briefly review this approach, we introduce three concepts of OS theory: *scaling*, *scaling level*, and *optimal scaling*. *Scaling* a variable means providing the original variable with a metric: if the original variable takes a set of K categories (or levels, or distinct values), a scaling operator replaces these modalities by a new set of numeric values ϕ_k ($k = 1, \dots, K$), named *scaling values* or *quantifications*. The set of suitable scaling values is constrained by a *scaling level*. The scaling level defines the properties of the original measurement scale that have to be kept in the new interval scale. For example, a variable observed on an ordinal scale has two properties: order (modalities have an inherent order from smaller to larger) and grouping (each modality defines a particular group of units); one can choose to retain the grouping property only: in this case, the variable will be scaled at a nominal level. On the other hand, if also the order property is preserved, the scaling level will be ordinal. The ordinal scaling level can also be applied to quantitative variables, to reject the hard assumption of linearity in favor of the milder assumption of monotonicity. To define this scaling process as *optimal*, the set of scaling values must be:

- Suitable, as it must respect the constraints defined by the scaling level;
- Optimal, as it must optimize the same criterion as the analysis in which the scaling process is involved.

In Non-Metric PLS Path Modeling (NM-PLSPM), scaling values are used to transform each original variable $\mathbf{x}_{pq}$ (the pth variable of the qth block, $p = 1, \dots, P_q$) into a new variable $\hat{\mathbf{x}}_{pq}$ with a metric structure (an interval measurement scale). The new variable is obtained as $\hat{\mathbf{x}}_{pq} = \tilde{\mathbf{X}}_{pq}\boldsymbol{\phi}_{pq}$, where the matrix $\tilde{\mathbf{X}}_{pq}$ defines a space in which constraints imposed by the scaling level are respected. Here, the scaling level imposed for ordinal quantifications is defined according to Kruskal's weak monotonicity approach (Kruskal 1964), which preserves the order of the categories of $\mathbf{x}$ as required by the conditions

$$(x_i \sim x_{i'}) \Rightarrow (\hat{x}_i = \hat{x}_{i'}) \text{ and } (x_i \prec x_{i'}) \Rightarrow (\hat{x}_i \leq \hat{x}_{i'}). \tag{12.7}$$

where i and i' are two distinct observations and the symbol $\prec$ indicates empirical order.

To guarantee the optimality of the scaling process in the PLS-PM framework, in NM-PLSPM the following optimization problem is considered:

$$\underset{\forall \mathbf{w}_q, \boldsymbol{\phi}_{pq}}{\text{Maximize}} \sum_{q \neq q'} c_{qq'} g(\text{cov}(\hat{\mathbf{X}}_q \mathbf{w}_q, \hat{\mathbf{X}}_{q'} \mathbf{w}_{q'})) \tag{12.8}$$

$$\text{s.t.} \quad \tau_q ||\mathbf{w}_q||^2 + (1 - \tau_q)\text{var}(\mathbf{X}_q \mathbf{w}_q) = 1, \quad \text{var}(\hat{\mathbf{x}}_{pq}) = 1$$
$$\tau_q = \{0, 1\}, \quad p = \{1, \ldots, P_q\}, \quad q = \{1, \ldots, Q\}$$

This criterion is similar to that optimized by PLS-PM, but it depends on two set of parameters: model parameters ($\mathbf{w}_q$) and scaling parameters ($\boldsymbol{\phi}_{pq}$). Moreover, it is equivalent to the criterion

$$\underset{\forall \mathbf{w}_q, \boldsymbol{\phi}_{pq}}{\text{Maximize}} \sum_{q} \text{cov}(\hat{\mathbf{X}}_q \mathbf{w}_q, \mathbf{z}_q) \tag{12.9}$$

where function $g()$ and $\mathbf{z}_q$ are defined in Sect. 12.2.1. Russolillo (2012) proposed the NM-PLSPM algorithm to solve this optimization problem when New Mode A is used for all the blocks. He then extended the algorithm to include the Mode B option (Russolillo 2014). The NM-PLSPM algorithm alternately optimizes criterion (12.9) with respect to each subset of parameters, keeping the other fixed. When PLS parameters $\mathbf{w}_q$ are optimized for fixed scaling parameters $\boldsymbol{\phi}_{pq}$, the usual PLS-PM iteration steps are used, while a quantification step is introduced in the iteration for optimizing $\boldsymbol{\phi}_{pq}$ for fixed $\mathbf{w}_q$. In particular, when either New Mode A or Mode A are used, the optimal solution for $\boldsymbol{\phi}_{pq}$ is given by the quantification function $\mathscr{Q}(\tilde{\mathbf{X}}_{pq}, \mathbf{z}_q)$ which orthogonally projects $\mathbf{z}_q$ into the space spanned by $\tilde{\mathbf{X}}_{pq}$; Russolillo (2014) proposed a backfitting procedure for optimizing $\boldsymbol{\phi}_{pq}$ for a given $\mathbf{w}_q$ when Mode B is used. This procedure yields the quantification function $\mathscr{Q}(\tilde{\mathbf{X}}_{pq}, \mathbf{z}_q^*)$, which orthogonally projects the vector $\mathbf{z}_q^* = (1/b_{pq})(\mathbf{z}_q - \sum_{j \neq p} b_{jq} \hat{\mathbf{x}}_{jq})$ into the space spanned by $\tilde{\mathbf{X}}_{pq}$, where b_{pq} is the pth element of the regression coefficient vector $\mathbf{b}_q = (\hat{\mathbf{X}}_q' \hat{\mathbf{X}}_q)^{-1} \hat{\mathbf{X}}_q' \mathbf{z}_q$.

The NM-PLSPM algorithm is expected to converge to at least a local solution of the general optimization problem presented in Eq. (12.8). Hence, whenever the PLS-PM algorithm is used to optimize a known criterion with respect to the model parameters, the corresponding Non-Metric version can be used to optimize to the same criterion with respect to both scaling and model parameters.

12.3 Data

The UNI.CO archive contains information about 21,782 Sapienza alumni who graduated between March 1st, 2008 and February 28th, 2009 and who signed a job contract in the subordinate or para-subordinate labor markets.[1] For more details about the UNI.CO archive and for preliminary statistical analyses based on a subset of this archive, see Alleva and Petrarca (2013), Gruppo UNI.CO (2015) and Petrarca (2014a,b).

We selected the 5602 alumni from the UNI.CO archive who graduated from a master program and "occurred" at least once in the CO archive. An occurrence is defined as working experience (paid contract) or professional experience (unpaid contract), either signed within 3 years of graduation or already in effect at the time of graduation. Job contracts in the UNI.CO archive are classified in five classes according to the amount of protection provided under Italian employment legislation as defined by the Italian National Statistical Institute (ISTAT). Class 1 corresponds to poor contractual protection, and Class 5 corresponds to the highest level of contractual protection. In addition, we classified the job contracts in nine groups of jobs according to tasks and duties undertaken, as indicated by the International Standard Classification of Occupations (ISCO) scale.[2] ISCO9 corresponds to least skilled jobs and ISCO1 corresponds to the highest skill level. For interpretation purposes we used the reversed ISCO scale in the model.

We focused on job contracts signed by each graduate within 3 years of graduation, and selected their first and last jobs together with the two jobs with the highest protection and skill levels. Note that for a given graduate one, two, three, or four jobs could be selected: e.g., they may only have had one job, and the highest level of protection/skill could overlap with the first and/or last jobs. Using this information we defined the trend in contractual protection as a binary variable indicating whether a graduate obtained more, or at least equal, contractual protection over time. In the same way, we defined the trend in job skill as a binary variable indicating whether a graduate obtained more skilled or equally skilled work over time. We used these two binary variables as indicators of career development in the three years after graduation. The list of variables used in the successive analyses is provided in Table 12.1.

[1] A para-subordinate job is a hybrid between a paid employment (subordinate job) and a self-employed economic activity.

[2] More details about this classification at http://www.ilo.org/public/english/bureau/stat/isco/docs/resol08.pdf.

Table 12.1 List of observed variables: for ordinal variables the number of levels is reported in parentheses

MV	Description	Scale
Age at graduation	Age at the time of graduation	Metric
Graduation grade	Graduation grade	Metric
Average grade	Average grade during the master degree	Metric
Supplementary year student	Dummy variable representing whether a student graduated on time (or not)	Binary
First ISCO	Skill level of first employment	Ordinal (9)
First protection	Contractual protection level of first employment	Ordinal (5)
Highest ISCO	Most skilled job obtained within three years of graduation	Ordinal (9)
Highest protection	Highest job contract protection obtained within three years of graduation	Ordinal (5)
Protection improvement	Dummy variable indicating whether a graduate obtained more, or at least equal, contractual protection over time	Binary
Skill level improvement	Dummy variable indicating whether a graduate obtained more skilled or equally skilled work over time	Binary

Table 12.2 Descriptive statistics for the metric variables in the model

MVs		Min	1st qu.	Median	Mean	3rd qu.	Max.
Age at graduation	Students	23	26	27	27.58	29	59
	Student workers	23	26	28	30.36	33	63
Graduation grade	Students	72	103	110	106.12	111	111
	Student workers	72	100	106	104.20	111	111
Average grade	Students	20.27	25.93	27.54	27.13	28.61	30
	Student workers	18.75	25.33	27.00	26.69	28.36	30

12.4 Structural Paths and Constructs Definition

Preliminary analyses of variables in Table 12.1 revealed a high degree of heterogeneity among the alumni. Consequently we decided to analyze two groups of alumni separately:

- STUDENTS: alumni who had not yet signed a job contract at the time of graduation (4605 units).
- STUDENT WORKERS: alumni with job contract in effect at the time of graduation (997 units).

We provide descriptive statistics for the quantitative variables in Table 12.2, and relative frequencies for the non-metric variables are provided in Table 12.3.

Table 12.3 Descriptive statistics for the non-metric variables in the model

MVs	Categories	STUDENTS (4605)	STUDENT WORKERS (997)	All (5602)
First protection	(1) Professional experience	0.04	0.18	0.08
	(2) Apprenticeships	0.04	0.21	0.07
	(3) Non-standard employment	0.66	0.47	0.62
	(4) Part-time permanent employment	0.05	0.06	0.05
	(5) Full-time permanent employment	0.21	0.08	0.18
Highest protection	(1) Professional experience	0.04	0.14	0.06
	(2) Apprenticeships	0.03	0.15	0.05
	(3) Non-standard employment	0.61	0.49	0.58
	(4) Part-time permanent employment	0.08	0.1	0.85
	(5) Full-time permanent employment	0.24	0.12	0.22
First ISCO	(1) ISCO 9	0.02	0.03	0.02
	(2) ISCO 8	0.00	0.00	0.00
	(3) ISCO 7	0.01	0.01	0.01
	(4) ISCO 6	0.00	0.0	0.00
	(5) ISCO 5	0.13	0.18	0.13
	(6) ISCO 4	0.22	0.31	0.24
	(7) ISCO 3	0.24	0.26	0.24
	(8) ISCO 2	0.38	0.20	0.35
	(9) ISCO 1	0.00	0.01	0.00
Highest ISCO	(1) ISCO 9	0.01	0.01	0.01
	(2) ISCO 8	0.00	0.00	0.00
	(3) ISCO 7	0.00	0.03	0.00
	(4) ISCO 6	0.00	0.00	0.00
	(5) ISCO 5	0.08	0.10	0.08
	(6) ISCO 4	0.18	0.25	0.19
	(7) ISCO 3	0.23	0.28	0.25
	(8) ISCO 2	0.50	0.34	0.47
	(9) ISCO 1	0.00	0.01	0.00
Supplementary year student	(1) No	0.64	0.74	0.66
	(0) Yes	0.36	0.26	0.34
Protection improvement	(0) No	0.83	0.91	0.84
	(1) Yes	0.17	0.09	0.16
Skill level improvement	(0) No	0.87	0.84	0.87
	(1) Yes	0.13	0.16	0.13

We assumed that the same latent variables would be involved in the analysis, but that the relations between them would be different. Because of the sequence of events characterizing each group of alumni and to the characteristics of the Italian university system we assumed a number of prior hypotheses that specify two different structural models. The first model applies to STUDENTS. The second model refers to alumni with a job contract in effect at the time of graduation. For both the STUDENTS and STUDENT WORKERS we expect that students who perform better graduate faster than others, thus at an earlier age. We also expect that brilliant, younger graduates find jobs with high protection and skill level; and that the best job graduates obtain within 3 years of graduation depends on their educational performance, their age at graduation, and their first employment. Furthermore we assumed that improvement in both contract protection and job skill level depends on the best and the first employments.

The main difference between the two models relies on the role of the first employment in the model: for STUDENTS we supposed the first employment to depend on age at graduation and on educational performance; however, in the STUDENT WORKERS model, the first employment is an exogenous variable impacting both age at graduation and educational performance. We summarize the postulated hypotheses in Table 12.4 and in Figs. 12.1 and 12.2.

Three of the six latent variables in the models are single-indicator latent variables, i.e., *Age at graduation*, *Protection improvement*, and *Skill level improvement*. Two are inwards constructs, i.e., *First employment* and *Best employment*, each formed by two observed variables. The last, i.e., *Educational performance* is approximated by a composite built on three indicators. We provide the list of the latent variables with associated MVs in Table 12.5.

Table 12.4 Structural model hypotheses. Hypotheses H1 and H3a to H5 refer to both STUDENTS and STUDENT WORKERS models

STUDENTS	STUDENT WORKERS
H1: *Age at graduation* depends on *Educational performance*	
H2a: *First employment* depends on *Educational performance*	**H2a**$_{\text{Reversed}}$: *First employment* impacts *Educational performance*
H2b:*First employment* depends on *Age at graduation*	**H2b**$_{\text{Reversed}}$: *First employment* impacts *Age at graduation*
H3a: *Best employment* depends on *Educational performance*	
H3b: *Best employment* depends on *First employment*	
H3c: *Best employment* depends on *Age at graduation*	
H4a: *Protection improvement* depends on *First employment*	
H4b: *Protection improvement* depends on *Best employment*	
H5a: *Skill level improvement* depends on *First employment*	
H5b: *Skill level improvement* depends on *Best employment*	

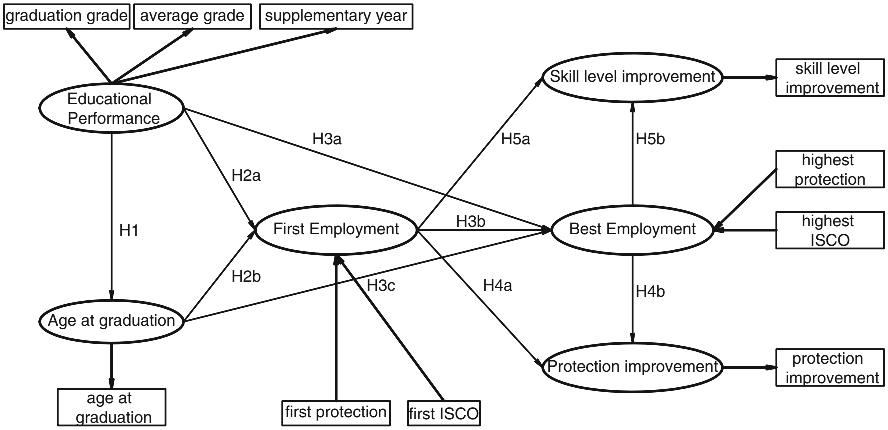


Fig. 12.1 Path diagram for STUDENTS

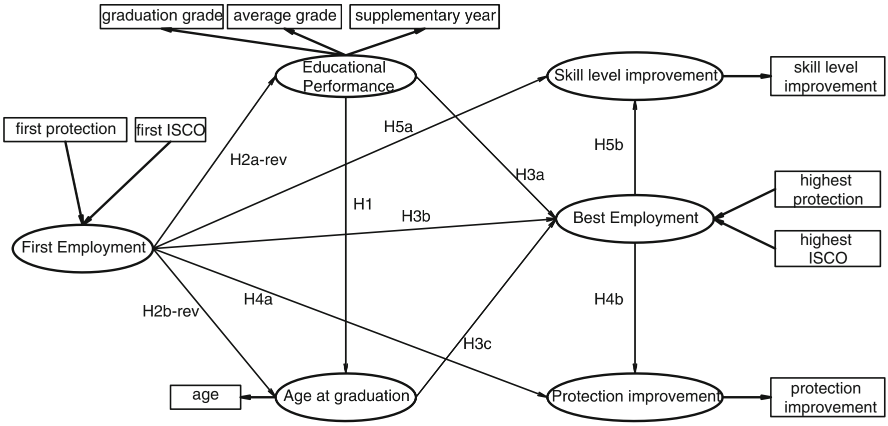


Fig. 12.2 Path diagram for STUDENT WORKERS

Table 12.5 Measurement model

LVs	Mode	MVs
Age at graduation	Single-indicator LV	Age at graduation
Educational performance	Mode A	Graduation grade
		Average grade
		Supplementary year
First employment	Mode B	First ISCO
		First protection
Best employment	Mode B	Highest ISCO
		Highest protection
Protection improvement	Binary single-indicator LV	Protection improvement
Skill level improvement	Binary single-indicator LV	Skill level improvement

We used the Non-Metric Partial Least Squares approach, as implemented in the *plspm* R-package (Sanchez et al. 2013), to estimate outer and inner model parameters referring to linear regressions. Afterwards, we used the *glm()* R function to estimate inner model parameters referring to the logistic regressions. To update the external weights, we used Mode A for *Educational performance*, because of the high collinearity between the corresponding MVs; we used Mode B for *First employment* and *Best employment*, as the corresponding MVs share very little variability. We used the centroid scheme for inner estimation and obtained confidence intervals for weights and path coefficients by percentile bootstrapping with $k = 1000$ re-samples.

12.5 Results

12.5.1 STUDENTS *Model*

A preliminary analysis of the STUDENTS dataset (4605 alumni) allowed us to remove the MV *first ISCO* from the model defined in Fig. 12.1, since bootstrapping did not validate the corresponding outer weight. When applying the NM-PLSPM we obtained the same kind of outputs as for standard PLS-PM, plus a scaling value for each category of ordinal/categorical variables in the model.

Figure 12.3 reports the optimal scaling values of the three ordinal MVs in the STUDENTS model. The more the dashed lines deviate from linearity, the more the new scale of the variable deviates from the assumption of equal distance between contiguous levels. Moreover, the optimal scaling values can be used to reduce the number of levels; for example, the quantifications obtained for the modalities of *highest ISCO* suggest that they could be grouped into three categories (ISCO9-ISCO6 for raw values 1–4, ISCO5-ISCO3 for raw values 5–7, ISCO2-ISCO1 for raw values 8–9). This result is consistent with the recent literature on occupational classification (Maselli 2012), which groups ISCO categories into the same categories, labeled low-skilled, medium-skilled, and high-skilled, respectively.

We report the measurement model results in Table 12.6. Along with the point estimates, we report 95% confidence intervals and the corresponding standard errors obtained using 1000 bootstrap samples. We have not reported the weights associated with single-indicator LVs because they equal 1, since the NM-PLSPM returns standardized manifest variables. As expected for an outward block, the three MVs associated with *Educational performance* show similar weights. High *Educational performance* is associated with no *supplementary year*, and high values of *graduation grade* and *average grade*. In other words, the best performing students graduated on time and with good final and average grades. When we compare the outer weights of the two indicators related to the *Best employment*, it is clear that the MV *highest protection* is associated with the highest impact. This is coherent with the Italian job market, in which people prefer to leave a poorly protected job to obtain better contactual protection than to obtain a more skilled job.

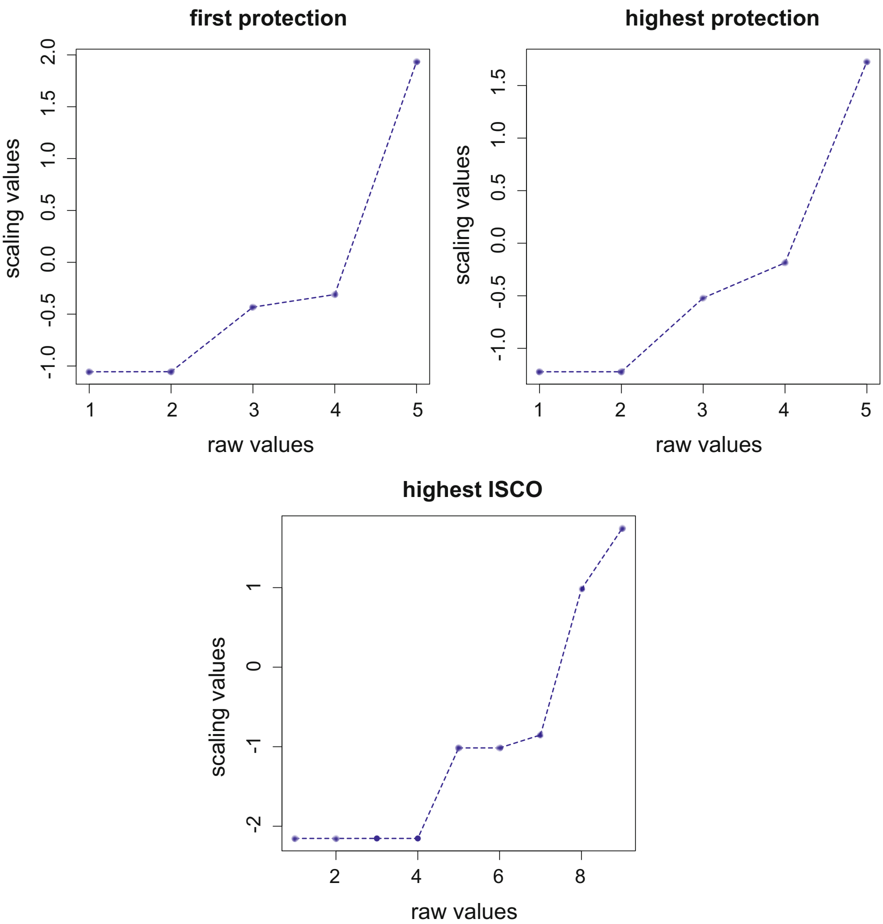


Fig. 12.3 STUDENT model results: optimal scaling values for ordinal MVs

Table 12.6 Measurement model results for STUDENTS: weights associated with single-indicator LVs are not shown

LVs	MVs	Weigths	Std. Error	C.I.
Educational performance	Supplementary year	0.44	0.02	[0.40, 0.47]
	Graduation grade	0.40	0.09	[0.39, 0.42]
	Average grade	0.37	0.09	[0.35, 0.38]
Best employment	Highest protection	0.98	0.01	[0.94, 0.99]
	Highest ISCO	0.14	0.06	[0.10, 0.29]

Confidence interval values were obtained by bootstrapping 1000 replicates

We report the results of our structural model, with the corresponding 95% confidence intervals, in Table 12.7 and Fig. 12.4. We used logistic regression to

Table 12.7 Structural model results for STUDENTS: we indicated pseudo-Nagelkerke R^2 with $*$ in the R^2 column; O.R. indicates the odds ratio for logistic regressions

Paths	R^2	$\hat{\beta}$	Std. error	C.I.	O.R.
Edu. perf. → Age graduation	0.17	−0.41	0.02	[−0.44, −0.38]	–
Edu. perf. → First empl.	0.03	0.08	0.02	[0.05, 0.13]	–
Age graduation → First empl.		−0.13	0.01	[−0.15, −0.10]	–
Edu. perf. → Best empl.	0.77	0.03	0.01	[0.01, 0.05]	–
Age graduation → Best empl.		−0.05	0.00	[−0.06, −0.03]	–
First empl. → Best empl.		0.87	0.00	[0.85, 0.88]	–
Best empl. → Protection impr.	0.77*	3.05	0.23	[2.64, 3.56]	21.11
First empl. → Protection impr.		0.57	0.07	[0.43, 0.72]	1.77
Best empl. → Skill level impr.	0.07*	0.77	0.07	[0.62, 0.91]	2.16
First empl. → Skill level impr.		−0.32	0.07	[−0.45, −0.18]	0.73

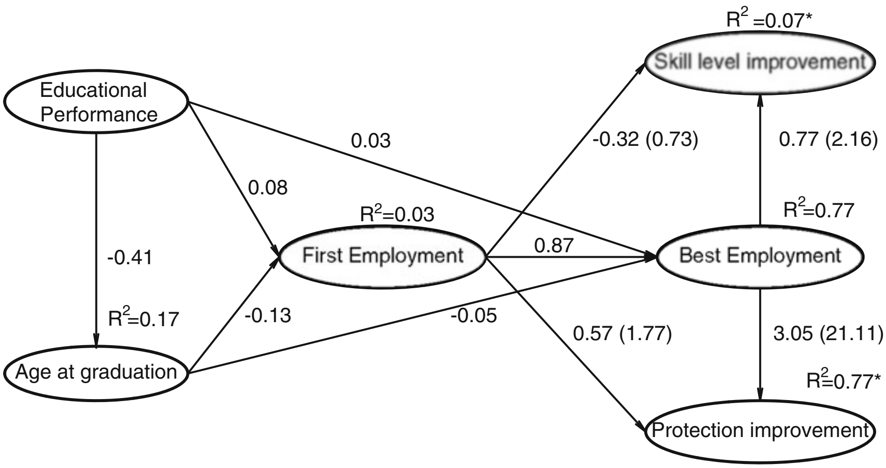


Fig. 12.4 Structural model for STUDENTS: values in parentheses indicate odds ratios; $*$ indicates pseudo-Nagelkerke R^2

predict the two binary single-indicator endogenous LV in the model, i.e., *Protection improvement* and *Skill level improvement*.

Our model does not explain *First employment* well, but predicts both *Best employment* and *Protection improvement* accurately. Figure 12.5 shows the ROC curves associated with the logistic regression in the structural model. The values of the AUC are 0.71 for the logistic model predicting *Skill level improvement* and 0.99 for that predicting *Protection improvement*.

Our results show that *Educational performance* has a significant negative impact on *Age at graduation*. H1 is validated: on average, high performing students graduate younger. *First employment* is significantly related to *Educational performance* and *Age at graduation*. Hypotheses 2a and 2b are validated: high

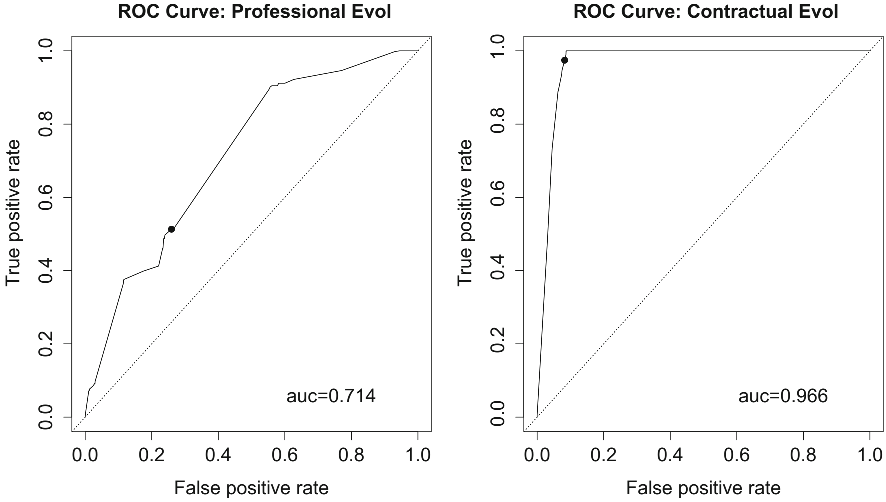


Fig. 12.5 Validation model for STUDENTS: the ROC curve. We report the value of the area under the curve (AUC)

performing graduates find better first jobs. *Best employment* mainly depends on *First employment* Hypotheses 3a, 3b, and 3c are validated. The two latent predictors *First employment* and *Best employment* positively influence the binary response *Protection improvement*, validating H4a and H4b. Holding the *Best employment* at a fixed value, for every one-unit increase in *First employment* score, the odds of observing increased contractual protection changes by a factor of 1.77 (with an average marginal effect of 0.004). Alumni who obtained a highly protected first contract are more likely to move to a more (or at least equally) protected job. Indeed, 74% of those who obtained a initial fulltime permanent contract changed it for a new contract of the same type. On the other hand, holding *First Employment* at a fixed value, for every one-unit increase in the *Best Employment* score, the odds of observing increased contractual protection changes by a factor of 21.11 (with an average marginal effect of 0.023). Therefore alumni with a highly skilled first employment are likely to improve their contractual protection.

First employment and *Best employment* also influence *Skill level improvement* significantly. As expected, *Best employment* positively influences *Skill level improvement*, as it results from individual efforts to obtain a highly skilled job. However the negative effect of *First employment* on *Skill level improvement* is less intuitive. Since the correlation between *First employment* and *Skill level improvement* is low ($r = 0.09$), while the correlation between *First employment* and *Best employment* is very high ($r = 0.88$), we have a *negative classical suppression* effect (see Cohen and Cohen (1975), Krus and Wilkinson (1986), Baron and Kenny (1986), Little et al. (2007) for details on suppression effects).

12.5.2 STUDENT WORKERS *Model*

We report measurement model results for STUDENT WORKERS in Table 12.8. The interpretation of outer weights is similar to the STUDENTS model. The only remarkable difference is that *highest ISCO* has a higher weight in this model than in the STUDENTS model, probably due to the fact that these alumni working longer.

We report structural model results in Table 12.9. The bootstrapping procedure validated all our hypotheses except for H1, H2b reversed, and H3c, so *Age at graduation* was excluded from the model. The path diagram representing the validated model is depicted in Fig. 12.6.

Hypothesis H2a reversed was validated, as we found a significant positive effect of *First employment* (for Student workers) on *Educational performance*: the better the job, the higher is student worker educational performance. Similarly to the previous model, *First employment* has a significant impact on *Best ermployment*, unlike *Educational performance*, whose impact is slight. *First employment* and *Best employment* explain Alumni's *Protection improvement* accurately. As for the STUDENTS, the impact of *Best employment* is much greater than that of *First employment*. Keeping *Best employment* at a fixed value, for every one-unit increase in the *First employment* score, the odds of observing increased contractual protection changes by a factor of 2.32 (with an average marginal effect of 0.001); on the other hand, keeping the *First employment* at a fixed value, for every one-unit increase in the *Best employment* score, the odds of observing increased contractual

Table 12.8 Measurement model results for STUDENT WORKERS: weights associated with single-indicator LVs are not shown

LVs	MVs	Weights	Std. error	C.I.
Educational performance	Supplementary year	0.40	0.04	[0.31, 0.48]
	Graduation grade	0.40	0.02	[0.37, 0.45]
	Average grade	0.40	0.02	[0.36, 0.46]
Best employment	Highest protection	0.91	0.01	[0.89, 0.94]
	Highest ISCO	0.29	0.03	[0.23, 0.34]

We obtained confidence interval values by bootstrapping 1000 replicates

Table 12.9 Structural model results for STUDENT WORKERS: we indicate pseudo-Nagelkerke R^2 with $*$ in the R^2 column; O.R. indicates the odds ratio

Paths	R^2	$\hat{\beta}$	Std. err	C.I.	O.R.
First empl. → Edu. perf.	0.06	0.24	0.03	[0.19, 0.32]	–
Edu. perf. → Best empl.	0.55	0.10	0.02	[0.05, 0.14]	–
First empl. → Best empl.		0.71	0.03	[0.66, 0.76]	–
First empl. → Protection impr.	0.84*	0.61	0.13	[0.36, 0.88]	2.32
Best empl. → Protection impr.		2.53	0.36	[1.94, 3.41]	12.55
First empl. → Skill level impr.	0.11*	−0.50	0.11	[−0.71, −0.28]	0.61
Best empl. → Skill level impr.		0.86	0.11	[0.64, 1.064]	2.36

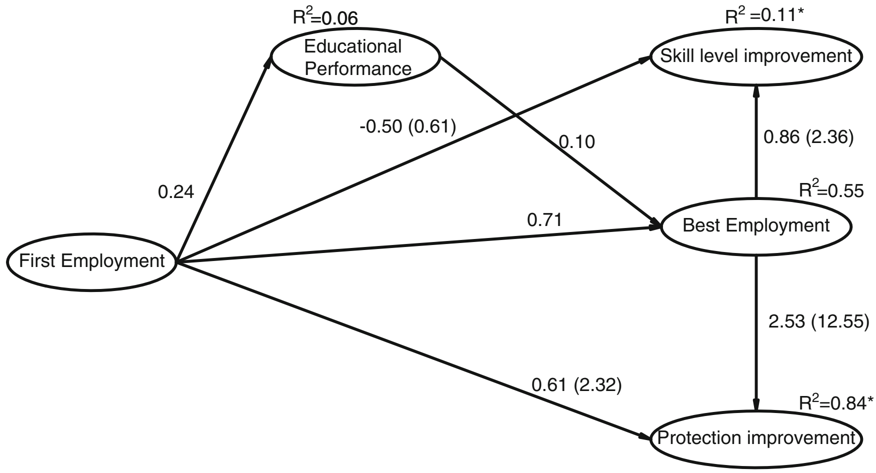


Fig. 12.6 Structural model for STUDENT WORKERS: values in parentheses indicate odds ratios; $*$ indicates pseudo-Nagelkerke R^2

protection changes by a factor of 12.55 (with an average marginal effect of 0.005). Confirming the findings for the STUDENTS, *Skill level improvement* negatively influences STUDENT WORKERS' *First employment*, but it is positively related to *Best employment*. Again, this is due to a suppression effect.

Figure 12.7 shows the ROC curves for the two logistic regressions. The values of the Area Under the Curve (AUC) are high in both cases: AUC equals 0.81 for the *Skill level improvement* logistic regression, and 0.99 for that of *Protection improvement*.

12.6 Conclusions

Analyzing real data from observational social science studies requires the use of soft modeling techniques, which are not based on rigid assumptions about the measurement scale and the distribution of the theoretical population from which the sample is supposed to be drawn. We exploited the flexibility of the PLS approach to model a predictive network of relationships between latent and manifest variables.

We enhanced the PLS analysis by introducing the Non-Metric approach to PLS-PM and integrating logistic regression in the structural model to manage ordinal manifest variables correctly and predict binary outcomes. These enhancements allowed us to measure the impact of the educational performance of Sapienza graduates on their career within 3 years of graduation. We analyzed a new database, UNI.CO, whose data combines the archives of the Sapienza University of Rome and the Italian Ministry of Labor and Social Policy. We proposed two different models

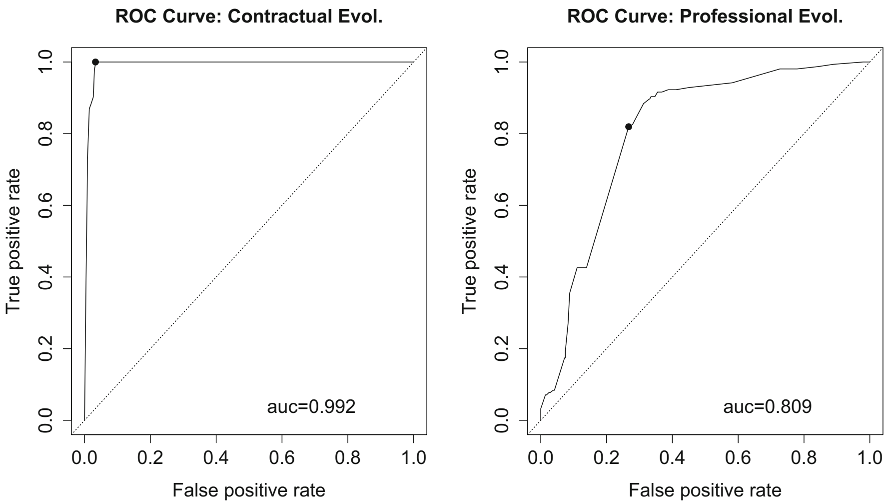


Fig. 12.7 Validation model for STUDENT WORKERS: the ROC curves. We report the value of the Area Under the Curve (AUC)

linking several latent concepts related to education and careers: one for STUDENTS and the other for STUDENT WORKERS.

Our analysis of UNI.CO dataset reveals three main insights:

- The academic performance of (non-worker) students strongly affects their age at graduation; this is typical in Italy, where students graduate when they pass all the required exams, no matter how long this takes. In such a context, students are allowed to take extra time to terminate their studies and only brilliant students tends to get their degree in due time.
- Academic performance and age at graduation only marginally affect careers, which mainly depends on alumni's performances at work, rather than at university.
- Graduates who obtain "good" first jobs tend to obtain better jobs more than those who obtain less advantageous first jobs, mostly in terms of contractual protection rather than of skill level.

We are aware that our analysis has at least two limits. First, our data do not represent a sample, but the whole Sapienza alumni population. The computational inference on model parameters aims to extend the conclusions of our model to all Italian State University graduates, under the hypothesis that Sapienza alumni provide a representative sample of Italian State University alumni. In our opinion this hypothesis is plausible, as Sapienza University is the biggest Italian University and attracts students from all over Italy. Second, we performed all our analyses on all Sapienza alumni whatever their field. Further studies are required to investigate whether different models hold for graduates in different fields.

As a final remark, in this application we proposed the use of NM-PLSPM to model composites. However, recent literature on consistent PLS-PM (Dijkstra and Henseler 2015) sketches new avenues for future research; in particular, we are currently investigating the use of NM-PLSPM to estimate parameters of factor models (Bollen 1989) involving categorical and/or ordinal MVs.

References

Alleva, G., & Petrarca, F. (2013). *New indicators for investigating the Integration of Sapienza graduates into the labor market*. Working papers n. 120/2013 del Dipartimento Memotef, ISSN 2239-608X.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*(6), 1173–1182.

Becker, J.-M., Rai, A., & Rigdon, E. E. (2013). Predictive validity and formative measurement in structural equation modeling: Embracing practical relevance. In *Proceedings of the International Conference on Information Systems (ICIS)*, Milan.

Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley.

Bollen, K. A., & Bauldry, S. (2011). Three Cs in measurement models: Causal indicators, composite indicators, and covariates. *Psychological Methods, 16*(3), 265–284.

Cantaluppi, G. (2012). A partial least squares algorithm handling ordinal variables also in presence of a small number of categories, arXiv:1212.5049.

Cantaluppi, G., & Boari, G. (2016). A partial least squares algorithm handling ordinal variables. In H. Abdi, V. E. Vinzi, G. Russolillo, G. Saporta & L. Trinchera (Eds.), The multiple facets of partial least squares methods, PLS, Paris, France 2014. In *Springer proceedings in mathematics and statistics*. (Vol. 173, pp 295–306). Cham: Springer.

Cohen, J., & Cohen, P. (1975). *Applied multiple regression/correlation analysis for the behavioral sciences*. New York: Wiley.

Dijkstra, T. K. (1981). *Latent variables in linear stochastic models: Reflections on "maximum likelihood" and "partial least squares" methods*. Ph.D. Thesis, Groningen University.

Dijkstra, T. K., & Henseler, J. (2015). Consistent and asymptotically normal PLS estimators for linear structural equations. *Computational Statistics and Data Analysis, 81*(1), 10–23. doi:10.1016/j.csda.2014.07.008.

Esposito, V. E., Trinchera, L., & Amato, S. (2010). PLS path modeling: From foundations to recent developments and open issues for model assessment and improvement. In V. E. Esposito, W. W. Chin, J. Henseler & H. Wang (Eds.), *Handbook of partial least squares: Concepts, methods and applications* (pp. 47–82). Berlin: Springer.

Gruppo UNI.CO. (2015). La Domanda di Lavoro per i laureati. I risultati dell'integrazione tra gli archivi amministrativi dell'Università Sapienza di Roma e del Ministero del Lavoro e delle Politiche Sociali, Edizioni Nuova Cultura- Roma. ISBN 9788868124816. doi:10.4458/4816.

Hanafi, M. (2007). PLS path modeling: Computation of latent variables with the estimation mode B. *Computational Statistics, 22*, 275–292.

Jakobowicz, E., & Derquenne, C. (2007). A modified PLS path modeling algorithm handling reflective categorical variables and a new model building strategy. *Computational Statistics and Data Analysis, 51*, 3666–3678.

Krämer, N. (2007). *Analysis of high-dimensional data with partial least squares and boosting*. Ph.D. Thesis, Technische Universität Berlin.

Krus, D. J., & Wilkinson, S. M. (1986). Demonstration of properties of a suppressor variable. *Behavior Research Methods, Instruments, & Computers, 18*(1), 21–24. doi:10.3758/BF03200988.

Kruskal, J. (1964). Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika, 29*(1), 1–27. doi:10.1007/BF02289565.

Little, T. D., Card, N. A., Bovaird, J. A., Preacher, K. J., & Crandall, C. S. (2007). Structural equation modeling of mediation and moderation with contextual factors. In T. D. Little, J. A. Bovaird & N. A. Card (Eds.), *Modeling contextual effects in longitudinal studies* (pp 207–230). Mahwah, NJ: Erlbaum.

Lohmöller, J. B. (1989). *Latent variables path modeling with partial least squares*. Berlin: Springer.

Maselli, I. (2012). The evolving supply and demand of skills in the labour market. *Intereconomics, 47*(1), 22–30. doi:10.1007/s10272-012-0402-2.

Noonan, R., & Wold, H. (1982). PLS path modeling with indirectly observed variables. In K. G. Jöreskog & H. Wold (Eds.), *Systems under indirect observation: Structure, prediction, causality* (Chap. 3, Vol. 2, pp. 75–94). Amsterdam: North-Holland.

Petrarca, F. (2014a). Non-metric PLS path modeling: Integration into the labour market of Sapienza graduates. In *Advances in latent variables - studies in theoretical and applied statistics* (pp. 159–170). Berlin: Springer. doi:10.1007/10104_2014_16.

Petrarca, F. (2014b). *Assessing Sapienza University alumni job careers: Enhanced partial least squares latent variable path models for the analysis of the UNI.CO administrative archive*. Ph.D. Thesis, Dipartimento di Economia dell'Università degli studi Roma Tre. http://hdl.handle.net/2307/4167.

Russolillo, G. (2012). Non-metric partial least squares. *Electronic Journal of Statistics, 6*, 1641–1669.

Russolillo, G. (2014). Extending the Non-metric PLS approach to inwards directed path models. In PLS'14 book of abstracts, *8th International Conference on Partial Least Squares and Related Methods*, Paris (pp. 159-160).

Sanchez, G., Trinchera, L., & Russolillo, G. (2013). R package of PLSPM. http://cran.r-project.org/web/packages/plspm/index.html.

Schuberth, F., Henseler, J., & Dijkstra, T. K. (2016). Partial least squares path modeling using ordinal categorical indicators. *Quality and Quantity, 1–27*. doi:10.1007/s11135-016-0401-7.

Stevens, S. S. (1946). On the theory of scales of measurement. *Science, 103*(2684), 677–680. New Series.

Tenenhaus, A., & Tenenhaus, M. (2011). Regularized generalized canonical correlation analysis. *Psychometrika, 76*(2), 257–284. doi:10.1007/S11336-011-9206-8.

Wold, H. (1985). Partial least squares. In S. Kotz & N. L. Johnson (Eds.), *Encyclopedia of statistical sciences* (Vol. 6, pp 581–591). New York: Wiley.

Chapter 13
Model Misspecifications and Bootstrap Parameter Recovery in PLS-SEM and CBSEM-Based Exploratory Modeling

Pratyush N. Sharma, Ryan T. Pohlig, and Kevin H. Kim

Abstract Theories are uncertain and evolving in exploratory research. This uncertainty can manifest itself in SEM studies either at the measurement or structural level, or both, and result in model misspecifications. Researchers often favor the use of PLS-SEM over CBSEM in exploratory research due to its tractability, flexibility, and its ability to avoid factor indeterminacy. While these strengths of PLS-SEM are undoubtedly appealing, empirical support regarding the robustness of model parameters under conditions of model misspecifications is lacking. This Monte Carlo study evaluates the efficiency and accuracy of bootstrap parameter recovery by PLS-SEM, CBSEM, and the Bollen-Stine methods under various conditions of measurement and structural misspecification effect sizes, sample sizes, and data distributions. Results point to the favorability of PLS-SEM in exploratory modeling when structural parameters are of interest, while CBSEM and Bollen-Stine methods are appealing when the focus is at the measurement level. A two-pronged strategy is advisable when theoretical uncertainty exists both at the measurement and structural levels.

Kevin H. Kim was deceased at the time of publication.

P.N. Sharma (✉)
Alfred Lerner College of Business and Economics, University of Delaware, Newark, DE, USA
e-mail: pnsharma@udel.edu

R.T. Pohlig
Biostatistics Core Facility, College of Health Sciences, University of Delaware, Newark, DE, USA
e-mail: rpohlig@udel.edu

K.H. Kim
School of Education, University of Pittsburgh, Pittsburgh, PA, USA

H. Latan, R. Noonan (eds.), *Partial Least Squares Path Modeling*,
DOI 10.1007/978-3-319-64069-3_13

13.1 Introduction

Structural equation modeling (SEM) is a widely used technique in the social and behavioral sciences to model complex relationships among observed and latent variables. Two of the most popular methods for model estimation are the covariance-based SEM (CBSEM) and partial least squares-based SEM (PLS-SEM). Despite the similarity in their function, the two methods differ in their respective goals. CBSEM attempts to estimate parameters that minimize the difference between the observed and model-based covariance matrix. PLS-SEM attempts to maximize the variance explained in the endogenous variables, rather than the true accuracy of the parameter estimates. The most common estimation method in CBSEM is the maximum likelihood (ML),[1] which assumes multivariate normality and large sample theory. PLS-SEM is a variance-based method that uses partial regression equations that minimize the residual variance in the dependent variables (Roldán and Sánchez-Franco 2012). PLS-SEM based models do not assume multivariate normality and are considered distribution-free; instead, the working assumption is that the sample distribution represents the population distribution (Lohmöller and Wold 1982; Hair et al. 2011). To help facilitate these methods, bootstrapping is often used to gather accurate estimates or to provide tests of inference.

Unlike the traditional parametric approaches, bootstrapping is a nonparametric approach to statistical inference that does not make any distributional assumptions for the parameters. Bootstrapping draws conclusions about the characteristics of a population strictly from the sample at hand, rather than making unrealistic assumptions about the population. That is, given the absence of information about a population, the sample is assumed to be the best estimate of the population. Hence, bootstrapping has advantages in situations where there is weak or no statistical theory about the distribution of a parameter or when the underlying distributional assumptions needed for valid parametric inference are violated (Mooney 1996). Bootstrapping allows for the possibility of conducting significance testing of a statistic such as a regression coefficient or factor loading, without distributional assumptions, by creating an empirical sampling distribution from the observed data. Such significance tests help analyze the probability of observing a statistic of that size or larger when the null hypothesis is true, i.e., H_0: $\theta = 0$. While allowing for a distribution free test, researchers have cautioned against a blind faith in bootstrapping, as it assumes that the sample approximates of the intended population (Yung and Bentler 1994).

Bootstrapping is widely applied in both PLS-SEM and ML-SEM. While PLS-SEM relies completely on bootstrapping to obtain standard errors for hypothesis testing, ML-SEM uses bootstrapping when distributional assumptions are violated. Significance testing is based on finding a test statistic that is considered extreme or highly unlikely given that the null hypothesis is true; therefore, it is important to

[1]We use the terms CBSEM and ML-SEM interchangeably there. Further, we use the terms PLS-SEM and PLS interchangeably whenever appropriate.

ensure that the test statistic comes from a distribution where the null hypothesis is true. Bollen and Stine (1992) showed that when using the Naïve bootstrap in ML, the bootstrap samples are taken from a population in which the null hypothesis is not true, leading to researchers incorrectly rejecting H_0 too often. It should be noted that this error is more likely for misspecified models, which are likely to arise in exploratory research. As a remedy, Bollen and Stine proposed a transformation of the data that makes the null hypothesis true, which forces H_0 to be true, resulting in fewer type-I errors (Bollen and Stine 1992).

In addition to its advantages over ML-SEM (Hair et al. 2011), a frequently stated assertion for favoring PLS-SEM over ML-SEM is its ability to provide robust results for exploratory modeling (Ringle et al. 2012) due to its flexibility (Noonan and Wold 1988) and the avoidance of inadmissible solutions and factor indeterminacy (Fornell and Bookstein 1982). Exploratory modeling is a process that involves guesswork and relatively loosely formulated hunches to refine existing theory that is currently weak and under development (Stebbins 2001). Through this process of exploration, and by focusing on the explanatory power of the models (i.e., R^2), researchers seek to arrive at a more accurate picture of the phenomena and in turn strengthen theory. However, performing exploratory analyses often leads to researchers creating misspecified or inaccurate models. Since PLS-SEM relies largely on the Naïve bootstrap, its use may lead to an increase in type-I errors. Yet, to the best of the authors' knowledge, this aspect of the oft-cited assertion, i.e., that PLS-SEM is preferred for exploratory modeling, has not been empirically tested in the PLS-SEM literature.

We suggest that researchers should have a better understanding of the bootstrap behavior, especially under conditions of model misspecifications and small sample sizes—conditions that are likely to arise in exploratory research. The goal of this Monte Carlo study is to evaluate the efficiency and accuracy of parameter recovery by PLS, ML, and the Bollen-Stine bootstrapping methods under various conditions of measurement and structural model misspecification effect sizes, sample sizes, and data distributions. In doing so, we extend the work by Sharma and Kim (2013), to provide researchers with guidelines when choosing to use either ML-SEM or PLS-SEM for exploratory research.

13.2 The Naïve Bootstrap

The bootstrap is a method for evaluating parameter estimates (Stine 1989) and a data-based approach for producing inferences (Efron and Tibshirani 1993). It is most commonly used to estimate quantities associated with a sampling distribution of an estimate or test statistic (Boos 2003). Conceptually, bootstrapping treats the observed data as a pseudo-population and repeatedly random samples with replacement from it. A random sample is defined as a selection of units of size *a*, selected at random with each unit's probability of being chosen equal to $1/a$. Sampling with replacement returns a unit to the population after being chosen, making the unit eligible to be selected again. The population from which the bootstrap samples (B) are taken come from the resample space (R), which is the original observed distribution (Efron and Tibshirani 1993). When the bootstrap sample size is equal to the observed sample size, the distribution of the test statistic from the bootstrap samples forms an empirical sampling distribution ($\widehat{F}_{\mathrm{ED}}$) (Rodgers 1999).

This empirical sampling distribution is an approximation of the statistic's true population distribution. When bootstrap samples are of the same size as the observed sample n, then the standard deviation of $\widehat{F}_{\mathrm{ED}}$ is an estimate of the population parameter's standard error. Bootstrapping has two main assumptions. The first, for accurate inferences to be made from the analysis, the original observed sample must be representative of the population of interest. Second, the relationship between a population and its sample can be modeled by the relationship between R and B (Efron and Tibshirani 1993; Yung and Bentler 1996). Bickel and Friedman (1981) conclude that the success of all bootstrap methods depend upon the assumption that the sampling behavior of a statistic is the same when it is taken from the empirical distribution and when it is taken from the original population.

13.2.1 Model-Based Bootstrap (Bollen-Stine)

As previously stated, the Naïve bootstrap in SEM causes the bootstrap samples to be taken from a population in which the null hypothesis is not true (Bollen and Stine 1992). Efron and Tibshirani (1993) and Yung and Bentler (1996) describe the steps for Naïve bootstrapping in SEM, after performing an analysis on the raw data. The steps are the same as outlined above, except the covariance matrix from each bootstrapped sample is used to perform the analyses. Then, the test statistic from the original data can be compared to the distribution from the bootstrapped test statistic values for significance. Bollen and Stine (1992) showed that the expected mean value and the variance of the observed test statistic and the bootstrapped test statistics are not equal. The expected values for the bootstrap estimates are larger than the ML methods causing the null hypothesis test to fail when using Naïve bootstrapping, regardless of whether it is true in the population or not. This violates the assumption of bootstrapping—the empirical distribution created by resampling is different from the population distribution and can lead to an increase in type-I errors.

To remedy this, Bollen and Stine (1992) proposed a semi-parametric bootstrap method as a solution. The sample covariance matrix is transformed to have a covariance structure specified by the null hypothesis, and the bootstrap samples are taken from the transformed data. Applying the transformation enables the observed sample test statistic to be compared to an empirical sampling distribution where the null hypothesis is true. After this transformation, the empirical sampling distribution behaves appropriately, meeting the bootstrap assumption. Then, similar to Naïve bootstrapping, the empirical sampling distribution can be used as a reference distribution for obtaining a test of significance (Finney and Distefano 2006). Nevitt and Hancock (2001) examined the performance of model-based bootstrapping for both normal and non-normal data, as well as correctly specified and misspecified models for a range of sample sizes. ML estimation performed well for correctly specified models when the assumption of normality was met. Regardless of the sample size, ML performed poorly for non-normal data. The Bollen and Stine model-based bootstrapping performed well regardless of normality and sample size conditions but was conservative with respect to type-I error and had overall less power (Nevitt and Hancock 2001). Similar results were found by Fouladi (2000).

13.3 Method

A Monte Carlo study was conducted to ascertain the impact of measurement and structural model misspecifications on the parameter recovery accuracy of three methods: Naïve and Bollen-Stine bootstrap in SEM and the Naïve PLS bootstrap. The latent variable model used in this study had two exogenous (ξ_1 and ξ_2) and two endogenous variables (η_1 and η_2) with three reflective indicators each. The factor loadings (lambdas) for the measurement model were set to 0.6, and the path loadings (betas) for the structural model were set at 0.3. For the structural model misspecification case (Fig. 13.1), the path between ξ_1 and η_2 (γ_{21}; highlighted in red) was misspecified, and its effect size was manipulated; $\gamma_{21} = \{0.0, 0.1, 0.2, 0.3, 0.5\}$. A cross-loading item was introduced in the model setup to simulate

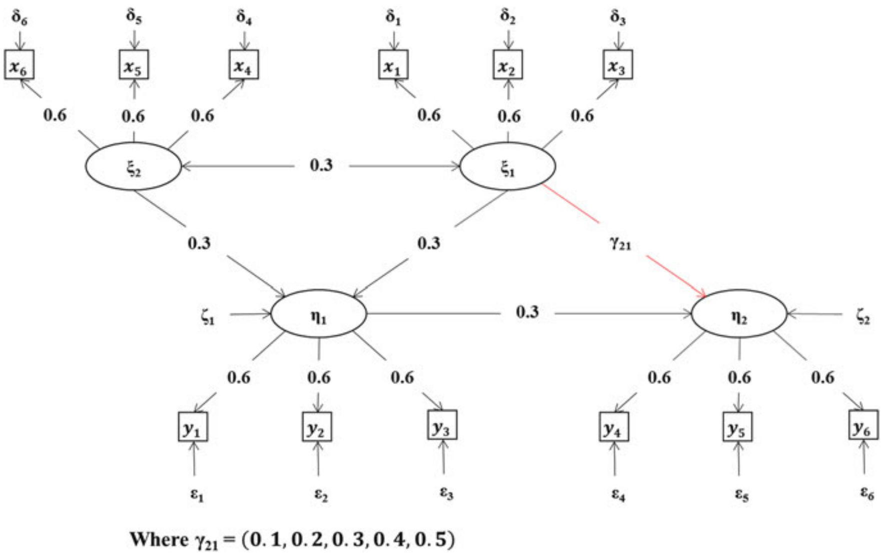


Fig. 13.1 Structural model misspecification

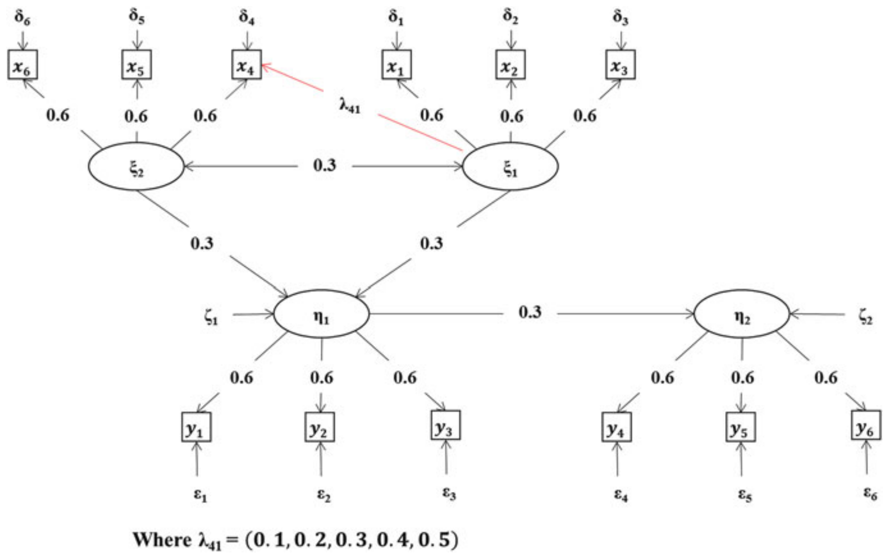


Fig. 13.2 Measurement model misspecification

measurement model misspecification (Fig. 13.2, highlighted in red); $\lambda_{41} = \{0.0, 0.1, 0.2, 0.3, 0.5\}$.

Data were generated using Fleishman and Vale-Maurelli's method (Fleishman 1978; Vale and Maurelli 1983) for five sample sizes (50, 100, 150, 200, and 500) and four distributions [$N(0, 1)$, $\chi^2_{\text{df}=3}$, $t_{\text{df}=5}$, and λ $U(0, 1)$]. The different distributions used cover a range of skewness and kurtosis. The $\chi^2_{\text{df}=3}$ has higher skew than the normal distribution and is more leptokurtic; the $t_{\text{df}=5}$ is more platykurtic and has more variability compared to a normal distribution; and the uniform distribution is highly platykurtic and with greater variability. One hundred replications were performed for each of the 100 conditions, with 250 bootstrap replications for each dataset. Two hundred and fifty was chosen as increasing the number of resamples taken would increase computing time for relatively small gain in efficiency (Nevitt and Hancock 2001). Standardized parameter estimates from PLS, ML, and ML Bollen-Stine bootstraps were compared. All simulations were run on the R computing environment (R Core Team 2014) using the *sem* (Fox 2006) and *semPLS* packages (Monecke and Leisch 2012). PLS parameters were estimated using path weighting scheme for inner weight computation and Mode-A (reflective) outer weight computation.

Both descriptive and inferential statistics were used to assess the methods' performances across all conditions for both types of misspecifications. A series of $3 \times 5 \times 4 \times 5$ mixed design ANOVAs were performed to examine the independent variable impact on estimation failure, bias, and root mean square difference (RMSD) for the parameter estimates. The first factor is the within subject factor of the bootstrapping method (Naïve, Bollen-Stine, and PLS). The between subject factors were sample size, data distribution, and misspecification size. Due to the fact that this is a simulation study with 100 replications per cell and 100 cells in the design each being tested with three methods, $n_{\text{total}} = 300{,}000$, examining p-values is not of interest since nearly every effect would be significant; therefore, only effects with at least a moderate effect size, $\eta_p^2 > .06$, were interpreted. A separate analysis was performed for the type-I error and power conditions, and the results were unchanged from the simultaneous analysis of all the conditions. Average bias is useful for assessing the direction of the errors (i.e., under or overestimation), whereas RMSD is an estimate of the error magnitudes which examines accuracy.

13.4 Results

13.4.1 Structural Misspecification Case

13.4.1.1 Convergence

Convergence rates of the bootstrap replications were tracked to test the stability of the methods being used. At sample sizes 150 or less, PLS bootstrap outperformed both the Bollen-Stine and Naïve methods in SEM by having less convergence failures. At sample sizes greater than 150, all methods performed comparably with negligible failures (Table 13.7 in Appendix).

13.4.1.2 Structural Parameters

Average bias and RMSD of the structural model regression coefficients, β and γ (elements of the B and Γ matrices), were tested. Both Bollen-Stine and the Naïve SEM bootstrap methods tended to overestimate the strength of the relationships among the coefficients, whereas PLS bootstrap underestimated the relationships (Table 13.1).

In order to assess the magnitude of error for the structural coefficients, the RMSD was examined. The pattern of differences for the methods differed among (1) the effect size averaged across the other factors, $F\,(8, 23760) = 301.64, p < .001,\ \eta_p^2 = .092$, and (2) the sample sizes averaged across the other factors, $F\,(10, 23760) = 202.61, p < .001,\ \eta_p^2 = .079$ (see Table 13.8 in the Appendix).

Table 13.1 Structural coefficient bias by effect size and method

	Effect size (γ_{21})				
	.00	.10	.20	.30	.50
Bollen-Stine	.003	.022	.047	.071	.137
Naïve	.001	.021	.045	.069	.134
PLS	−.076	−.070	−.062	−.055	−.039

Table 13.2 Structural coefficient RMSD by effect size and method

	Effect size (γ_{21})				
	.00	.10	.20	.30	.50
Bollen-Stine	.133	.138	.151	.172	.230
Naïve	.127	.135	.146	.163	.224
PLS	.106	.103	.099	.096	.095

Table 13.3 Structural coefficient RMSD by sample size and method

	Sample size					
	50	100	150	200	250	500
Bollen-Stine	.270	.190	.155	.138	.128	.107
Naïve	.241	.185	.155	.138	.129	.107
PLS	.133	.101	.094	.092	.091	.089

PLS had a smaller RMSD (i.e., less error) than both the Naïve and Bollen-Stine bootstrapping for all misspecification effect sizes (all $p < .05$). PLS also had a stable RMSD compared to the Naïve and Bollen-Stine methods, whose RMSDs increased as the size of the misspecification increased (Table 13.2).

Additionally, PLS had a smaller RMSD (i.e., less error) than both the Naïve and Bollen-Stine bootstrapping for all sample sizes, (all $p < .05$). While PLS had a smaller RMSD, as sample size increased, all the methods had similar results in the largest sample size condition (Table 13.3).

13.4.1.3 Measurement Parameters

Average bias and RMSD of the measurement model factor loadings, λ_x and λ_y (elements of the Λ_x and Λ_y matrices), were tested. Bollen-Stine and the Naïve bootstrap tended to underestimate the strength of the relationships among the coefficients, whereas PLS overestimated the relationships (see Table 13.9 in the Appendix).

The RMSD for the factor loadings differed (1) among methods $F(2, 23760) = 44402.83, p < .001,\ \eta_p^2 = .789$ averaged across the other factors, and (2) differed among sample sizes averaged across the other factors, $F(5, 11880) = 4093.97, p < .001,\ \eta_p^2 = .633$. For presentation purposes, we will discuss the results of the method by sample size interaction, although the effect just misses our interpretation criteria, $F(10, 23760) = 106.10, p < .001,\ \eta_p^2 = .043$ (see Table 13.10 in the Appendix). PLS had a larger RMSD (i.e., more error) than both the Naïve and Bollen-Stine bootstrap for all sample sizes, (all $p < .05$). The Naïve and Bollen-Stine bootstrap methods had smaller RMSDs across the sample sizes; as sample size increased, their RMSDs continued to shrink, while PLS did not change (Table 13.4).

Table 13.4 Factor loading RMSD by sample size and method

	Sample size					
	50	100	150	200	250	500
Bollen-Stine	.156	.108	.085	.072	.065	.045
Naïve	.142	.106	.085	.073	.065	.045
PLS	.246	.192	.175	.170	.165	.161

13.4.2 Measurement Misspecification Case

13.4.2.1 Convergence

Convergence rates of the bootstrap replications were tracked to show stability of the methods being used. Again, PLS outperformed both the Bollen-Stine and Naïve methods at sample sizes 150 or less and had less convergence failures. At sample sizes greater than 150, all methods performed comparably (Table 13.11 in Appendix).

13.4.2.2 Structural Parameters

Average bias and RMSD of the structural model regression coefficients, β and γ (elements of the B and Γ matrices), were tested. Bollen-Stine ($M = 0.001$, $SE = 0.001$) and the Naïve ($M = 0.001$, $SE = 0.001$) bootstrap methods tended to overestimate the strength of the relationships among the coefficients; on the other hand, PLS ($M = -0.077$, $SE = 0.001$) again underestimated the relationships.

In order to assess the magnitude of error for the structural coefficients, the RMSD was examined. The pattern of differences on the RMSD for the structural coefficients among the methods differed among the sample sizes averaged across the other factors, $F\,(10, 23760) = 249.89, p < .001,\ \ \eta_p^2 = .095$ (see Table 13.12 in the Appendix). A more nuanced relationship existed for the RMSD of the structural parameters in the presence of measurement misspecification. PLS had a smaller RMSD (i.e., less error) than both the Naïve and Bollen-Stine bootstrapping for samples sizes below 200, (all $p < .05$). There were no differences at sample size 200 ($p > .05$), but for larger sample sizes, the Naïve and Bollen-Stine bootstrapping exhibited smaller RMSDs (all $p < .05$) (Table 13.5).

13.4.2.3 Measurement Parameters

Average bias and RMSD of the measurement model factor loadings, λ_x and λ_y (elements of the Λ_x and Λ_y matrices), were tested. Again, Bollen-Stine and the Naïve bootstrapping tended to underestimate the strength of the relationships among the coefficients, whereas PLS overestimated the relationships (see Table 13.13 in the Appendix).

Table 13.5 Structural coefficient RMSD by sample size and method

	Sample size					
	50	100	150	200	250	500
Bollen-Stine	.252	.165	.126	.106	.093	.065
Naïve	.232	.162	.127	.105	.093	.065
PLS	.137	.107	.102	.101	.100	.100

Table 13.6 Factor loading RMSD by sample size and method

	Sample size					
	50	100	150	200	250	500
Bollen-Stine	.157	.110	.086	.073	.065	.045
Naïve	.143	.108	.087	.074	.066	.045
PLS	.258	.201	.181	.173	.167	.161

The RMSD for the factor loadings differed (1) among methods $F(2, 23760) = 33186.03, p < .001,\ \eta_p^2 = .736$ averaged across the other factors, and (2) differed among sample sizes averaged across the other factors, $F(5, 11880) = 3727.30, p < .001,\ \eta_p^2 = .611$. For presentation purposes, we will discuss the results of the method by sample size interaction, although the effect just misses our interpretation criteria, $F(10, 23760) = 56.01, p < .001,\ \eta_p^2 = .023$ (see Table 13.14 in the Appendix). PLS had a larger RMSD (i.e., more error) than both the Naïve and Bollen-Stine bootstrapping for all sample sizes (all $p < .05$). And as sample size increased, their RMSDs continued to get smaller (Table 13.6).

13.5 Discussion

Theories are uncertain and evolving in exploratory research (Stebbins 2001). This uncertainty can manifest itself in SEM studies either at the measurement or structural level, or both. Researchers often favor PLS-SEM over ML-SEM (CBSEM) for exploratory research based on its tractability, flexibility, and its ability to avoid factor indeterminacy (Ringle et al. 2012; Noonan and Wold 1988; Fornell and Bookstein 1982). While these strengths of PLS-SEM are undoubtedly appealing, empirical support regarding the robustness of model parameters under conditions of model misspecifications is lacking in the literature. The crux of this chapter is that exploratory research may often create conditions that give rise to misspecified models, both at the structural and measurement levels. Yet, to the authors' best knowledge, no existing studies have tested the accuracy and efficiency of parameter recovery by PLS bootstrap and compared them with ML under such conditions. Our goal was to fill this gap by testing the bootstrap parameter recoveries of PLS, ML, and Bollen-Stine methods under various conditions of model misspecification effect sizes, sample sizes, and distributional conditions. Building on the previous work by Sharma and Kim (2013), our results largely validate the researchers' assertions regarding the strength of PLS-SEM in exploratory modeling, while also pointing to the utility of CBSEM in certain conditions.

Specifically, under conditions where structural misspecification is suspected and the theory is evolving at the structural level, PLS bootstrap had better parameter accuracy and efficiency than ML and Bollen-Stine methods for structural parameters. On the other hand, under the same conditions of structural misspecifications, ML and Bollen-Stine had better measurement model recoveries. Under conditions where measurement misspecification is suspected, i.e., theory is uncertain at the construct level, PLS had less error at smaller sample sizes (250 or less) beyond which ML and Bollen-Stine methods performed better in recovering structural estimates. ML and Bollen-Stine also had better measurement model recoveries in this case. In terms of the experimental conditions, smaller sample sizes and larger misspecified effect sizes had a detrimental effect on the accuracy of parameter recovery. On the other hand, data distributions did not have any appreciable detrimental effect on any of the three methods.

With established tools and instruments, it might be argued that measurement misspecifications are less likely to occur. In this case, model exploration is based at the structural level, and PLS-SEM might be the preferred method. If the constructs are not well defined by the previous literature, and the focus is on construct reliability, measurement misspecifications might be more likely to occur, and ML-SEM is the preferred approach. Studies that consider the refinement of measurement models as their goal often validate the construct reliabilities using ML-SEM (Garver and Mentzer 1999). In comparison, PLS-SEM-based studies are often used in exploratory research with the goal of investigating structural links. Our results bear this practice out. We find that, in general, the Naïve PLS bootstrap was more accurate and efficient than ML and Bollen-Stine SEM bootstraps for estimating structural model parameters. However, the reverse was true for measurement model estimates. These results point to the favorability of PLS when the structural parameters are of interest. On the other hand, ML and Bollen-Stine methods are appealing when the focus of the study is at the measurement level. When both structural and measurement model misspecifications are suspected, our suggestion is that the researchers use a two-pronged strategy by first using ML or Bollen-Stine methods to clarify the measurement model and then continue with PLS to estimate structural paths.

Finally, we note that these results and recommendations are valid only under the specific context considered in this study—in exploratory research, where structural and measurement model misspecifications are likely and where parameter accuracy is of interest. This study used a common factor-based data model where PLS-SEM is known to be at a relative disadvantage as compared to CBSEM (Sarstedt et al. 2016). Further analyses based on a composite model may provide more insights regarding the strengths and weaknesses of PLS-SEM vis-à-vis CBSEM, including the recently proposed consistent version of PLS (Dijkstra and Henseler 2015). While we maintained a narrow focus on model misspecifications, there are other aspects of exploratory research (e.g., data considerations, such as its availability and appropriateness) that have not been considered. Also not considered were the differences between exploratory and confirmatory approaches, and there exists a vast literature on this topic. The onus remains on the researcher to correctly specify their study as exploratory or confirmatory. This decision depends to a large extent on the goal and the context of the study and the general state of theoretical and methodological developments in the field. In his influential essay, Tukey asserted that exploratory research is an "attitude" and that both confirmatory and exploratory approaches are necessary for theory development (Tukey 1980). PLS- and ML-based methods are complementary in this regard, and their judicious application can help in creating stronger theories.

Appendix

Table 13.7 Convergence rates by sample size and method

	Sample size					
	50	100	150	200	250	500
Bollen-Stine	90.80%	98.78%	99.84%	99.98%	100.00%	100.00%
Naïve	81.38%	93.83%	98.33%	99.56%	99.87%	100.00%
PLS	99.88%	99.97%	99.99%	100.00%	100.00%	100.00%

Table 13.8 Structural coefficients RMSD

Effect	Df	F	p	η_p^2
Method	2	3115.27	<.001	.208
Method × sample size	10	202.61	<.001	.079
Method × effect size	8	301.64	<.001	.092
Method × distribution	6	.950	.458	.000
Method × sample size × effect size	40	4.94	<.001	.008
Method × sample size × distribution	30	.51	.989	.001
Method × effect size × distribution	24	1.2	.228	.001
Method × sample size × effect size × distribution	120	.82	.927	.004
Sample size	5	867.85	<.001	.268
Effect size	4	385.13	<.001	.115
Distribution	3	.12	.951	.000
Sample size × effect size	20	6.82	<.001	.011
Sample size × distribution	15	.52	.930	.001
Effect size × distribution	12	1.14	.319	.001
Sample size × effect size × distribution	60	.74	.931	.004

Table 13.9 Measurement model (factor loading) average bias

Effect	Df	F	p	η_p^2
Method	2	122, 745.60	<.001	.912
Method × sample size	10	645.53	<.001	.214
Method × effect size	8	51.87	<.001	.107
Method × distribution	6	1.39	.213	.000
Method × sample size × effect size	40	3.20	<.001	.005
Method × sample size × distribution	30	1.57	.024	.002
Method × effect size × distribution	24	.90	.608	.001
Method × sample size × effect size × distribution	120	1.17	.104	.006
Sample size	5	318.35	<.001	.118
Effect size	4	4.33	.002	.001
Distribution	3	4.42	.004	.001
Sample size × effect size	20	.79	.734	.001
Sample size × distribution	15	1.86	.022	.002
Effect size × distribution	12	.74	.709	.001
Sample size × effect size × distribution	60	.86	.771	.004

Table 13.10 Measurement model (factor loading) RMSD

Effect	Df	F	p	η_p^2
Method	2	44, 402.83	<.001	.789
Method × sample size	10	106.10	<.001	.043
Method × effect size	8	10.51	<.001	.004
Method × distribution	6	.87	.514	.000
Method × sample size × effect size	40	1.74	.003	.003
Method × sample size × distribution	30	1.30	.125	.002
Method × effect size × distribution	24	1.00	.467	.001
Method × sample size × effect size × distribution	120	1.24	.038	.006
Sample size	5	4093.97	<.001	.633
Effect size	4	9.68	<.001	.003
Distribution	3	4.29	.005	.001
Sample size × effect size	20	2.13	.002	.004
Sample size × distribution	15	1.86	.023	.002
Effect size × distribution	12	1.13	.327	.001
Sample size × effect size × distribution	60	1.20	.140	.006

Table 13.11 Convergence rates by sample size and method

	Sample size					
	50	100	150	200	250	500
Bollen-Stine	90.65%	98.36%	99.68%	99.65%	99.72%	99.60%
Naïve	82.02%	93.92%	98.19%	99.41%	99.52%	99.66%
PLS	99.82%	99.95%	99.99%	100.00%	100.00%	100.00%

Table 13.12 Structural coefficient RMSD

Effect	Df	F	p	η_p^2
Method	2	368.45	<.001	.030
Method × sample size	10	249.89	<.001	.095
Method × effect size	8	.51	.849	.000
Method × distribution	6	1.54	.160	.000
Method × sample size × effect size	40	.48	.998	.001
Method × sample size × distribution	30	1.10	.319	.001
Method × effect size × distribution	24	1.23	.200	.001
Method × sample size × effect size × distribution	120	1.44	.001	.007
Sample size	5	921.63	<.001	.280
Effect size	4	1.18	.320	.000
Distribution	3	.50	.682	.000
Sample size × effect size	20	.77	.734	.001
Sample size × distribution	15	1.66	.052	.002
Effect size × distribution	12	1.58	.089	.002
Sample size × effect size × distribution	60	.87	.759	.004

Table 13.13 Measurement model (factor loading) average bias

Effect	Df	F	p	η_p^2
Method	2	86, 873.07	<.001	.880
Method × sample size	10	666.01	<.001	.219
Method × effect size	8	1.05	.394	.000
Method × distribution	6	1.00	.423	.000
Method × sample size × effect size	40	1.22	.163	.002
Method × sample size × distribution	30	1.37	.084	.002
Method × effect size × distribution	24	.94	.549	.001
Method × sample size × effect size × distribution	120	1.35	.006	.007
Sample size	5	353.86	<.001	.130
Effect size	4	1.17	.323	.000
Distribution	3	4.20	.006	.001
Sample size × effect size	20	.69	.838	.001
Sample size × distribution	15	2.01	.012	.003
Effect size × distribution	12	.77	.687	.001
Sample size × effect size × distribution	60	.90	.704	.004

Table 13.14 Measurement model (factor loading) RMSD

Effect	Df	F	p	η_p^2
Method	2	33, 186.03	<.001	.736
Method × sample size	10	56.01	<.001	.023
Method × effect size	8	.89	.526	.000
Method × distribution	6	.84	.541	.000
Method × sample size × effect size	40	1.05	.385	.002
Method × sample size × distribution	30	1.08	.347	.001
Method × effect size × distribution	24	.99	.473	.001
Method × sample size × effect size × distribution	120	1.29	.018	.006
Sample size	5	3727.30	<.001	.611
Effect size	4	1.18	.318	.000
Distribution	3	2.71	.043	.001
Sample size × effect size	20	1.10	.339	.002
Sample size × distribution	15	1.73	.039	.002
Effect size × distribution	12	.78	.676	.001
Sample size × effect size × distribution	60	1.39	.025	.007

References

Bickel, P. J., & Freedman, D. A. (1981). Some asymptotic theory for the bootstrap. *The Annals of Statistics, 9*, 1196–1217.

Bollen, K. A., & Stine, R. A. (1992). Bootstrapping goodness of fit measures in structural equation modeling. *Sociological Methods & Research, 2*, 205–229. doi:10.1177/0049124192021002004.

Boos, D. D. (2003). Introduction to the bootstrap world. *Statistical Science, 18*, 168–174.

Dijkstra, T., & Henseler, J. (2015). Consistent partial least squares path modeling. *MIS Quarterly, 39*(2), 297–316.

Efron, B., & Tibshirani, R. (1993). *An introduction to the bootstrap*. New York: Chapman Hall.

Finney, S. J., & Distefano, C. (2006). Non-normal and categorical data in structural equation modeling. In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: A second course* (pp. 269–314). Greenwich, CT: Information Age.

Fleishman, A. I. (1978). A method for simulating non-normal distributions. *Psychometrika, 43*(4), 521–532.

Fornell, C., & Bookstein, F. (1982). Two structural equation models: LISREL and PLS applied to consumer exit-voice theory. *Journal of Marketing Research, 19*, 440–452.

Fouladi, R. T. (2000). Performance of modified test statistics in covariance and correlation structure analysis under conditions of multivariate nonnormality. *Structural Equation Modeling, 7*, 356–410.

Fox, J. (2006). Structural equation modeling with the sem package in R. *Structural Equation Modeling, 13*(3), 465–486.

Garver, M. S., & Mentzer, J. T. (1999). Logistics research methods: Employing structural equation modeling to test for construct validity. *Journal of Business Logistics, 20*(1), 33–57.

Hair, J., Ringle, C., & Sarstedt, M. (2011). PLS-SEM: Indeed a silver bullet. *Journal of Marketing Theory and Practice, 19*(2), 139–152.

Lohmöller, J. B., & Wold, H. (1982). *Introduction to PLS estimation of path models with latent variables, including some recent developments on mixed scales variables*. Forschungsberichte. 82.02. Hochschule der Bundeswehr, München.

Monecke, A., & Leisch, F. (2012). semPLS: Structural equation modeling using partial least squares. *Journal of Statistical Software, 48*(3), 1–32.
Mooney, C. (1996). Bootstrap statistical inference: Examples and evaluations for political science. *American Journal of Political Science, 40*(2), 570–602. doi:10.2307/2111639.
Nevitt, J., & Hancock, G. R. (2001). Performance of bootstrapping approaches to model test statistics and parameter standard error estimation in structural equation modeling. *Structural Equation Modeling, 8*(3), 353–377.
Noonan, R., & Wold, H. (1988). Partial least squares path analysis. In J. P. Keeves (Ed.) *Educational research, methodology and measurement: An international handbook*. Oxford: Pergamon.
R Core Team. (2014). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.
Ringle, C., Sarstedt, M., & Straub, D. (2012). Editor's comments: A critical look at the use of PLS-SEM in MIS quarterly. *MIS Quarterly, 36*(1), iii–xiv.
Rodgers, J. L. (1999). The bootstrap, the jackknife, and the randomization test: A sampling taxonomy. *Multivariate Behavioral Research, 34*(4), 441–456.
Roldán J. L., & Sánchez-Franco M. J. (2012). Variance-based structural equation modeling: guidelines for using partial least squares. In M. Mora et al. (Eds.), *Research methodologies, innovations and philosophies in software systems engineering and information systems*. IGI Global. doi:10.4018/978-1-4666-0179-6
Sarstedt, M., Hair, J. F., Ringle, C. M., Thiele, K. O., & Gudergan, S. P. (2016). Estimation issues with PLS and CBSEM: Where the bias lies! *Journal of Business Research, 69*(10), 3998–4010. doi:10.1016/j.jbusres.2016.06.007.
Sharma, P. N., & Kim, K. H. (2013). A comparison of PLS and ML bootstrapping techniques in SEM. In H. Abdi et al. (Eds.), *New perspectives in partial least squares and related methods, Springer series in proceedings in mathematics and statistics* (Vol. 56). New York: Springer. doi:10.1007/978-1-4614-8283-3_13.
Stebbins, R. A. (2001). *Exploratory research in the social sciences* (Vol. 48). Thousand Oaks, CA: Sage.
Stine, R. (1989). An introduction to bootstrap methods: Examples and ideas. *Sociological Methods & Research, 18*(2–3), 243–291. doi:10.1177/0049124189018002003.
Tukey, J. (1980). We need both exploratory and confirmatory. *The American Statistician, 34*(1), 23–25.
Vale, C. D., & Maurelli, V. A. (1983). Simulating multivariate non-normal distributions. *Psychometrika, 48*, 465–471.
Yung, Y. F., & Bentler, P. M. (1994). Bootstrap corrected ADF test statistics in covariance structure analysis. *The British Journal of Mathematical and Statistical Psychology, 47*(1), 63–84.
Yung, Y. F., & Bentler, P. M. (1996). Bootstrapping techniques in analysis of mean and covariance structures. In G. A. Marcoulides & R. E. Schumacker (Eds.), *Advanced structural equation modeling issues and techniques* (pp. 195–226). Mahwah, NJ: Erlbaum.

Part III
Applications

Chapter 14
Personality, Intellectual Ability, and the Self-Concept of Gifted Children: An Application of PLS-SEM

R. Frank Falk

Abstract The latent variable path analysis program LVPLS was based on Herman Wold's nonlinear iterative partial least squares (NIPALS) approach to theory construction and data analysis. Current developments derived from NIPALS have formed partial least squares structural equation modeling (PLS-SEM). Both serve as appropriate techniques for data analysis under varying conditions. The study described in this chapter uses PLS-SEM to explore the predictive relationships among personality, intellectual ability, and self-concept in a sample of gifted youth. In the model, intellectual ability and introversion accounted for 24% of the variance in self-concept. Calculations and presentation of results are courtesy of Christian M. Ringle and the SmartPLS 3 computer program (http://www.smartpls.com).

Herman Wold's nonlinear iterative partial least squares (NIPALS) approach to theory construction and data analysis, first introduced in the late 1960s, later become known as soft modeling (Wold 1980) and latent variable path analysis (LVPLS). My introduction to Dr. Wold and his work occurred in the 1980s at the University of Denver. Dr. John Horn, professor in psychology, brought Dr. Wold to the University for a number of presentations. Wold's soft modeling represented a new way of bringing together measurement and theory. This approach had a profound effect on the way I began to conceptualize data analysis.

Several years later as a visiting scholar at the University of Virginia, I was able to work with J. Jack McArdle, who introduced me to Jan-Bernd Lohmöller. Lohmöller's program latent variable path analysis with partial least squares estimation (LVPLS 1989) provided the means to calculate PLS. His program and my interest in soft modeling eventually led to the publication of a book titled *A Primer*

Special thanks to Nancy B. Miller, Linda K. Silverman, the editor Hengky Latan, and the two anonymous reviewers for their critical comments on the manuscript.

R. Frank Falk, Ph.D. (✉)
Institute for the Study of Advanced Development, Westminster, CO, USA
e-mail: rfalk@uakron.edu

H. Latan, R. Noonan (eds.), *Partial Least Squares Path Modeling*,
DOI 10.1007/978-3-319-64069-3_14

for Soft Modeling in the early 1990s (Falk and Miller 1992). Thus, my work falls under the heading of what Henseler et al. (2016) refer to as the “traditional” method as opposed to the “modern” method. The modern method is the result of work by many in the international business disciplines, beginning in the late 1990s and early 2000s and continuing to the present (e.g., Chin 1998; Haenlein and Kaplan 2004; Hair et al. 2017; Tenenhaus et al. 2005).

Conceptually, part of the difference between the traditional and modern method is explained by the terms latent variable path modeling (LVPLS), also known as PLS regression, and structural equation modeling (PLS-SEM). In LVPLS, the emphasis is on prediction of the relationships between latent variables in the structural model. In PLS-SEM, the emphasis is on explanation. The fundamental concern in the manifest variable model in LVPLS is between mode A and mode B, known as reflective and formative modes. The original assumption was that the manifest variables were reliable and represented either a single latent variable or separate aspects of a latent variable. Today in PLS-SEM there is much greater emphasis on the measurement model in which higher reliability standards are sought and measures of convergent and discriminate validity are employed. Finally, in PLS-SEM a statistical technique called bootstrapping is used to test the probability that the reflective loadings and composite weights are non-zero. Additionally, the predictive power or relevance of the model can be examined, in certain cases, using the blindfolding procedure. Thus, PLS-SEM represents a causality model, based on highly reliable and valid instrumentation and well-established theoretical relationships.

In LVPLS, soft modeling refers to the notion that measurement and theory are at a formative stage. The essential idea is to gain some degree of predictability as a way to move measurement and theory forward. While achieving high predictability is the goal in science, having some degree of predictability is better than having none. In other words, although “10% predictability means that only 10% [of the variance] is understood . . . this may be a valuable piece of information” (Falk and Miller 1992, p. 5).

In principle, this study belongs more in the traditional than the modern method. The theoretical relationships are speculative and the data relatively soft. In the model to be examined, only one of the manifest variables meets normality and linearity assumptions. However, well-established measures are used for the measurement of all latent variables, although some probably do not meet the interval level of measurement. My interest in using PLS-SEM techniques in this study is to explore the procedures of the modern era with available measures and provisional theory.

14.1 The Theoretical Model

Parents and clinicians tend to see a positive view of the self as an indicator of healthy child development. For this reason, self-concept is measured in all children assessed at the Gifted Development Center. In cases where self-appraisal is low, "the discovery of giftedness is often ameliorative, kindling aspirations and healing self-concepts" (Silverman 2013, p. 200).

According to William James (1915), one's self is unique in that the person is both the knower and the known. Charles Cooley (1902) referred to the self as the reflective self and called it the reflected appraisal of others or "the looking-glass self." Research indicates that gifted children may be even more sensitive to the feedback of others (Falk and Miller 1998), especially adolescent girls (Gross et al. 2007).

Self-concept, or the way we see ourselves, is based on the way we think others see and evaluate us. However, not all others have the same impact on our self-concept; some are more significant than others (Mead 1934).

The goal of this study is to understand the factors that contribute to the self-concept of the gifted child. Parents usually exert the greatest influence on the child's view of the self. Two measures of parents' observations of the child's personality are introversion-extraversion and overexcitability (OE). These factors are believed to be related to self-concept (Silverman 2013).

Introversion/Extroversion In the early 1930s, Carl Jung (1971) developed the idea of introversion and extraversion as a continuum of personality type. In Jung's view introverts conserve energy, while extraverts expend energy. He asserted that individuals are born with an introvert or extravert personality type.

Moving from the clinical perspective of personality, researchers in the area of temperament sought to establish the biological features of personality. This culminated in the work of Buss and Plomin (1984) on temperament. For them, extraversion and introversion were understood as sociability and shyness. Sociability was measured by the degree of preference for being with others; shyness was understood as a person's behavior in the presence of casual others and strangers. Extraverts are outgoing, prefer being with others, and are not self-conscious in the presence of strangers. Introverts prefer doing things alone and do not initiate social behaviors. In the Big Five personality model, the extraversion scale is on a continuum from solitary/reserved to outgoing/energetic (Gallagher 2013).

In discussing extraversion and introversion in the gifted, Silverman (2013) emphasizes Jung's conception.

> Introverts are oriented inward toward the subjective world of thoughts and concepts; they get their energy from inside themselves; and they are inclined toward reflection. Extraverts are oriented outward, become energized through interaction with people and things, and are directed toward action. Whereas introverts feel drained by too much interaction with people, extraverts are energized by interaction—the more, the merrier. (p. 151)

Overexcitability OE is an innate characteristic that describes a person's reaction to external stimuli. It represents a heightened sensitivity and intensity of the central nervous system. There are five areas of OE: psychomotor (high energy levels), sensual (sensory enjoyment), imaginational (rich fantasy), intellectual (high curiosity), and emotional (strong feelings). The OEs were introduced in 1937 by Kazimierz Dabrowski, a Polish psychiatrist and theorist, based on clinical observation (Piechowski 2014a).

OEs have been explored extensively in the field of gifted education as characteristics of gifted children, adolescents, and adults (see, e.g., Falk and Miller 2009; Mendaglio 2008; Piechowski 2014b; Silverman 2013). In cross-cultural research of gifted children, Kuo et al. (2012) found each OE was correlated with a specific area of the brain, supporting the notion of its inherent quality.

Four of the OEs are associated with facets of the openness to experience factor in the Big Five model of personality—sensual, imaginational, intellectual, and emotional. These OEs have been shown to correlate with esthetics, fantasy, ideas, and feelings, respectively (Gallagher 2013). A dimension of the perception of self, verbal self-concept, was found to be positively correlated with sensual, imaginational, intellectual, and emotional OE (Gross et al. 2007). Two studies have found relationships between intellectual OE and self-concept (Gross et al. 2007; Rinn et al. 2010).

Intellectual ability is a term that captures many cognitive capabilities, such as reasoning, thinking, and remembering. Commons (1985) suggested that the possible number of intellectual abilities could be as high as 800,000 or more. The question of how many separate factors are required to account for the multitude of individual abilities varies from 1 to 9, according to Horn and Noll (1994). Verbal comprehension and perceptional reasoning, two components of abstract reasoning, are used in this study because they are the best indicators of intellectual giftedness.

14.2 Hypotheses

The general proposition of this study is that intellectual ability is associated with self-concept, introversion/extraversion, and OE. Intellectual ability influences the way others respond to the child and, in turn, his or her self-concept. The latent variable hypotheses are as follows: (1) intellectual ability has a direct positive effect on self-concept, OE, and introversion/extraversion; (2) introversion-extraversion has a direct positive effect on self-concept; and (3) OE positively affects self-concept.

14.3 Method

14.3.1 Subjects

Data used in the study was collected at the Gifted Development Center, a testing and counseling center in Westminster, Colorado, USA. Subjects were children tested at the Center between 1994 and 2014. Only those cases where parents gave permission for their children to participate in research and those cases having no missing data on the variables in the study were included. These criteria were met by 242 cases with children between the ages of 9 and 12. The gender distribution was as follows: 64% male and 36% female.

14.3.2 Manifest Variables

In the preliminary analysis, 17 manifest variables were included in the model that formed 4 latent variables—intellectual ability, introversion/extraversion, OE, and self-concept. The manifest variables included two measures of intellectual ability (Verbal Comprehension and Perceptual Reasoning), five measures of OE (psychomotor, sensual, imaginational, intellectual, and emotional), four indicants of introversion and extroversion (mother's and father's scores), and six dimensions of self-concept. Bootstrap tests that show no loadings, weights, and paths were statistically significant. Thus a different analytic approach was pursued that included all 17 manifest variables.

The second tactic was to create two hierarchical latent variables by combining first-order latent variables (Becker et al. 2012). One hierarchical latent variable "overexcitability" was created by combining one latent variable defined by psychomotor and sensual OE and another made up of imaginational, intellectual, and emotional OE. A second hierarchical latent variable was created for introversion-extraversion from two first-level latent variables—fathers' scores for child's introversion and extraversion and mothers' scores for their child's introversion and extraversion. Unfortunately, the bootstrap procedure produced nonsignificant results. It was also not possible to predict the variance in the hierarchical variables (OE and introversion/extraversion) with intellectual ability since the total variance was already accounted for by first-order latent variables. Using these findings, the following 11 manifest variables were included in a reduced manifest variable model.

Self-Concept The Self-Perception Profile for Children—ages 8–12, "What I Am Like"—was used as a measure of the child's self-concept (Harter 1982). It was designed for children in what Harter has described as middle childhood. It is in this period that children have evaluations of self in different areas of their life experiences. This instrument assesses the following areas: scholastic competence, social acceptance, athletic competence, physical appear, behavioral conduct, and global self-worth. Global self-worth, a sense of overall self-esteem, is a separate dimension and not a summary of the other five. All six dimensions are independent factors both theoretically and empirically. There are 36 items on the Self-Perception Inventory, a self-report questionnaire. The inventory, which was validated using factor analytic techniques, has been widely used to assess children's self-concepts (Harter 1982).

Introversion-Extraversion In 2011, Silverman revised her original introversion-extraversion scale (Silverman 2002) to include 12 additional items. The revised scale is used in this study. While the scale has been useful for clinical assessment, its reliability and other psychometric properties have not been established.

The introversion-extraversion scale contains items representing sociability and shyness, as well as cognitive and psychological traits. Parents are asked to check the column indicating how close their child is to one of the descriptors in each pair, e.g., "Can focus on many ideas at once" versus "Likes to concentrate on one activity at a time." Introversion and extroversion are two ends of a continuum. Only mothers' scores were used. (See Appendix for other examples of the items used in measuring introversion and extraversion.)

Overexcitability The Overexcitability Inventory for Parents (OIP-II) is a 25-item parent report questionnaire that measures the parents' perception, or appraisal, of their child's five independent overexcitability factors. Its level of measurement is ordered-metric. The psychometric properties in the general population are unknown. Reliability for OEs in this instrument has been reported as high (ranging from 0.77 to 0.89) (Falk and Silverman 2016). Intellectual OE (mean = 4.08) has been found to correlate with giftedness in many studies (e.g., Bouchet and Falk 2001; Wirthwein and Rost 2011; Van den Broeck et al. 2014); it is the only measure of OE included in the model. (See Appendix for examples of the items used in measuring intellectual OE.)

Intellectual Ability To assess intellectual ability, subtests from the Wechsler Intelligence Scale for Children, Fourth Edition (WISC-IV), were used. Six unique abilities form two of the four Composites—Verbal Comprehension and Perceptual Reasoning. These two Composites create measures of general intelligence that are less influenced by a child's performance on auditory memory tasks (Working Memory) and speed (Processing Speed).

Greater discrepancies among WISC-IV Composite scores have been found for gifted children than for the non-gifted (Rimm, cited in Silverman 2013). In the WISC-IV manual, Composite scores for the gifted in a normative sample are similar and highest for Verbal Comprehension and Perceptual Reasoning (124.7 and 120.4, respectively) and lowest for Working Memory and Processing Speed (112.5 and 110.6, respectively) (Wechsler 2003). Therefore, the child's Verbal Comprehension and Perceptual Reasoning Composites were selected to assess intellectual ability in this study of gifted children.

Verbal Comprehension is composed of three subtests: Vocabulary, Similarities, and Comprehension. Information or Word Reasoning may, at appropriate times, be substituted for any one of these. Perceptual Reasoning includes Block Design, Matrix Reasoning, and Picture Concepts. Picture Completion is an acceptable alternative for one of these in certain cases (Flanagan and Kaufman 2004). Verbal Comprehension Composite scores in this study ranged from 121 to 182, with a mean of 141 and a standard deviation of 12. Scores higher than 160 were calculated with extended norms (Zhu et al. 2008). This demonstrates that the sample represents a gifted population.

14.4 Analysis

Data were entered into the SmartPLS 3 program (Ringle et al. 2015). The output from this program consists of a graphic presentation of the model as well as a complete set of calculations. *A primer on partial least square structural equation modeling* is also available for assistance in using and interpreting the program (Hair et al. 2017).

14.5 Results

The initial model had 11 manifest variables and 4 latent variables. The latent variables introversion and self-concept had inner directed arrows from the manifest variables to the latent construct, sometimes referred to as mode B or formative measurements. Thus, the values of the arrows represented composite weights used to create the latent variables. Intellectual ability is a first principal component factor. The values of the manifest variables represented first latent variable loadings; their residuals were quite small. The test for collinearity, called the variance inflation factor (VIF) for self-concept, was 1.05 and indicated no significant collinearity. Convergent validity using redundancy analysis was 0.70, reaching the minimum standard.

Verbal Comprehension and Perceptual Reasoning Composites from the WISC-IV both had high loadings on the latent variable intellectual ability (Verbal Comprehension = 0.86 and Perceptual Reasoning = 0.80). The measurement of these Composites has well-established reliability and validity (Wechsler 2003). The composite weights on introversion were approximately equal, with extraversion being negative (−0.49) and introversion being positive (0.54). The label of this latent variable was now "introversion" as its positive loading defined the variable. The composite reliability for intellectual ability was 0.8; AVE was 0.68, indicating reliable and convergent validity.

The weight of parents' report of the child's intellectual OE on the latent variable (intellectual OE) was 1.00. This measure was a single manifest variable; therefore the latent variable and the manifest variable were equivalent. Earlier attempts to include other OE measures produced insignificant weights.

The composite weights on self-concept were as follows: athletic competence, 0.63; behavioral conduct, 0.36; global self-worth, 0.25; physical appearance, 0.14; scholastic competence, 0.60; and social competence, 0.31. By the traditional LVPLS method, all variables in the measurement model would be retained; but in the modern PLS-SEM method, bootstrap statistical tests are applied to all reflective loadings and composite weights.

All of the loadings on intellectual ability had a probability of 0.05 and were therefore significant. The weight for extroversion on the latent variable introversion was not significant, and the remaining manifest variable is mother's score on

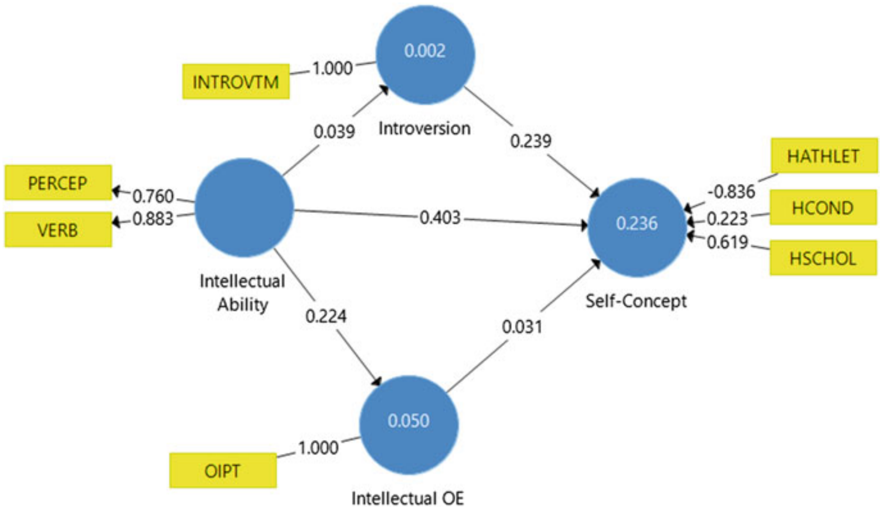


Fig. 14.1 PLS model

introversion. Three manifest variables on self-concept had acceptable levels of significance: athletic competence 0.00, behavioral conduct 0.00, and social competence 0.00. Recalculating the entire model (See Fig. 14.1) with the newly specified measurements shows all reflective loadings and composite weights with a significant probability of 0.05 or less.

The discriminant validity for the remaining manifest variables in the model was measured using the heterotrait-monotrait ratio (HTMT). The results show a ratio of 0.30 between intellectual ability and intellectual OE and 0.06 between intellectual ability and introversion. The ratio for intellectual OE and introversion was 0.14. Thus, discriminant validity is established.

The standardized root mean square residual (SRMR) was used to assess the model fit. A value of below 0.08 is considered a good fit in covariance-based structural equation modeling but may be too restrictive for PLS (Hair et al. 2017). For this model the SRMR = 0.07, well below the more restrictive 0.08 criteria, indicating a good fit of the model for the data.

Paths were positive between latent variables intellectual ability and intellectual OE (0.22) and between intellectual ability and self-concept (0.40). The path between intellectual ability and introversion was 0.04. The path from introversion to self-concept was 0.24, while the path between intellectual OE and self-concept was small (0.03), indicating that intellectual OE had almost no direct influence on a child's self-concept in this model.

Twenty-four percent of the variance ($R^2 = 0.24$, $p = 0.00$) in the child's self-concept was accounted for by intellectual ability and introversion, with intellectual ability having the largest contribution. The indirect effect for intellectual ability was 0.016. For intellectual OE, variance explained was small ($R^2 = 0.05$, $p = 0.06$).

Therefore little variance in intellectual OE was accounted for by intellectual ability since no other predictors were present. This is a strong indication that intellectual OE is different from intellectual ability. The variance explained for introversion was very small ($R^2 = 0.002$, $p = 0.51$) Statistical significance was established by the bootstrap procedure in SmartPLS 3.

14.6 Discussion

The strongest predictor of self-concept in the model was the child's intellectual ability. Findings indicate that the higher a child's general intelligence, the more positive his or her self-concept. This should be good news to parents of gifted children who may worry about their children being out of sync with age peers and how this might affect their self-appraisal. However, the discovery of giftedness can often lead to feelings of self-worth. Likewise, the opportunity to associate with other gifted youth, either in school or in special programs where true peers are discovered, is likely to increase positive views of oneself (Silverman 2013).

Introversion was also a strong predictor of self-concept, indicating that gifted introverts have positive self-views. This suggests that children who need time alone, learn by observing, like to concentrate on one activity at a time, prefer one close friend, and are more cautious and reflective have higher self-concepts. Therefore, children who exhibit introversion should be valued and supported at home and in the school. Their self-reflection should be appreciated and the benefits of their depth of thought and preference for listening recognized (Helgoe 2013; Olwen Laney 2002).

No relationship was found between intellectual OE and self-concept in this study. Two previous studies found that psychomotor OE related to academic self-concept; however, both of those studies used different measures of self-concept and assessed older children (Gross et al. 2007; Rinn et al. 2010). Further research is needed to examine the effect of other OEs on the self-concept of children.

One limitation of this study is the lack of children's self-report for variables such as introversion-extraversion and overexcitability. Although parents' perceptions are frequently the best measures of the child, and certainly reveal the way the child is seen by significant others, information about the child's perception may be important.

It is clear from the results of this study that more predictive variables are needed to account for the variance in the self-concept of gifted children. For example, the child's perception of his or her own overexcitability could be assessed using a new instrument called the Overexcitability Questionnaire-Two-C for children ages 6–11 (Falk et al. 2016). This instrument was designed for use with children too young to complete the Overexcitability Questionnaire-II, a self-report measure for those 12 years old and older.

Having explored the modern PLS-SEM, there are some measurement features that I can appreciate. These include a reliability measure for reflective measurement models, called average variance extracted, and a variance inflation factor for examining the collinearity in formative measurement models. Additionally, the bootstrapping procedure for assessing the statistical significance of loadings, weights, and paths can be useful as long as appropriate significance values are chosen. These must be consistent with the psychometric properties of the manifest variables; values of 0.05 or smaller may not be appropriate.

NIPLS and LVPLS were intended to provide researchers with a tool to explore measurement and latent variable relationships when the data and the theoretical relationships are soft. The push for greater reliability of variables and stronger theory, while commendable, should not exclude meaningful attempts to better understand the relationship between concepts with less than ideal specifications.

Appendix

Instruments used in this study have copyright privileges that apply to distribution of test items as follows. Publishers of the Wechsler Intelligence Scale for Children forbid that any test items be distributed. The Self-Perception Profile and the Overexcitability Instrument for Parents (OIP-II) allow example items to be shown (see below).

Responses to the Self-Perception Profile are along a four-category continuum of "Very True," "True," "Sometimes," and "Neutral." There are a total of 25 items measuring introversion and extroversion in facing pairs. Two examples of these items are "Needs time alone" versus "Needs social interaction" and "Reflective" versus "Impulsive."

The Overexcitability Inventory for Parents-II (OIP-II) has a total of 25 items. Five items measure each of the five OE's: psychomotor, sensual, imaginational, intellectual, and emotional. There are six response categories ranging from "Not at all like my child" to "Very much like my child." Two examples of intellectual OE are "My child observes and analyzes everything" and "My child loves to solve problems and develop new concepts."

References

Becker, J.-M., Klein, K., & Wetzel, M. (2012). Hierarchical latent variable models in PLS-SEM: Guidelines for using reflective-formative type models. *Long Range Planning, 45*, 359–394. doi:10.1016/j.lrp.2012.10.001.

Bouchet, N. M., & Falk, R. F. (2001). The relationship among giftedness, gender, and overexcitability. *The Gifted Child Quarterly, 45*, 260–267.

Buss, A. H., & Plomin, R. (1984). *Temperament: Early developing personality traits*. Hillsdale, NJ: Lawrence Erlbaum.

Chin, W. W. (1998). The partial least squares approach to structural equation modeling. In G. A. Marcoulides (Ed.), *Modern methods for business research* (pp. 295–358). London: Erlbaum.

Commons, M. (1985). *How novelty produced continuity in cognitive development within a domain and accounts for unequal development across domains*. Toronto: Society for Research of Child Development.

Cooley, C. H. (1902). *Human nature and the social order*. New York: Scribner.

Falk, R. F., & Miller, N. B. (1992). *A primer for soft modeling*. Akron, OH: University of Akron Press. (Now available on ResearchGate).

Falk, R. F., & Miller, N. B. (1998). The reflective self: A sociological perspective. *Roeper Review, 20*, 150–152.

Falk, R. F., & Miller, N. B. (2009). Building firm foundations: Research and assessments. In S. Daniels & M. M. Piechowski (Eds.), *Living with intensity: Understanding the sensitivity, excitability, and emotional development in gifted children, adolescents, and adults* (pp. 239–259). Scottsdale, AZ: Great Potential Press.

Falk, R. F., & Silverman, L. K. (2016). *The overexcitability inventory for parents (OIP-II): 28-item, six-factor version manual, scoring system, and questionnaire*. Westminster, CO: Institute for the Study of Advanced Development.

Falk, R. F., Daniels, S., & Silverman, L. K. (2016). *Overexcitability questionnaire-two for children: Manual, scoring system, and questionnaire*. Westminster, CO: Institute for the Study of Advanced Development.

Flanagan, D. F., & Kaufman, A. S. (2004). *Essentials of WISC-IV assessment*. Hoboken, NJ: Wiley.

Gallagher, S. A. (2013). Building bridges: Research on gifted children's personalities from three psychological theories. In C. S. Neville, M. M. Piechowski, & S. S. Tonal (Eds.), *Off the charts: Asynchrony and the gifted child* (pp. 48–98). Unionville, NY: Royal Fireworks Press.

Gross, C. M., Rinn, A. N., & Jamieson, K. M. (2007). The relationship between gifted adolescents' overexcitabilities and self-concepts: An analysis of gender and grade level. *Roeper Review, 29*, 240–248.

Haenlein, M., & Kaplan, A. M. (2004). A beginner's guide to partial least squares analysis. *Understanding Statistics, 3*(4), 283–297. doi:10.1207/s15328031us0304_4.

Hair, J. F., Hult, G. T., Ringle, C. M., & Sarstedt, M. (2017). *A primer on partial least squares structural equation modeling (PLS-SEM)* (2nd ed.). Thousand Oaks, CA: Sage.

Harter, S. (1982). The perceived competence scale for children. *Child Development, 53*, 87–97. doi:10.2307/1129640.

Henseler, J., Hubona, G. S., & Ray, P. A. (2016). Using PLS path modeling in technology research: Updated guidelines. *Industrial Management & Data Systems, 116*, 1–19.

Helgoe, L. (2013). *Introvert power: Why your inner life is your hidden strength*. Naperville, IL: Sourcebooks.

Horn, J., & Noll, J. (1994). A system for understanding cognitive capabilities: A theory and the evidence on which it is based. In D. K. Detterman (Ed.), *Current topics in human intellectual ability. Theories of intellectual ability* (Vol. 4, pp. 151–203). Norwood, NJ: Ablex.

James, W. (1915). *Psychology*. New York: Holt, Rinehart and Winston.

Jung, C. G. (1971). *Personality types. The collected works of G. G. Jung, Bollingen Series XX, Vol. 6*. Princeton, NJ: Princeton University Press.

Kuo, C. C., Chang, H. J., Chow, K. H., Chang, Y. P., Lin, Y. H., & Lin, C. P. (2012). Linking psychological traits with physical brain structure: A MRI study of talented and average ability high school students. *Bulletin of Educational Psychology, 43*(3), 805–832.

Lohmoeller, J. (1989). *Latent variable path modeling with partial least squares*. New York: Springer.

Mead, G. H. (1934). *Mind self and society from the standpoint of a social behaviorist*. Chicago: University of Chicago Press.

Mendaglio, S. (Ed.). (2008). *Dabrowski's theory of positive disintegration*. Scottsdale, AZ: Great Potential Press.

Olwen Laney, M. (2002). *The introvert advantage: How to thrive in an extrovert world*. New York: Workman.

Piechowski, M. M. (2014a). The roots of Dabrowski'stheory. *Advanced Development, 14*, 28–41.

Piechowski, M. M. (2014b). *"Mellow out," they say. If I only could: Intensities and sensitivities of the young and bright*. Unionville, NY: Royal Fireworks Press.

Ringle, C. M., Wende, S., & Becker, J. (2015). *SmartPLS 3*. Bönningstedt: SmartPLS GmbH. Retrieved from http://www.smartpls.com.

Rinn, A. N., Mendagilio, S., Rudasill, K. M., & McQueen, K. S. (2010). Examining the relationship between the overexcitabilities and self-concepts of gifted adolescents via multivariate cluster analysis. *The Gifted Child Quarterly, 54*(1), 3–7.

Silverman, L. K. (2002). *Upside-down brilliance: The visual-spatial learner*. Denver, CO: Deleon.

Silverman, L. K. (2011). *Introversion/extraversion scale*. Westminster, CO: Gifted Development Center. Retrieved from www.gifteddevelopment.com.

Silverman, L. K. (2013). *Giftedness 101*. New York: Springer.

Tenenhaus, M., EspositoVinzi, V., Chatelin, Y. M., & Lauro, C. (2005). PLS path modeling. *Computational Statistics and Data Analysis, 48*(1), 159–2005. doi:10.1016/j.csda.2004.03.005.

Van den Broeck, W., Hofmans, J., Cooremans, S., & Staels, E. (2014). Factorial validity and measurement invariance across intelligence levels and gender of the overexcitabilities questionnaire-II (OEQ-II). *Psychological Assessment, 26*(1), 55–68. doi:10.1037/a0034475.

Wechsler, D. (2003). *The WISC-IV technical and interpretive manual*. San Antonio, TX: Pearson Assessments.

Wirthwein, L., & Rost, D. H. (2011). Focusing on overexcitabilities: Studies with intellectually gifted and academically talented adults. *Personality and Individual Differences, 51*, 337–342. doi:10.1016/j.paid.2011.03.041.

Wold, H. (1980). Soft modeling: Intermediate between traditional model building and data analysis. *Mathematical Statistics, 6*, 333–346.

Zhu, J., Cayton, T., Weiss, L., & Gabel, A. (2008). *Wechsler intelligence scale for children* (4th ed.). Technical Report #7. Upper Saddle River, NJ: Pearson Education.

Chapter 15
Ethical Awareness, Ethical Judgment, and Whistleblowing: A Moderated Mediation Analysis

Hengky Latan, Charbel Jose Chiappetta Jabbour, and Ana Beatriz Lopes de Sousa Jabbour

Abstract This study aims to examine the ethical decision-making (EDM) model proposed by Schwartz (J Bus Ethics, 2015. doi:10.1007/s10551-015-2886-8), where we consider the factors of nonrationality and aspects that affect ethical judgments of auditors to make the decision to blow the whistle. In this chapter, we argue that the intention of whistleblowing depends on ethical awareness (EAW) and ethical judgment (EJW) as well as there is a mediation-moderation due to emotion (EMT) and perceived moral intensity (PMI) of auditors. Data was collected using an online survey with 162 external auditors who worked in audit firms in Indonesia as well as 173 internal auditors working in the manufacturing and financial services. The result of multigroup analysis shows that emotion (EMT) can mediate the relationship between EAW and EJW. The nature of this relationship is more complex, so we then tested it by adding moderating variables using consistent partial least squares (PLSc) approach. We found that EMT and PMI can improve the relationship between ethical judgments and whistleblowing intentions. These findings indicate

H. Latan (✉)
Department of Accounting, STIE Bank BPD Jateng, Jl. Pemuda No 4 A, Semarang, 50139, Indonesia

Department of Accounting, Petra Christian University, Jl. Siwalankerto No 121-131 Surabaya 60228, Indonesia
e-mail: latanhengky@gmail.com

C.J.C. Jabbour
Montpellier Business School, Montpellier Research in Management, 2300, avenue des Moulins, 34185 Montpellier, Cédex4, France

Centre for Advanced Management Education (CAME), Stirling Management School, University of Stirling, Stirling, FK9 4LA, Scotland, UK
e-mail: c.chiappetta-jabbour@montpellier-bs.com; c.j.chiappettajabbour@stir.ac.uk

A.B.Lopes de Sousa Jabbour
Montpellier Business School, Montpellier Research in Management, 2300, avenue des Moulins, 34185 Montpellier, Cédex4, France

DMEM Department, University of Strathclyde, 75 Montrose St, Glasgow, G1 1XJ, UK
e-mail: a.sousa-jabbour@montpellier-bs.com; ana.jabbour@strath.ac.uk

H. Latan, R. Noonan (eds.), *Partial Least Squares Path Modeling*,
DOI 10.1007/978-3-319-64069-3_15

that internal auditors are more likely to blow the whistle than external auditors, and reporting wrongdoing internally and anonymously is the preferred way of professional accountants to blow the whistle in Indonesia.

15.1 Introduction

Whistleblowing has gained the attention of the global community and the media in recent years, partly because of large awards offered by the Dodd-Frank Act of 2010 and partly due to a case of fraud involving Olympus corporation and Michael Woodford who was fired when he revealed payment irregularities (Archambeault and Webber 2015; Rao et al. 2011; MacGregor and Stuebs 2014). This indicates that a whistleblower does not only arise from inside the organization, but it can also come from outside, referred to as an external whistleblower (Maroun and Atkins 2014b; Maroun and Gowar 2013).

An internal whistleblower can observe the various violations that occur within an organization such as discrimination, corruption, cronyism, or other unethical behavior. Meanwhile, an external whistleblower can observe noncompliance with the fulfillment of corporate social responsibility and the environment (Culiberg and Mihelic 2016; Vandekerckhove and Lewis 2012). Thus, the important role of whistleblowers in detecting wrongdoing cannot be denied (Latan et al. 2016). However, being a whistleblower is not easy, because one must consider the positive and negative impacts caused, and it also involves the complicated process of ethical decision-making (EDM) (Ponemon 1994; Shawver et al. 2015; Webber and Archambeault 2015; O'Sullivan and Ngau 2014). EDM can be understood as deciding or judging whether the action or decision is ethical (Lehnert et al. 2015). Given that the internal control system is designed to minimize risks such as financial fraud, it will rely heavily on moral reasoning which is conducted by auditors (both internal and external). However, an auditor is often faced with ethical issues that pit ethics and professional codes against ethical decisions.[1]

The critical reviews conducted by Culiberg and Mihelic (2016) and Vandekerckhove and Lewis (2012) showed that there is still an empirical gap in this area that requires further testing. For example, most previous studies have focused too much on internal whistleblowers (such as employees, managers, internal auditors, and management accountants) and ignored external whistleblowers (Latan et al. 2016; Alleyne et al. 2016; Miceli et al. 2014).[2] In this context, subjects such as how to protect external whistleblowers (Maroun and Gowar 2013) and how they are perceived need to be further addressed (Maroun and Atkins 2014a, b). At the same time, the body of literature currently offers little insight into how a person reacts to wrongdoings to make a decision to blow the whistle. This relates to the ethical decision-making (EDM) model proposed by Rest (1986), where there are

[1] Jubb (2000) gives a detailed explanation of the roles of auditors as whistleblowers.

[2] Miceli et al. (2014) provide a detailed distinction between internal and external whistleblowers. In this chapter, we use the term external whistleblower compared to the "bell-ringers" proposed by them to be more familiar.

four stages that must be passed, namely, awareness, judgment, intent, and actual behavior.

As stated by Culiberg and Mihelic (2016), most of the existing empirical research related to whistleblowing has examined the relationship between judgment and intent (Zhang et al. 2009; Chiu 2002, 2003; Liyanarachchi and Newdick 2009) and supports it fully. However, there are no studies that extend this testing to other stages, such as considering the influence of ethical awareness on ethical judgment. In addition, some studies also show that there are other factors that can affect this process, such as moral intensity (Jones 1991) and emotion (Henik 2008, 2015; Hollings 2013). Schwartz (2015) showed an EDM model of integration, combining the factors of rationality and nonrationality. This model assumes that ethical behavior depends on people who face ethical biases (related to mood or moral intensity) and the environmental situation at the time. Jones (1991) defines moral intensity as a measure of moral imperatives-related problems in certain situations.

Perceived moral intensity will help auditors when facing an ethical dilemma, while emotions are feelings that arise (such as anger or fear) when encountering wrongdoing. These influence the auditor's decision to blow the whistle (Jones 1991; Henik 2008, 2015; Latan et al. 2016). Both of these factors play an important role and are a key element in the EDM model of whistleblowing. Therefore, the purpose of this study was to extend the EDM model of testing for whistleblowing by considering the role of two whistleblower groups (internal and external) in the Indonesian context.

Indonesia provides a good setting to test this model because it offers an interesting phenomenon to study. For example, according to a report from global fraud study conducted by the Association of Certified Fraud Examiners (ACFE) in 2016, Southeast Asia was in first position for cases of fraud, and Indonesia is one of five countries in the world experiencing higher levels of fraud after South Africa, India, Nigeria, and China. This is an indication that auditors in Indonesia (internal and external) may be still reluctant to become whistleblowers (Latan et al. 2016). As stated by Jubb (2000), internal or external auditors are often faced with an ethical dilemma when wanting to reveal wrongdoings in the workplace: they have conflicts of loyalty and professionalism. Hence, the decision to blow the whistle is complicated. However, research in Southeast Asia and Indonesia is rare, and there is still an empirical gap (Culiberg and Mihelic 2016; Latan et al. 2016). Thus, it is important to examine what factors are instrumental to the auditor's decision to blow the whistle.

Our study contributes to the current literature in several ways. First, this is the first study to extend the testing EDM model to whistleblowing, where there are many factors and relationships between variables that have not been tested in previous research.[3] Thus, this study answers the call of Culiberg and Mihelic (2016) to extend

[3]This study provides empirical evidence of EDM theoretical models developed by Schwartz (2015). Although not all of the variables considered, this provides sufficient preliminary evidence.

the testing of these models in the context of accounting and ethics. Although some previous studies have discussed this model (Zhang et al. 2009; Chiu 2003; Arnold et al. 2013; Yu 2015), they can be developed further. Second, this is the first study to compare two groups of whistleblowers—internal and external auditors—which is helpful in explaining which group is more prone to blowing the whistle. Until now, no previous empirical studies have fully considered testing the two whistleblower groups together in a single model. Although Shawver et al. (2015) used professional accountants as samples (including internal and external auditors) in testing the EDM model for whistleblowing, they did not test the samples separately.[4]

Third, this study extends state-of-the-art research on whistleblowing by providing evidence from Indonesia. To the best of our knowledge, no study conducted in Indonesia has tested EDM models of decisions to blow the whistle. As there are no empirical results available from Indonesia on whistleblowing in the context of accounting, this study provides initial evidence of the importance of individual and nonrationality factors in favor of EDM model proposed by Schwartz (2015) which have been the focus of research lately. Finally, it is important to conduct this study with experienced professionals such as auditors, who experience real-life ethical dilemmas that maybe different from those outside professional organizations (e.g., employees, consultants, customers, shareholders). However, few studies use the auditor as a sample (Curtis and Taylor 2009; Latan et al. 2016; Culiberg and Mihelic 2016; Alleyne et al. 2013).

The remainder of the chapter is organized as follows. The next section presents the development of the hypotheses, followed by the research methodology. Next, we discuss our results. Finally, we further analyze our results and provide important implications of our study as well as its limitations.

15.2 Literature Review and Hypothesis Development

15.2.1 The Ethical Decision-Making Model

EDM is one of the issues that have attracted the attention of researchers in the field of business ethics but also in other disciplines such as marketing, moral psychology, organizational behavior, philosophy, and social economics. The extent of illegal and unethical behavior that occurs in organizations and society in general has motivated researchers to develop an EDM model on an ongoing basis. The main assumption among all bodies of knowledge in the literature on EDM is a rationality-based process. One of the most widely cited and tested EDM models was proposed by Rest (1986) which consists of four components, namely, awareness, judgment, intent, and actual behavior. Until now, there have been several theoretical models of EDM that have been proposed, including a model of the contingency by Ferrell and Gresham

[4]Shawver et al. (2015) combine the two groups into a single dataset. This makes the results of the analysis become inaccurate and biased.

(1985), a situational interactionist model by Trevino (1986), the general theory of ethics and its modifications (Hunt and Vitell 1986, 2006), modified Rest model by Jones (1991), and the integrated EDM model by Schwartz (2015). The main purpose of building these models is to explain and predict the process by which a person makes ethical decisions and the factors underlying such decisions.

Ferrell and Gresham (1985) adopted a framework for contingency aiming to explain the processes of EDM that influence ethical decisions of marketers. In this model, they propose three contingency factors: individual factors (e.g., knowledge, values, attitudes, and intentions), organizational factors (e.g., organizational pressures and opportunities), and environmental factors (e.g., company policies and interactions between groups) that directly affect the ethical decisions of individuals. Trevino (1986) developed a situational interactionist model by combining individual factors (such as moral development) with the situational factors to explain and predict the EDM of individuals within an organization. More specifically, the model shows that the relationship between the individual's cognitive moral development and ethical behavior will be moderated by the two factors. Individual factors include the strength of the ego, field dependence, and locus of control, whereas situational factors include the immediate context of work, organizational culture, and nature of work.[5] In addition, Trevino (1986) also adopted the six stages of cognitive moral development developed by Kohlberg which becomes operative in the EDM process. Hunt and Vitell (1986) proposed a general theory of ethics that is more comprehensive in explaining the process of EDM and widely accepted in the field of marketing. According to their theory, once a person is faced with an ethical dilemma, where there are alternatives and consequences (influenced by cultural, environmental, professional, organizational, industrial, and personal characteristics), they will make an evaluation (both deontological and teleological), before making ethical judgments. After that, the ethical judgment will directly affect the ethical intentions which in turn affect the actual behavior (Hunt and Vitell 2006). The Hunt-Vitell model also added feedback generated from the actual consequences of people behavior to make personal experience in the future.

Unlike the previous three competing models, Jones (1991) built an EDM model considering Rest (1986)'s model. According to Jones (1991), the literature does not have a model which shows the characteristics of a moral problem itself which affects the EDM process, and he proposes an issues-contingent model of EDM. This combines the concept of moral intensity and organizational factors in the Rest's model, which is a new paradigm in EDM models. In addition, it considers that individuals who have a superior position in the organization, as a routine, more often faced ethical issues in decision-making and vice versa. Thus, the stronger the intensity of the ethical issues, the more likely the decision-makers are to lean toward ethical behavior. Therefore, Jones (1991) hypothesizes that moral intensity and

[5] There is a similarity between the model of Trevino (1986) with the model of Ferrell and Gresham (1985), which consider individual and situational factors. The difference is the role of both, one as a predictor and the other as moderator.

organizational factors play a role as predictor variables that directly and separately contribute to the EDM process.

Most recently, Schwartz (2015) conducted a synthesis of all existing EDM models and previous studies and proposed a new model called the "integrated EDM model." This combines all theoretical and empirical models into a single comprehensive model. This present study adopts the perspective of the framework proposed by Schwartz (2015), where we consider factors of nonrationality (such as emotions) as well as individual factors (such as moral intensity) as the mediation-moderation effects in the relationship between the variables that affect the decision-making process of an auditor to blow the whistle (see Fig. 15.1). As stated by Schwartz (2015), EDM is a complex process that involves many variables that are interrelated (neurocognitive-affective processes) and influence each other. For example, in the EDM model described earlier, nonrationality factors were not fully discussed, and for this reason the rationalist approach seems to have limitations and shortcomings, especially in conditions that are unpredictable and dynamic. We chose the nonrationality factors to be tested because they are more dominant in the process of moral judgment, in which rationality plays a secondary role after "a fact" is clear. In other words, when someone finds wrongdoing, but it is outside of the organization's ethical code of conduct, for example, the nonrationality factor will dominate the EDM process. Conversely, when the wrongdoing is common and has been agreed upon, then the rationality factor will dominate. If the rights and duties of the auditor as a whistleblower have not been set out clearly in the law on protection, then the nonrationality of factors tends to be more important in the

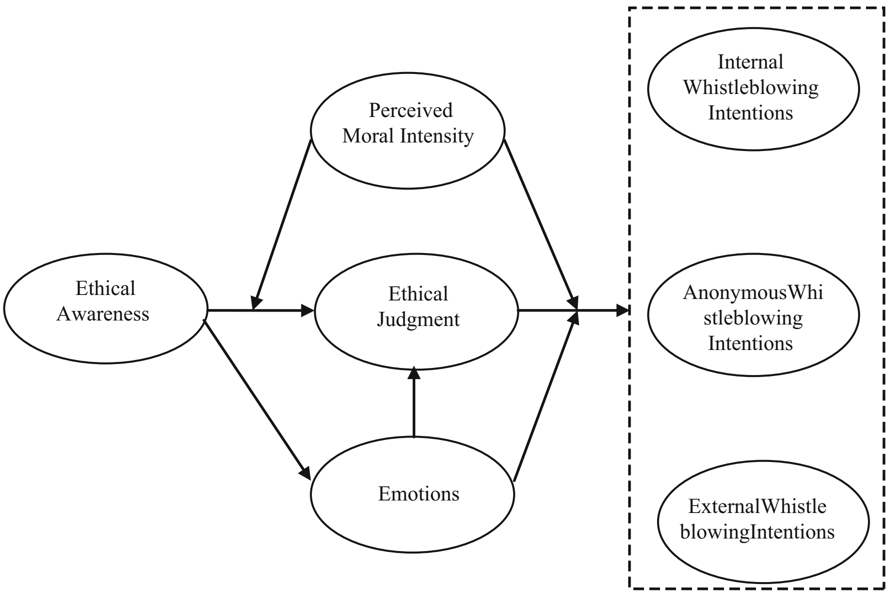


Fig. 15.1 Conceptual model of the whistleblowing decision-making process

EDM process to blow the whistle.[6] The EDM model proposed by Schwartz (2015) is also built on the model of Rest but with additional modification factors of rational and nonrational as intermediaries as well as individual and situational factors as moderating variables. This model has not been widely tested in comparison with previous models, especially in decision-making for whistleblowing.

15.2.2 Ethical Awareness, Emotions, and Ethical Judgment

Butterfield et al. (2000) define ethical awareness as consciousness owned by an individual at a certain time point when faced with ethical dilemmas that require a decision or action that may affect the interests of themselves or others in a way that may conflict with one or more of moral standards. Classical theory of EDM found ethical awareness is a strong predictor of ethical judgment (Rest 1986; Jones 1991) and mediated by nonrationality factors (affective) such as emotions (Lehnert et al. 2015; Henik 2008; Schwartz 2015). As proposed by Henik (2008) and developed further by Schwartz (2015), emotions (such as fear or anger) are also able to mediate the relationship between ethical awareness and ethical judgment for whistleblowing. Emotions can form prosocial or antisocial behavior which can affect a person's decision to reveal any wrongdoing. Previous research has found a significant relationship between ethical awareness and ethical judgment among marketing professionals (Singhapakdi et al. 1996), among upper-division business students (Haines et al. 2008), and in formal infrastructure (Rottig et al. 2011) and mediated by emotion (Connelly et al. 2004; Singh et al. 2016; Henik 2015). From the above discussion, the following hypotheses can be derived:

H1a *Ethical awareness has a positive direct effect on ethical judgment.*

H1b *Ethical awareness has a positive indirect effect on ethical judgment through emotions.*

15.2.3 Moderating Effect of Perceived Moral Intensity on Ethical Awareness and Ethical Judgment

Jones (1991) defines moral intensity as a measure of moral imperative-related problems in certain situations. According to Jones (1991), EDM models should place emphasis on the characteristics of ethical issues themselves. Based on the issues-contingency perspective, Jones placed moral intensity as a predictor variable that affects every phase of the EDM process. Many previous studies have examined

[6] See Leys and Vandekerckhove (2014) for an explanation of the rights and duties of a whistleblower for some types of wrongdoings.

this variable in the context of business ethics (Lehnert et al. 2015; Craft 2013; O'Fallon and Butterfield 2013) and provide results that can be developed further. We adopt this perspective that assumes individuals more easily identify ethical issues when they have high moral intensity. Moral intensity consists of six components (see Jones 1991), but according to Curtis and Taylor (2009), only three factors are relevant in the context of the audit, namely, magnitude of consequences, probability of effect, and proximity; these three factors can affect the auditor's ethical judgment to blow the whistle (p. 198).[7]

Magnitude of consequences is how much loss will result from the wrongdoings and affect the ethical judgment of the auditor. Probability of effect is the impact of that loss in the future (such as retaliation or job loss) and also how it will influence the ethical judgment of the auditor and the intention to blow the whistle. Finally, proximity is a direct influence caused by unethical behavior which harms one of the group members (such as co-workers or family members) and how it affects the ethical judgment of auditors to blow the whistle. In other words, if the impact of the one act does not directly affect the lives of people nearby, the auditor may be reluctant to disclose the error. Previous research has found a significant relationship between moral intensity and ethical judgments (Singer et al. 1998; Valentine and Hollingworth 2012; Yu 2015; McMahon and Harvey 2007; Leitsch 2004). Other studies of Beu et al. (2003) and Singh et al. (2016) showed that moral intensity moderates the relationship between several independent variables to ethical judgments. From the above discussion, the following hypothesis can be derived:

H2 *Moral intensity moderates the relationship between ethical awareness and ethical judgment.*

15.2.4 Moderating Effect of Emotions on Ethical Judgment and Whistleblowing Intentions

By recognizing that decisions can be divided into (a) rationalist based (i.e., reason) and (b) nonrationalist based (i.e., intuition and emotion) (Schwartz 2015), several previous studies have realized the importance of the role of emotions in influencing ethical decisions (Connelly et al. 2004; Curtis 2006). Emotions are feelings that arise (such as anger or fear) when encountering wrongdoing and also influence the auditors' ethical judgment to arrive at the decision to blow the whistle (Henik 2008). Emotions can directly affect the ethical judgment and moral reasoning (Singh et al. 2016). For example, negative mood can be associated with lower intentions to report the unethical actions of others to a superior within an organization (Curtis 2006). According to Schwartz (2015), emotions can also serve as a moderating variable

[7] Alleyne et al. (2016) and Latan et al. (2016) have used the moral intensity as a moderating variable in research related to the whistleblowing intentions.

on the relationship between ethical judgments and whistleblowing intentions. When the auditor is making ethical judgments on specific cases, for example, feelings like anger or fear will continue to be part of a subsequent decision, whether to reveal wrongdoing through internal routes (IWB), external (EWB), or anonymous (AWB) whistleblowing. If the auditor is quite afraid of revealing errors found, because it will affect personal and professional dimensions in the future, then the internal and anonymous route of whistleblowing is usually selected. Conversely, when the auditors ignore the risks, because wrongdoing affects the lives of many people (e.g., Edward Snowden who leaked secret documents from the NSA), they will probably choose the route of external whistleblowing. Previous research has found a significant relationship between emotion and ethical judgments (Connelly et al. 2004; Curtis 2006) and the role of emotions as a moderator in the relationship between ethical judgments and whistleblowing intentions (Hollings 2013; Henik 2015; Schwartz 2015). From the above discussion the following hypothesis can be derived:

H3a *Emotions moderate the relationship between ethical judgment and IWB.*

H3b *Emotions moderate the relationship between ethical judgment and EWB.*

H3c *Emotions moderate the relationship between ethical judgment and AWB.*

15.2.5 Moderating Effect of Perceived Moral Intensity on Ethical Judgment and Whistleblowing Intentions

Recent research shows that high moral intensity can affect ethical judgments of auditors (Yu 2015) and will have a positive impact on the intention to blow the whistle (Alleyne et al. 2013). The model proposed by Jones (1991) placed moral intensity as a predictor variable in influencing every stage of the EDM process. We revise the role of the moral intensity variable by placing it as a moderating variable in line with the integrated EDM model proposed by Schwartz (2015). Ethical judgments made by individuals will be better when matched with high moral intensity and interaction, which in turn have a positive influence on the intention to report wrongdoings. In other words, the higher the perceived moral intensity of an issue, the more likely the person is to make ethical decisions, which in turn affects the intention to blow the whistle. Previous research has shown that ethical judgment has a positive influence on whistleblowing intentions (Zhang et al. 2009; Chiu 2003) and is moderated by moral intensity (Alleyne et al. 2013; Latan et al. 2016). From the above discussion, the following hypothesis can be derived:

H4a *Ethical judgment has a positive direct effect on IWB.*

H4b *Ethical judgment has a positive direct effect on EWB.*

H4c *Ethical judgment has a positive direct effect on AWB.*

H5a *Moral intensity moderates the relationship between EJW and IWB.*

H5b *Moral intensity moderates the relationship between EJW and EWB.*

H5c *Moral intensity moderates the relationship between EJW and AWB.*

15.3 Research Method

15.3.1 Sample Selection and Data Collection

The respondents in our survey are professional accountants working for audit, manufacturing, and financial service companies listed on the Indonesia Stock Exchange (BEI). We chose companies in manufacturing and financial services because, as reported by ACFE 2016, these sectors have the most cases of wrongdoing in Southeast Asia. We also ensure that external auditors who audited the companies were used as a sample and matched with the internal auditor of the companies. The data collection was done using a questionnaire placed on an online platform. A Web link to the questionnaire was then sent by email to the firms. Email addresses from the audit firms were obtained from the directory of the Indonesian Institute of Certified Public Accountants (IAPI) for 2015. Email addresses of manufacturing and financial service companies were extracted from each company's website. Based on the directory and the information available, approximately 74 audit firms were contacted with 400 total respondents from external auditors. Furthermore, 223 manufacturing and financial service companies were contacted. These companies had, in total, 560 internal auditors. After sending a request to participate in the survey, we sent three subsequent emails as a reminder. To ensure data quality control, we checked the collected data, to verify whether there was missing data, straight line responses, or similarity of answers. We found a few problematic cases that were removed from the data before further analysis. Finally, we made additional efforts to increase the response rate, by directly calling the target respondents. To convince the respondents, we conceal their identity (such as name and address of the company), and they remain anonymous. Furthermore, we determine the cutoff time for the return of the questionnaire, which was 3 months, for the purpose of testing nonresponse bias, as suggested by Dillman et al. (2014).

Between July and October 2016, we obtained 179 questionnaire responses from external auditors and 194 questionnaires from internal auditors, of which 38 were incomplete, so the number of questionnaires that were valid and could be used in this study was 335 with a 34.89% response rate. Of the total questionnaires collected, 48.35% came from audit firms and the rest, respectively, 36.09% and 15.56%, came from manufacturing and financial services (see Table 15.1).

Results of the t-test showed that there was no difference in statistical significance of responses ($p < 0.05$) between public accountants who came from the Big 4 and non-Big 4 and also for the social desirability response bias problems (Randall and

Table 15.1 Response rate and profile of respondents

Survey Result	Frequency	Percent
A. Response Rate		
External auditors , Initial = 400	179	18.64 %
Internal auditors, Initial = 560	194	20.21 %
Incomplete questionnaires	38	3.96 %
Response Rate	**335**	**34.89 %**
B. Profile of Respondents		
Gender		
Male	212	63.28 %
Female	123	36.72 %
Total	**335**	**100 %**
Organizational position		
Senior audit staff	143	42.7 %
Junior audit staff	192	57.3 %
Total	**335**	**100 %**
Academic qualifications (education)		
Bachelor's degree	207	61.8 %
Master's degree and doctorate	128	38.2 %
Total	**335**	**100 %**
Professional qualifications	147	43.9 %
CPA	145	43.3 %
QIA and CIA	43	12.8 %
Unqualified	**335**	**100 %**
Total		

Fernandes 2013). This indicates that the size of the audit firm will not affect the results of analysis and there are no problems in social desirability response bias of the respondent's own reporting of whistleblowing intentions.[8] These results also indicate that there is no problem of selection bias that causes the auditor not to take part in the survey (Randall and Fernandes 2013). In addition, the statistical test results also showed that there was no significant difference between respondents who answered in the beginning of data collection, compared with respondents who answered at the end, which means there is no problem of nonresponse bias that occurs systematically (Dillman et al. 2014). To ensure there is no common method bias (Podsakoff et al. 2012), we use the full collinearity approach (Kock 2015). The AVIF (acquired value of in-force) value obtained from analysis is less than 3.3, thus indicating that there is no common method bias problem in this study.

Table 15.1 presents the profile of respondents in this study. The 335 completed questionnaires were divided into two subsamples: 162 external auditors and 173 internal auditors; 63.3% were male and 36.7% were female, with an average age of 37.2 years. In terms of positions, 42.7% of the sample comprised senior audit staff, and 57.3% comprised junior audit staff. As for qualifications, 61.8%

[8]Social desirability response bias is broadly understood as the tendency of individuals to deny socially undesirable traits and behaviors and to admit to socially desirable ones.

held a bachelor's degree, and 38.2% held a master's degree or doctorate, while 87.2% of the sample had professional qualifications, with 43.9% of the sample having completed a professional qualification CPA and 43.3% having completed the Qualified Internal Auditor (QIA) and Certified Internal Auditor (CIA) examinations.

15.3.2 The Survey Structure

The survey used to measure each of the variables in this study consists of three parts. The first section described the purpose and objectives of this research, by asking the respondent's willingness to participate in the survey. The second section asked for the respondents' demographic information such as gender, age, education level, occupation, and qualifications. The third section presented scenarios and questions related to the variables to be studied. Given the difficulty in gaining access to the object in order to observe real unethical behavior, a scenario approach is commonly used in research in the field of accounting and ethics (e.g., Alleyne et al. 2016; Arnold et al. 2013; Chan and Leung 2006; Curtis and Taylor 2009; Shawver et al. 2015). This approach illustrates a specific case, and the respondents are asked to respond and put themselves as an actor in such situations. The scenario used in this study was adopted from the scenario used by Bagdasarov et al. (2016), Clements and Shawver (2011), Curtis and Taylor (2009), Kaplan and Whitecotton (2001), and Schultz et al. (1993) with modifications, which highlights the numerous violations of professional ethics and wrongdoings in a company.[9]

To create a scale able to measure the intentions to blow the whistle, we used a total of ten items of questions based on the internal, external, and anonymous reporting routes adopted by Park et al. (2008). The survey respondents were asked about reporting routes that they use to select when they find wrongdoings that occur (hypothetical scenario). The variable ethical awareness was measured by three questions adopted from Arnold et al. (2013). Respondents were asked about whether an action in the case scenario is ethical or unethical behavior. The variable ethical judgment for whistleblowing was measured through four items inspired by Reidenbach and Robin (2013). Respondents were asked about whether an action in the scenario is moral or not morally right, just or unjust, acceptable or unacceptable, and so on. Tables 15.2 and 15.3 show indicators and outcome measurement models for variables of ethical awareness, ethical judgment, and intentions of whistleblowing.

The moral intensity variable is measured by six questions adopted from Clements and Shawver (2011). Respondents were asked to provide feedback on the scenarios to assess the intensity level of their morals. Finally, emotional variables measured four items of questions adopted from Connelly et al. (2004). Respondents were

[9]The use of scenarios is more effective to give stimuli to the auditor in making ethical decisions when faced with certain situations.

Table 15.2 Construct indicators and measurement model of whistleblowing intentions

Indicators/items	Code	FL	AVE	rho_A
Internal whistleblowing (IWB)				
Report it to the appropriate persons within the firm	IWB1	0.864	0.608	0.875
Use the reporting channels inside of the firm	IWB2	0.738		
Let upper-level management know about it	IWB3	0.880		
Tell my supervisor about it	IWB4	0.604		
External whistleblowing (EWB)				
Report it to the appropriate authorities outside of the firm	EWB1	0.800	0.578	0.849
Use the reporting channels outside of the firm	EWB2	0.800		
Provide information to outside agencies	EWB3	0.762		
Inform the public about it	EWB4	0.671		
Anonymous whistleblowing (AWB)				
Reports it using an assumed name	AWB1	0.783	0.668	0.803
Reports the wrongdoing but doesn't give any information about himself	AWB2	0.850		

FL factor loading

Table 15.3 Construct indicators and measurement model of EAW and WBJ

Indicators/items	Code	FL	AVE	rho_A
Ethical awareness (EAW)				
To what extent do you regard the action as unethical	EAW1	0.918	0.622	0.863
To what extent would the "typical" [internal] auditor at your level in your firm [company] regard this action as unethical	EAW2	0.562		
To what extent would the "typical" [external] auditor at your level in your firm [company] regard this action as unethical	EAW3	0.841		
Ethical judgment whistleblowing (EJW)				
Fair/unfair	EJW1	0.925	0.809	0.945
Just/unjust	EJW2	0.848		
Acceptable/unacceptable	EJW3	0.892		
Morally/not morally right	EJW4	0.929		

FL factor loading

asked to provide feedback on the scenarios to assess the level of their emotions. The value of the loading factor, average variance extracted (AVE), and reliability derived from the analysis of the measurement model for all variables are loading factor >0.60, composite reliability/rho_A > 0.70, and AVE > 0.50, so it meets the recommended requirements (Hair et al. 2017; Henseler et al. 2018). However, there are some indicators of measurement models that were retained, with the value of the loading factor being >0.5. As stated by Hair et al. (2017, p. 114), the value of the loading factor shows the explained variance in a construct. So, if the value AVE is already more than 0.5, the indicator with low loading values can be kept to maintain the content validity. Table 15.4 shows the indicators and outcome measurement model for moral intensity and emotional variables.

In addition, we tested the discriminant validity or divergent validity for all latent variables in the model using the heterotrait-monotrait ratio (HTMT). As stated by Henseler et al. (2015), HTMT is a new procedure to test the discriminant validity and is more appropriate than the Fornell-Larcker criterion. The HTMT approach has reliable performance and overcomes bias in the estimation of parameters of the structural model. In Table 15.5, it is shown that the value of HTMT was smaller than 0.90, which means that it meets the recommended rule of thumb (Hair et al. 2017; Henseler et al. 2015).

Table 15.4 Construct indicators and measurement model of PMI and emotions

Indicators/items	Code	FL	AVE	rho_A
Perceived moral intensity (PMI)				
Should not do the proposed action	PMI1	0.660	0.619	0.911
Approving the bad debt adjustment is wrong	PMI2	0.750		
Approving the bad debt adjustment will cause harm	PMI3	0.826		
Approving the bad debt adjustment will not cause any harm	PMI4	0.875		
If the CEO is a personal friend, approving the bad debt adjustment is wrong	PMI5	0.829		
Approving the bad debt adjustment will harm very few people if any	PMI6	0.761		
Emotions (EMT)				
Feel that you have really accomplished something significant	EMT1	0.803	0.515	0.826
Find it incredible how you have had an influence in others' lives	EMT2	0.835		
Think that a change will not necessarily improve your situation	EMT3	0.643		
Feel like there was nothing you could do	EMT4	0.553		

FL factor loading

Table 15.5 Correlations and discriminant validity results

Construct	Mean	SD	1	2	3	4	5	6	7
AWB	4.73	1.32	1	0.766	0.443	0.651	0.681	0.762	0.567
EMT	4.86	1.36	0.615*	1	0.701	0.801	0.711	0.822	0.721
EAW	5.55	1.14	0.353*	0.563*	1	0.514	0.633	0.647	0.595
EJW	4.93	1.44	0.562*	0.697*	0.446*	1	0.658	0.655	0.812
EWB	5.42	1.21	0.564*	0.587*	0.549*	0.589*	1	0.826	0.697
IWB	4.94	1.22	0.628*	0.684*	0.539*	0.591*	0.707*	1	0.628
PMI	5.19	1.46	0.481*	0.615*	0.508*	0.754*	0.609*	0.555*	1

Notes: Below the diagonal elements are the correlations between the construct values
Above the diagonal elements are the HTMT values
Correlation is significant at the 0.05 level (2-tailed)

15.3.3 Data Analysis

Before we analyzed the overall model, we ensured that the adequacy of the sample size for estimation of the model had been fulfilled. Because the data analysis in this study uses the consistent partial least squares (PLSc) approach, a sample needs to have at least 100 cases (Latan and Ghozali 2015). The main purpose of PLSc is to mimic the covariance-based SEM approach to test or confirm the theory (Dijkstra and Henseler 2015). By using PLSc, the estimator of the model will be consistent for the loading and the correlation between latent variables and allows us to access the goodness of fit (Dijkstra 2014). We chose PLSc with the consideration that it is more appropriate to test complex models, where the CB-SEM approach would be difficult to apply (Richter et al. 2016; Rigdon 2016). Previous research in this area already uses PLS-SEM as an analytical tool (Buchan 2005; Haines et al. 2008). In contrast to other SEM techniques, PLS-SEM does not rely on the assumption of normality (distribution free) because it is nonparametric. However, some assumptions, such as multicollinearity and goodness of fit for the local models assessment, need to be considered. Overall, the data analysis in this study goes through three stages. First, we analyze the measurement model to ensure indicator constructs are valid and reliable using the full sample. Second, we examine multigroup analysis to compare the two subsamples for each path coefficient. Third, we examine the effect of mediation-moderation to determine the role of moral intensity and emotional variables.

15.4 Results

In this study, data analysis and hypotheses testing were conducted by using variance-based SEM. One of the techniques available today is PLS-SEM, which is the most fully developed and has become a vital tool for researchers to examine various issues of social science. PLS-SEM was developed with the main purpose of prediction and then extended to test the theory with consistent results for the factor models. We chose to use PLSc (on selection algorithms and bootstrapping) considering that it will provide similar results to CB-SEM.[10] We use the SmartPLS 3 program (Ringle et al. 2015) to analyze these models by using PLSc.

PLS-SEM analysis proceeds through two stages, namely, the measurement model and the structural model. Assessment of the measurement model is intended to test the validity (convergent and discriminant) and reliability of each indicator forming latent constructs. After we make sure that all the indicators' constructs are valid and reliable (see Fig. 15.2), we continue the analysis to the second stage of assessing the quality of the structural model and run multigroup analysis to test the

[10]Dijkstra and Henseler (2015) give a detailed explanation related to PLSc.

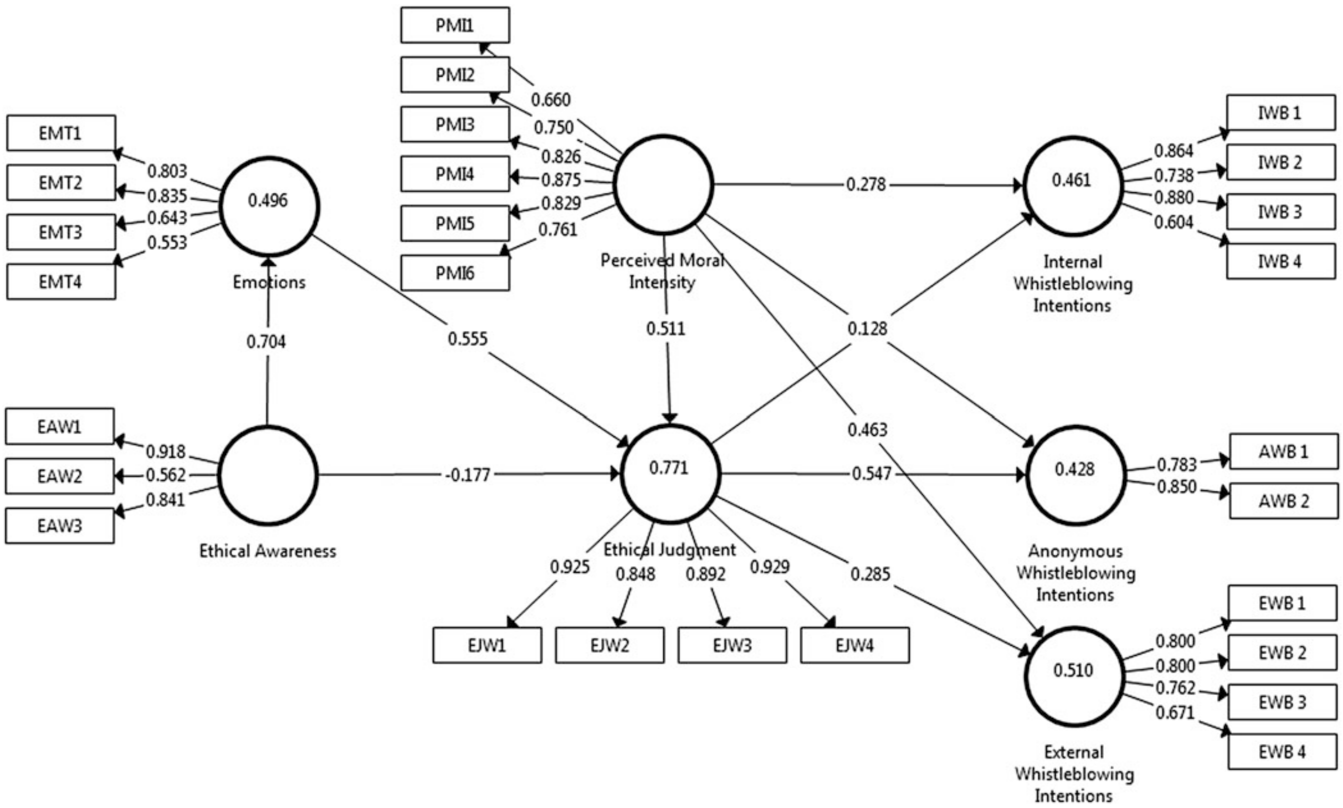


Fig. 15.2 Evaluation of the measurement model with the full sample

Table 15.6 Structural model results

Constructs	R^2	Adj. R^2	f^2	Q^2	VIF	SRMR	NFI	AFVIF
Ethical awareness (EAW)	–	–	0.067–0.494	–	2.030	–	–	–
Ethical judgment (EJW)	0.771	0.769	0.056–0.178	0.764	2.945	–	–	–
Moral intensity (PMI)	–	–	0.010–0.520	–	2.193	–	–	–
Emotions (EMT)	0.496	0.494	0.049–0.472	0.491	2.845	–	–	–
Internal whistleblowing (IWB)	0.461	0.458	–	0.453	–	0.049	0.837	2.503
Anonymous whistleblowing (AWB)	0.428	0.425	–	0.423	–	0.049	0.837	2.503
External whistleblowing (EWB)	0.510	0.507	–	0.502	–	0.049	0.837	2.503

hypothesis. The results of the quality assessment for the structural model are given in Table 15.6.

In Table 15.6 it is shown that the whistleblowing intention (IWB, AWB, and EWB) can be explained by the predictor variables with adjusted R^2 of 0.425–0.507. This value indicates that the ability of the predictor variables to explain the outcome variables was approaching substantial (Latan and Ghozali 2015). The resulting effect size value of each predictor variable in the model ranged from 0.01 to 0.520, which is included in the category of small to large. The value of variance inflation factor (VIF) generated for all the independent variables in the model is <3.3, which means that there was no collinearity problem between the predictor variables. The Q^2 predictive relevance value generated excellent endogenous variables, i.e., >0, which means that the model has predictive relevance. The value of goodness of fit is generated through the standardized root mean squared residual (SRMR) that is equal to 0.049 <0.080 and the normed fix index (NFI) 0.837 >0.80, which means that our model fits the empirical data.

15.4.1 Multigroup Analysis (PLS-MGA)

We run multigroup analysis to compare the two subsamples of internal whistleblower (internal auditor) and external whistleblower (external auditor) for each path coefficients using the PLS-MGA approach. The purpose of the analysis of PLS-MGA was to compare two groups of samples to determine statistically significant differences in group-specific parameter estimates (Matthews 2018; Sarstedt et al. 2011) and in this case which group is more prone or unlikely to blow the whistle. Before running the PLS-MGA, we consider it to test the measurement invariance of composite models (MICOM) using a permutation procedure.[11] We test measurement invariance to ensure that the specific group difference of the estimation model does not affect the results for latent variables in the whole group (Henseler et al. 2016; Solovida and Latan 2017). The analysis showed that there was no significant difference between variance and average values for the two groups (see Table 15.7), which means no invariance problem that will affect the outcome.

Based on the analysis in Table 15.7, it can be seen that the ethical awareness (EAW) has no effect on ethical judgment (EJW) for either internal or external group auditors. The analysis for each group obtained coefficient (β) values for the relationship EAW $\rightarrow$ EJW of 0.070 and -0.057, with 95% bias corrected and accelerated (BCa) > 0.05. This means that the hypothesis 1a (H1a) was rejected. These results support previous studies (Chan and Leung 2006; Valentine and Fleischman 2004). EAW cannot be a direct predictor of the EJW, and this is consistent with the integrated EDM model by Schwartz (2015), where there is another factor that mediates both. EAW of professional accountants in this study also found variance in their ability to respond to a case scenario. Furthermore, the values of the coefficient (β) for the relationship EAW $\rightarrow$ EMT are 0.670 and 0.553, and EMT $\rightarrow$ EJW are 0.482 and 0.386, with 95% bias corrected and accelerated (BCa) < 0.01, respectively. This means that the hypothesis 1b (H1b) is supported.

We also tested the indirect effect by using the method proposed by Cepeda et al. (2018) and obtained the same results.[12] These results support previous studies (Henik 2015; Connelly et al. 2004; Singh et al. 2016; Curtis 2006). This suggests that emotions may serve as indirect-only mediation or full mediation of the relationship between EAW and EJW. When someone finds wrongdoing, it will affect their emotions prior to making ethical judgments. From these findings, it can be concluded that the internal auditors have more intense EAW, EMT, and EJW than the external auditors.

Finally, from Table 15.7, it can be seen that the values of the coefficient (β) for the relationship EJW $\rightarrow$ IWB are 0.446 and 0.317; EJW $\rightarrow$ AWB is 0.400; 0.419 and EJW $\rightarrow$ EWB is 0.315; and 0.309 for each group of samples with 95% bias corrected and accelerated (BCa) < 0.01, respectively. This means that the hypothesis 4 (H4a, H4b, and H4c) is supported. These results support previous studies (Zhang et al. 2009; Chiu 2003; Arnold et al. 2013; Buchan 2005). As stated by Culiberg and Mihelic (2016), most of the research in this area has provided conclusive results for the relationship between EJW and whistleblowing intentions. A professional accountant who has made ethical judgments can report wrongdoing found through one of these three route options available: internal, external, or anonymous. The results showed that the internal route is the most preferred by the internal auditor followed by an anonymous and external route. In contrast, for external auditors, the anonymous route is the most preferred, followed by internal and external. This

[11]Conceptually, measurement invariance expresses the idea that the measurement properties of X in relation to the target latent trait Wt are the same across populations.

[12]Cepeda et al. (2018) propose to use a spreadsheet to calculate the indirect effects.

Table 15.7 PLS-MGA results (direct effect)

Structural path	Internal (β)	External (β)	PLS-MGA	95% BCa CI permutation	MICOM	Equal variances	Conclusion
EAW → EJW	$0.070^{n.s.}$	$-0.057^{n.s.}$	$0.068^{n.s.}$	$0.133^{n.s.}$	$(-0.038;-0.102)^{n.s.}$	Yes	H1a not supported
EAW → EMT	0.670**	0.553**	$0.052^{n.s.}$	$0.112^{n.s.}$	$(-0.038; -0.191)^{n.s.}$	Yes	H1b supported
EMT → EJW	0.482**	0.386**	$0.175^{n.s.}$	$0.195^{n.s.}$	$(-0.191; -0.102)^{n.s.}$	Yes	H1b supported
EJW→ IWB	0.446**	0.317**	$0.178^{n.s.}$	$0.226^{n.s.}$	$(-0.102; 0.007)^{n.s.}$	Yes	H4a supported
EJW → AWB	0.400**	0.419**	$0.838^{n.s.}$	$0.219^{n.s.}$	$(-0.102; -0.206)^{n.s.}$	Yes	H4b supported
EJW → EWB	0.315**	0.309**	$0.751^{n.s.}$	$0.260^{n.s.}$	$(-0.102; -0.084)^{n.s.}$	Yes	H4c supported

n.s. not significant
$p < 0.05$ (one-tailed test)
*$p < 0.01$ (one-tailed test)

indicates that professional accountants of both groups in the cases of Indonesia chose an external route to blow the whistle as the last option. They are more likely to disclose an error discovered through internal and anonymous routes. One reason that might affect their decisions is fear of retaliation and the various risks that arise when using an external route for whistleblowing.

These findings indicate that internal auditors have a higher (more likely) intention to report any act than external auditors and blowing the whistle internally and anonymously can be more useful for professional accountants. Findings are aligned with the general statement that employees are not the only ones with privileged information about a company, and consequently outsiders may observe various wrongdoings (Culiberg and Mihelic 2016). However, the present study adds a more detailed suggestion that internal auditors are more likely to report than external auditors. Although the literature has suggested that there is not, a priori, a profile of whistleblowers that organizations can attempt to screen out during recruitment (Henik 2015), our findings suggest that internal auditors are more likely to blow the whistle than external ones. While the literature recognizes that there are challenges in fully protecting external whistleblowers (Maroun and Gowar 2013), our findings suggest that discussing how to fully protect internal auditors should also be a priority.

However, as discussed by Maroun and Atkins (2014a, b), there is an upward trend of increasing the availability of information to stakeholders and enhancing the level of expectation that the public have on auditors, in terms of transparency and accountability and in terms of relevance of audit reports (Maroun and Atkins 2014a, b). If this were reinforced in Indonesia, our results would be different. This scenario will need to further consider the challenges in fully protecting external whistleblowers (Maroun and Gowar 2013).

15.4.2 Importance Performance-Map Analysis

We tested importance-performance map analysis (IPMA). Ringle and Sarstedt (2016) stated that the IPMA gives researchers the opportunity to enrich their PLS-SEM analysis and, thereby, gain additional results and findings. Nevertheless, PLS-SEM has several key advantages over traditional IPMA that typically relies on multiple regression analysis. First, in determining the importance scores, PLS-SEM is a valuable analytical tool as it is capable of integrally assessing a complex network of relationships connecting drivers to a target construct of interest. Second, it can incorporate latent constructs (see Streukens et al. 2018). The IPMA analysis results are shown in Table 15.8.

From the above analysis (Table 15.8), it can be seen that the EMT has a relatively low performance value of 64.45. If matched by other constructs, EMT's performance is slightly below average. On the other hand, with a total effect of 0.595, this construct's importance is high enough. Therefore, a one-unit increase in EMT's performance from 64.45 to 65.45 would increase the performance of IWB,

Table 15.8 The IPMA for construct IWB, AWB, and EWB

Constructs	IWB		AWB		EWB	
	Importance	Performance	Importance	Performance	Importance	Performance
EAW	0.377	71.60	0.369	71.60	0.283	71.60
EMT	0.595	64.45	0.590	64.45	0.460	64.45
EJW	0.170	65.52	0.237	65.52	0.259	65.52
PMI	0.098	69.71	0.137	69.71	0.150	69.71

Table 15.9 Relationships between variables (interaction effect)

Structural path	Coef (β)	SD	95\% BCa CI	Conclusion
EAW × PMI → EJW	0.031	0.044	(0.140, 0.046)	H2 not supported
EJW × EMT → IWB	0.181	0.051	(0.015, 0.262)*	H3a supported
EJW × EMT → AWB	0.115	0.046	(0.049, 0.133)*	H3b supported
EJW × EMT → EWB	0.151	0.049	(0.034, 0.200)*	H3c supported
EJW × PMI → IWB	0.176	0.049	(0.011, 0.257)*	H5a supported
EJW × PMI → AWB	0.103	0.050	(0.043, 0.108)*	H5b supported
EJW × PMI → EWB	0.098	0.044	(0.042, 0.125)*	H5c supported

Note: **, * Statistically significant at the 1\% and 5\% levels, respectively

AWB, and EWB to 0.595, 0.590, and 0.460, respectively. Therefore, if companies aim to improve the IWB, AWB, and EWB of internal and external auditors, their first priority must be to improve the performance aspects of EMT. Furthermore, aspects related to EJW, PMI, and EAW follow as second, third, and fourth priorities.

15.4.3 Interaction Effect Analysis

We tested the interactions using the orthogonalization approach. This approach was chosen because it produces an accurate estimate, has a high predictive accuracy, and is able to minimize the collinearity problem. The results of the analysis of interactions are shown in Table 15.9.

In Table 15.9 it is shown that H3a, H3b, and H3c and H5a, H5b, and H5c are fully supported, showing that moral intensity and emotional may moderate the relationship between EJW and whistleblowing intentions. The relationship EAW × PMI → EJW obtained insignificant results, with coefficient (β) = 0.031 and 95% bias corrected and accelerated (BCa) = 0.140 >0.05. This suggests that emotions or feelings of auditors themselves play an important role in improving the ethical assessment of auditors with the consequence that they have a higher whistleblowing intention to report any wrongdoing that occurs, reinforcing the discussion on nonrationalist-based decision-making (Schwartz 2015). This finding can be understood by taking into account a broader discussion on how mood and emotions can influence whistleblowing (Curtis 2006).

While the moral intensity that comes from the experience of auditors would assist in considering any magnitude of consequences, the possibility of future losses and the proximity to the organization influence actions to blow the whistle. Emotions felt would assist the auditor in considering the various risks arising from actions taken.

From the results of this analysis, we reached the same conclusion, that the internal and anonymous route is a favorite choice for professional accountants in Indonesia to reporting wrongdoing. These results support previous studies (Hollings 2013; Henik 2015; Alleyne et al. 2016; Latan et al. 2016). Given the cultural and social norms' strength in Indonesia, the freedom to act and speak out becomes a supporting factor for professional accountants in improving the intention to report wrongdoing without fear of reprisal. Nevertheless, it is important to further develop institutional mechanisms capable of fully protecting whistleblowing (Maroun and Gowar 2013).

15.5 Conclusion

This study aims to examine the integrated EDM model proposed by Schwartz (2015), where we consider the factors of individual nonrationality that affect ethical judgments of the auditor to arrive at the decision to blow the whistle. We answered the call of Culiberg and Mihelic (2016) to extend the testing of EDM models in the whistleblowing context. In this chapter, we argue that the intention of whistleblowing depends on EAW and EJW as well as on emotion and perceived moral intensity.

We support the hypothesis that EAW cannot directly affect the EJW but must go through the nonrationality of factors such as emotion. We also found that internal and anonymous whistleblowing routes were used by professional accountants in the case of Indonesia. In terms of practical implications, these findings provide a deep understanding of how audit firms, manufacturing, and financial services should be selective in choosing audit staff who uphold professional and ethical standards of behavior. In addition, companies need to make strong efforts to implement a comprehensive ethics program including training in ethics and codes of conduct, which provide guidance to staff auditors to resolve ethical conflicts and increase professional responsibility to report wrongdoing. Companies also need to apply the right strategy to enhance the auditor's whistleblowing intentions and reduce the fear of retaliation, for example, by providing a whistleblowing hotline or reporting of anonymity, which was a favorite choice for the Indonesian context.

15.6 Limitation and Future Research

There are several limitations to this study which need the attention of the reader. First, this study did not consider cultural factors that may affect the EDM process. Some cultural factors such as nationality, patriotism, religion, and political system may affect the EAW and EJW of auditors. These findings may differ in other countries. Second, this study only considers the factors of nonrationality in the integrated EDM model proposed by Schwartz (2015), without examining the factors of rationality. Different results may be obtained when considering both. Third, this study only used two variables as mediation-moderation in the model. Lehnert et al. (2015) showed that there are still many relevant variables (moderation and mediation) more important to be considered and tested in the EDM model. Fourth, this study did not consider the effect of extraneous variables (such as age, gender, education, or total tenure) and unobserved heterogeneity that might interfere with the results. However, several previous studies showed inconsistency in the role of extraneous variables in the EDM model (Chan and Leung 2006; Cagle and Baucus 2006; Ebrahimi et al. 2005; Shafer et al. 2001; Marques and Azevedo-Pereira 2009). In addition, the selection bias could have been handled more carefully. Finally, this study only tested the whistleblowing intentions without testing actual behavior.

Further research can follow up the testing of integrated EDM model by Schwartz (2015) for whistleblowing by considering factors of rationality and nonrationality as intuition, reason, and confirmation. Cultural factors also need to be considered for further study. This is a call for research to provide empirical evidence of the model. Furthermore, future research may use other moderating variables such as intrinsic religiosity, personal spirituality, moral obligation, retaliation, intelligence, and others which have an important role in the EDM process (Liyanarachchi and Newdick 2009; Haines et al. 2008; Bloodgood et al. 2008). Replication studies on the other subject group (e.g., consumers vs shareholder) and other organizations (e.g., government and public administration) will also allow access to generalize the findings of this study. Overall, the researchers feel that it is necessary to replicate this study by using qualitative approaches such as case studies or fuzzy-set qualitative comparative analysis (Ragin 2008), taking into account unobserved heterogeneity testing (Hair et al. 2012; Schlittgen et al. 2016), which might be fruitful for new avenues for future study,[13] as there are not many studies have used a qualitative approach to test the EDM model for whistleblowing.

[13]Lehnert et al. (2015) were surprised to find only two studies using qualitative approach in their literature review.

References

Alleyne, P., Hudaib, M., & Pike, R. (2013). Towards a conceptual model of whistle-blowing intentions among external auditors. *The British Accounting Review, 45*(1), 10–23.

Alleyne, P., Hudaib, M., & Haniffa, R. (2016). The moderating role of perceived organisational support in breaking the silence of public accountants. *Journal of Business Ethics*, 1–19. doi:10.1007/s10551-015-2946-0.

Archambeault, D. S., & Webber, S. (2015). Whistleblowing 101. *CPA Journal, 85*(7), 62–68.

Arnold, D. F., Dorminey, J. W., Neidermeyer, A. A., & Neidermeyer, P. E. (2013). Internal and external auditor ethical decision-making. *Managerial Auditing Journal, 28*(4), 300–322.

Bagdasarov, Z., Johnson, J. F., MacDougall, A. E., Steele, L. M., Connelly, S., & Mumford, M. D. (2016). Mental models and ethical decision making: The mediating role of sensemaking. *Journal of Business Ethics, 138*, 133–144.

Beu, D. S., Buckley, M. R., & Harvey, M. G. (2003). Ethical decision-making: A multidimensional construct. *Business Ethics: A European Review, 12*(1), 88–107.

Bloodgood, J. M., Turnley, W. H., & Mudrack, P. (2008). The influence of ethics instruction, religiosity, and intelligence on cheating behavior. *Journal of Business Ethics, 82*(3), 557–571.

Buchan, H. F. (2005). Ethical decision making in the public accounting profession: An extension of Ajzen's theory of planned behavior. *Journal of Business Ethics, 61*(2), 165–181.

Butterfield, K. D., Trevino, L. K., & Weaver, G. R. (2000). Moral awareness in business organizations: Influences of issue-related and social context factors. *Human Relations, 53*(7), 981–1018.

Cagle, J. A. B., & Baucus, M. S. (2006). Case studies of ethics scandals: Effects on ethical perceptions of finance students. *Journal of Business Ethics, 64*, 213–229.

Cepeda, G., Nitzl, C., & Roldán, J. L. (2018). Mediation analyses in partial least squares structural equation modeling: Guidelines and empirical examples. In H. Latan, H., & R. Noonan (Eds.), *Partial least squares path modeling: Basic concepts, methodological issues and applications*. New York: Springer.

Chan, S. Y. S., & Leung, P. (2006). The effects of accounting students' ethical reasoning and personal factors on their ethical sensitivity. *Managerial Auditing Journal, 21*(4), 436–457.

Chiu, R. K. (2002). Ethical judgement, locus of control, and whistleblowing intention: A case study of mainland Chinese MBA students. *Managerial Auditing Journal, 17*(9), 581–587.

Chiu, R. K. (2003). Ethical judgment and whistleblowing intention: Examining the moderating role of locus of control. *Journal of Business Ethics, 43*, 65–74.

Clements, L. H., & Shawver, T. (2011). Moral intensity and intentions of accounting professionals to whistleblow internally. *Journal of Forensic Studies in Accounting and Business, 3*(1), 67–82.

Connelly, S., Helton-Fauth, W., & Mumford, M. D. (2004). A managerial in-basket study of the impact of trait emotions on ethical choice. *Journal of Business Ethics, 51*, 245–267.

Craft, J. L. (2013). A review of the empirical ethical decision-making literature: 2004–2011. *Journal of Business Ethics, 117*, 221–259.

Culiberg, B., & Mihelic, K. K. (2016). The evolution of whistleblowing studies: A critical review and research agenda. *Journal of Business Ethics*. doi:10.1007/s10551-016-3237-0.

Curtis, M. B. (2006). Are audit-related ethical decisions dependent upon mood? *Journal of Business Ethics, 68*, 191–209.

Curtis, M. B., & Taylor, E. Z. (2009). Whistleblowing in public accounting: Influence of identity disclosure, situational context, and personal characteristics. *Accounting and the Public Interest, 9*(1), 191–220.

Dijkstra, T. K. (2014). PLS' janus face – Response to professor Rigdon's 'rethinking partial least squares modeling: In praise of simple methods'. *Long Range Planning, 47*(3), 146–153.

Dijkstra, T. K., & Henseler, J. (2015). Consistent partial least squares path modeling. *MIS Quarterly, 39*(2), 297–316.

Dillman, D. A., Smyth, J. D., & Christian, L. M. (2014). *Internet, phone, mail, and mixed mode surveys: The tailored design method* (4th ed.). Hoboken, NJ: Wiley.

Ebrahimi, B. P., Petrick, J. A., & Young, S. A. (2005). Managerial role motivation and role-related ethical orientation in Hong Kong. *Journal of Business Ethics, 60*(1), 29–45.

Ferrell, O. C., & Gresham, L. G. (1985). A contingency framework for understanding ethical decision making in marketing. *Journal of Marketing, 49*(3), 87–96.

Haines, R., Street, M. D., & Haines, D. (2008). The influence of perceived importance of an ethical issue on moral judgment, moral obligation, and moral intent. *Journal of Business Ethics, 81*(2), 387–399.

Hair, J. F., Sarstedt, M., Ringle, C. M., & Mena, J. A. (2012). An assessment of the use of partial least squares structural equation modeling in marketing research. *Journal of the Academy of Marketing Science, 40*(3), 414–433. doi:10.1007/s11747-011-0261-6.

Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2017). *A primer on partial least squares structural equation modeling (PLS-SEM)* (2nd ed.). Thousand Oaks, CA: Sage.

Henik, E. (2008). Mad as hell or scared stiff? The effects of value conflict and emotions on potential whistle-blowers. *Journal of Business Ethics, 80*, 111–119.

Henik, E. (2015). Understanding whistle-blowing: A set-theoretic approach. *Journal of Business Research, 68*(2), 442–450.

Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science, 43*(1), 115–135. doi:10.1007/s11747-014-0403-8.

Henseler, J., Ringle, C. M., & Sarstedt, M. (2016). Testing measurement invariance of composites using partial least squares. *International Marketing Review, 33*(3), 405–431. doi:10.1108/IMR-09-2014-0304.

Henseler, J., Hubona, G., & Ray, P. A. (2018). Partial least squares path modeling: Updated guidelines. In H. Latan & R. Noonan (Eds.), *Partial least squares structural equation modeling: Basic concepts, methodological issues and applications*. New York: Springer.

Hollings, J. (2013). Let the story go: The role of emotion in the decision-making process of the reluctant, vulnerable witness or whistle-blower. *Journal of Business Ethics, 114*, 501–512.

Hunt, S. D., & Vitell, S. J. (1986). A general theory of marketing ethics. *Journal of Macromarketing, 6*(1), 5–16.

Hunt, S. D., & Vitell, S. J. (2006). The general theory of marketing ethics: A revision and three questions. *Journal of Macromarketing, 26*(2), 143–153.

Jones, T. M. (1991). Ethical decision making by individuals in organizations: An issue-contingent model. *The Academy of Management Review, 16*(2), 366–395.

Jubb, P. B. (2000). Auditors as whistleblowers. *International Journal of Auditing, 4*, 153–167.

Kaplan, S. E., & Whitecotton, S. M. (2001). An examination of auditors' reporting intentions when another auditor is offered client employment. *Auditing: A Journal of Practice & Theory, 20*(1), 45–63.

Kock, N. (2015). Common method bias in PLS-SEM: A full collinearity assessment approach. *International Journal of e-Collaboration, 11*(4), 1–10.

Latan, H., & Ghozali, I. (2015). *Partial least squares: Concepts, techniques and application using program SmartPLS 3.0* (2nd ed.). Semarang: Diponegoro University Press.

Latan, H., Ringle, C. M., & Jabbour, C. J. C. (2016). Whistleblowing intentions among public accountants in Indonesia: Testing for the moderation effects. *Journal of Business Ethics*. doi:10.1007/s10551-016-3318-0.

Lehnert, K., Park, Y.-H., & Singh, N. (2015). Research note and review of the empirical ethical decision-making literature: Boundary conditions and extensions. *Journal of Business Ethics, 129*, 195–219.

Leitsch, D. L. (2004). Differences in the perceptions of moral intensity in the moral decision process: An empirical examination of accounting students. *Journal of Business Ethics, 53*(3), 313–323.

Leys, J., & Vandekerckhove, W. (2014). Whistleblowing duties. In A. J. Brown, D. Lewis, R. Moberly, & W. Vandekerckhove (Eds.), *International handbook of whistleblowing research* (pp. 115–132). Cheltenham: Edward Elgar.

Liyanarachchi, G. A., & Newdick, C. (2009). The impact of moral reasoning and retaliation on whistle-blowing: New Zealand evidence. *Journal of Business Ethics, 89*(1), 37–57.

MacGregor, J., & Stuebs, M. (2014). Whistle while you work: Whistleblowing in the presence of competing incentives and pressures. *Accounting Perspectives, 13*(4), 309–324.

Maroun, W., & Atkins, J. (2014a). Section 45 of the auditing profession act: Blowing the whistle for audit quality? *The British Accounting Review, 46*(3), 248–263.

Maroun, W., & Atkins, J. (2014b). Whistle-blowing by external auditors in South Africa: Enclosure, efficient bodies and disciplinary power. *Accounting, Auditing & Accountability Journal, 27*(5), 834–862.

Maroun, W., & Gowar, C. (2013). South African auditors blowing the whistle without protection: A challenge for trust and legitimacy. *International Journal of Auditing, 17*(2), 177–189.

Marques, P. A., & Azevedo-Pereira, J. (2009). Ethical ideology and ethical judgments in the Portuguese accounting profession. *Journal of Business Ethics, 86*, 227–242.

Matthews, L. (2018). Applying multigroup-analysis in PLS-SEM: A step-by-step process. In H. Latan & R. Noonan (Eds.), *Partial least squares structural equation modeling: Basic concepts, methodological issues and applications*. New York: Springer.

McMahon, J. M., & Harvey, R. J. (2007). The effect of moral intensity on ethical judgment. *Journal of Business Ethics, 72*(4), 335–357.

Miceli, M. P., Dreyfus, S., & Near, J. P. (2014). Outsider whistleblowers: Conceptualizing and distinguishing "bell-ringing" behavior. In A. J. Brown, D. Lewis, R. Moberly, & W. Vandekerckhove (Eds.), *International handbook on whistleblowing research* (pp. 71–94). Cheltenham: Edward Elgar.

O'Fallon, M. J., & Butterfield, K. D. (2013). A review of the empirical ethical decision-making literature: 1996–2003. In A. C. Michalos & D. C. Poff (Eds.), *Citation classics from the Journal of Business Ethics: Celebrating the first thirty years of publication* (pp. 213–263). New York: Springer.

O'Sullivan, P., & Ngau, O. (2014). Whistleblowing: A critical philosophical analysis of the component moral decisions of the act and some new perspectives on its moral significance. *Business Ethics: A European Review, 23*(4), 401–415.

Park, H., Blenkinsopp, J., Oktem, M. K., & Omurgonulsen, U. (2008). Cultural orientation and attitudes toward different forms of whistleblowing: A comparison of South Korea, Turkey, and the U.K. *Journal of Business Ethics, 82*, 929–939.

Podsakoff, P. M., MacKenzie, S. B., & Podsakoff, N. P. (2012). Sources of method bias in social science research and recommendations on how to control it. *Annual Review of Psychology, 63*(4), 539–569.

Ponemon, L. A. (1994). Whistle-blowing as an internal control mechanism: Individual and organizational considerations. *Auditing: A Journal of Practice & Theory, 13*(2), 118–130.

Ragin, C. C. (2008). *Redesigning social inquiry fuzzy sets and beyond*. Chicago: The University of Chicago Press.

Randall, D. M., & Fernandes, M. F. (2013). The social desirability response bias in ethics research. In A. C. Michalos & D. C. Poff (Eds.), *Citation classics from the Journal of Business Ethics: Celebrating the first thirty years of publication* (pp. 173–190). New York: Springer.

Rao, H., MacDonald, J., & Crawford, D. (2011). The Dodd-Frank wall street reform and consumer protection act of 2010. *CPA Journal, 81*(3), 15–25.

Reidenbach, R. E., & Robin, D. P. (2013). Toward the development of a multidimensional scale for improving evaluations of business ethics. In A. C. Michalos & D. C. Poff (Eds.), *Citation classics from the Journal of Business Ethics: Celebrating the first thirty years of publication* (pp. 45–67). New York: Springer.

Rest, J. R. (1986). *Moral development: Advances in research and theory*. New York: Praeger.

Richter, N. F., Cepeda, G., Roldan, J. L., & Ringle, C. M. (2016). European management research using partial least squares structural equation modeling (PLS-SEM). *European Management Journal, 34*(6), 589–597.

Rigdon, E. E. (2016). Choosing PLS path modeling as analytical method in European management research: A realist perspective. *European Management Journal, 34*(6), 598–605.

Ringle, C. M., & Sarstedt, M. (2016). Gain more insight from your PLS-SEM results: The importance-performance map analysis. *Industrial Management & Data Systems, 116*(9), 1865–1886.

Ringle, C. M., Wende, S., & Becker, J.-M. (2015). *SmartPLS 3*. Boenningstedet: SmartPLS GmbH. Retrieved from http://www.smartpls.com.

Rottig, D., Koufteros, X., & Umphress, E. (2011). Formal infrastructure and ethical decision making: An empirical investigation and implications for supply management. *Decision Sciences, 42*(1), 163–204.

Sarstedt, M., Henseler, J., & Ringle, C. M. (2011). Multi-group analysis in partial least squares (PLS) path modeling: Alternative methods and empirical results. In M. Sarstedt, M. Schwaiger, & C. R. Taylor (Eds.), *Advances in international marketing* (Vol. 22, pp. 195–218). Bingley: Emerald. doi:10.1108/S1474-7979(2011)0000022012.

Schlittgen, R., Ringle, C. M., Sarstedt, M., & Becker, J.-M. (2016). Segmentation of PLS path models by iterative reweighted regressions. *Journal of Business Research, 69*(10), 4583–4592.

Schultz, J. J., Johnson, D. A., Morris, D., & Dyrnes, S. (1993). An investigation of the reporting of questionable acts in an international setting. *Journal of Accounting Research, 31*(3), 75–103.

Schwartz, M. S. (2015). Ethical decision-making theory: An integrated approach. *Journal of Business Ethics*. doi:10.1007/s10551-015-2886-8.

Shafer, W. E., Morris, R. E., & Ketchand, A. A. (2001). Effects of personal values on auditors' ethical decisions. *Accounting, Auditing & Accountability Journal, 14*(3), 254–277.

Shawver, T. J., Clements, L. H., & Sennetti, J. T. (2015). How does moral intensity impact the moral judgments and whistleblowing intentions of professional accountants? In C. Jeffrey (Ed.), *Research on professional responsibility and ethics in accounting* (Vol. 19, pp. 27–60). Bingley: Emerald Group.

Singer, M., Mitchell, S., & Turner, J. (1998). Consideration of moral intensity in ethicality judgements: Its relationship with whistle-blowing and need-for-cognition. *Journal of Business Ethics, 17*(5), 527–541.

Singh, J. J., Garg, N., Govind, R., & Vitell, S. J. (2016). Anger strays, fear refrains: The differential effect of negative emotions on consumers' ethical judgments. *Journal of Business Ethics*. doi:10.1007/s10551-016-3248-x.

Singhapakdi, A., Vitell, S. J., & Kraft, K. L. (1996). Moral intensity and ethical decision-making of marketing professionals. *Journal of Business Research, 36*, 245–255.

Solovida, G. T., & Latan, H. (2017). Linking environmental strategy to environmental performance: Mediation role of environmental management accounting. *Sustainability Accounting, Management and Policy Journal* (fortcoming).

Streukens, S., Leroi-Werelds, S., & Willems, K. (2018). Dealing with non-linearity in importance performance-map analysis (IPMA): An integrative framework in a PLS-SEM context. In H. Latan & R. Noonan (Eds.), *Partial least path modeling: Basic concepts, methodological issues and applications*. New York: Springer.

Trevino, L. K. (1986). Ethical decision making in organizations: A person-situation interactionist model. *The Academy of Management Review, 11*(3), 601–617.

Valentine, S., & Fleischman, G. (2004). Ethical reasoning in an equitable relief innocent spouse context. *Journal of Business Ethics, 45*(4), 325–339.

Valentine, S., & Hollingworth, D. (2012). Moral intensity, issue importance, and ethical reasoning in operations situations. *Journal of Business Ethics, 108*(4), 509–523.

Vandekerckhove, W., & Lewis, D. (2012). The content of whistleblowing procedures: A critical review of recent official guidelines. *Journal of Business Ethics, 108*(2), 253–264.

Webber, S., & Archambeault, D. S. (2015). Whistleblowing: Not so simple for accountants. *CPA Journal, 85*(8), 62–68.

Yu, Y.-M. (2015). Comparative analysis of Jones' and Kelley's ethical decision-making models. *Journal of Business Ethics, 130*, 573–583.

Zhang, J., Chiu, R., & Wei, L. (2009). Decision-making process of internal whistleblowing behavior in China: Empirical evidence and implications. *Journal of Business Ethics, 88*, 25–41.

Chapter 16
Latent Variable Regression for Laboratory Hyperspectral Images

Paul Geladi, Hans Grahn, and Kim H. Esbensen

Abstract This chapter is about the application of latent variable-based regression methods on hyperspectral images. It is an applied chapter, and no new PLS algorithms are presented. The emphasis is on visual diagnostics and interpretation by showing how these work for the examples given. Section 16.1 of this chapter introduces the basic concepts of multivariate regression and of multivariate and hyperspectral images. In Sect. 16.2 the hyperspectral imaging technique used and the two examples (cheese and textile) are explained. Also some sampling issues are discussed here. Principal component analysis (PCA) is a powerful latent variable-based tool for cleaning images. Section 16.3 describes PLS quantitative model building and diagnostics, both numerical and visual for the cheese example, and finishes with PLSDA qualitative modeling for the textile example.

16.1 Multivariate and Hyperspectral Images and Their Relation to Regression

16.1.1 Multivariate Regression

It is good to start with some definitions and nomenclature. Multivariate regression in its simplest format can be written as:

$$\mathbf{y}_{\mathrm{cal}} = \mathbf{X}_{\mathrm{cal}}\mathbf{b} + \mathbf{f}_{\mathrm{cal}}{}^{*} \quad \text{or} \quad \mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{f} \tag{16.1}$$

P. Geladi (✉)
SBT, SLU, Umeå, Sweden
e-mail: paul.geladi@slu.se

H. Grahn
Örebro University, Örebro, Sweden
e-mail: hans@corpus-datamining.com

K.H. Esbensen
KHE Consult, Copenhagen, Denmark
e-mail: khe.consult@gmail.com

H. Latan, R. Noonan (eds.), *Partial Least Squares Path Modeling*,
DOI 10.1007/978-3-319-64069-3_16

$\mathbf{y}_{cal}$: a vector** $I \times 1$ of 1 variable measured on I objects "dependent variable" used for calibration
$\mathbf{X}_{cal}$: a data matrix ($I \times K$) of K variables measured on I objects (calibration samples) "independent variables or predictors" used for calibration
$\mathbf{b}$: a vector of regression coefficients ($K \times 1$)
$\mathbf{f}_{cal}$: a vector of residuals ($I \times 1$)

$\mathbf{X}_{cal}\mathbf{b}$ can also be called $\mathbf{y}_{hat}$. It is the part of $\mathbf{y}_{cal}$ that can be modeled.

**The index cal helps in differentiating calibration and test set.*
***All vectors are column vectors; if row vectors are needed, column vectors can be transposed, and this is indicated by the transpose operator superscript T* (see Eq. 16.5).

Equation (16.1) can also be written for more than one y.

$$\mathbf{Y}_{cal} = \mathbf{X}_{cal}\mathbf{B} + \mathbf{F}_{cal} \tag{16.2}$$

$\mathbf{Y}_{cal}$: a matrix $I \times J$ of "dependent" variables
$\mathbf{B}$: a matrix $K \times J$ of regression coefficients
$\mathbf{F}_{cal}$: a residual matrix of size $I \times J$

This equation is not used in the remainder of the text because of space limitations.

The K variables in $\mathbf{X}_{cal}$ can, e.g., be from a spectroscopic technique, and the variables in $\mathbf{y}_{cal}$ could typically be chemical composition. The solution for obtaining $\mathbf{b}$ can be MLR (OLS), a latent variable solution (like PLS), or some other regression technique. An example would be protein content (in a food commodity material) in $\mathbf{y}_{cal}$ and NIR spectra (with many hundreds of wavelengths responses) for the calibration materials in $\mathbf{X}_{cal}$. When a useful $\mathbf{b}$ vector is found, it can be used for calculating or predicting protein contents from new NIR spectra according to:

$$\mathbf{y}_{pred} = \mathbf{X}_{test}\mathbf{b} \tag{16.3}$$

This is done for a test set that contains different objects from the same population of the calibration set in Eq. (16.1).

$\mathbf{y}_{pred}$: ($J \times 1$) predicted values for test objects
$\mathbf{X}_{test}$: ($J \times K$) spectra measured for new objects (test samples)

$\mathbf{f}_{test}$ has to be defined:

$$\mathbf{f}_{test} = \mathbf{y}_{test} - \mathbf{y}_{pred} \tag{16.4}$$

$\mathbf{y}_{test}$: ($J \times 1$) known values for the objects in the test set

The very welcome property of such a solution is that inexpensive and quick NIR spectral measurements, $\mathbf{X}$, can *replace* slow and expensive (and not always environmentally friendly) protein determinations, $\mathbf{y}$. Finding a useful $\mathbf{b}$ vector

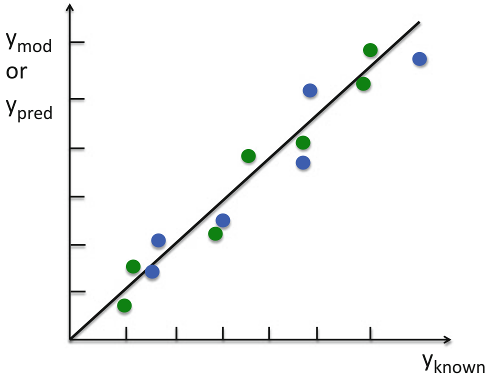


Fig. 16.1 Generic relationship between $\mathbf{y}$ and $\mathbf{y}_{hat}$. *Green points* for calibration model and *blue points* for test pred

has shown to be a very tricky undertaking, especially if $K > I$, and this is why latent variable regression methods such a PLS were introduced. The chemometrics literature from the late 1980s onwards contains many descriptions of how PLS regression is applied in such situations (Martens and Næs 1989; Næs et al. 2002; Beebe et al. 1998; Brereton 2003; Varmuza and Filzmoser 2009; Brown 1993). The PLS algorithms used in this chapter are the same as those explained in the references.

Figure 16.1 shows some properties of Eqs. (16.1) and (16.3).

The horizontal axis gives the values of the dependent variables $\mathbf{y}_{cal}$ and $\mathbf{y}_{test}$ in some relevant measurement unit. The vertical axis shows $\mathbf{y}_{hat}$ and $\mathbf{y}_{pred}$ in the same units. The green points represent the data (values) used to calibrate and to find $\mathbf{b}$ as in Eq. (16.1). The blue points are the test data (values) as in Eq. (16.3). Ideally both green and blue points should be close to or on the black diagonal (slope 45 degrees). One also wants the green data to have a realistic spread with respect to the future prediction scenarios, and the blue points should have a similar spread. With real-world samples/analytical measurements, there will always be some uncertainty, and the points never fall exactly on the diagonal line. This is due to sampling, preparation, and analytical errors, among other things.

16.1.2 Images, Multivariate, and Hyperspectral Images

An image representing the surface of a target material is organized as an array of $L \times M$, where a measured intensity is registered for each pixel. Photographic images are monochrome (an average intensity over all colors) or RGB (intensities for red, green, and blue). Color images are thus arrays of fixed size $L \times M \times 3$ (see Fig. 16.2).

While standard RGB color images are very useful for scientific documentation, and sometimes for quantitative measurements as well (but restricted to three intensities), new methods of constructing images have also emerged.

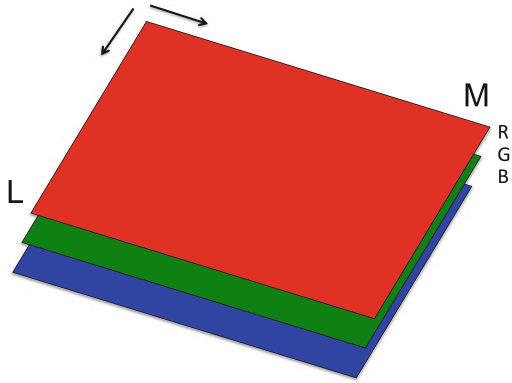


Fig. 16.2 An RGB image consisting of three planes

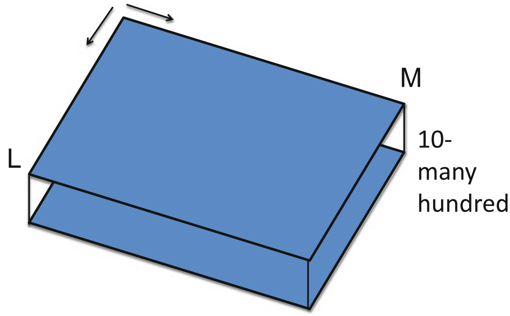


Fig. 16.3 A multivariate image (10 variables)

This leads to the concept of the *multivariate image*, size ($L \times M \times K$) where K is now a number of variables, often significantly larger than 3. From airborne imaging comes the term hyperspectral, meaning that K is usually over 100 and possibly in the range of thousands, which Fig. 16.3 illustrates.

In a hyperspectral image, each pixel position is now assigned a complete spectrum vector; see Fig. 16.4 for a graphical presentation.

Good review descriptions are given in the literature (Goetz and Curtiss 1996; Roggo et al. 2005; Winson et al. 1997; Fernandez-Pierna et al. 2006; Geladi et al. 2010; Grahn and Geladi 2007; Cloutis 1996; Ghiyamat and Shafri 2010; Van der Meer 2012; Van der Meer et al. 2012) not the least in the remote sensing community.

Multivariate and hyperspectral images have a 2-D spatial extension and a spectral dimension. They occur in a number of measurement situations. Very popular are satellite and clinical imaging, but also laboratory, field, and industrial applications are emerging. Imaged scenes or objects occur in all sizes from astronomical down to atomic. A large number of physical principles are used to acquire these types of imagery. Many of the physical imaging principles give multivariate information per pixel. A typical example in the optical region would be images of size 2050 × 2050 with 19 wavelength bands or images in the near infrared of 384 × 500 with 288 wavelengths.

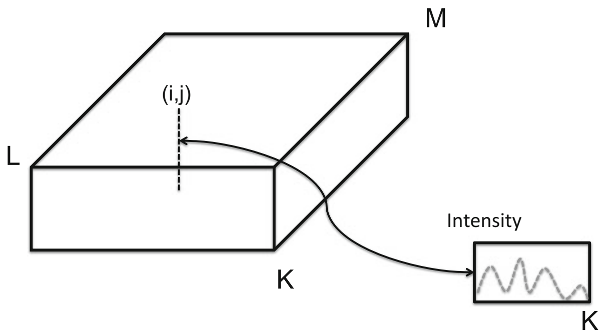


Fig. 16.4 A hyperspectral image of *K* variables, *K* is usually >100. Every pixel (coordinates *i*,*j*) is assigned with a complete intensity spectrum of *K* variables

Among the spectral methods leading to hyperspectral imaging are ultraviolet transmission, ultraviolet fluorescence, visual transmission, reflection or fluorescence, near-infrared transmission or reflectance, infrared transmission or reflectance, Raman reflectance or transmission, and infrared emission. Also acoustic, magnetic resonance, electron, and ion microscopy methods in all their modalities are able to produce images having many variables per pixel (Geladi and Grahn 1996; Grahn et al. 2016; Li et al. 2016).

16.2 Multivariate Analysis for Hyperspectral Images

16.2.1 The Example and Image Recording

The easiest way to explain the key principles of analyzing multivariate and hyperspectral images is by a master example: 15 cheeses with different composition regarding fat, protein, and energy content are used for this purpose. The cheeses were bought in the supermarket. From each cheese a small 20 mm × 20 mm × 20 mm bit was cut out to make a mosaic with a paper plate as background. This mosaic was made in three *replicates*. Figure 16.5 shows the layout of the experimental design.

In Table 16.1, it is shown the composition of all these calibration cheeses. These data are read off from the cheese packages and do not represent actual analysis but very likely close enough for the present demonstration purposes. Figure 16.5 is just a standard RGB color image. However, the visual color of the cheeses has only a very small role with respect to the compositional contents, and therefore scanning in the near-infrared region becomes more useful regarding quantitative prediction purposes.

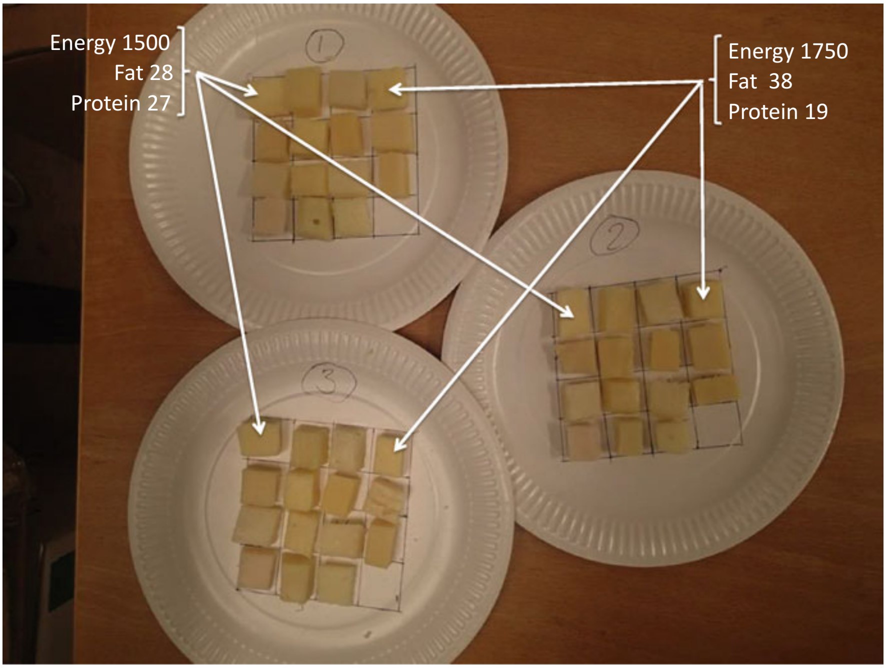


Fig. 16.5 Experimental layout of 15 cheese cubes, in three replicates. Counting on each plate is 1–4 *upper left* to *right*, then 5–8 one row *lower*, 9–12 one row *below*, and 13–15 last row. The numbers are explained in Table 16.1. Three bulk analysis results are shown as examples

Table 16.1 ID numbers, names, and composition data of 15 calibration cheeses

ID number	Name	Energy (kJ/100 g)	Fat (g/100 g)	Protein (g/100 g)
1	Herrgård	1500	28	27
2	Västan	895	10	30
3	Fontana 6	1604	30	24.9
4	Boxholm	1750	38	19
5	Greve	1200	17	33
6	Munken	1782	38	21
7	Old Dutch Master	1811	36	28
8	Emmental	1605	30	25
9	Fontana 12	1315	30	25
10	Billinge	1142	17	29
11	FB QuesoIberico	1737	25.8	23.3
12	FB Gruyere	1660	32	27
13	FB Chevre	1819	37	26
14	FB Appenzeller	1690	32	25
15	FB Manchego	1820	37	25

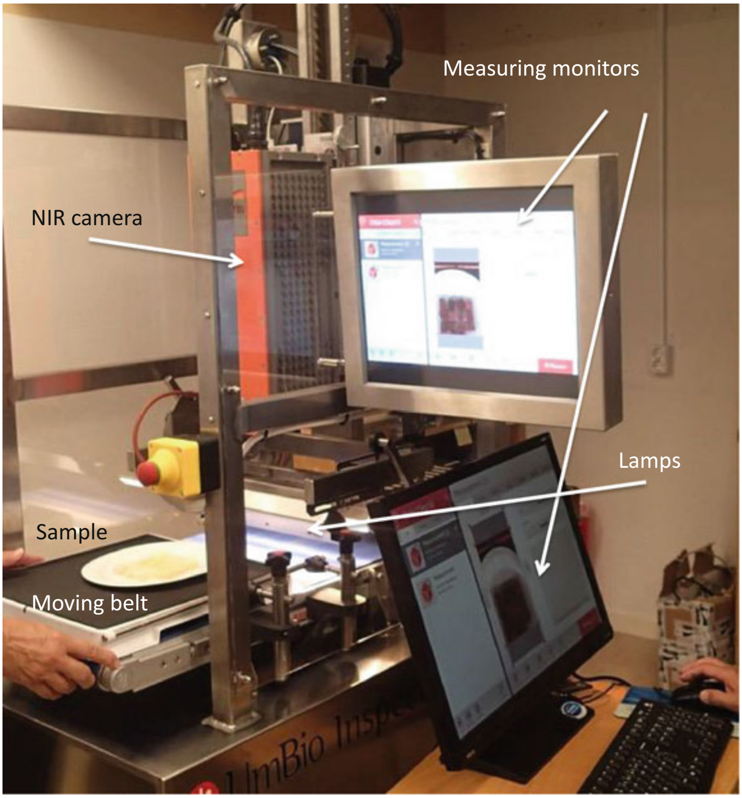


Fig. 16.6 The linescan equipment. Moving belt, linescan camera, screens, and keyboard

The paper plates and the cheeses were measured using a hyperspectral camera, a linescan Specim model SWIR 3.0 (Fig. 16.6).

The camera is based on an HgCdTe detector and a prism-grating-prism (PGP) monochromator. A 15 mm lens was used giving a line width of 150 mm. An exemplar image can be seen in Fig. 16.7.

The specific camera setup results in lines of width: 384 pixels, with 288 intensities (variables, K) for the different NIR wavelengths employed (nominal 970–2500 nm, 5.6 nm for each variable bin) in each pixel. By scanning the samples, lines are added to each other to make an elongated rectangular image. This is done by a synchronized belt transport. The resulting images were approximately 500 lines by 384 pixels per line and 288 wavelengths per pixel. For each image, a white reference sample (Spectralon) was also measured as was a dark current "image." These were used to calculate corrected absorbance values. Exactly the same setup has been used for the textile example in Sect. 16.3.4.

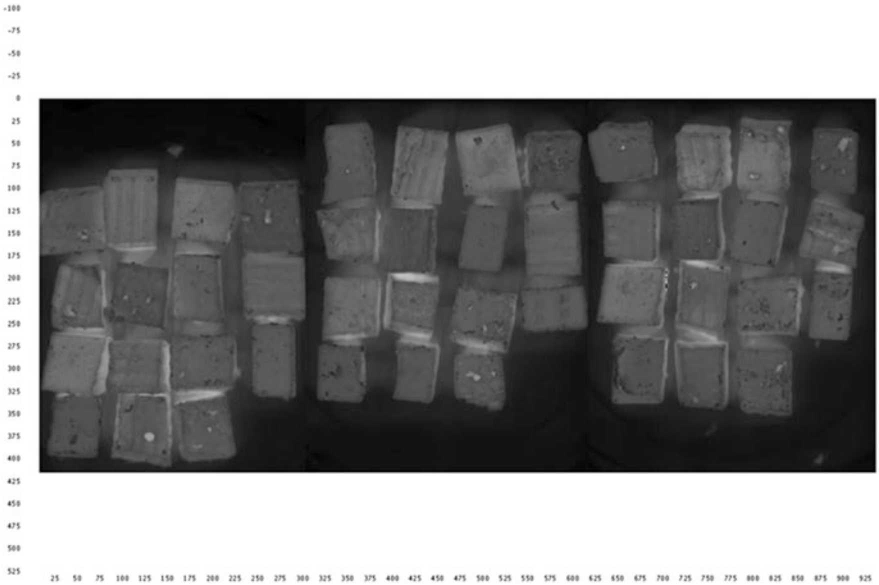


Fig. 16.7 The mosaic of the three replicate images at 1800 nm. The total mosaic is 415 lines and 929 pixels per line (385,535 pixels). A scale for size is given. Replicate 1 is *left*, replicate 2 is *middle*, and replicate 3 is *right*

16.2.2 Sampling Issues

The cheese cubes in Fig. 16.5 are cutouts from larger pieces of 50–300 g right off the shelves in the supermarket. These pieces are in their turn representative of a whole daily, weekly, or monthly production in a diary. It is of course important that all samples are representative of the target material, which can be many orders of magnitude larger than what is actually imaged. In a typical real-world industrial situation, the measured cubes could represent, for example, a daily or weekly production batch. For the present purpose, these issues are not fatal because the example is only used for a technical demonstration of the imaging and its regression modeling.

However there are a few sampling issues that still matter:

- Do replicate cubes cut from the same original sample of cheese have the same properties? Three replicate cubes were prepared to obtain some information in this matter.
- Are the cubes themselves homogeneous? Here imaging is a good solution, because it will be able to show some aspects of inhomogeneity. Not every pixel will have an identical spectrum, and there may be spatial patterns (2-D) to this inhomogeneity. Bulk analysis does not give information on this sampling aspect.

– Is multispectral imaging producing identical results for every pixel in a cube? Imaging suffers from various shading and edge effects, and these should be kept in control or, if detected, compensated for, or removed.
– Pixels are square and the objects imaged may have round edges. This always gives edge pixels containing information from both object and background.

Some of these issues are illustrated in the next section. Some discussion of sampling issues can be found in Esbensen and Geladi (2009) and in Esbensen and Julius (2009).

Finally there is the archetypical regression issue. Is there a composition value for every pixel in the image? And, of course there is not—it would be physically and economically impossible to do a chemical analysis for *each pixel* in the image. This means that for every cube of material, there is only *one* bulk composition available. In order to make a sensible regression model, a cube of material should then be represented by an average spectrum.

16.2.3 Image Cleaning and Exploratory Analysis

Figure 16.7 was prepared by removing the basic background regions. In the obtained image, it was also necessary to remove some more detailed background patches together with regions of shading or faulty illumination (specular reflection).This was done effectively by calculating a number of principal component images and by studying score and image plots by interactive brushing (see below).

Principal component analysis on images is carried out by first reorganizing ("unfolding or matricizing") them into a data matrix.

$$\mathbf{X} = \mathbf{T}\mathbf{P}^{\mathrm{T}} + \mathbf{E} \tag{16.5}$$

X: the unfolded matrix the image is of size $(L \times M) \times K$. PCA is usually done after variable-wise mean-centering.
T: the score matrix $(L \times M) \times A$ where A is the number of components calculated.
P: the loading matrix size $K \times A$.
E: the residual matrix size $(L \times M) \times K$.
T: transposition symbol.

A very important point needs to be mentioned here. PCA modeling shown in Eq. (16.5) is performed iteratively to identify and remove pixels that are not expected to conform to the PLS model in Eq. (16.1).Cleaning is almost never just a simple one-step procedure.

The $L \times M$ elements of each of the score vectors in **T** (Eq. 16.5) have pixel coordinates. This means that they can be used to construct an image; this may be called back projection. The result of back projection is A score images. The elements of **T** can also be plotted against each other to form score scatter plots, i.e., traditional score plots. In traditional score plots, each pixel is represented by

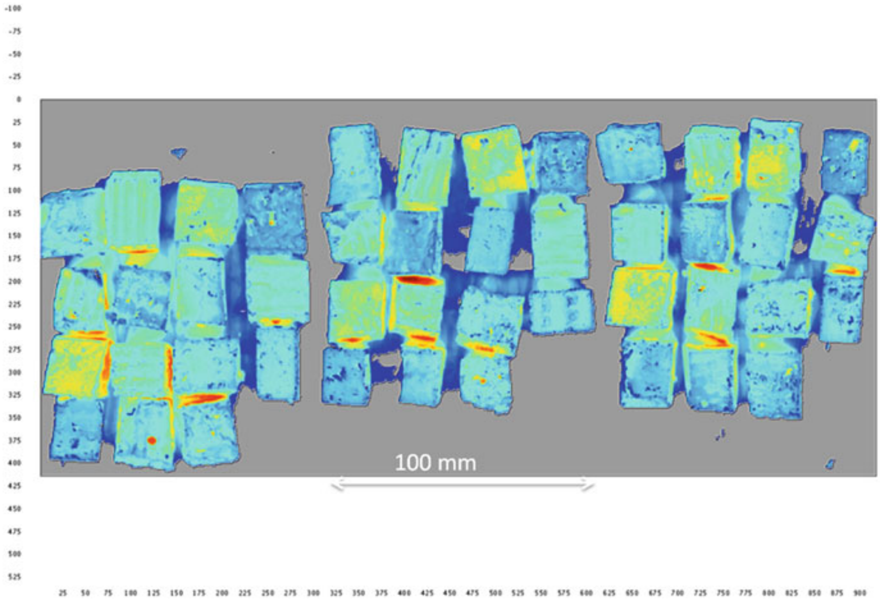


Fig. 16.8 Interactive cleaning of the image using PCA and brushing used to remove the cardboard background

a point. For the many thousands of pixels in images, the points would fall on top of each other and make the score plot unreadable. Therefore, score plots are shown as pixel density contours. Both score images and score plots can be used to find outliers, background, illumination errors, etc. that should be removed.

The first operation in the present example is removing the paper background. The result of this is seen in Fig. 16.8, which still retains shadows between the cheese bits as well as irregular reflections at the object edges, and these are removed by further cleaning in the same manner as before.

Figure 16.9 shows the final image result that will be used for regression calibration purposes.

In Fig. 16.10 it is shown how region of interest (ROI) selection can be done on the cheeses. The regions of interest are selected in such a way that edge errors are avoided. The regions of interest can then be used when calculating average spectra.

Figure 16.11 shows such a set of average spectra. Each spectrum now represents a specific cheese. The number of wavelengths was at the same time reduced from 288 to 276 to leave out some noisy wavelengths below 1000 nm, a mild form of variable selection. The range is 1000–2547 nm with one variable still supporting 5.6 nm.

Figure 16.12 shows very interesting properties of the cheeses in two calculated principal components: the first and fifth component score plot.

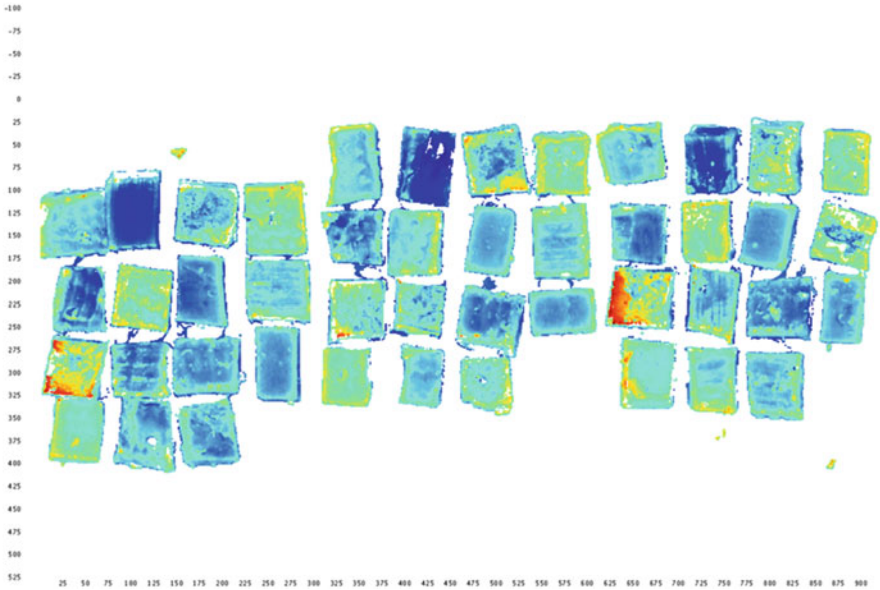


Fig. 16.9 The final cleaned image where only cheese is left. There are still imaging errors such as *shades* in holes and edge effects

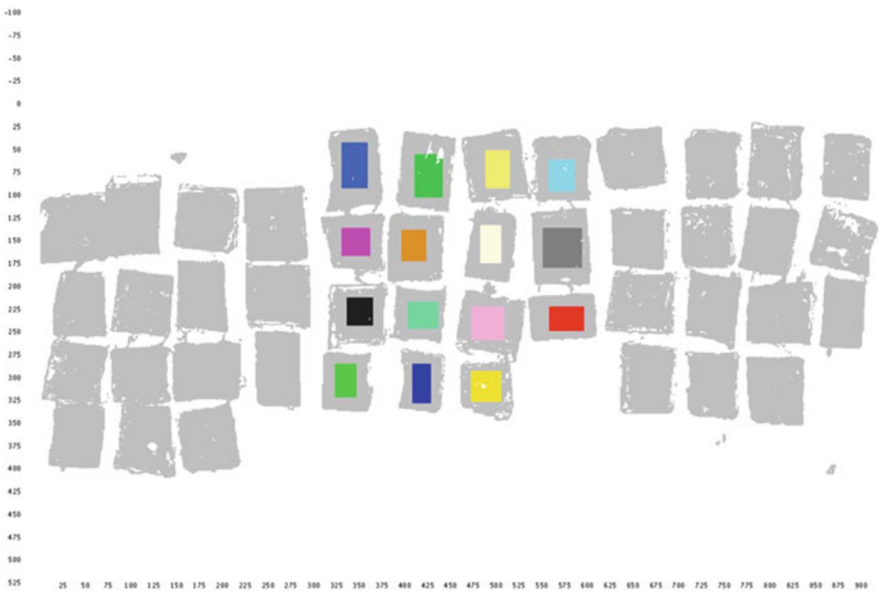


Fig. 16.10 Region of interest (ROI) selection in the cheeses to make average spectra for regression

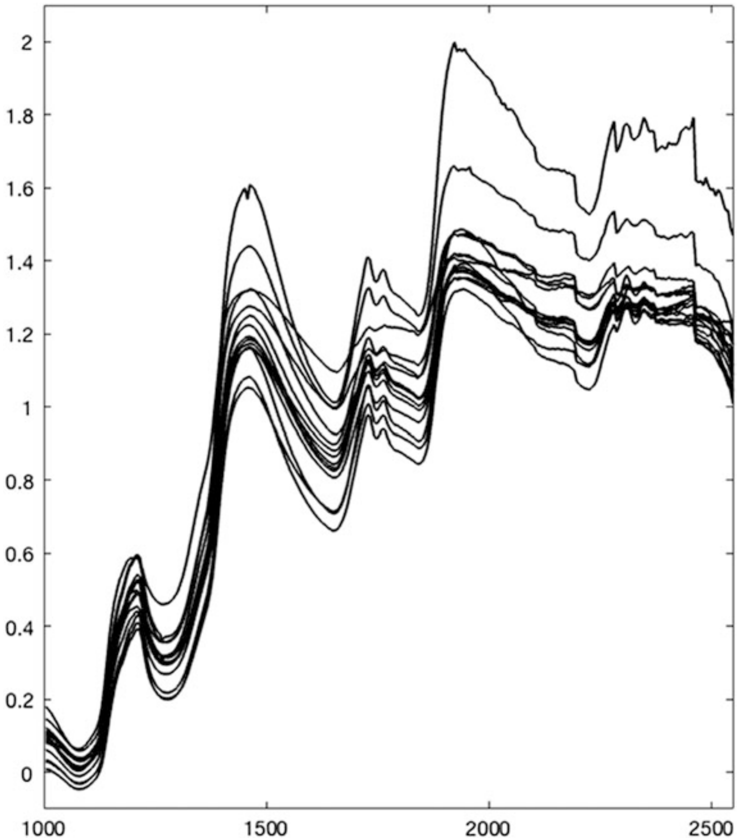


Fig. 16.11 A set of average spectra of 276 wavelengths representing the15 different cheeses

The color annotation shows that there are some differences between the cheeses, and heterogeneities within cheese blocks can also be seen. There are also still illumination effects at the edges. This illustrates that imaging in the near infrared is quite unique in revealing such properties, some of which are beneficial for analytical credibility, and others are method artifacts which must be controlled or eliminated. This also shows the importance of producing average spectra in the correct object regions to represent each cheese properly. Some reviews about hyperspectral imaging and analysis of the data have appeared in the literature in recent years: in agricultural and food industry, pharmaceutical, nanomaterials, and forensic traces (Grahn et al. 2016).

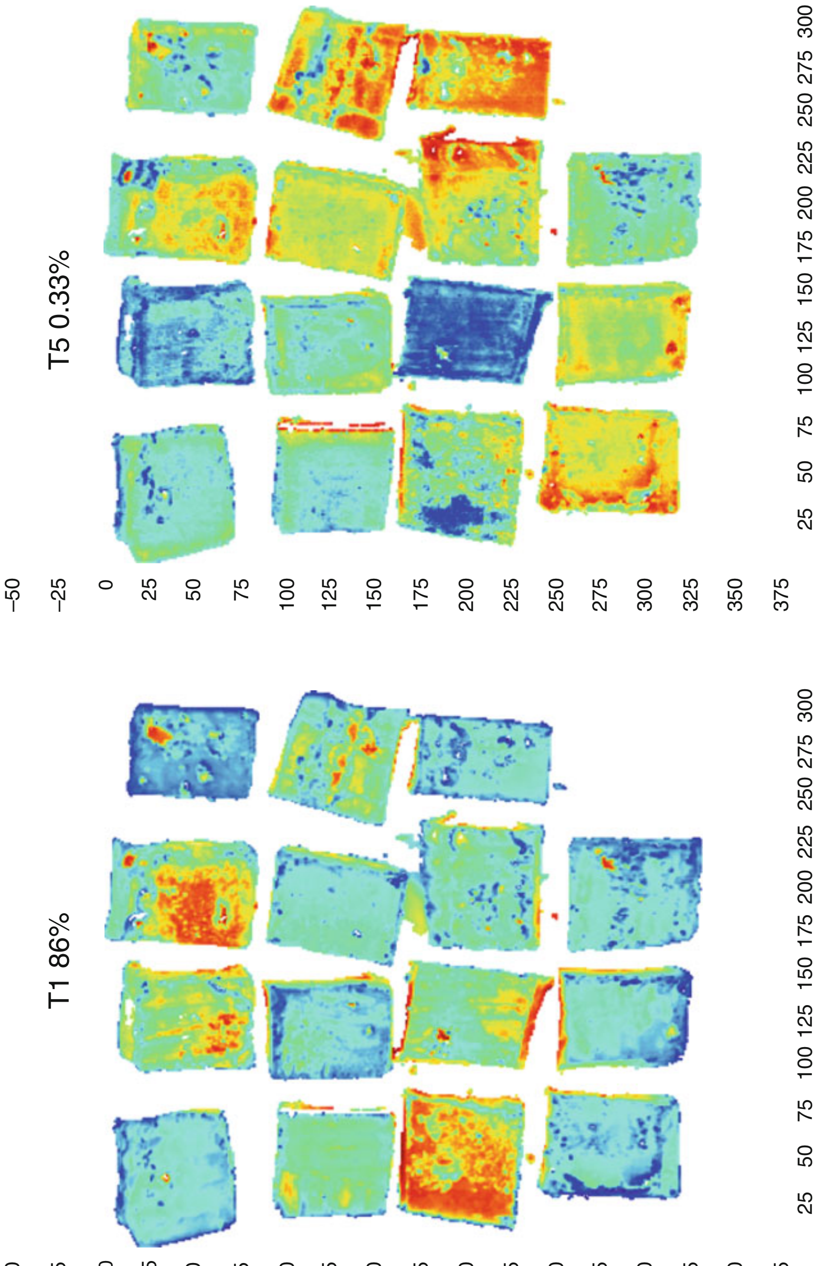


Fig. 16.12 Score images for T1 and T5 of one of the replicates. The color is score value

16.3 Latent Variable (PLS) Regression for Hyperspectral Images

16.3.1 PLS Regression Models Between Mean Spectra and Data

There are regions in the image where an average concentration would seem a reasonable target (see Fig. 16.13).

It is then possible to make a regression model between the *average* spectrum of such a region and the concentration. For mean spectral characterization, a spatial region in the middle of the cheese cuts was outlined to avoid peripheral shade and edge effects.

The three replicate hyperspectral images (Fig. 16.9) were all used to extract mean spectra. This gives 15 spectra $\times$ 3 replicates or 45 spectra. These were used as X variables in a PLS model with the composition (fat, protein) and energy content as y variables, respectively. Given enough data, at least two-thirds should be used for modeling. The resulting data sets were 45 $\times$ 276 for $\mathbf{X}$ and 45 $\times$ 3 for $\mathbf{Y}$. On this basis a split can easily be made into a calibration and a test set. In Fig. 16.9, the left and middle parts were used for extracting calibration vectors, and the right one was used for extracting test vectors.

This gives 30 $\times$ 276 and 30 $\times$ 3 calibration data sets ($\mathbf{X}_{\text{cal}}$ and $\mathbf{Y}_{\text{cal}}$) and 15 $\times$ 276 and 15 $\times$ 3 test sets ($\mathbf{X}_{\text{test}}$ and $\mathbf{Y}_{\text{test}}$). The test set is used to check whether the calibration really works by calculating diagnostics in a validation step. $\mathbf{Y}_{\text{cal}}$ and $\mathbf{Y}_{\text{test}}$ were made into three vectors each for separate modeling of energy, fat, and protein. This means that Eq. (16.1) was used, not Eq. (16.2).

Instead of using exactly one average for each cheese bit, it would have been possible to make a number of averages for each bit using different subsets. This trick is sometimes used to create more objects and therefore more degrees of freedom for the regression equation.

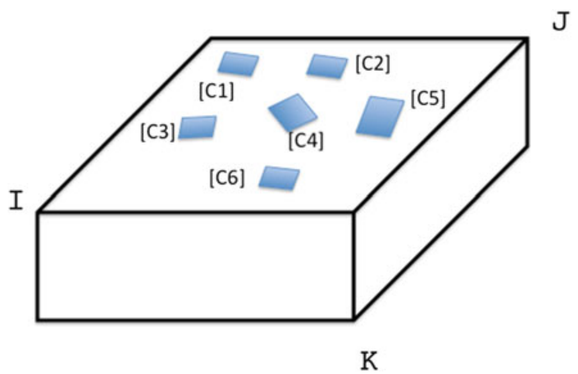


Fig. 16.13 For certain regions in the hyperspectral image, a bulk concentration is known

Table 16.2 % X variance and % y variance used by each component and cumulative

Component	X variance	X varcumul	y variance	y varcumul
1	85.12	85.12	37.16	37.16
2	12.18	97.3	21.06	58.21
3	0.74	98.04	31.03	89.24
4	0.90	98.94	2.98	92.23
5	0.73	99.67	1.22	93.45
6	0.17	99.84	1.61	95.06
7	0.04	99.88	1.58	96.64
8	0.01	99.89	1.65	98.29
9	0.06	99.95	0.29	98.59

The model was made for only one y variable: energy content, so $\mathbf{y}$ is a vector of 30 × 1, not a matrix

An intrinsic issue regarding any regression model is that the appropriate number of components in the prediction model to be used has to be determined. This is the critically important validation issue, which has been the topic of many discussions in the dedicated literature in statistics and data analysis in general and in chemometrics in particular. How can the future prediction performance be ascertained? Can this question be reliable answered based only on the existing calibration data set? Or is it necessary to have access to at least some data from said "future"? This critical issue can only be raised here … reference is made to the extensive literature which can be found starting with the key entries below:

Important issues are:

– How should a split in calibration and test set be done?
– How should test set validation be carried out?
– What diagnostic parameter needs to be optimized for getting the optimal model?

According to Table 16.2, at least 3–4 PLS components should be used to build an adequate model, but up to eight components *could* be used. Figure 16.14 shows this graphically.

In the case of energy content, it was decided to stay with four components. A typical error that is often made is over-fitting a PLS model. It is also clear that very small amounts of X variance are used for the higher components. Such components would be rather unreasonable to include in the predictive model.

16.3.2 Some Important PLS Regression Diagnostics

Table 16.3 shows the diagnostic results for all y variables as an overview. For explaining the diagnostics, Eqs. (16.1) and (16.3) can be repeated, remembering that for PLS models $\mathbf{X}$ and $\mathbf{y}$ are mean-centered. $\mathbf{y}_{test}$ represents the measured values for

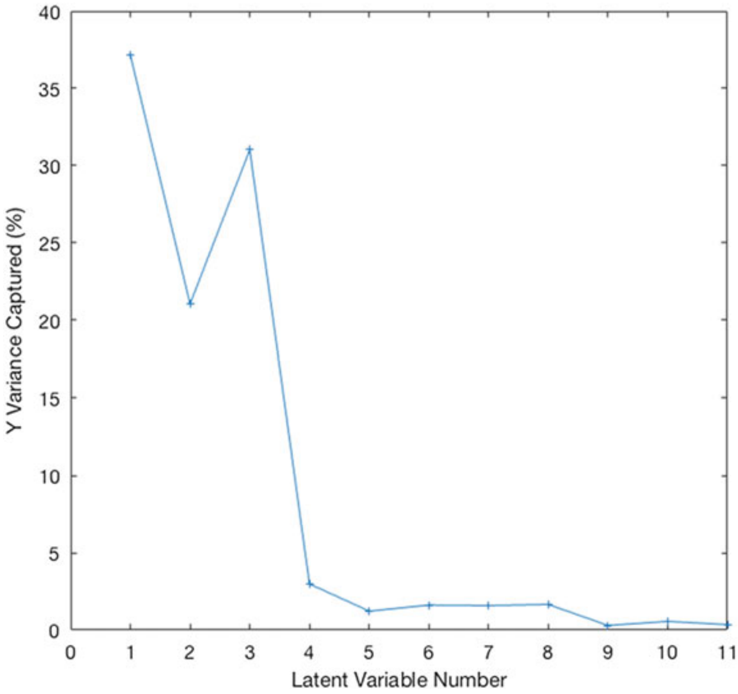


Fig. 16.14 Percentage of the total variance of **y** explained against component number

Table 16.3 PLS modeling and test results for the *y* variables: energy, fat, and protein

Name	# comp	R2cal	R2test	RMSEC	RMSEP	Bias	Range
Energy	4	0.92	0.93	77	76	−5.5	895–1820
Fat	4	0.89	0.89	2.7	2.9	0.76	10–38
Fat	7	0.93	0.91	2.2	2.6	−0.70	10–38
Protein	4	0.84	0.79	1.3	1.7	−0.44	19–33
Protein	7	0.95	0.85	0.76	1.5	−0.18	19–33

the *y*-variables in the test set.

$$\mathbf{y}_{\mathrm{cal}} = \mathbf{X}_{\mathrm{cal}}\mathbf{b} + \mathbf{f}_{\mathrm{cal}} \tag{16.1}$$

$$\mathbf{y}_{\mathrm{pred}} = \mathbf{X}_{\mathrm{test}}\mathbf{b} \tag{16.3}$$

The equations for the diagnostics are (SS means sum of squares):

$$\mathrm{R2cal} = 1 - \mathrm{SS}\left(\mathbf{f}_{\mathrm{cal}}\right)/\mathrm{SS}\left(\mathbf{y}_{\mathrm{cal}}\right) \tag{16.6}$$

$$\mathrm{R2test} = 1 - \mathrm{SS}\left(\mathbf{f}_{\mathrm{test}}\right)/\mathrm{SS}\left(\mathbf{y}_{\mathrm{test}}\right) \tag{16.7}$$

R2cal and R2test are similar to variance ratios. They can vary between 0 and 1. 1 means a perfect model (zero residual). 0.5 means that the modeled part and the residual part are equal, so this means no model at all can be made. Usually, an R2cal of above 0.9 and R2test of above 0.85 is necessary for a useful PLS prediction to be possible.

RMSEC is the root mean squared error of calibration, and RMSEP is the root mean squared error of prediction. They are given by:

$$\mathrm{RMSEC} = \left[\mathbf{f}_{\mathrm{cal}}{}^{\mathrm{T}}\mathbf{f}_{\mathrm{cal}}(\mathrm{df})^{-1}\right]^{0.5} \tag{16.8}$$

$$\mathrm{RMSEP} = \left[\mathbf{f}_{\mathrm{test}}{}^{\mathrm{T}}\mathbf{f}_{\mathrm{test}}J^{-1}\right]^{0.5} \tag{16.9}$$

where df is some number of degrees of freedom, usually number of calibration objects I minus number of components used. These numbers are expected to be small compared to the range of the data.

The bias is given as:

$$\mathrm{bias} = \mathbf{1}^{\mathrm{T}}\mathbf{f}_{\mathrm{test}}J^{-1} \tag{16.10}$$

where J is the number of test objects. The bias should be much less than RMSEP, ideally 0.

For interpreting Table 16.3 this means:

Energy: R2cal and R2test are above 0.9 with only four PLS components, which is very good. The bias is only 7% of the RMSEP. The RMSEP is 12 times smaller than the range, and that is also good.

Fat: R2cal and R2test are above 0.9 with seven components. The bias is 27% of the RMSEP, and this is clearly open for improvement. The RMSEP is ten times smaller than the range. With only four components, the model does not become much worse.

Protein: R2cal is high, but R2test is only 0.85 with seven components. The bias is 12% of the RMSEP, and the RMSEP fits nine times in the range. With only four components, the model does not become too much worse. For both fat and protein, there are definitely more calibration samples needed in order to define the most useful number of PLS components.

As a general conclusion, PLS calibrations with only four components for energy work rather well, PLS calibrations for fat and protein may need more PLS components, but then more calibration objects than the present complement will have to be used. One should also be aware of the fact that no real analyses were carried out; only the label information was used; this must certainly be considered

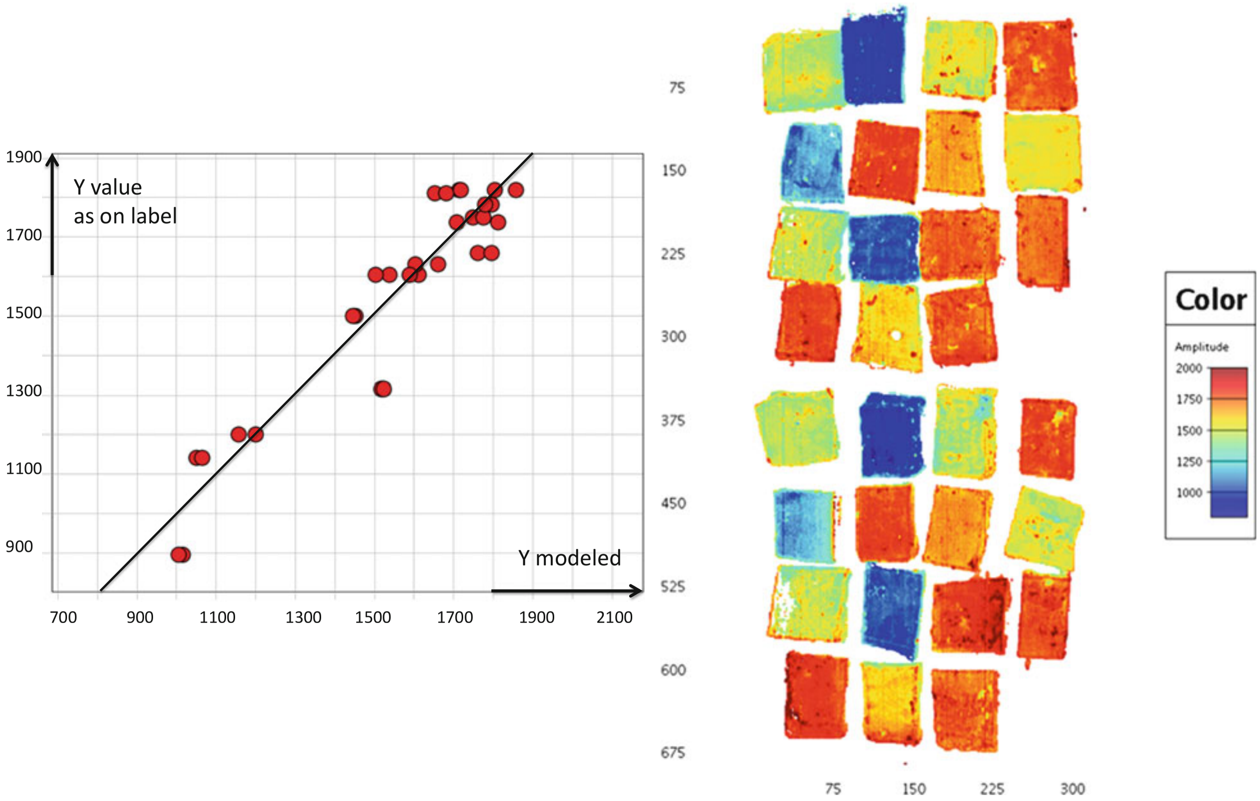


Fig. 16.15 *Left*: calibration plot for average spectra for energy (four PLS components). Vertical axis: energy read on the label. Horizontal axis: predicted energy from PLS. *Right*: PLS prediction showing $\mathbf{y}_{\text{hat}}$ for energy (four PLS components) for all pixels

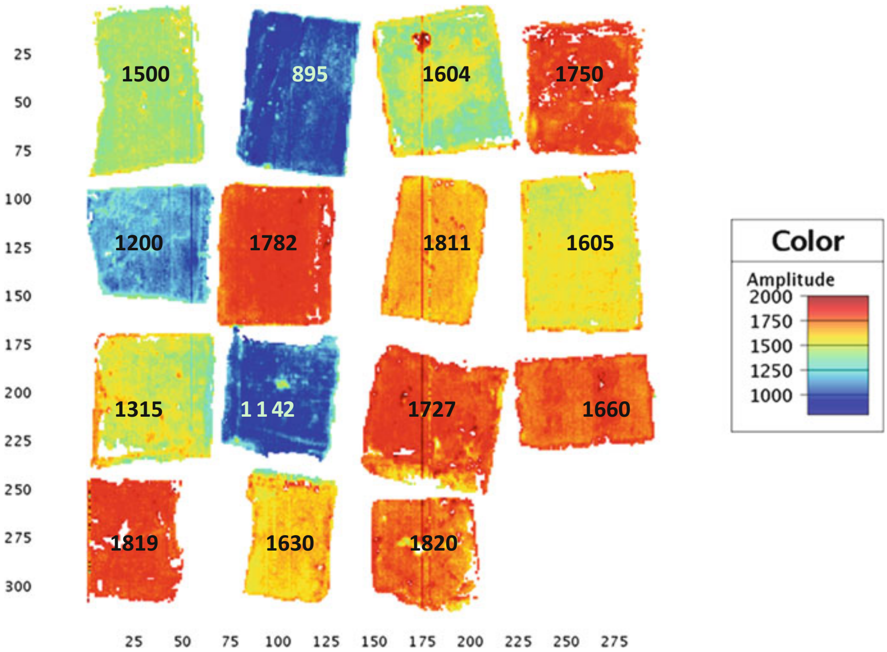


Fig. 16.16 Prediction for the test set (replicate 3, energy, four PLS components) with average energy values indicated

Fig. 16.17 Textile materials (socks) of different composition measured for the example

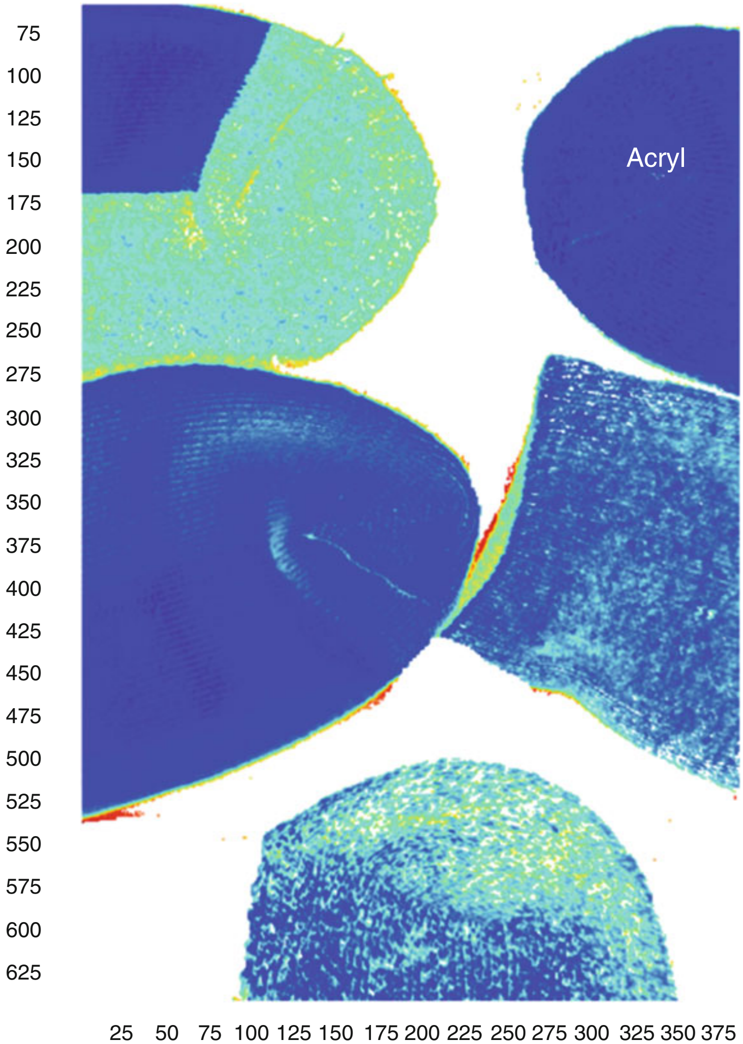


Fig. 16.18 The first PCA score component of the "textile image." The acryl material is indicated

as an extra source of error between *X* and *y*. Also the design of the test set is open to criticism. The test set used here is probably not the ideal one, and many other choices could have been made. But the present choice serves well for illustrating the principles behind validation and the resulting diagnostics. See Esbensen and Lied (2007), Esbensen and Geladi (2010), and Esbensen (2012).

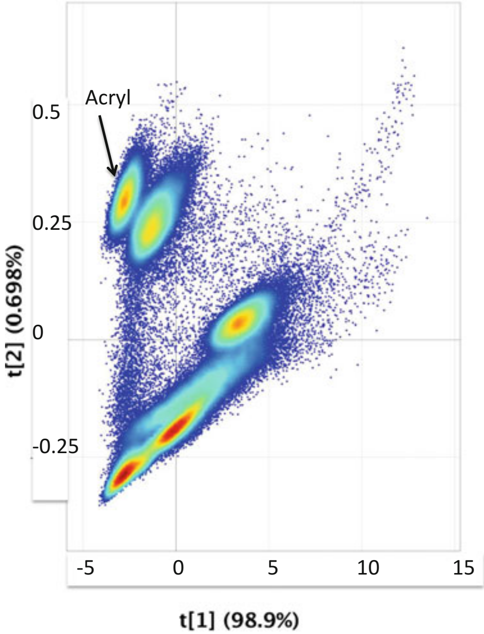


Fig. 16.19 The score plot of the textile image for components 1 and 2. The cluster for acryl pixels is indicated

16.3.3 PLS Prediction Visualized in Images

A calibration image was made using replicates 1 and 2 (2 × 15 cheeses) as shown in Figs. 16.7–16.9.For this image, average spectra were calculated, and regression was carried out against the energy values in Table 16.1.

Image preprocessing was by dividing each spectrum by its standard deviation and removing the mean value. This often corrects many optical errors. Mean centering of $\mathbf{X}_{\text{cal}}$ and $\mathbf{y}_{\text{cal}}$ was also used. A four component PLS model was found to be good enough.

Figure 16.15 left shows the PLS calculated y-values $\mathbf{y}_{\text{hat}}$ against the ones from the cheese packages $\mathbf{y}_{\text{cal}}$. The linear relation is reasonable good. The PLS regression coefficients were then used for calculating energy values for every pixel in the image. The result is shown in Fig. 16.15 right panel. The energies are color coded. One may notice a good correspondence between the colors for replicate 1 and that for replicate 2. The PLS model was also used for predicting the energy values for replicate image 3 for each pixel. The result is given in Fig. 16.16, and the average energy $\mathbf{y}_{\text{test}}$ values are also given in the figure.

It is easily seen by the color coding that average energy values are well predicted. One may also observe that the cheeses are not homogeneous, because color contrasts

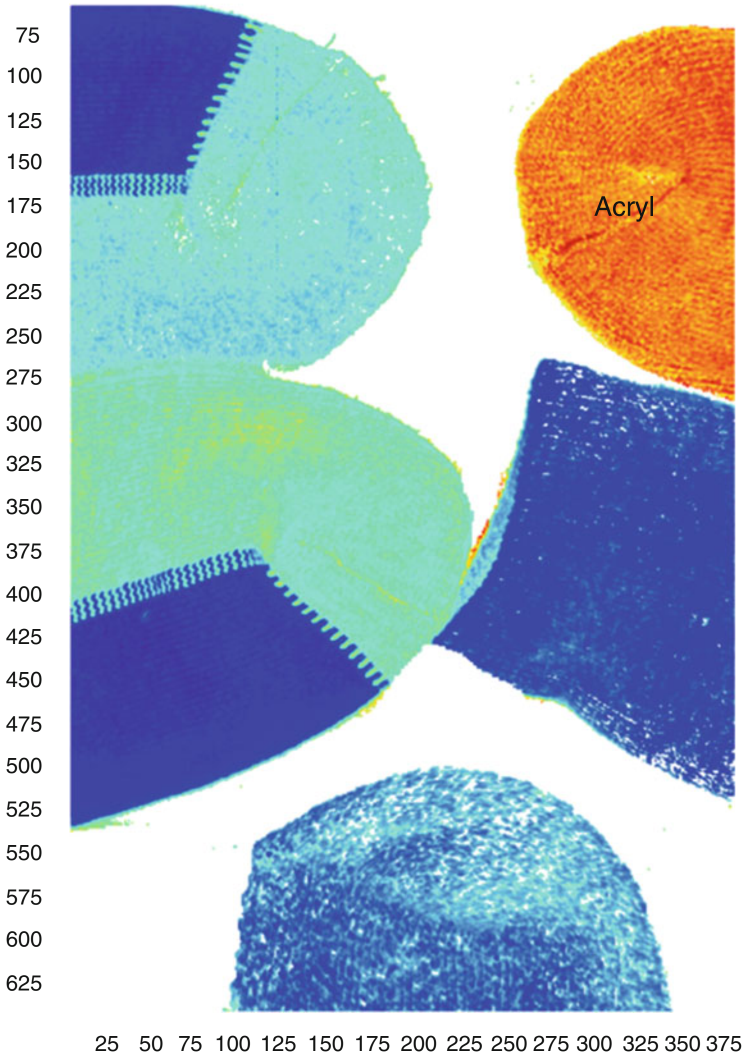


Fig. 16.20 The first score image of the PLSDA model

are seen within each cheese. For some of the cheeses, this is worse than for others. This is the real advantage of using regression on images. All interpretation can be visualized where statistical diagnostics only give a number calculated by averaging over many pixels.

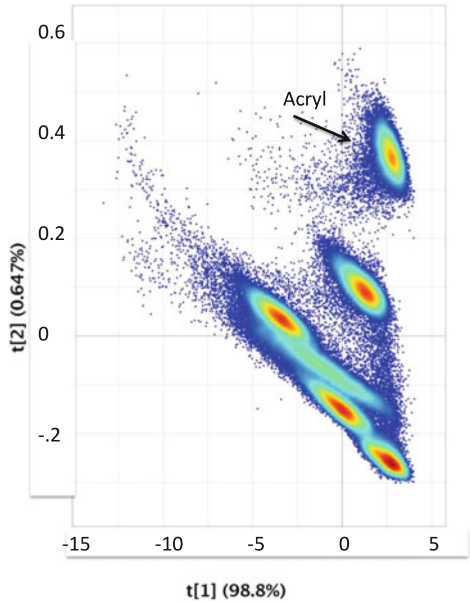


Fig. 16.21 The score plot of first and second score of the PLSDA model

16.3.4 PLS Discriminant Analysis PLSDA

PLSDA uses the Eqs. (16.1) and (16.3), but instead of a continuous response variable **y**, there is a categorical 0–1 variable, e.g., indicating composition or class membership. The response variable **y** only contains ones and zeroes.

The PLSDA model is constructed just as a usual PLS model with all diagnostics and problems of how many components are needed and how a calibration and test set can be constructed. This was described in Sect. 16.3.2. Once a good PLSDA model is found, it is possible to use it to predict class membership or whatever was used as the response variable **y**.

For explaining PLSDA and its advantages, textile composition was used. Figure 16.17 gives a number of socks and their composition as can be read on the packaging.

One thing about textiles is that the color has very little to do with the composition. Color is often determined by a minute amount of textile dye; see Fig. 16.17. The hyperspectral image was made as described in Sect. 16.2.2, and only the wavelength range 2001–2493 nm (89 variables) was used.

The only preprocessing was mean-centering. Figure 16.18 gives a PCA result after cleaning for background and error pixels. The remaining pixels are 159,558.

The scatter plot of scores 1 and 2 is given in Fig. 16.19. The cluster for the acryl material is not well separated from other clusters.

For doing PLSDA, a new variable with ones and zeros is created: one for not acryl and zero for acryl. Using this new *y* variable, a PLSDA model can be made. The model explains 97% of the *y*-variance with three PLS components.

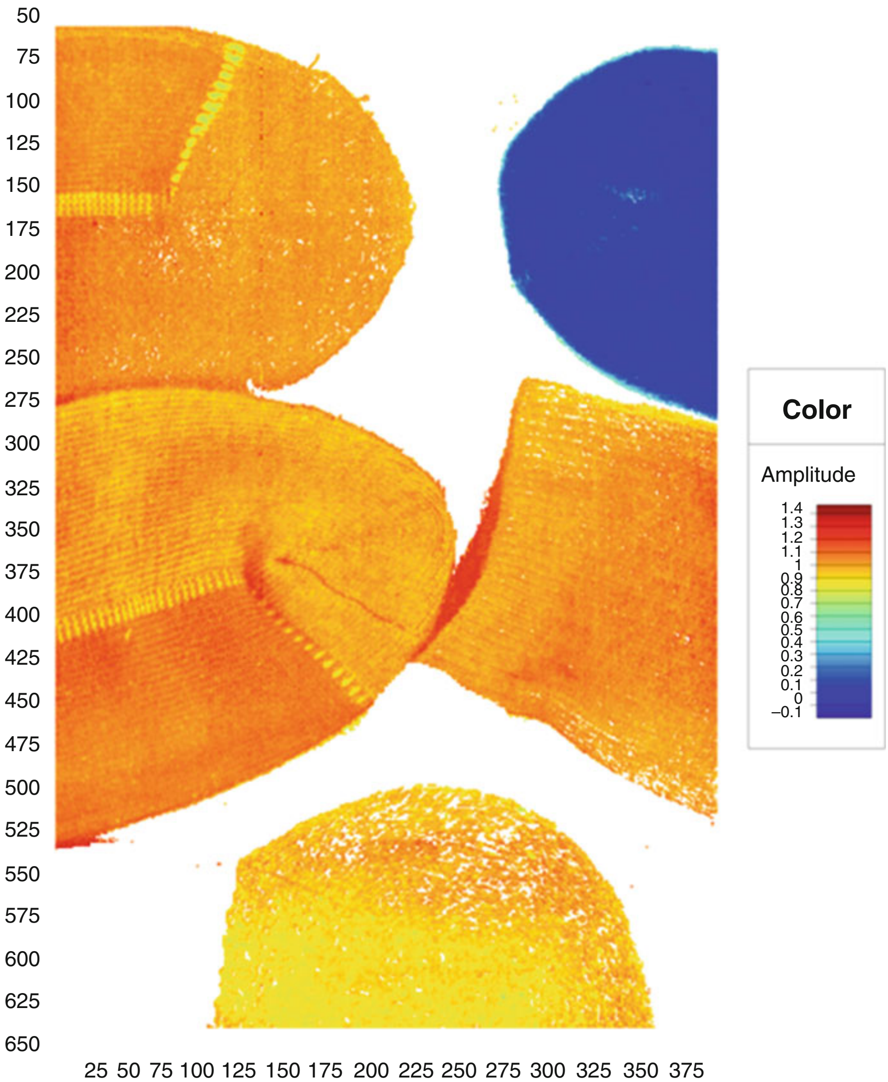


Fig. 16.22 The modeled responses of the PLSDA model in color coding [acryl = 0 (*blue*), non-acryl = 1 (*orange*)]

Figure 16.20 gives the first PLSDA component. It is clear that the acryl is quite different in color (PLSDA score) from the other materials. This is a huge improvement over Fig. 16.18. Figure 16.21 gives the score scatter plot of components 1 and 2.

Also here the acryl is better separated from the other materials than it was in Fig. 16.19. A conclusion is that PLSDA helps in making more meaningful clusters

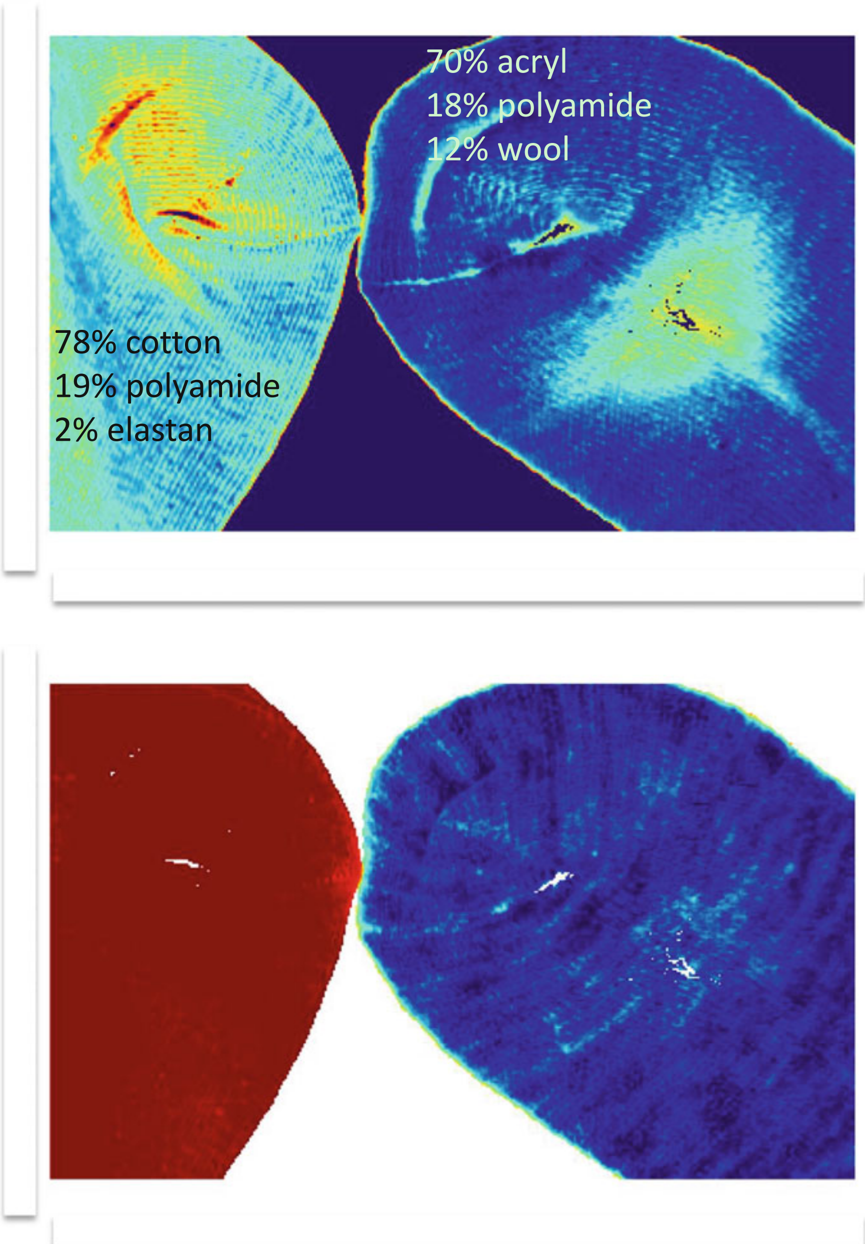


Fig. 16.23 The test set image at 2068 nm (*upper panel*) and the PLSDA prediction (*lower panel*). The *left* sock is mainly cotton; the *right* sock is mainly acryl

of pixels that PCA sometimes cannot make or separate. Figure 16.22 shows the PLSDA response $\mathbf{y}_{hat}$ as modeled in color code. Figures like this can be used to quickly visually locate the presence of acryl based textile materials.

A test set was made with two new socks (bought in a different store and different brands, measured on different days). The image was pretreated by taking away background pixels. The prediction image is given in Fig. 16.23 upper part, with the composition in the figure. Below in Fig. 16.23 is the PLSDA prediction $\mathbf{y}_{pred}$ using the model used for Fig. 16.22. The 70% acryl sock is predicted as high in acryl (blue), and the other sock is predicted as low in acryl (dark red).

A more thorough exposé (and methodological critique) of the pitfalls of ill-reflected PLSDA application can be found in Brereton and Lloyd (2014), greatly recommended.

In recent years a number of review articles on using regression models for hyperspectral imaging have been published, mainly for food processing (Chen et al. 2016; Pang et al. 2016; Gowen et al. 2015; Dai et al. 2014; Tao et al. 2013; Prats-Montalban et al. 2012; Nicolai et al. 2007). Furthermore, there are many applications of a similar nature in the remote sensing literature, but these are not mentioned here.

Acknowledgment The analysis was done using the software products Evince ver. 2.7.2 (Prediktera AB, Umeå, Sweden), Matlabver R2016a (Mathworks, Natick MA), and PLS_Toolboxver 8.1.1 (Eigenvector, Manson, WA). We thank Gunilla Nordström of Corpus Data and Image Analysis AB for assistance with running the cheese example.

References

Beebe, K., Pell, R., & Seasholtz, M. B. (1998). *Chemometrics. A practical guide*. New York: Wiley.

Brereton, R. (2003). *Chemometrics. Data analysis for the laboratory and chemical plant*. Chichester: Wiley.

Brereton, R., & Lloyd, G. (2014). Partial least squares discriminant analysis: Taking the magic away. *Journal of Chemometrics, 28*, 213–225.

Brown, P. (1993). *Measurement, regression and calibration*. Oxford: Clarendon Press.

Chen, Y., Sun, D., Cheng, J., & Gao, W. (2016). Recent advances for rapid identification of chemical information of muscle foods by hyperspectral imaging analysis. *Food Engineering Reviews, 8*, 336–350.

Cloutis, E. (1996). Review article, Hyperspectral geological remote sensing: Evaluation of analytical techniques. *International Journal of Remote Sensing, 17*, 2215–2242.

Dai, Q., Sun, D., Xiong, Z., Cheng, J., & Zeng, X. (2014). Recent advances in data mining techniques and their applications in hyperspectral image processing for the food industry. *Comprehensive Reviews in Food Science and Food Safety, 13*, 891–905.

Esbensen, K. (2012). *Multivariate data analysis – in practice*. Oslo: CAMO Publisher.

Esbensen, K., & Geladi, P. (2009). Principal component analysis (PCA): Concept, geometrical interpretation, mathematical background, algorithms, history. In S. Brown, R. Tauler, & B. Walczak (Eds.), *Comprehensive chemometrics: Chemical and biochemical data analysis* (Vol. 2, pp. 211–226). Oxford: Elsevier Science.

Esbensen, K., & Geladi, P. (2010). Principles of proper validation: Use and abuse of re-sampling for validation. *Journal of Chemometrics, 24*, 168–187.

Esbensen, K., & Julius, L. (2009). Representative sampling, data quality, validation – a necessary trinity in chemometrics. In S. Brown, R. Tauler, & B. Walczak (Eds.), *Comprehensive chemometrics: Chemical and biochemical data analysis* (Vol. 2, pp. 1–20). Oxford: Elsevier Science.

Esbensen, K., & Lied, T. (2007). Principles of image cross-validation (ICV): Representative segmentation of image data structures. In H. Grahn & P. Geladi (Eds.), *Techniques and applications of hyperspectral image analysis* (pp. 155–180). New York: Wiley.

Fernandez-Pierna, J., Baeten, V., & Dardenne, P. (2006). Screening of compound feeds using NIR hyperspectral data. *Chemometrics and Intelligent Laboratory Systems, 84*, 114–118.

Geladi, P., & Grahn, H. (1996). *Multivariate image analysis*. Chichester: Wiley.

Ghiyamat, A., & Shafri, H. (2010). A review on hyperspectral remote sensing for homogeneous and heterogeneous forest biodiversity assessment. *International Journal of Remote Sensing, 31*, 1837–1856.

Goetz, A., & Curtiss, B. (1996). Hyperspectral imaging of the earth: Remote analytical chemistry in an uncontrolled environment. *Field Analytical Chemistry and Technology, 1*, 67–76.

Gowen, A., Feng, Y., Gaston, E., & Valdramidis, V. (2015). Recent applications of hyperspectral imaging in microbiology. *Talanta, 137*, 43–54.

Grahn, H., & Geladi, P. (Eds.). (2007). *Techniques and applications of hyperspectral image analysis*. Chichester: Wiley.

Grahn, H., Geladi, P., & Esbensen, K. (2016). *Multivariate and hyperspectral image analysis, encyclopedia of analytical chemistry* (pp. 1–22). New York: Wiley.

Li, J., Sun, D., & Cheng, J. (2016). Recent advances in nondestructive analytical techniques for determining the total soluble solids in fruits: A review. *Comprehensive Reviews in Food Science and Food Safety, 15*, 897–911.

Martens, H., & Næs, T. (1989). *Multivariate calibration*. Chichester: Wiley.

Næs, T., Isaksson, T., Fearn, T., & Davies, T. (2002). *A user-friendly guide to multivariate calibration and classification*. Chichester: NIR Publications.

Nicolai, B., Buellens, K., Bobelyn, E., Peirs, A., et al. (2007). Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review. *Postharvest Biology and Technology, 46*, 99–118.

Pang, T., Sun, D., Cheng, J., & Pu, H. (2016). Regression algorithms in hyperspectral data analysis for meat quality detection and evaluation. *Comprehensive Reviews in Food Science and Food Safety, 15*, 529–541.

Prats-Montalban, J., de Juan, A., & Ferrer, A. (2012). Multivariate image analysis: A review with applications. *Chemometrics and Intelligent Laboratory Systems, 107*, 1–23.

Roggo, Y., Edmond, A., Chalus, P., & Ulmscheider, M. (2005). Infrared hyperspectral imaging for qualitative analysis of pharmaceutical solid forms. *Analytica Chimica Acta, 535*(1–2), 79–87.

Tao, L., Yang, X., Deng, J., & Zhang, X. (2013). Application of near infrared reflectance spectroscopy to predict meat chemical compositions: A review. *Spectroscopy and Spectral Analysis, 33*, 3008–3015.

Van der Meer, F. (2012). Remote sensing image analysis and geostatistics. *International Journal of Remote Sensing, 33*, 5644–5677.

Van der Meer, F., van der Werff, H., van Ruitenbeek, F., et al. (2012). Multi-and hyperspectral geologic remote sensing. A review. *International Journal of Applied Earth Observation and Geoinformation, 14*, 112–128.

Varmuza, K., & Filzmoser, P. (2009). *Introduction to multivariate statistical analysis in chemometrics*. Boca Raton, FL: CRC Press.

Winson, M., Goodacre, R., Timmins, E., et al. (1997). Diffuse reflectance absorbance spectroscopy taking in chemometrics (DRASTIC). A hyperspectral FT-IR based approach to rapid screening for metabolite overproduction. *Analytica Chimica Acta, 348*(1–3), 273–282.

Chapter 17
Dealing with Nonlinearity in Importance-Performance Map Analysis (IPMA): An Integrative Framework in a PLS-SEM Context

Sandra Streukens, Sara Leroi-Werelds, and Kim Willems

Abstract Importance-performance map analysis (IPMA) combines PLS-SEM estimates, indicating the importance of an exogenous construct's influence on another endogenous construct of interest, with an additional dimension comprising the exogenous construct's performance in a two-dimensional map. From a practical point of view, IPMA contributes to more rigorous management decision-making. The basic principles of IPMA are well understood, yet the inter-construct relationships are typically modeled as being linear. An abundance of empirical literature indicates that this may lead to erroneous conclusions. In an IPMA context, this can lead to false conclusions regarding an exogenous construct's importance. Although several approaches exist to account for nonlinear inter-construct relationships, these approaches are characterized by drawbacks impeding their applications in practice. Overall, this serves as a backdrop for the current chapter which aims to contribute to (PLS-SEM) IPMA theory in the following ways. First, we provide an integrative framework to guide IPMAs using PLS-SEM. Second and synergistically with the first contribution, we introduce a so-called log-log model that allows to capture the most common functional forms (i.e., both linear and nonlinear) without the need to make a priori assumptions about the correct functional form specification. Third, a comprehensive empirical application is provided that illustrates our proposed IPMA framework as well as the proposed log-log model to more adequately capture the nature of the PLS-SEM relationships ultimately defining the IPMA's importance dimension.

S. Streukens (✉) • S. Leroi-Werelds
Department of Marketing and Strategy, Hasselt University, Martelarenlaan 42, 3500 Hasselt, Belgium
e-mail: sandra.streukens@uhasselt.be; sara.leroiwerelds@uhasselt.be

K. Willems
Department of Business, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium
e-mail: kim.willems@vub.ac.be

H. Latan, R. Noonan (eds.), *Partial Least Squares Path Modeling*,
DOI 10.1007/978-3-319-64069-3_17

17.1 Introduction

The central tenet of the satisfaction-profit chain (SPC) is that effectively managing drivers of customer satisfaction is the road to enhanced business performance (Anderson and Mittal 2000; Kamakura et al. 2002). From a decision-making point of view, and building on the relationships put forward in the SPC, importance-performance map analysis (IPMA) can be considered an effective tool in satisfaction and thus business performance and management (cf. Matzler et al. 2004). The fact that managers are increasingly held financially accountable for their decisions further underscores the relevance of IPMA (Seggie et al. 2007).

In a nutshell and in general terms, IPMA contrasts the impact of key exogenous constructs or indicators (both are also referred to as drivers or input constructs or variables) in shaping a certain endogenous target construct with the average value, representing performance of the driver construct or indicator (Ringle and Sarstedt 2016). Combining the importance and performance measures of the drivers in a two-dimensional map then allows managers to identify key drivers of the target construct, to formulate improvement priorities, to find areas of possible overkill, and to pinpoint areas of “acceptable” disadvantages (see also Matzler et al. 2004).

IPMA per se is not restricted to a partial least squares structural equation modeling (PLS-SEM) context. However, IPMA in combination with PLS-SEM offers several key advantages, such as PLS-SEM’s ability to model a comprehensive nomological web of interrelations among constructs and its ability to include latent constructs. In line with these advantages, the focus in this chapter will be on IPMA in a PLS-SEM context.

Although several IPMA applications in a PLS-SEM context have emerged in the literature in recent years (e.g., Hock et al. 2010; Völckner et al. 2010; Rigdon et al. 2011), a common characteristic of these applications is that they all assume linear relationships. Inspection of the literature reveals that linear relationships are not necessarily appropriate in modeling business phenomena. Examples include Narasimhan and Kim (2002) in operations research, Langfred (2004) and Sciascia and Mazzola (2008) in management, Titah and Barki (2009) in management information systems, Lu and Beamish (2004) in strategy, and Seiders et al. (2005) in marketing.

Also regarding customer satisfaction management, which is the substantive domain central to this chapter, the possibility of nonlinearity needs to be taken into account when modeling the nomological web of relationships put forward in the SPC (Mittal et al. 1998; Dong et al. 2011; Anderson and Mittal 2000). In terms of IPMA, failing to take into account possible nonlinearities might lead to erroneous strategic decision-making regarding the management of customer satisfaction drivers.

Ignoring the customer satisfaction management background for the moment, the two general objectives of this chapter are twofold: first, to provide an integrative framework on how to conduct an IPMA that is both managerially relevant and theoretically valid and, second, to compare and contrast different analytical approaches to account for nonlinear structural model effects that determine importance scores in the IPMA. Given the added value of PLS-SEM in conducting an IPMA, the proposed IPMA framework and the approaches to deal with nonlinear relationships are, without loss of generalizability, discussed in a PLS-SEM context.

In relation to the work by Ringle and Sarstedt (2016) on IPMA as an extension of basic PLS-SEM, addressing the abovementioned research objectives yields the following contributions to theory and practice. First, our framework extends the work by Ringle and Sarstedt (2016) by paying attention to key decisions that need to be made in the pre-analytical stages of IPMA. That is, in order to arrive at an IPMA that is truly relevant both practically and theoretically, it is critical to keep in mind that IPMA is more than just a data analytical tool. Rather, IPMA involves making several interrelated decisions that start already at the research design phase, such as measurement model specification and questionnaire design. Second, although the current research focuses on nonlinearities in the SPC, the different methods to account for nonlinearity in structural model relationships discussed are generally applicable. As such, this study responds to the recent surge of interest in modeling nonlinear effects in PLS-SEM (see also Henseler et al. 2012). Third, in terms of IPMA, being able to adequately capture the functional form of structural relationships leads to an increased likelihood of improved strategic decision-making.

The remainder of this chapter is structured as follows. The subsequent two sections, Sects. 17.2 and 17.3, serve as a preparatory basis for the actual contributions as outlined above. Section 17.2 explains the basic principles of IPMA. In Sect. 17.3, the SPC is discussed as well as the need for taking into account nonlinearities in the SPC. Section 17.4 introduces the integrative IPMA framework and discusses the various stages involved. Again, without loss of generalizability, the framework discusses IPMA in combination with PLS-SEM. In Sect. 17.4, attention is particularly devoted to the modeling of nonlinear relationships in PLS-SEM. Section 17.5 describes the application of the proposed IPMA framework using real-life data. Sections 17.6 and 17.7 conclude this chapter.

17.2 IPMA: The Basics

17.2.1 What Is IPMA?

IPMA, originally introduced by Martilla and James (1977), yields insight into which drivers must be prioritized to achieve superior levels of a target construct of interest (e.g., satisfaction). In general, data derived from (satisfaction) surveys are used to construct a two-dimensional map, where performance is depicted along the horizontal axis (i.e., x-axis) and importance on the vertical axis (i.e., y-axis). As is shown in Fig. 17.1, for both axes, a cutoff value is specified to split each axis in a low and a high segment, dividing the matrix into the following four quadrants.

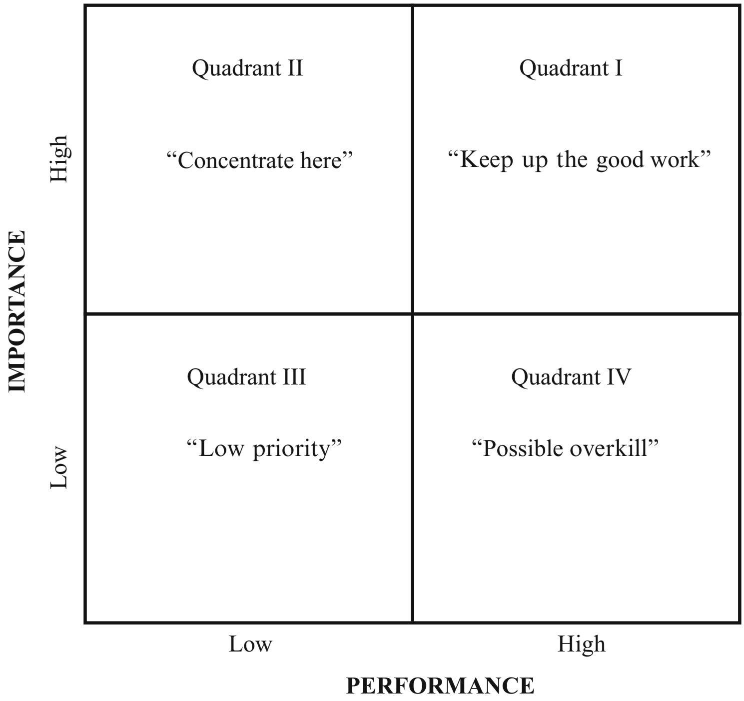


Fig. 17.1 Basic PLS-SEM IPMA chart

Drivers in quadrant I, named "Keep up the good work," are characterized by both a high importance level and a high performance level. These drivers represent opportunities for gaining or sustaining a superior level of the target construct. The drivers in quadrant II, named "Concentrate here," are key elements for improvement, as these drivers are considered important by respondents, while the perceived level of performance leaves things to be desired. The two quadrants at the bottom of the matrix are characterized by a low importance level. Hence, assuming equal costs, improvement initiatives concerning drivers located here can be expected to provide the lowest return on investment. Quadrant III, referred to as "Low priority," combines low importance with low performance. Drivers in this quadrant do not merit special attention or additional effort. Finally, quadrant IV, named "Possible overkill," represents drivers on which the respondents perceive a high level of performance but do not considered them very important. Similar to quadrant III, these drivers do not represent feasible alternatives for improving target construct performance. Rather, to avoid the risk of possible overkill, resources committed to these drivers would be better employed elsewhere.

17.2.2 IPMA and PLS-SEM

In a PLS-SEM context, the basic idea of IPMA as described above and shown in Fig. 17.1 remains unaltered. Nevertheless, PLS-SEM has several key advantages over traditional IPMA, which typically relies on multiple regression analysis. First, in determining the importance scores, PLS-SEM is a valuable analytical tool as it is capable of integrally assessing a complex network of relationships connecting drivers to a target construct of interest. Second, it can incorporate latent constructs. This is particularly relevant as in many research contexts, key constructs can only be validly measured by a set of indicators. These key constructs may concern both drivers and target constructs in the IPMA.

Consistent with Ringle and Sarstedt (2016), it should be stressed that in a PLS-SEM context, the IPMA may be conducted at either the latent variable or indicator level. The IPMA principles are not influenced by whether the analysis is done at the latent variable or indicator level. However, analysis at the indicator level leads to improved actionability, as indicators describe elements that shape the corresponding construct.

17.2.3 Performance Scores

As can be seen in Fig. 17.1, the horizontal axis captures driver performance. There are several ways to express these performance scores, depending on whether the IPMA is done at the indicator or latent variable level.

For IPMA at the indicator level, the mean or median score of the indicator represents the relevant performance score. Alternatively, for IPMA at the latent variable level, the average latent variable score represents perceived performance. With regard to the latter, it needs to be stressed that, in order to be meaningful, the latent variable's indicator weights all need to be in the same direction (Tenenhaus et al. 2005; Ringle and Sarstedt 2016).

Furthermore, to enhance the interpretability of the results, the performance scores are usually rescaled on 0–100 scale. Equation (17.1) shows how the latent variable scores are rescaled using the procedure suggested by Fornell et al. (1996):

$$\mathrm{LV}^{\mathrm{rescaled}} = \left(\frac{\sum_{i=1}^{j} w_i \overline{x}_i - \sum_{i=1}^{j} w_i \min[x_i]}{\sum_{i=1}^{j} w_i \max[x_i] - \sum_{i=1}^{j} w_i \min[x_i]} \right) \times 100 \tag{17.1}$$

In Eq. (17.1), w_i is the indicator weights associated with indicator x_i, while $\max[x_i]$ and $\min[x_i]$ denote, respectively, the maximum and minimum possible values for indicator x_i. Again, all weights need to be in the same direction. Equation (17.1) is also applicable at the indicator level. In this case, the calculation needs to be done at the indicator level using a weight equal to 1. Note that this rescaling is done automatically by SmartPLS 3 (Ringle et al. 2015).

17.2.4 Importance Scores

As can be seen in Fig. 17.1, the vertical axis of the chart captures the drivers' importance scores. Analytically, the importance score reflects the total effect of a predictor construct (i.e., driver) on a particular target construct. In general terms, and assuming a recursive structural model, let parameter $\delta_{kl}(k \neq l)$ reflect the total effect of construct k on construct l. Thus, parameter δ_{kl} reflects the entire set of relationships in a structural model connecting latent construct k to latent construct l. Parameter δ_{kl} can be calculated from the empirical results describing the set of relationships connecting latent construct k to latent construct l, as shown below in Eq. (17.2):

$$\delta_{kl} = \sum_{P:(\mathrm{LV}_{ki} \to \mathrm{LV}_l)} \left(\prod_{(\mathrm{LV}_k, \mathrm{LV}_l) \in P} \beta_{ij} \right) \tag{17.2}$$

In Eq. (17.2), and in the case of linear relationships, β_{ij} are the structural model coefficients belonging to the paths that connect latent construct k to latent construct l. In words, Eq. (17.2) states that the total effect of construct k on construct l can be computed by calculating the product of the structural model coefficients β_{ij} belonging to each of the separate direct relationships connecting construct k and construct l and subsequently summing these products' overall relevant paths connecting construct k and construct l. Or equivalently, as Nitzl et al. (2016) put it, the total effect is the sum of the relevant direct and indirect effects.

The idea expressed by Eq. (17.2) can be extended to include measurement model parameters as well, which is relevant for IPMAs conducted at the indicator level. In this case, the weight of indicator needs to be included in Eq. (17.2) as well. A graphical illustration of this can be found in Streukens and Leroi-Werelds (2016). Note that Ringle and Sarstedt (2016, p. 1869) indicate that this extension to the indicator level by including the indicator's weight in Eq. (17.2) can be done for both formative and reflective indicators.

Finally, the statistical significance of the importance scores can be assessed by means of bootstrap confidence intervals as outlined in detail by Streukens and Leroi-Werelds (2016).

17.2.5 Defining the Quadrants

Determining the cutoff values concerning what constitutes low performance/importance and high performance/importance is rather arbitrary. Multiple ways exist to specify these values. A commonly used way to specify these cutoff values is to use the mean (e.g., see Matzler et al. 2004) or median (e.g., see Berghman et al. 2013) value accompanying each of the axes.

17.3 The SPC and Nonlinear Relationships

17.3.1 Introduction of the SPC

As mentioned in the introduction, the IPMA and the SPC are a golden combination for many (marketing) managers. This stems from the fact that both the IMPA and the SPC are concerned with making effective resource allocation decisions in order to enhance business performance. To integrate the strengths of both models, the relationships put forward in the SPC (see also Fig. 17.2) can be estimated using PLS-SEM. Subsequently, the PLS-SEM estimation results can be used to calculate the importance scores (see also Eq. 17.2).

Starting at the back end of the chain, business profitability is positively influenced by customer loyalty, which reflects a customer's overall attachment to an offering, brand, or organization (Oliver 1999). In customer research, behavioral intentions are typically used as a proxy for customer loyalty. Customer satisfaction is the customers' cumulative evaluation that is based on all experiences with the company's offering over time (Anderson et al. 1994), and ample research supports the positive impact of this construct on customer loyalty (e.g., see Leroi-Werelds et al. 2014; Streukens et al. 2011). Consistent with Fishbein's (1967) multi-attribute model, customer satisfaction is a function of perceived attribute performance. Here, attributes are specific and measurable characteristics associated with a particular offering (see also Streukens et al. 2011). For example, in Gomez et al.'s (2004) study, attributes related to customer service, quality, and value for money were included as predictors of overall satisfaction with a supermarket.

Note that for the current study, the focus will be on the relationship between attribute performance and overall satisfaction and the relationship between overall satisfaction and loyalty. This is indicated in Fig. 17.2 by the dotted rectangle.

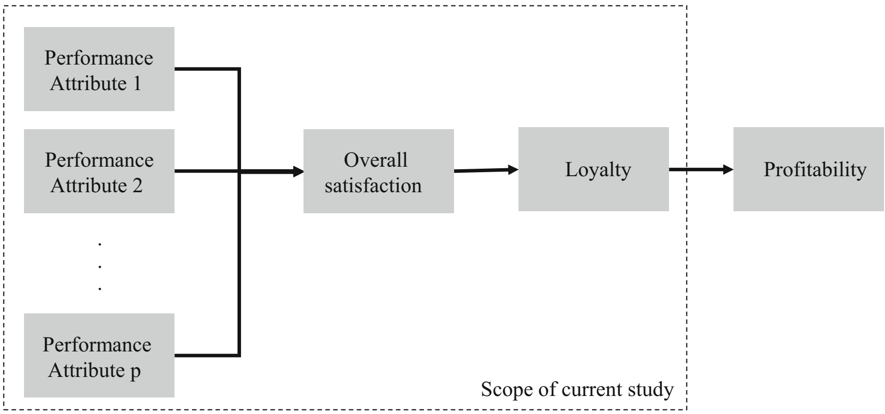


Fig. 17.2 The satisfaction-profit chain (adapted from Anderson and Mittal 2000)

17.3.2 Nonlinear Relationships in the SPC

Although the significance and direction of the relationships put forward in the SPC are well supported, an increasing amount of research suggests that the functional form of these relationships is not necessarily linear. Failing to account for possible nonlinearity in the links comprising the SPC may result in not finding support for expected linkages and/or incorrectly prioritize efforts to improve performance. Put differently, the danger of conducting PLS-SEM and the subsequent IPMA through a linear lens could be the misallocation of resources (see also Anderson and Mittal 2000).

In Sect. 17.3.2.1 the Kano-model (Kano et al. 1984) will be used to discuss the different (non)linear functional forms that may describe the attribute-overall satisfaction relationship. After that, Sect. 17.3.2.2 outlines five possible (non)linear functional forms regarding the relationship between overall satisfaction and loyalty.

17.3.2.1 Attribute Performance and Overall Satisfaction: Kano's Model

The Kano model distinguishes among three different types of attributes: *attractive* attributes ("delighters"), *must-be* attributes ("dissatisfiers"), and *one-dimensional* attributes ("satisfiers"). Figure 17.3 provides a graphical overview of Kano's model.

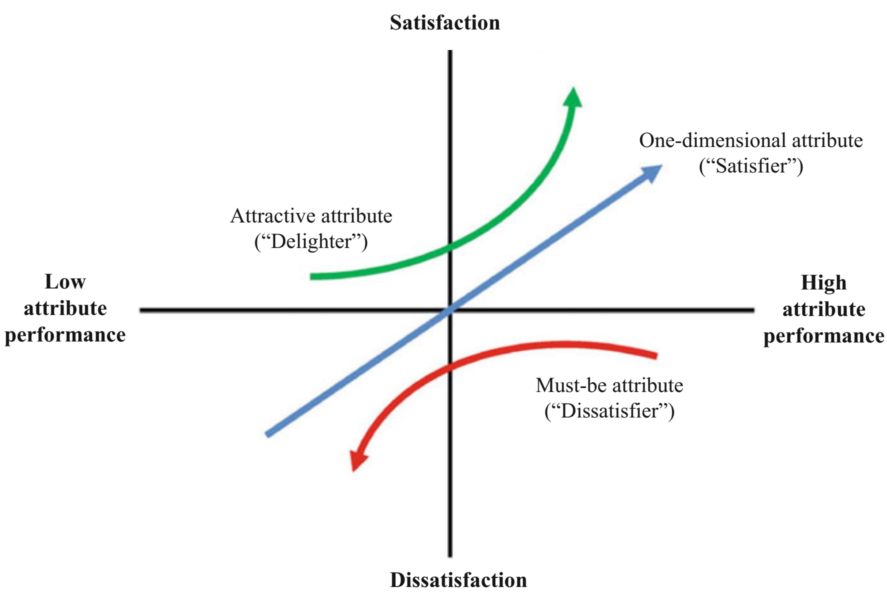


Fig. 17.3 Kano's model

Attractive attributes are attributes that increase satisfaction when fulfilled and have little influence on satisfaction ratings even when not fulfilled. Theoretically, *attractive* attributes are associated with customer delight, which is a positive emotion generally resulting from a surprisingly positive experience (Rust and Oliver 2000). In terms of functional form, *attractive* attributes are characterized by increasing returns. *Must-be* attributes cause dissatisfaction when absent but do not have an impact on satisfaction when present. Prospect theory (Kahneman and Tversky 1979), according to which "losses loom larger than gains," provides a theoretical rationale for this pattern. The functional form associated with *must-be* attributes is characterized by decreasing returns. The third and final category of attributes consists of so-called *one-dimensional* attributes or performance attributes. These attributes lead to satisfaction if performance is high and to dissatisfaction if performance is low (Matzler et al. 2004). *One-dimensional* attributes are characterized by a linear functional form, implying that a change in attribute performance has a constant impact on satisfaction.

17.3.2.2 Nonlinearities in the Satisfaction-Loyalty Link

According to Anderson and Mittal (2000), the relationship between satisfaction and loyalty may also exhibit nonlinearity. Based on an overview of the literature, Dong et al. (2011) discern among five different functional forms that have been put forward and empirically tested regarding the satisfaction-loyalty relationship. What Dong et al. (2011) refer to as a linear, concave, and convex relationship coincides with the functional form associated with Kano's *one-dimensional*, *attractive*, and *must-be* attributes, respectively. The two other functional forms discussed by Dong et al. (2011) include the S-shaped and inverse S-shaped function. An S-shaped function suggests decreasing returns for customers that are highly satisfied but increasing returns for customers who are less satisfied. The inverse S-shaped function implies increasing returns for customers who are highly satisfied and decreasing returns for customers with low satisfaction. Figure 17.4 summarizes the five functional forms of the satisfaction-loyalty relationship as presented by Dong et al. (2011).

17.4 IPMA: An Integrative Framework

This section outlines a three-stage integrative framework to conduct a strategically relevant IPMA. It is important to note that this framework, which is graphically presented in Fig. 17.5, implies that IPMA is more than just analytical approach.

That is, in order to be of true value, a strategically relevant IPMA requires addressing several key questions prior to the actual data analysis (e.g., the identification of attributes). This is reflected by stage 1 in Fig. 17.5. As indicated by stage 2 in Fig. 17.5, and in line with the discussion of possible nonlinearity in the relationships

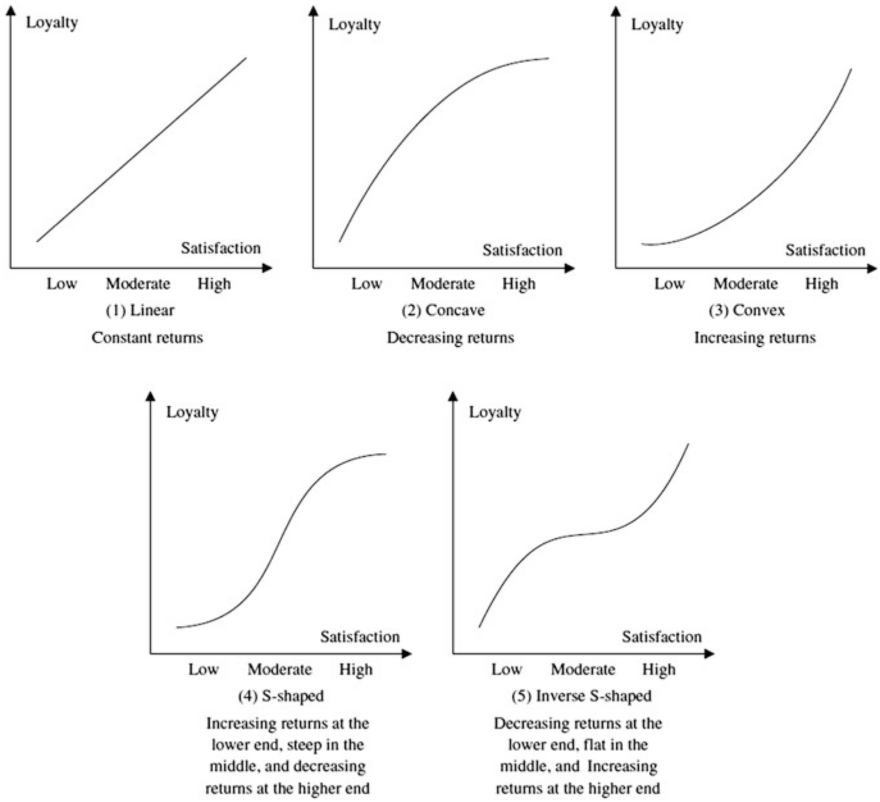


Fig. 17.4 Functional forms satisfaction-loyalty relationship (cf. Dong et al. 2011)

comprising a nomological network (i.e., the SPC), modeling the appropriate functional form is pivotal to adequately making strategic decisions involving resource allocation. As can be concluded from Fig. 17.5, several approaches are proposed to model nonlinear relationships. Finally, in stage 3, to make the transition from the statistical results to actionable practical implication, attention needs to be devoted to the interpretation of the results. Whereas the interpretation of the performance scores is rather straightforward, the interpretation of the importance scores will be relatively complex in case of nonlinear functional forms. Furthermore, stage 3 also includes (external) validation of the results. This is particularly relevant given the prediction-oriented nature of PLS-SEM (Shmueli et al. 2016; Carrión et al. 2016).

Two additional remarks concerning the framework in Fig. 17.5 need to be made. First, the steps in Fig. 17.5 are presented sequentially. However, for some decisions, these steps are interrelated. Whenever that is the case, it is explicitly mentioned. Second, although this chapter focuses in particular on the modeling of nonlinear relationships, the presented framework is also applicable to models that consist of only linear relationships. Rather, the basic IPMA, which assumes only linear

Stage 1: Research Design

- **Attribute identification**
  - Qualitative research (Interviews, CIT)
- **Questionnaire design**
  - Include all relevant attributes (content valdity)
  - Actionability / managerial relevance
  - Hierarchical design to reduce multicollinearity
  - Validated scales for higher order evaluative judgments
- **Response format**
  - Itemized rating scales or continuous rating scales
- **Sample size**
  - Power considerations and representativeness

Stage 2: Functional forms

- **Different possibilities**
  - Linear model
  - Penalty-reward contrast analysis
  - Transformed variables
    Linear-log model, Log-linear model, Log-log model
  - Polynomial model
- **Theoretical validity**

Stage 3: Interpretation results

- **Requirements check**
- **Importance scores**
  - Marginal effects / first-order (partial) derivatives
  - Total effects
- **Performance scores**
  - Rescale
  - Latent variable or indicator level
- **Cut-off values**
  - Arbitrary
  - Means / medians
- **Set priorities**
  - Create chart
- **Validation**
  - Split-sample method
  - Data from other setting
  - k-fold Cross-validation

Fig. 17.5 IPMA: an integrative framework relationships, can be considered as a special, very restricted case of the more advanced model that accounts for nonlinearity.

17.4.1 Stage 1: Research Design

A truly actionable IPMA is more than an analytical approach and requires several key considerations in the research design stage, such as attribute identification and questionnaire design.

17.4.1.1 Attribute[1] Identification

Conducting a series of interviews is often an excellent first step in identifying the relevant attributes, as the validity of the results of the IPMA depends on the completeness of the set of attributes under consideration. Or as Anderson and Mittal (2000) put it: the set of attributes should be as distinct and as broad as possible. An alternative method, preferably to be used in conjunction with interviews, is to use critical incident technique (CIT) data to identify active and/or salient sources of (dis)satisfaction. Moreover, the process of attribute identification needs to be repeated from time to time, as the list of attributes predicting overall satisfaction is likely to change due to a change in customer needs.

A frequently encountered issue involves the number of interviews that needs to be conducted. There is no fixed number that can be put forward here, but the saturation rule of Strauss and Corbin (1990) is a useful guideline. According to this rule, the researcher needs to continue conducting interviews until no new information comes to light (i.e., saturation point). Furthermore, in contrast to quantitative research, the samples used in interviews and other forms of qualitative research are not meant to be representative of the underlying population (see also Malhotra et al. 2012).

17.4.1.2 Modeling Attributes

Closely related to identification of the attributes is the specification of the accompanying measurement model. Several aspects need to be considered here. First, and also mentioned previously, is that attributes may be defined at the indicator or latent variable level. Second, although Fishbein's (1967) attribute model that is typically used in customer satisfaction research (see also Sect. 17.3.1) implies a formative

[1] Attribute identification is a term that is often used in the customer satisfaction research/SPC. More generally, what we refer to as attributes can be considered as the drivers or input variables of various kinds in an IPMA.

measurement model, Ringle and Sarstedt (2016) state that reflective measurement models can also be used to measure the latent constructs underlying an IPMA.

Although formatively and reflectively specified constructs can act as exogenous constructs (i.e., drivers) in an IPMA, formatively specified constructs are preferred for reasons of actionability and thus strategical relevance of the IPMA (see also Ringle and Sarstedt 2016, p. 1869).

17.4.1.3 Questionnaire Design

Data for an IPMA typically stem from (satisfaction) surveys. Consistent with the need to identify the complete set of relevant attributes (see Sect. 17.4.1.1), the key term for the design of the questionnaire is content validity. This means that all elements relevant to the customer in a particular situation need to be included. To achieve this, the output from the qualitative research conducted to identify the relevant attributes serves as input for the design of the questionnaire. Furthermore, whenever possible one is advised to use validated scales. This is typically the case for constructs that serve as target or intermediary constructs in an IPMA such as, in casu, loyalty intentions, and satisfaction.

Multicollinearity is a common issue in analyzing IPMA data, especially for formatively measured constructs. Although there are several technical ways to deal with this problem, the multicollinearity problem can be minimized through the design of the questionnaire. To do so, Mikulić and Prebežac (2009) suggest the following guidelines for questionnaire design. First, there should not be any overlapping between the conceptual domains of attributes, and the predictors should be on the same level of abstraction. Second, employ a so-called hierarchical design in which the different hierarchical levels measure attribute performance at various levels of abstraction. For an example of this latter recommendation, see the work of Dagger et al. (2007). The use of a hierarchical level model reduces multicollinearity as it decreases the number of independent variables per equation. Guidelines on how to model higher-order constructs in a PLS-SEM context can be found in Becker et al. (2012).

17.4.1.4 Response Formats

The estimation of (non)linear models requires metric data. Rating scales such as Likert scales are most often used in marketing research and are generally considered metric when at least five response categories are employed (Weijters et al. 2010). From a strict methodological point of view, Dawes (2008) concludes that the number of categories (i.e., five, seven, or ten) is trivial in case of regression-based techniques. However, Preston and Colman (2000) suggest that researchers should opt for seven, nine, or ten categories to warrant optimal levels of reliability, validity, discrimination power, and respondent preference. In terms of the analytical approaches to be discussed in the subsequent section, the use of Likert scales is a feasible option in the penalty-reward analysis, the linear model, the log-linear model, the linear-log model, and the log-log model.

In contrast, for a polynomial model, the use of Likert-type scales is not suitable (Finn 2011; Russell and Bobko 1992; Carte and Russell 2003) as it causes information loss that may result in an unknown systematic error. Hence, for polynomial models, the use of continuous rating scales is needed (see also Russell and Bobko 1992).

17.4.1.5 Sample Size

Similar to other analytical approaches, the needed sample size for a (PLS-SEM based) IPMA needs to be driven by statistical power considerations (Marcoulides et al. 2009) as well as representativeness (Streukens and Leroi-Werelds 2016).

17.4.2 Stage 2: The Functional Forms of the Relationships

Several approaches exist to model nonlinear relationships in a PLS-SEM context. Three alternative methods are discussed and compared below: the penalty-reward contrast analysis (PRCA), polynomial model, and analysis with transformed variables (i.e., linear-log, log-linear, and log-log models).

17.4.2.1 Penalty-Reward Contrast Analysis

Probably the most frequently used analytical approach in modeling the different types of attributes as implied by the Kano model is PRCA. Examples of studies employing PRCA include the work of Matzler et al. (2004), Busacca and Padula (2005), and Conklin et al. (2004).

PRCA analyzes the impact of high and low attribute performance on satisfaction by using two dummies, say D_i^{L} and D_i^{H}, for each attribute i (Mikulić and Prebežac 2011). The coding of the two dummies for attribute$i$ is done as follows: $\left(D_i^{\mathrm{L}}, D_i^{\mathrm{H}}\right) = (1, 0)$ indicates "low attribute performance," $\left(D_i^{\mathrm{L}}, D_i^{\mathrm{H}}\right) = (0, 1)$ indicates "high attribute performance," and $\left(D_i^{\mathrm{L}}, D_i^{\mathrm{H}}\right) = (0, 0)$ indicates "average attribute performance."

Running a PRCA thus implies that for each attribute, two coefficients are obtained: one to reflect the impact when performance is low, say β_i^{L}, and one to reflect the impact when performance is high, say β_i^{H}. Figure 17.6 presents the PRCA model in a PLS-SEM context. The model in Fig. 17.6 focuses on a single attribute measured by a single item. Moreover, satisfaction acts as target construct and is, consistent with the literature, measured by two reflective items.For academic research, it may be of interest to formally assess the nature of the attributes involved (e.g., see Mittal et al. 1998; Streukens and De Ruyter 2004). This can be done by

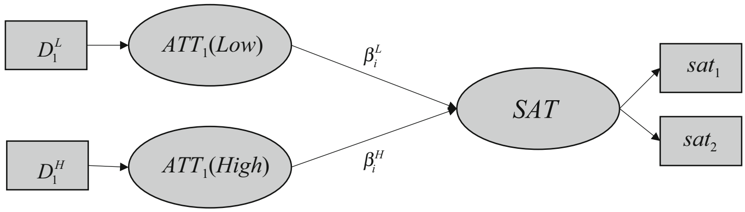


Fig. 17.6 PRCA in PLS-SEM

testing the following null hypothesis using the procedure outlined in Streukens and Leroi-Werelds (2016):

$$H_0 : \beta_i^{\mathrm{L}} = \beta_i^{\mathrm{H}}$$

If this null hypothesis cannot be rejected, the attribute can be regarded as a *one-dimensional* or linear attribute. Conversely, if the null hypothesis is rejected, then the attribute is either an *attractive* attribute (i.e., $\beta_i^{\mathrm{L}} < \beta_i^{\mathrm{H}}$) or a *must-be* attribute (i.e., $\beta_i^{\mathrm{H}} < \beta_i^{\mathrm{L}}$).

Despite its simplicity, several key disadvantages are associated with the use of PRCA. First, PRCA does not really capture nonlinearity. At best, it addresses whether the relationship between attribute performance and overall satisfaction is (a)symmetric. In the case of a symmetric relationship, low and high attribute performances have an equal impact on overall satisfaction. In the case of an asymmetric relationship, low and high attribute performances have a different impact on overall satisfaction. For more information on (a)symmetric relationships, the reader is referred to the work of Matzler et al. (2004). Second, as pointed out by Mikulić and Prebežac (2011), standardized coefficients for the PRCA dummies (as in PLS-SEM) may yield misleading results if the dummies have unequal distributions of zeros and ones, which is typically the case. Although a notable disadvantage, it can be circumvented by using the unstandardized model coefficients for model interpretation.

17.4.2.2 Polynomial Models

Polynomial models offer a flexible approach to capture a wide variety of functional forms without having to specify a specific form of nonlinearity a priori. Likewise, assessing whether nonlinearity is significant is straightforward in the case of polynomial models as it can be directly concluded from the statistical significance of the higher-order terms (Finn 2011; Dong et al. 2011). When modeling the link between attribute performance and overall satisfaction as implied by the Kano model, the quadratic model presented in Eq. (17.3) would apply:

$$\mathrm{SAT} = \beta_1 \mathrm{ATT}_i + \beta_2 \mathrm{ATT}_i^2 \tag{17.3}$$

Assessing the statistical significance of the model parameters reveals information about the functional form of the relationship between attribute performance and overall satisfaction. That is, in the case $\beta_1 > 0$ and $\beta_2 = 0$, a linear functional form is implied ("one-dimensional attribute"); if $\beta_2 < 0$, the relationship displays decreasing returns (*must-be* attribute); if $\beta_2 > 0$, the relationship exhibits increasing returns (*attractive* attribute). To capture the five different functional forms for the satisfaction-loyalty relationship as proposed by Dong et al. (2011), a cubic model, as presented in Eq. (17.4), is needed.

$$\mathrm{LOY} = \beta_1 \mathrm{SAT} + \beta_2 \mathrm{SAT}^2 + \beta_3 \mathrm{SAT}^3 \tag{17.4}$$

This cubic model is able to capture the five functional forms displayed in Fig. 17.4. The pattern of the model coefficients contains information about the functional form. If $\beta_1 > 0$ and $\beta_2 = \beta_3 = 0$, a linear functional form is implied; if $\beta_2 < 0$ and $\beta_3 = 0$, the relationship is concave; if $\beta_2 > 0$ and $\beta_3 = 0$, the relationship is convex; if $\beta_3 < 0$, an S-shaped relationship exists; and if $\beta_3 > 0$, the relationships are characterized by an inverse S-shaped pattern.

The approach to estimate a polynomial PLS-SEM model depends on the nature of the exogenous' construct measurement model (see also Henseler et al. 2012; Fassott et al. 2016). In case of a formative exogenous construct, the two-stage approach needs to be used.[2]

In the first stage, the PLS-SEM model without the quadratic effect is estimated to obtain estimates for the latent variable scores. These latent variable scores are saved for further analysis. In the second stage, the quadratic term is added to the model. The indicator of this quadratic term is the squared value of the relevant latent variable score obtained in stage 1. Furthermore, the latent variable scores obtained in the previous stage act also as indicators of the first-order term in the model and the endogenous construct, respectively. Figure 17.7 graphically presents the two-stage approach for a formative exogenous construct assessing attribute performance. Note that, although in Fig. 17.7 the formative exogenous construct only has a single indicator, the approach is readily applicable to multi-item formative exogenous constructs. Likewise, the approach can be easily extended to higher-order polynomial effects.

As evidenced by Henseler and Chin (2010), the analytical approach to be used for reflective exogenous construct is less straightforward, as the optimal choice depends on the objective of the researcher. Henseler and Chin (2010) distinguish three different types of research objectives: assessing the significance of the polynomial effect (case 1), finding an estimate for the true parameter of the polynomial effect

[2]Technically, the hybrid approach is also a feasible approach (see also Henseler et al. 2012). However, as this approach is not available in any of the standard PLS-SEM software packages, this approach will not be discussed in this chapter.

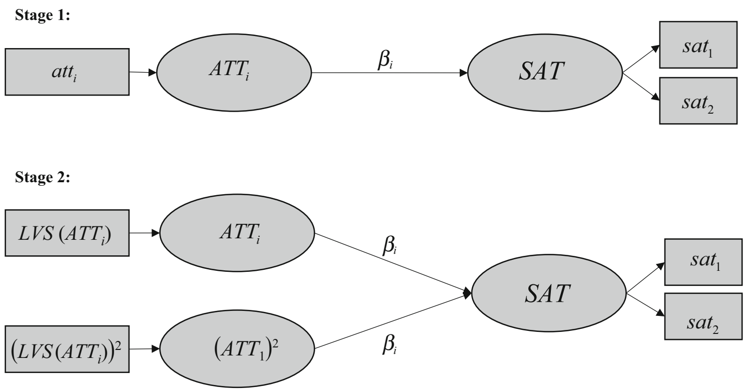


Fig. 17.7 Two-stage approach PLS-SEM (quadratic effect)

(case 2), and predicting the endogenous construct (case 3). All three cases will be discussed below.

Case 1: Assessing the Statistical Significance of the Polynomial Effect If the researcher is primarily interested in the significance of the polynomial effect, the two-stage approach is the method of choice. Here, a similar approach as outlined above for formative exogenous constructs applies.

Case 2: Finding Estimate of the True Parameter of the Polynomial Effect If the researcher is interested in finding an estimate for the true parameter of a polynomial effect, Henseler and Chin (2010) recommend using the orthogonalizing approach. Drawing upon Lance's (1988) residual centering technique, orthogonalizing essentially involves creating polynomial terms that are uncorrelated with (i.e., orthogonal to) its lower-order constituents. In the relationship between satisfaction and loyalty, satisfaction is measured by two reflective indicators (i.e., sat_1 and sat_2). We follow the procedure outlined by Henseler and Chin (2010) and Little et al. (2006). The computations for the quadratic and cubic term and the accompanying PLS-SEM models for the orthogonalizing approach are presented in Figs. 17.8 and 17.9, respectively.

Case 3: Prediction of Endogenous Construct The third and final case that needs to be distinguished with reflective exogenous constructs occurs when a researcher strives for optimal prediction accuracy of the target construct. In this case, Henseler and Chin (2010) recommend using either the orthogonalizing approach (see also Figs. 17.8 and 17.9) or the product-indicator approach. For the relationship between satisfaction (two reflective items) and loyalty (three reflective items), the product-indicator approach is shown in Fig. 17.10.

As mentioned above, key advantages of the polynomial approach are that it is flexible and the user does not need to specify the exact functional form in advance. Moreover, the often mentioned disadvantage of multicollinearity can be relatively easily resolved by orthogonalizing the involved terms. An important drawback of

Computations prior to PLS-SEM estimation

a) Compute all unique product indicators involving sat_1 and sat_2, which results in the following three products: $sat_1 \cdot sat_1$, $sat_1 \cdot sat_2$, and $sat_2 \cdot sat_2$.

b) Each of the three products obtained under a) are then regressed on all indicators. This implies the following three regression equations

$$sat_1 \cdot sat_1 = \gamma_1 sat_1 + \gamma_2 sat_2 + \delta_1$$

$$sat_1 \cdot sat_2 = \gamma_3 sat_1 + \gamma_4 sat_2 + \delta_2$$

$$sat_2 \cdot sat_2 = \gamma_5 sat_1 + \gamma_6 sat_2 + \delta_3$$

c) The residuals of the three regression equations, denoted by δ_1 to δ_3, serve as indicators for the quadratic term SAT^2.

PLS-SEM model

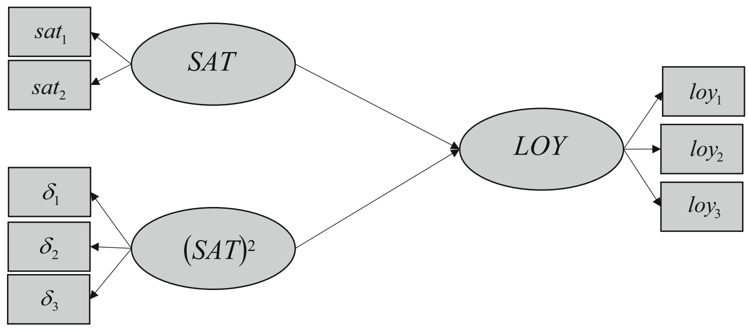


Fig. 17.8 Orthogonalization approach quadratic term

the polynomial model is the need for ratio-scaled data for the exogenous constructs (Carte and Russell 2003). Although Russell and Bobko (1992) recommend using continuous rating scales to solve this issue (see also Sect. 17.4.1.3), it needs to be acknowledged that for many constructs in business research, ratio scales do not exist. Although not impossible to create, it may represent a considerable challenge for researcher to develop ratio scales. Another disadvantage of the polynomial model is that it does not capture consistently decreasing or increasing returns (see also Streukens and De Ruyter 2004). For instance, the quadratic function describing the relationship between a *must-be* attribute and overall satisfaction may reach a maximum value after which the relationship under consideration becomes negative, which conflicts with the theory underlying the SPC.

Computations prior to PLS-SEM estimation

d) Compute all unique product indicators involving sat_1 and sat_2, which results in the following four products: $sat_1 \cdot sat_1 \cdot sat_1$ (or equivalently sat_1^3); $sat_1 \cdot sat_1 \cdot sat_2$ (or equivalently $sat_1^2 \cdot sat_2$); $sat_1 \cdot sat_2 \cdot sat_2$ (or equivalently $sat_1 \cdot sat_2^2$); and $sat_2 \cdot sat_2 \cdot sat_2$ (or equivalently sat_2^3).

e) Each of the three products obtained under a) are then regressed on all indicators. This implies the following four regression equations

$$sat_1^3 = \gamma_7 sat_1 + \gamma_8 sat_2 + \gamma_9 sat_1^2 + \gamma_{10} sat_2^2 + \gamma_{11} sat_1 \cdot sat_2 + \delta_4$$

$$sat_1^2 \cdot sat_2 = \gamma_{12} sat_1 + \gamma_{13} sat_2 + \gamma_{14} sat_1^2 + \gamma_{15} sat_2^2 + \gamma_{16} sat_1 \cdot sat_2 + \delta_5$$

$$sat_1 \cdot sat_2^2 = \gamma_{17} sat_1 + \gamma_{18} sat_2 + \gamma_{19} sat_1^2 + \gamma_{20} sat_2^2 + \gamma_{21} sat_1 \cdot sat_2 + \delta_6$$

$$sat_2^3 = \gamma_{22} sat_1 + \gamma_{23} sat_2 + \gamma_{24} sat_1^2 + \gamma_{25} sat_2^2 + \gamma_{26} sat_1 \cdot sat_2 + \delta_7$$

f) The residuals of the four regression equations, denoted by δ_4 to δ_7, serve as indicators for the cubic term SAT^3

For the quadratic terms see approach in Figure 8

PLS-SEM model

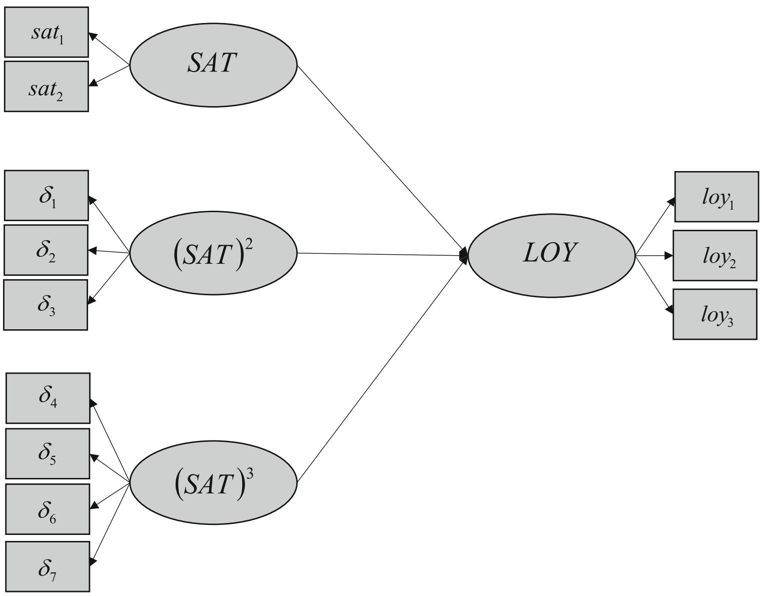


Fig. 17.9 Orthogonalization approach cubic term

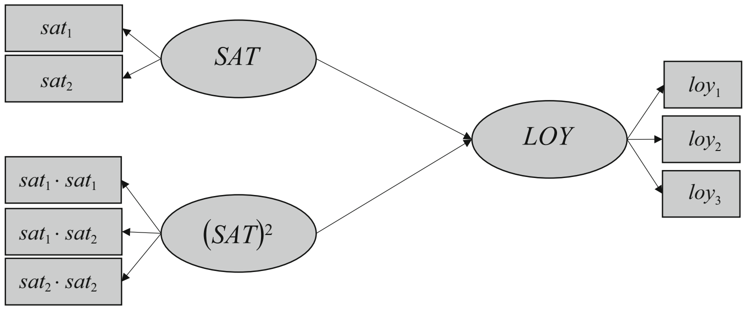


Fig. 17.10 Product-indicator approach (quadratic model)

17.4.2.3 PLS-SEM with Transformed Variables: The Linear-Log Model and the Log-Linear Model

This section introduces two models with transformed variables, namely, the linear-log model and the log-linear model, that are directly related to the basic linear PLS-SEM model. The name "transformed variables" stems from the fact that by transforming either the criterion or predictor variable, the standard linear PLS-SEM model can be used to capture a nonlinear functional form. Given the log-linear and linear-log models' kinship to the linear PLS-SEM model, the discussion of PLS-SEM with transformed variables will use the commonly-used linear model as a starting point. Although the focus in the sections below is on the relationship between attribute performance and overall satisfaction, the models can be readily applied to other relationships as well, e.g., the satisfaction-loyalty link.

The Linear Model For the simplest case of the relationship between overall satisfaction and performance of a single attribute, the linear model is defined as shown in Eq. (17.5):

$$\mathrm{SAT} = \beta_i \mathrm{ATT}_i \tag{17.5}$$

In terms of Kano's framework, this linear model is suitable for capturing the relationship between performance on a *one-dimensional* attribute and a higher-order evaluative judgment such as customer satisfaction.

The Linear-Log Model Relationships which are characterized by decreasing returns can be captured by means of a so-called linear-log model, which is shown in Eq. (17.6). Compared to the linear model (see also Eq. 17.5), the linear-log model uses the natural logarithm of the original variable as predictor. In terms of Kano's model, the linear-log model is suitable for modeling *must-be* attributes.

$$\mathrm{SAT} = \beta_i \ln \mathrm{ATT}_i \tag{17.6}$$

The Log-Linear Model In contrast to *must-be* attributes, the relationship between *attractive* attributes and a higher-order customer evaluative judgment displays increasing returns. Also in this case, a simple variation of the basic model presented in Eq. (17.5) can be used to adequately capture this specific relationship. That is, increasing returns can be modeled by means of a log-linear model. Compared to the basic linear model, the log-linear model uses the natural logarithm of the original dependent variable as criterion. The general structure of a log-linear model is shown in Eq. (17.7):

$$\ln \mathrm{SAT} = \beta_i \mathrm{ATT}_i \tag{17.7}$$

Their obvious kinship to the linear model which is typically used and the ease with which the linear-log model and log-linear model can be implemented in a PLS-SEM context are key advantages of these models. To learn more about the implementation in a PLS-SEM context of the models implied by Eqs. (17.5)–(17.7), see the three upper panels in Fig. 17.11. Consistent with marketing literature and practice, single-item attribute performance measures are assumed, and a reflective multiple-item scale is assumed for the satisfaction construct in the models shown in Fig. 17.11.

It is important to stress that the models outlined in Eqs. (17.5)–(17.7) and shown in Fig. 17.11 require the specification of the functional form in advance. That is, one needs to know prior to the estimation whether one is dealing with a *one-*

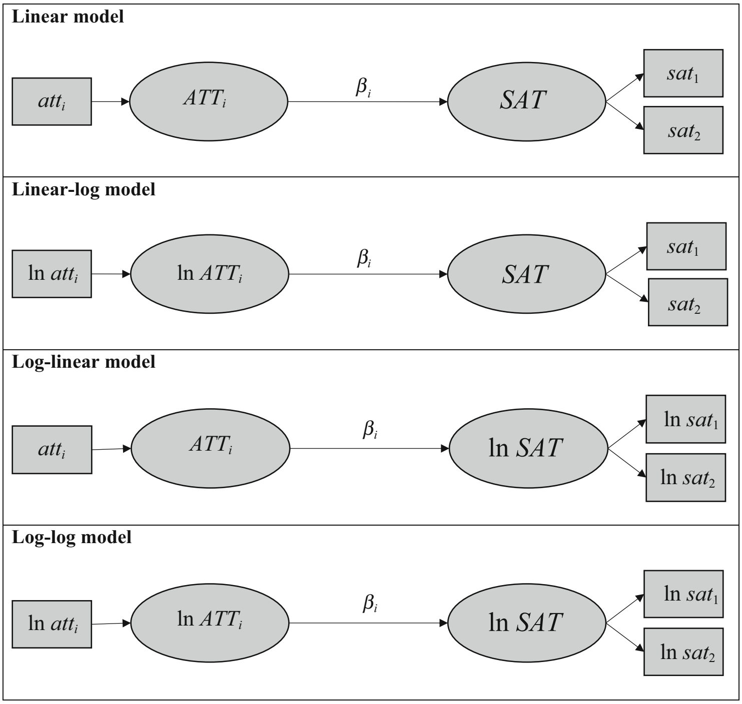


Fig. 17.11 PLS-SEM with transformed variables

dimensional, *must-be*, or *attractive* attribute. Although this may be perceived as a disadvantage, this drawback can be solved by carrying out an additional study in which Kano et al.'s (1984) approach to classifying attributes is applied. See the work of Mikulić and Prebežac (2011; pp. 48–51) for a detailed description of this attribute classification approach. Another problem arises when a relationship with increasing returns needs to be included in a single model that also needs to account for a linear relationship and/or a relationship displaying decreasing returns. Because the dependent variable of the log-linear model is different than the dependent variable of the linear and linear-log model, they cannot be included in a single model. For example, suppose attribute 1 (ATT_1) is a *one-dimensional* attribute, attribute 2 (ATT_2) is a *must-be* attribute, and attribute 3 (ATT_3) is an *attractive* attribute. Including *attractive* attribute ATT_3 in the same equation as ATT_1 and ATT_2 is not possible, as the dependent variable for the attribute ATT_3 differs from the dependent variable in modeling the impact of the latter two attributes (i.e., lnSAT vs. SAT). One possible way to overcome this problem is to estimate separate equations. However, this possibility is undesirable as it leads to biased coefficient estimates due to the omission of relevant variables.

17.4.2.4 PLS-SEM with Transformed Variables: The Log-Log Model

In line with its name, the log-log model involves taking the natural logarithm of both the dependent and independent variable. Departing from the basic model (see also Eq. 17.5), the log-log model is defined as shown in Eq. (17.8):

$$\ln \mathrm{SAT} = \beta_i \ln \mathrm{ATT}_i \tag{17.8}$$

The log-log model can account for all the functional forms implied by the Kano model without having to specify in advance the nature of the relationship. The magnitude of the model parameter β_i reveals the functional form best describing the relationship between variables involved. More specifically, $\beta_i = 1$ implies a linear relationship (i.e., *one-dimensional* attribute), $\beta_i > 1$ corresponds with increasing returns (i.e., *attractive* attribute), and $0 < \beta_i < 1$ reflects decreasing returns (i.e., *must-be* attribute). The fourth panel in Fig. 17.11 indicates how the log-log model can be applied in a PLS-SEM context. Just as for the other models related to the attribute performance-satisfaction link, a single-item attribute performance measurement model and a reflective multiple-item scale for satisfaction are assumed.

To understand the exact nature of the functional form (see, for instance, the work of Streukens and De Ruyter 2004), bias-corrected and accelerated bootstrap confidence intervals (see also Streukens and Leroi-Werelds 2016) need to be constructed to test whether $\beta_i = 1$, $\beta_i > 1$, or $0 < \beta_i < 1$.

Overall, in comparison to other models with transformed variables (i.e., log-linear model and linear-log model), the log-log model offers an elegant solution to the two drawbacks associated with the former type of models, namely, a priori specification of functional form and incorporating *attractive* attributes in single models with *must-be* and/or *one-dimensional* attributes.

To conclude this part on modeling different functional forms, Table 17.1 summarizes the advantages and disadvantages associated with the different approaches available to account for nonlinear functional forms. It is important to stress the role of theory in opting for the appropriate functional forms.

17.4.3 Stage 3: Interpretation of Results

The third and final stage of our framework focuses on the interpretation of the analytical results. This stage is critical in making the transition from statistical output to actionable results.

17.4.3.1 Requirement Check

Regardless of the measurement model being reflectively or formatively specified, an important requirement to determine the performance scores on the latent variable level is that all outer weights are positive (Ringle and Sarstedt 2016; Tenenhaus et al. 2005). Cenfetelli and Bassellier (2009) and Ringle and Sarstedt (2016) acknowledge the problem of co-occurrence of both positive and negative weights and warn for the adverse effects this may have on the interpretation of the results. Different reasons, conceptual as well as methodological, may underlie the co-occurrence of positive and negative weights. Depending on the cause, different courses of action are needed.

Conceptually, a negative indicator weight may result from the opposite formulation of the corresponding questionnaire item. In such instances, the researcher has to reverse the scale on which the item was measured.

Methodologically, negative indicator weights may be the result of multicollinearity. Inspection of the VIF values provides evidence of the existence of multicollinearity. In case of substantial multicollinearity (i.e., VIF ≥ 5), the researcher may consider deleting indicators, constructing higher-order constructs, or combining collinear indicators into a single new composite indicator (see Hair et al. 2014, p. 125 for more information).

Dijkstra and Henseler (2011) propose the use of best-fitting proper indices (BFPI) to assure that the indicator weights (as well as the loading) for a formative measurement model are nonnegative (i.e., proper). Despite its apparent attractive features, a current drawback of using BFPI is that it is not, to the best of our knowledge, available in software packages such as SmartPLS 3 (Ringle et al. 2015).

Table 17.1 (Dis)advantages of the different approaches to model nonlinear effects

	Penalty-reward contrast analysis	Linear-log and log-linear model	Polynomial model	Log-log model
Advantages	• Easy to use	• Models nonlinearity • Monotone function	• Models nonlinearity Flexibility, allows also for modeled (inverse) S-shaped functions • Does not require a priori specification of functional form	• Models nonlinearity • Monotone function • Does not require a priori specification of functional form
Disadvantages	• Can model asymmetry, but not nonlinearity	• A priori specification of functional form • Increasing returns cannot be modeled in a single equation with decreasing returns or a linear functional form	• Requires ratio-scaled variables • Non-monotone function (theoretical validity) • Multicollinearity (can be solved by orthogonalization)	• Only models linear, monotonically increasing and decreasing functions (i.e., no (inverted) S-shaped functions)

17.4.3.2 Importance Scores: Marginal Effects

To assess the impact or importance of a driver on the target, the structural model parameters play a pivotal role. In this respect, it is not necessarily the structural model parameters per se we are interested but rather the rate of change in a target construct for a driver or the driver's marginal effect. Mathematically, the first (partial) derivative of any type of equation gives the formula for the rate of change in the criterion variable for any predictor variable (Johnson et al. 1978; Stolzenberg 1980).

For a simple linear model (e.g., in the format of $\text{SAT} = \beta_i \text{ATT}_i$), the first derivative boils down to β_i and thus equals the structural model coefficient that can be readily derived from the PLS-SEM output. For all other functional forms discussed above, the marginal effect of a particular driver on a target construct is less straightforward but still can be derived from the PLS-SEM output. Just as for a simple linear function, the first derivative equals the marginal effect, but it yields a formula rather than a constant (see also Stolzenberg 1980). Table 17.2 summarizes the first derivatives for the different functions discussed in this chapter.

To arrive at the actual value reflecting the importance of a certain driver, the mean score of the latent or indicator variable is plugged into the equations listed in Table 17.2 (see also Roncek 1991, 1993).

In line with PLS-SEM's capability to integrally assess a nomological network of relationships, the importance rating equals the total effect of the driver on an endogenous construct of interest. That is, the idea of the total effects as expressed by Eq. (17.2) still applies when the nomological network contains one or more nonlinear effects. However, in this latter case, the formulae presented in Table 17.2 represent the different marginal effects that need to be used in combination with Eq. (17.2) to compute the total effect (see also Streukens et al. 2011).

17.4.3.3 Performance Scores

The determination of the performance scores in an IPMA remains unaffected by the nature of the relationships connecting the different constructs. Hence, the same guidelines as discussed in Sect. 17.2.2 apply. It should be noted the performance scores need to be calculated for the untransformed variables.

17.4.3.4 Cutoff Values

The definition of the quadrants remains arbitrary. See Sect. 17.2.4 for alternative ways to set the cutoff values that define the different quadrants.

Table 17.2 Marginal effects for the different functional forms

Approach	Accompanying function	Shape	Marginal effect
Penalty-reward contrast analysis	$\mathrm{SAT} = \beta_i^{\mathrm{L}} D_i^{\mathrm{L}} + \beta_i^{\mathrm{H}} D_i^{\mathrm{H}}$	Piecewise linear	β_i^{L} or β_i^{H}
Linear model	$\mathrm{SAT} = \beta_i \mathrm{ATT}_i$	Linear	β_i
Linear-log model	$\mathrm{SAT} = \beta_i \ln \mathrm{ATT}_i$	Decreasing returns	β_i / ATT_i
Log-linear model	$\ln \mathrm{SAT} = \beta_i \mathrm{ATT}_i$	Increasing returns	$\beta_i \cdot \mathrm{e}^{\beta_i \mathrm{ATT}_i}$
Log-log model	$\ln \mathrm{SAT} = \beta_i \ln \mathrm{ATT}_i$	Various	$(\mathrm{SAT}/\mathrm{ATT}_i)\ \beta_i$
Quadratic model	$\mathrm{SAT} = \beta_{1i} \mathrm{ATT}_i + \beta_{2i} {\mathrm{ATT}_i}^2$	Various	$\beta_{1i} + 2\beta_{2i} \mathrm{ATT}_i$
Cubic model	$\mathrm{LOY} = \beta_1 \mathrm{SAT} + \beta_2 \mathrm{SAT}^2 + \beta_3 \mathrm{SAT}^3$	Various	$\beta_{1i} + 2\beta_{2i} \mathrm{SAT} + 3\beta_3 \mathrm{SAT}^2$

17.4.3.5 Setting the Right Priorities

Information on each driver in terms of performance and importance can be graphically presented in a two-dimensional map (see also Fig. 17.1). The subsequent resource allocation decisions are straightforward in case of linear functional forms but may be more complicated for nonlinear relationships. We summarize below the resource allocation guidelines proposed by Anderson and Mittal (2000) for *attractive* (log-linear model) and *must-be* (linear-log model) attributes. An important assumption regarding the guidelines below is that the relationships among subsequent constructs (e.g., the link between satisfaction and loyalty) are linear.

For *attractive* attributes that are characterized by both a high performance and a high importance score, the current investments should be maintained ("Keep up the good work"). When *attractive* attributes are characterized by low performance and low importance scores, no improvement initiatives need to be undertaken as an increase in performance will yield only a very limited increase in importance (i.e., flat part of the nonlinear relationship). *Must-be* attributes display diminishing returns, which means that the performance increases have a diminishing effect as the level of performance rises. Anderson and Mittal (2000) therefore recommend investments to improve the performance of *must-be* attributes which are important but on which the performance is still low.

17.4.3.6 Validation

In line with the prediction-oriented nature of PLS-SEM, being able to generate out-of-sample predictions from a model and to evaluate its predictive power is vital when conducting an IPMA (cf. Shmueli et al. 2016; Carrión et al. 2016). As outlined below, several approaches are available to assess the model's predictive performance.

Carrión et al. (2016) strongly recommend the use of a split-sample approach ("holdout samples") as a means for establishing whether or not a model has an adequate level of predictive performance. In their work, they provide an excellent overview of how to implement this in a PLS-SEM context.

In case of small- to medium-sized samples, the abovementioned split-sample approach may not be feasible as partitioning the already relatively small sample size into a training and validation sample can introduce too much bias. For these situations, Shmueli et al. (2016) propose the use of a k-fold cross-validation approach. With regard to this k-fold cross-validation approach, which is available in SmartPLS 3 (Ringle et al. 2015) under the name "blindfolding," Shmueli et al. (2016) and Rigdon (2014) rightly point out that the traditionally reported Stone-Geisser Q^2 statistics suffer from serious limitations and should only be used with the greatest care.

Another way of assessing predictive performance is to use samples from other contexts as validation samples. In a PLS-SEM context, this is demonstrated in the work of Miltgen et al. (2016). In a similar vein as the split-sample approach suggested by Carrión et al. (2016), this approach involves using parameter estimates from the original sample to predict the outcomes in another external sample.

17.5 Empirical Illustration[3]

This fifth section demonstrates how the log-log model can be used in an IPMA. Data were collected from customers of a European DIY company whose business consisted of providing customers with the resources needed to install ventilation systems by themselves. All analyses were performed using SmartPLS 3 (Ringle et al. 2015).

17.5.1 Stage 1: Research Design

Based on interviews with a set of customers ($n = 10$), six attributes were identified as being drivers of customer satisfaction. These six items can be found in Table 17.3.

For each attribute, the questionnaire included a single item to tap customers' quality perceptions. The respondents were asked to rate each attribute's quality on a 9-point Likert scale (1 = very low quality, 9 = very high quality). In a similar vein, overall satisfaction was measured using single item in combination with a 9-point Likert scale (1 = very dissatisfied, 9 = very satisfied). The accompanying descriptive statistics can be found in Table 17.3. In total, an effective sample size of $n = 149$ was obtained.

As can be observed in subsequent figures, we employed a single-item measurement model for each attribute. Technically this implies that the IPMA at the latent construct level and IPMA at the indicator level coincide.

17.5.2 Stage 2: The Functional Forms of the Relationships

The log-log model was used to account for possible nonlinearity in the attribute-satisfaction relationships. The key advantage of the chosen approach is that one does not have to specify the functional form in advance. Prior to running the model in SmartPLS 3 (Ringle et al. 2015), the natural log transformation was performed on all variables involved. For reasons to be explained later on in Sect. 17.5.3, the researcher is advised to construct a data file that contains both the original variables and the transformed variables.

[3] All files pertaining to this application are available upon request from the first author.

Table 17.3 Results of empirical application

	Mean	SD	Performance (0–100%)	Unstandardized structural model coefficient	BCa percentile bootstrap CI	Marginal effect	BCa percentile bootstrap CI shape	Shape
Attribute: quality products (IQ01)	7.74	1.15	84.312	0.610	[−0.069;0.576]	–	–	–
Attribute: quality-provided tools (IQ02)	7.53	1.39	81.628	0.146	[−0.296;0.329]	–	–	–
Attribute: quality working plan (IQ03)	6.11	2.13	63.926	0.255	[0.028;0.451]	0.281	[−0.963;−0.364]	Decreasing returns
Attribute: quality delivery (PQ01)	5.79	2.38	59.815	0.314	[0.030;0.627]	0.370	[−0.964;−0.562]	Decreasing returns
Attribute: quality invoicing (PQ02)	5.68	2.43	58.473	0.084	[−0.151;0.374]	–	–	–
Attribute: quality-finished installation (OQ01)	7.96	0.99	82.681	0.023	[−0.099;0.151]	–	–	–
Overall satisfaction with company (SAT)	6.64	2.13	70.554					

Note: Marginal effects and the bootstrap confidence interval regarding functional form (shape) are only computed for significant attributes

To estimate the log-log model, the structural model as shown in Fig. 17.12 panel A is run. The unstandardized coefficients (cf. Hock et al. 2010) and the accompanying bias-corrected and accelerated percentile bootstrap confidence intervals based on 10,000 samples (cf. Streukens and Leroi-Werelds 2016) are presented in Table 17.3. Note that in SmartPLS 3 (Ringle et al. 2015), the option "importance-performance matrix analysis (IPMA)" was selected as this automatically yields the unstandardized coefficients and the rescaled performance scores.

Note that for the significant attributes, bootstrap confidence intervals were constructed to determine the shape of the functional forms describing the relationship between the attribute's performance and overall satisfaction. Based on the results listed in Table 17.3, it can be concluded that the two significant attributes are so-called *must-be* attributes.

17.5.3 Stage 3: Interpretation

As evidenced in Table 17.1, the marginal effects for the log-log model depend on the value of the (in)dependent constructs involved. Hence, in order to obtain the appropriate importance scores for the situation at hand, the performance scores of the relevant variables need to be imputed in equations describing the marginal effects for the log-log model (see Table 17.2). For this aim, either the original *untransformed* variables or the *rescaled* performance of the *untransformed* variables, both of which can be found in Table 17.3, can be used. A quick way to obtain the *rescaled* performance scores of the *untransformed* variables is to run the PLS-SEM model with the untransformed variables (see also Fig. 17.12, panel B). This is the reason why the researcher was advised to store both the transformed and untransformed variables in a single data file. Table 17.2 contains the marginal effects for the different attributes. Also for this second estimation run, the option "importance-performance matrix analysis" was selected in the program. It is important to stress that the second estimation (estimation of model shown in Fig. 17.12, panel B) is technically not required to determine the performance scores. Rather, it is merely a handy shortcut to obtain the rescaled performance levels.

The actual IPMA chart can be created using Microsoft Excel. Although SmartPLS 3 (Ringle et al. 2015) can produce these charts automatically, the nonconstant marginal effects in the case of nonlinear functional forms cannot (yet) be taken into account by the program. The only data needed to construct the chart are the importance and performance scores described in the previous section and listed in Table 17.3. In Table 17.3, the entries in the column named "Performance (0–100%)" provide the x-axis coordinates (i.e., driver performance), and the entries in the columns named "Marginal effects" provide the y-axis coordinates (i.e., driver importance). The resulting chart, shown in Fig. 17.13, only contains two data points as we only include drivers that have a significant impact on the target construct.

The cutoff values for the axis are determined as follows. For the importance score, a cutoff value of 0.30 is used, which corresponds to a medium effect size

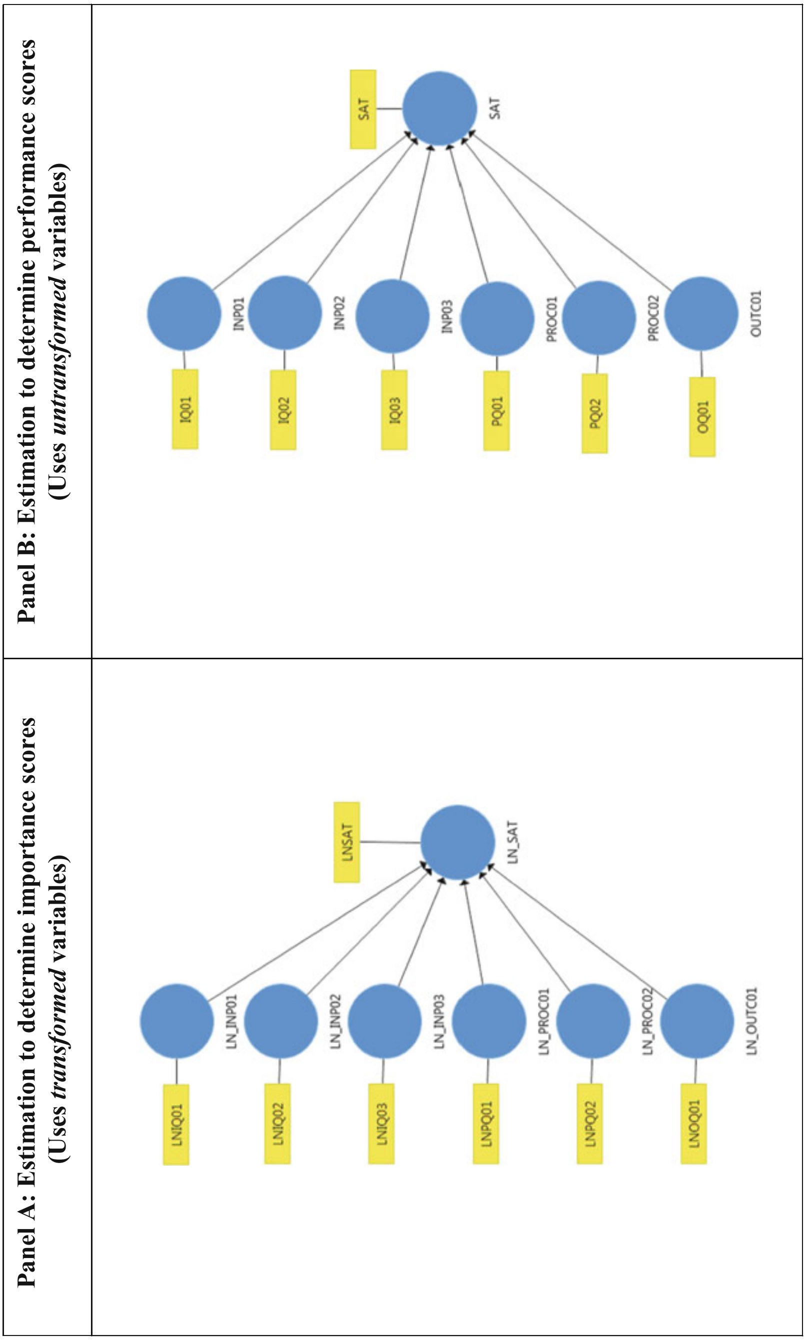


Fig. 17.12 PLS-SEM models accompanying empirical demonstration IPMA

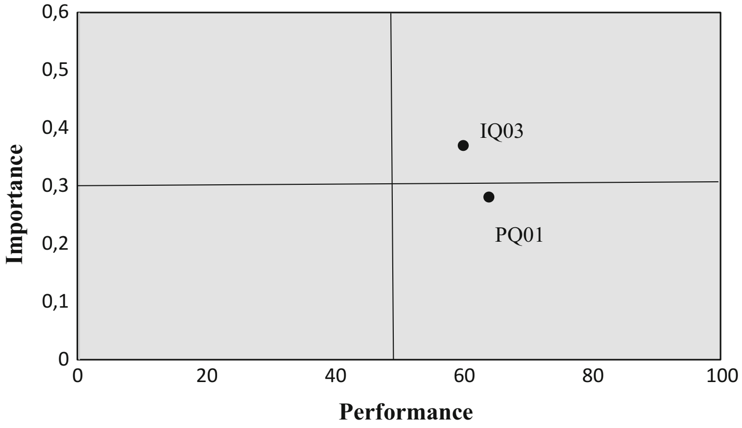


Fig. 17.13 IPMA chart empirical illustration

for correlation coefficients. For the performance score, the cutoff value equals 50, which is the midpoint of the 0–100 range to which the drivers are rescaled.

Based on the chosen cutoff values, attribute IQ03 (quality working plan) falls in the quadrant "Keep up the good work," whereas attribute PQ01 (quality delivery) falls in the "Possible overdrive" quadrant. This would imply that investments to at least maintain the perceived performance of attribute IQ03 are warranted, whereas investments to maintain the perceived performance of attribute PQ01 are less relevant. Given that resources are scarce, one could even argue that resources currently spent on PQ01 could be (partly) reallocated to maintaining or even improving the performance on attribute IQ03.

Ideally, the final step involves the validation of the findings. Unfortunately, the sample size was too small to apply a split-sample procedure, nor was an external validation sample available.

17.6 Conclusion

In times when managers are increasingly held accountable for the financial performance implications of their strategic actions, IPMA is an effective method to help set priorities and to optimally allocate scarce resources. As convincingly argued by Ringle and Sarstedt (2016), IPMA adds valuable additional insights above and beyond the results of a standard PLS-SEM.

Consistent with the view expressed in the academic literature across various domains that the structural model relationships do not necessarily exhibit a linear functional form, a strategically relevant IPMA must be able to take into account possible nonlinearities in the inter-construct relations.

In response to the need for IPMAs that will contribute to optimal strategic decision-making, this chapter proposes an integrative IPMA framework in which the possibility of nonlinear functional structural relationships is explicitly accounted for. Although the suggested IPMA framework is applicable regardless of the analytical approach to model the involved relationships that ultimately yield the importance score, we have deliberately chosen to discuss the framework through a PLS-SEM lens. The value of PLS-SEM as a basis for an IPMA stems from the fact that PLS-SEM is capable of integrally estimating complex nomological networks of relationships and can take into account latent constructs as well.

Compared to previous IPMA guidelines, the proposed framework takes a broader perspective on IPMA than just a data analytical tool. Rather, designing and conducting an IPMA that is both managerially relevant and theoretically sound start already at the research design stage. Moreover, in line with PLS-SEM's prediction-oriented nature, validation is suggested to constitute an essential element of IPMA.

Within the proposed IPMA framework, the estimation of (possible) nonlinear relationships plays a dominant role. We compare and contrast several approaches to account for a variety of functional forms in a PLS-SEM context. Similar to the overall IPMA framework, the focus on PLS-SEM in modeling nonlinearities in no way limits the generalizability of our work.

Complementing existing work on modeling nonlinear function in a PLS-SEM context, this study introduces the log-log model as a flexible approach to capture all possible functional forms without having to specify the functional form in advance. Being a direct extension of the commonly used linear model, the log-log model will be easy to apply for most PLS-SEM users.

We drew upon the SPC to illuminate and empirically illustrate the ideas put forward in the suggested framework. Also regarding this choice, it is pivotal to stress that this does not limit the generalizability of our work in any way. Despite the focus on a specific substantive domain, we sincerely believe that access to the data and other files associated with our empirical application may have a positive impact on the practical implementation of our ideas. Hence, feel free to ask for them, use them, and adapt them to your own IPMA in the hope to arrive at a truly valuable IPMA and to get even more from your PLS-SEM results than ever before.

Regarding the link between overall satisfaction and loyalty, the literature proposes even a larger variety of functional forms that may describe this relationship. The log-log model is not capable of capturing all of the suggested functional forms by Dong et al. (2011). To account for all the possible functional forms of the satisfaction-loyalty link, a cubic model offers the most flexible approach. Similar to the log-log model, the cubic model does not require to specify the functional form in advance. However, a key drawback of the cubic model is that it requires the use of ratio data (Carte and Russell 2003), which may not always be possible to collect. Finally, although the current study focuses on customer satisfaction and PLS-SEM, the proposed IPMA framework and the approaches to capture nonlinear functional forms are also applicable in other contexts.

17.7 Limitations and Suggestions for Further Research

This study has not taken into account possible dynamic effects. Past research (e.g., Bolton 1998) has shown that changes in performance have a different effect on satisfaction depending on the duration of the relationship. In a similar vein, Mittal et al. (1999) provide empirical evidence that the nature of the attribute performance-satisfaction link may change over time. More research is needed on how these effects can be incorporated in an IPMA to optimize strategic decisions.

Closely related to the previous point is the recognition that the relationships in the SPC vary as a function of customer characteristics (e.g., see Garbarino and Johnson 1999; Mittal and Katrichis 2000). If customer characteristics that influence the relationships determining the importance scores in an IPMA can be identified beforehand, this problem may be solved by conducting a separate IPMA for each segment. If customer heterogeneity is unobserved, as is often the case, possible solutions may be found along the lines of FIMIX PLS-SEM. Hence, a promising avenue for future research is how to make strategic resource allocation decisions while taking into account observed customer heterogeneity.

Hopefully the integrative IPMA framework proposed in this chapter advances the existing knowledge and skills to make well-informed strategic decisions. The focus of the current chapter is on the increase in attribute performance and other constructs which ultimately translate into enhanced revenues (i.e., SPC). To fully understand the impact of investments associated with strategic decisions, both the revenues and the costs of the investment need to be taken into account (Zhu et al. 2004). As such, research that places the results from an IPMA in a larger investment framework, such as proposed by Streukens et al. (2011), might further increase the managerial relevance of the suggested framework.

References

Anderson, E. W., & Mittal, V. (2000). Strengthening the satisfaction-profit chain. *Journal of Service Research, 3*(2), 107–120.

Anderson, E. W., Fornell, C., & Lehmann, D. R. (1994). Customer satisfaction, market share, and profitability: Findings from Sweden. *Journal of Marketing, 58*(3), 53–66.

Becker, J. M., Klein, K., & Wetzels, M. (2012). Hierarchical latent variable models in PLS-SEM: Guidelines for using reflective-formative type models. *Long Range Planning, 45*(5–6), 359–394. doi:10.1016/j.lrp.2012.10.001.

Berghman, L., Matthyssens, P., Streukens, S., & Vandenbempt, K. (2013). Deliberate learning mechanisms for stimulating strategic innovation capacity. *Long Range Planning, 46*(1), 39–71.

Bolton, R. N. (1998). A dynamic model of the duration of the customer's relationship with a continuous service provider: The role of satisfaction. *Marketing Science, 17*(1), 45–65.

Busacca, B., & Padula, G. (2005). Understanding the relationship between attribute performance and overall satisfaction: Theory, measurement and implications. *Marketing Intelligence & Planning, 23*(6), 543–561. doi:10.1108/02634500510624110.

Carrión, G. C., Henseler, J., Ringle, C. M., & Roldán, J. L. (2016). Prediction-oriented modeling in business research by means of PLS path modeling: Introduction to a JBR special section. *Journal of Business Research, 69*(10), 4545–4551.

Carte, T. A., & Russell, C. J. (2003). In pursuit of moderation: Nine common errors and their solutions. *MIS Quarterly, 27*(3), 479–501.

Cenfetelli, R. T., & Bassellier, G. (2009). Interpretation of formative measurement in information systems research. *MIS Quarterly, 33*(4), 689–707.

Conklin, M., Powaga, K., & Lipovetsky, S. (2004). Customer satisfaction analysis: Identification of key drivers. *European Journal of Operational Research, 154*(3), 819–827.

Dagger, T. S., Sweeney, J. C., & Johnson, L. W. (2007). A hierarchical model of health service quality scale development and investigation of an integrated model. *Journal of Service Research, 10*(2), 123–142.

Dawes, J. G. (2008). Do data characteristics change according to the number of scale points used? An experiment using 5 point, 7 point and 10 point scales. *International Journal of Market Research, 51*(1), 61–104.

Dijkstra, T. K., & Henseler, J. (2011). Linear indices in nonlinear structural equation models: Best fitting proper indices and other composites. *Quality and Quantity, 45*(6), 1505–1518. doi:10.1007/s11135-010-9359-z.

Dong, S., Ding, M., Grewal, R., & Zhao, P. (2011). Functional forms of the satisfaction–loyalty relationship. *International Journal of Research in Marketing, 28*(1), 38–50. doi:10.1016/j.ijresmar.2010.09.002.

Fassott, G., Henseler, J., & Coelho, P. S. (2016). Testing moderating effects in PLS path models with composite variables. *Industrial Management & Data Systems, 116*(9), 1887–1900. doi:10.1108/IMDS-06-2016-0248.

Finn, A. (2011). Investigating the non-linear effects of e-service quality dimensions on customer satisfaction. *Journal of Retailing and Consumer Services, 18*(1), 27–37.

Fishbein, M. (1967). A consideration of beliefs and their role in attitude measurement. In M. Fishbein (Ed.), *Readings in attitude theory and measurement* (pp. 257–266). New York: Wiley.

Fornell, C., Johnson, M. D., Anderson, E. W., Cha, J., & Bryant, B. E. (1996). The American customer satisfaction index: Nature, purpose, and findings. *Journal of Marketing, 60*(4), 7–18.

Garbarino, E., & Johnson, M. S. (1999). The different roles of satisfaction, trust, and commitment in customer relationships. *Journal of Marketing, 63*(2), 70–87.

Gomez, M. I., McLaughlin, E. W., & Wittink, D. R. (2004). Customer satisfaction and retail sales performance: An empirical investigation. *Journal of Retailing, 80*(4), 265–278.

Hair, J. F., Jr., Hult, G. T. M., Ringle, C., & Sarstedt, M. (2014). *A primer on partial least squares structural equation modeling (PLS-SEM)*. Thousand Oaks, CA: Sage.

Henseler, J., & Chin, W. W. (2010). A comparison of approaches for the analysis of interaction effects between latent variables using partial least squares path modeling. *Structural Equation Modeling, 17*(1), 82–109. doi:10.1080/10705510903439003.

Henseler, J., Fassott, G., Dijkstra, T. K., & Wilson, B. (2012). Analysing quadratic effects of formative constructs by means of variance-based structural equation modelling. *European Journal of Information Systems, 21*(1), 99–112. doi:10.1057/ejis.2011.36.

Hock, C., Ringle, C. M., & Sarstedt, M. (2010). Management of multi-purpose stadiums: Importance and performance measurement of service interfaces. *International Journal of Services, Technology and Management, 14*(2–3), 188–207.

Johnson, R. E., Kiokemeister, F. L., & Wolk, E. S. (1978). *Calculus with analytic geometry* (6th ed.). Boston: Allyn and Bacon.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*(2), 263–291.

Kamakura, W. A., Mittal, V., De Rosa, F., & Mazzon, J. A. (2002). Assessing the service-profit chain. *Marketing Science, 21*(3), 294–317.

Kano, N., Seraku, N., Takahashi, F., & Tsuji, S. (1984). Attractive quality and must-be quality. *Hinshitsu: The Journal of the Japanese Society for Quality Control, 14*(2), 39–48.

Lance, C. E. (1988). Residual centering, exploratory and confirmatory moderator analysis, and decomposition of effects in path models containing interactions. *Applied Psychological Measurement, 12*(2), 163–175.

Langfred, C. W. (2004). Too much of a good thing? Negative effects of high trust and individual autonomy in self-managing teams. *The Academy of Management Journal, 47*(3), 385–399.

Leroi-Werelds, S., Streukens, S., Brady, M. K., & Swinnen, G. (2014). Assessing the value of commonly used methods for measuring customer value: A multi-setting empirical study. *Journal of the Academy of Marketing Science, 42*(4), 430–451. doi:10.1007/s11747-013-0363-4.

Little, T. D., Bovaird, J. A., & Widaman, K. F. (2006). On the merits of orthogonalizing powered and product terms: Implications for modeling interactions among latent variables. *Structural Equation Modeling, 13*(4), 497–519.

Lu, J. W., & Beamish, P. W. (2004). International diversification and firm performance: The S-curve hypothesis. *The Academy of Management Journal, 47*(4), 598–609.

Malhotra, N. K., Birks, D. F., & Wills, P. (2012). *Marketing research: An applied approach*. Harlow: Financial Times.

Marcoulides, G. A., Chin, W. W., & Saunders, C. (2009). A critical look at partial least squares modeling. *MIS Quarterly, 33*(1), 171–175.

Martilla, J. A., & James, J. C. (1977). Importance-performance analysis. *Journal of Marketing, 41*, 77–79.

Matzler, K., Bailom, F., Hinterhuber, H. H., Renzl, B., & Pichler, J. (2004). The asymmetric relationship between attribute-level performance and overall customer satisfaction: A reconsideration of the importance–performance analysis. *Industrial Marketing Management, 33*(4), 271–277.

Mikulić, J., & Prebežac, D. (2009). *Developing service improvement strategies underconsideration of multicollinearity and asymmetries in loyalty intentions – A study from the airline industry* (pp. 414–423). Proceedings of the 11th International Conference QUIS – The Service Conference, Wolfsburg, Germany.

Mikulić, J., & Prebežac, D. (2011). A critical review of techniques for classifying quality attributes in the Kano model. *Managing Service Quality, 21*(1), 46–66.

Miltgen, C. L., Henseler, J., Gelhard, C., & Popovič, A. (2016). Introducing new products that affect consumer privacy: A mediation model. *Journal of Business Research, 69*(10), 4659–4666.

Mittal, V., & Katrichis, J. M. (2000). Distinctions between new and loyal customers. *Marketing Research, 12*(1), 26–32.

Mittal, V., Ross, W. T., & Baldasare, P. M. (1998). The asymmetric impact of negative and positive attribute-level performance on overall satisfaction and repurchase intentions. *Journal of Marketing, 62*, 33–47.

Mittal, V., Kumar, P., & Tsiros, M. (1999). Attribute-level performance, satisfaction, and behavioral intentions over time: A consumption-system approach. *Journal of Marketing, 63*(2), 88–101.

Narasimhan, R., & Kim, S. W. (2002). Effect of supply chain integration on the relationship between diversification and performance: Evidence from Japanese and Korean firms. *Journal of Operations Management, 20*(3), 303–323.

Nitzl, C., Roldan, J. L., & Cepeda, G. (2016). Mediation analysis in partial least squares path modeling. *Industrial Management & Data Systems, 116*(9), 1849–1864. doi:10.1108/IMDS-07-2015-0302.

Oliver, R. L. (1999). Whence consumer loyalty? *Journal of Marketing, 63*, 33–44. doi:10.2307/1252099.

Preston, C. C., & Colman, A. M. (2000). Optimal number of response categories in rating scales: Reliability, validity, discriminating power, and respondent preferences. *Acta Psychologica, 104*(1), 1–15.

Roncek, D. W. (1991). Using logit coefficients to obtain the effects of independent variables on changes in probabilities. *Social Forces, 70*(2), 509–518.

Roncek, D. W. (1993). When will they ever learn that first derivatives identify the effects of continuous independent variables or "Officer, you can't give me a Ticket, I wasn't Speeding for an entire hour". *Social Forces, 71*(4), 1067–1078.

Rigdon, E. E. (2014). Rethinking partial least squares path modeling: Breaking chains and forging ahead. *Long Range Planning, 47*(3), 161–167. doi:10.1016/j.lrp.2014.02.003.

Rigdon, E. E., Ringle, C. M., Sarstedt, M., & Gudergan, S. P. (2011). Assessing heterogeneity in customer satisfaction studies: Across industry similarities and within industry differences. *Advances in International Marketing, 22*(1), 169–194.

Ringle, C. M., & Sarstedt, M. (2016). Gain more insight from your PLS-SEM results: The importance-performance map analysis. *Industrial Management & Data Systems, 116*(9), 1865–1886.

Ringle, C. M., Wende, S., & Becker, J. (2015). *SmartPLS 3*. Bönningstedt: SmartPLS. Retrieved from http://www.smartpls.com

Russell, C. J., & Bobko, P. (1992). Moderated regression analysis and Likert scales: Too coarse for comfort. *The Journal of Applied Psychology, 77*(3), 336–342.

Rust, R. T., & Oliver, R. L. (2000). Should we delight the customer? *Journal of the Academy of Marketing Science, 28*(1), 86–94.

Sciascia, S., & Mazzola, P. (2008). Family involvement in ownership and management: Exploring nonlinear effects on performance. *Family Business Review, 21*(4), 331–345. doi:10.1177/08944865080210040105.

Seggie, S. H., Cavusgil, E., & Phelan, S. E. (2007). Measurement of return on marketing investment: A conceptual framework and the future of marketing metrics. *Industrial Marketing Management, 36*(6), 834–841.

Seiders, K., Voss, G. B., Grewal, D., & Godfrey, A. L. (2005). Do satisfied customers buy more? Examining moderating influences in a retailing context. *Journal of Marketing, 69*(4), 26–43. doi:10.1509/jmkg.2005.69.4.26.

Shmueli, G., Ray, S., Estrada, J. M. V., & Chatla, S. B. (2016). The elephant in the room: Predictive performance of PLS models. *Journal of Business Research, 69*(10), 4552–4564.

Stolzenberg, R. M. (1980). The measurement and decomposition of causal effects in nonlinear and nonadditive models. *Sociological Methodology, 11*, 459–488. doi:10.2307/270872.

Strauss, A., & Corbin, J. (1990). *Basics of qualitative research* (Vol. 15). Newbury Park, CA: Sage.

Streukens, S., & De Ruyter, K. (2004). Reconsidering nonlinearity and asymmetry in customer satisfaction and loyalty models: An empirical study in three retail service settings. *Marketing Letters, 15*(2–3), 99–111.

Streukens, S., & Leroi-Werelds, S. (2016). Bootstrapping and PLS-SEM: A step-by-step guide to get more out of your bootstrap results. *European Management Journal, 34*, 618–632.

Streukens, S., van Hoesel, S., & de Ruyter, K. (2011). Return on marketing investments in B2B customer relationships: A decision-making and optimization approach. *Industrial Marketing Management, 40*(1), 149–161.

Tenenhaus, M., Esposito Vinzi, V., Chatelin, Y.-M., & Lauro, C. (2005). PLS path modeling. *Computational Statistics and Data Analysis, 48*(1), 159–205. doi:10.1016/j.csda.2004.03.005.

Titah, R., & Barki, H. (2009). Nonlinearities between attitude and subjective norms in information technology acceptance: A negative synergy? *MIS Quarterly, 33*(4), 827–844.

Völckner, F., Sattler, H., Hennig-Thurau, T., & Ringle, C. M. (2010). The role of parent brand quality for service brand extension success. *Journal of Service Research, 13*(4), 379–396.

Weijters, B., Cabooter, E., & Schillewaert, N. (2010). The effect of rating scale format on response styles: The number of response categories and response category labels. *International Journal of Research in Marketing, 27*(3), 236–247.

Zhu, Z., Sivakumar, K., & Parasuraman, A. (2004). A mathematical model of service failure and recovery strategies. *Decision Sciences, 35*(3), 493–525.

Author Index

A
Aguirre-Urreta, M., 27
Albers, S., 20, 31, 85
Amato, S., 29
Antonakis, J., 24

B
Beal, D. J., 22
Becker, J.-M., 33, 34, 87, 88, 153, 156–159, 198, 199, 206, 207, 211, 221, 261, 303, 379
Bendahan, S., 24
Benitez-Amado, J., 26
Bentler, P. M., 23, 28, 29, 55, 75, 124, 127, 282, 284
Bernstein, I. H., 30, 46, 47
Bollen, K. A., 28, 113, 124, 260, 283–285
Bonett, D. G., 29
Bookstein, F. L., 157, 283, 290
Borgelt, C., 30
Brady, M. K., 30
Braojos-Gomez, J., 26
Brosseau-Liard, P. E., 26
Buckler, F., 33
Byrne, B. M., 28, 29, 198

C
Calantone, R. J., 87
Cepeda Carrión, G., 27
Chatelin, Y.-M., 22, 23, 27, 29, 33, 67, 87, 111–114, 126, 127, 135, 165, 300, 589
Chen, Q., 31, 178, 179
Chen, Y., 26, 364
Chin, W. W., 32, 35, 90, 126, 157, 160, 161, 163, 167, 181, 188, 211, 219, 220, 223, 300, 382, 383
Cohen, J., 31, 116, 246, 274

D
Daryanto, A., 26
De Ruyter, K., 380, 384, 388
Diamantopoulos, A., 25, 85, 198
Dijkstra, T. K., 20–23, 25, 28, 30, 32, 33, 55–77, 110, 112, 116, 117, 124, 126, 127, 133, 134, 136, 156, 240, 264, 291, 325, 389

E
Esposito Vinzi, V., 29, 90, 153–170
Evermann, J., 153, 154, 156, 160, 163, 164, 167

F
Fahrettin Demirel, O., 33
Fassott, G., 32, 382
Finch, W. H., 33
Fink, A., 33
Fornell, C., 30, 125, 126, 157, 283, 290, 371
French, B. F., 33
Fuchs, C., 25, 85, 198

G
Gaul, W., 30
Gill, R. D., 23
Goodhue, D. L., 27

H. Latan, R. Noonan (eds.), *Partial Least Squares Path Modeling*,
DOI 10.1007/978-3-319-64069-3

H
Haenlein, M., 300
Hair, J. F., 20, 34, 88, 110, 111, 153, 155, 157, 174, 183, 197–214, 220, 221, 223, 224, 226, 227, 232, 235, 239, 240, 282, 283, 300, 305, 306, 323, 324, 389
Halpin, P. F., 21, 113
Hanafi, M., 158, 261
Hennig-Thurau, T., 33
Henseler, Jörg, 19–35, 55, 56, 69, 71, 72, 87, 88, 91, 110–112, 116, 117, 124–127, 130, 133, 136, 156, 157, 173–175, 183, 210, 221, 223, 224, 227, 232–235, 239, 240, 264, 278, 291, 300, 323–325, 328, 369, 382, 383, 389
Höök, K., 20, 24
Hu, L.-T., 28, 124, 127
Huang, W., 23
Hubona, G., 19–35
Hwang, H., 33, 136

J
Jacquart, P, 24

K
Kaiser, S., 25, 85
Kettenring, J. R., 23, 63–65, 112
Ketterlinus, R. D., 21, 113
Koppius, O. R., 27, 154
Kou, G., 26
Krijnen, W. P., 23
Kruse, R., 30

L
Lalive, R., 24
Lamb, M. E., 21
Lancelot-Miltgen, C., 27
Landis, R. S., 22
Larcker, D. F., 30, 125, 130
Lauro, C., 85, 153–170
Lausen, B., 33
Lewis, W., 312
Llorens-Montes, F. J., 26
Lohmöller, J.-B., 12, 29, 42, 47, 87, 90, 110, 114, 156, 157, 159, 161, 162, 198, 201, 225, 239, 260, 282, 299
Löwgren, J., 20, 24
Lynch, J. G., 31, 178, 179

M
Maraun, M. D., 21, 113
Marcoulides, G. A., 20, 380
McDonald, R. P., 20, 24
Mena, J. A., 20, 34, 88, 110, 111, 153, 155, 157, 174, 178, 183, 198, 199, 201, 203–205, 209, 211, 212, 220, 221, 223, 224, 226, 227, 232, 235, 239, 240, 283, 300, 305, 306, 323, 324, 333, 389
Mooi, E. A., 203, 209, 239

N
Nevo, S., 26
Noonan, R. D., 3–17, 114, 283, 290
Nunnally, J. C., 30, 46, 47
Nürnberger, A., 30

O
Oztekin, A., 33

P
Pieper, T. M., 20, 34, 88, 110, 111
Popovič, A., 27

Q
Qannari, M., 30

R
Rai, A., 33, 34
Ramirez, E., 30
Ray, P. A., 19–35
Reinartz, W. J., 27
Rhemtulla, M., 26
Rigdon, E. E., 20–22, 24, 25, 29, 88, 90, 110, 113, 154, 156–159, 193, 198, 204, 206, 210, 212, 325, 368
Rindskopf, D., 25
Ringle, C. M., 87, 90, 111, 197–212, 221, 283, 290, 305, 325, 330, 369, 371, 389, 393, 394, 396, 398
Roldán, J. L., 173–193, 282
Rönkkö, M., 27, 127, 156, 157

S
Sahmer, K., 30
Sampson, P. D., 21

Sarstedt, M., 20, 29, 33, 88, 90, 154–158, 197–212, 219, 220, 224, 225, 239, 240, 291, 327, 330, 368, 369, 371, 372, 379, 389, 398
Saunders, C., 20
Savalei, V., 127
Schermelleh-Engel, K., 32, 77, 117
Schlittgen, R., 33, 199, 207
Schuberth, F., 26, 109–146, 264
Seidel, W., 33
Shmueli, G., 27, 74, 125, 136, 154, 155, 157, 160, 163, 165, 170, 175, 376, 393
Sijtsma, K., 30, 46
Spiliopoulou, M., 30
Stine, R. A., 28, 124, 283–285, 288
Straub, D. W., 87, 90, 111
Streukens, S., 26, 330, 367–400

T
Takane, Y., 33
Tenenhaus, A., 33, 158, 262
Tenenhaus, M., 22, 23, 27, 29, 33, 67, 87, 112–114, 126, 127, 135, 158, 165, 262, 300, 371, 389
Tesluk, P. E., 22
Thompson, R. L., 27
Trinchera, L., 117, 259–278
Turkyilmaz, A., 33

U
Ultsch, A., 33

V
Völckner, F., 198, 368
Voorhees, C. M., 30

W
Wang, H., 20, 29, 31
Wang, Y., 26
Wende, S., 33, 34, 90
Wetzels, M., 87
Wilczynski, P., 25, 85
Will, A., 33
Wilson, B., 87, 88, 240
Wold, H. O. A., 4, 10, 12, 13, 42, 47, 55, 56, 58, 62, 63, 76, 77, 87, 112–114, 125, 127, 157, 159, 163, 164, 198, 260, 261, 282, 283, 290, 299

Z
Zaim, S., 33
Zhao, X., 31, 178
Ziggers, G.-W., 29

Subject Index

A

Advanced analysis of composites (ADANCO), 19, 33, 55, 127
Akaike's information criterion (AIC)
  consistent, 204, 205
  modified AIC with factor 3, 204, 205
Alternating least squares (ALS) 21, 56
Analysis of variance (ANOVA), 117, 286
Average variance extracted (AVE), 30, 34, 125, 132, 144, 226–228, 253, 305, 307, 323, 324

B

Bayesian information criterion (BIC), 204, 205
Benessere Equo e Sostenibile (BES), 82–94, 97, 99, 105
Bentler-Bonett index, 29
Best fitting proper indices (BFPI), 389
Bias
  common method bias (CMB), 51, 245–256, 321
  small sample bias, 73
Blindfolding, 27, 34, 163, 300, 393
  Stone-Geisser Q2, 163
Bootstrap
  Bollen-Stine, 28, 283, 285, 286, 288–290
  Naïve, 283–284, 289
  sample, 28, 31, 129, 131, 142, 143, 177, 271, 283, 284
  semi-parametric, 284
  test, 124, 176, 303

C

Canonical correlation
  generalized canonical variables, 57, 63
  variable, 56, 63
  variables analysis, 56, 57
Categorical variables, 21, 26, 110, 117–120, 221, 222, 225, 227, 231, 235, 240, 271
CBSEM. *See* Covariance-based SEM (CBSEM)
Centroid scheme, 98, 114, 127, 260, 261, 271
Chemometrics, 341, 353
Classification, 83, 201, 209, 266, 271, 388
  probabilistic, 201
Clustering, 198
Clusterwise regression, 207
Collinearity
  multicollinearity, 9, 31, 85, 159, 161, 325, 377, 379, 383, 389, 390
  test, 246, 253–256
Commonality analysis
  joint contribution, 6, 8
  unique contribution, 6, 16
Common factor model, 42, 47, 50–52, 113, 121, 128, 130, 135, 156, 264
Component-based path modeling, 154–157, 160, 166–169
Composite model, 21, 22, 24–27, 30, 33, 34, 87–90, 113, 115, 116, 121, 201, 210, 221, 222, 291, 328
  factor, 21, 25, 27

H. Latan, R. Noonan (eds.), *Partial Least Squares Path Modeling*,
DOI 10.1007/978-3-319-64069-3

Compositional invariance, 210, 222, 232, 233, 240
 permutation test, 232
Configural invariance, 210, 222, 227
Consistency
 consistent and asymptotically normal (CAN) estimator, 62, 63, 156
 at large, 56
Corporate social responsibility, 312
Correction for attenuation, 22, 23, 116, 120–122, 136
Correlation
 Bravais-Pearson (BP), 118, 130, 133, 134
 polychoric, 118, 120, 126, 127, 133, 136, 264
 polyserial, 118–120, 136
Correspondence analysis (CA), 117
Covariance-based SEM (CBSEM), 20, 27, 28, 42, 156, 256, 282, 290, 325
Critical incident technique (CIT), 37, 378
Cronbach's alpha, 30, 34, 46
 ordinal, 133
Cross tabs, 209

D
Diagnostics
 regression diagnostics, 353–361
 visual diagnostics, 339
Discrepancy
 geodesic discrepancy (dG), 28
 maximum likelihood discrepancy, 28
 unweighted least squares discrepancy (d_{ULS}), 28, 32

E
Econometric methods, 57, 69
EEOS. *See* Equality of educational opportunity survey (EEOS)
Efficiency, 24, 44, 50, 56, 60, 62, 70, 77, 283, 286, 290, 291
Emergent constructs, 24
Emotion, 313, 316–319, 322–325, 327, 328, 331, 332, 375
Entropy statistic EN, 205
EQS program, 56, 75
Equality of educational opportunity survey (EEOS), 6–9
Error
 type I, 177, 224, 251, 283, 285, 287
 type II, 177, 223, 251
Ethical decision making, 312–317, 319, 328, 332, 333
Euclidean distance, 28, 116
Expectation-maximization (EM)
 algorithm, 202, 205, 206, 246, 255, 256
 imputation, 203
Explanation
 explanatory modeling, 154
 explanatory variables, 10, 60, 67, 198, 199, 201, 203, 208, 209, 211
Ex post analysis, 199, 208–209

F
Factor analysis
 confirmation, 254
 confirmatory, 29, 251–253
 model, 29, 32
Factor-factor links, 254
Factorial scheme, 98, 260, 261
Factor-indicator links, 247
FIML. *See* Full information maximum likelihood (FIML)
First International Mathematics Study (FIMS), 4, 7, 8
Fix-Point (FP) approach, 10
Fornell-Larcker criterion, 30, 32, 125, 130, 133, 324
Full information maximum likelihood (FIML), 72
Functional form
 concave relationship, 375
 convex relationship, 375
 inverse S-shaped, 375
 linear, 374–376, 382, 386, 389, 390, 393, 396
 linear-log, 380
 log-linear, 380, 386
 log-log, 388
 polynomial, 380, 382
 S-shaped, 375
Fuzzy segment membership, 205

G
Generalized canonical correlation analysis (GCCA), 112, 261, 262
Generalized least squares (GLS), 75
GENVAR program, 65
Geodesic distance, 75
 discrepancy, 28
Goodness of fit (GoF), 28, 29, 57, 75–76, 96, 97, 99, 161, 162, 325, 327
 relative Goodness of fit (GoFrel), 29

H
Heterogeneity
identifying unobserved, 90, 198
treating unobserved, 197–212
unobserved, 33, 90–91, 101, 197–213, 239, 240, 333
Heterotrait-monotrait ratio (HTMT), 19–21, 30, 32, 34, 110, 125, 133, 141, 227, 230, 306, 324
Heuristics, 206
Heywood case, 23
Hierarchical models
composite, 87–89
construct, 82
structured latent variables, 11–13, 15
HTMT. *See* Heterotrait-monotrait ratio (HTMT)
Hyperspectral
images, 339–364
imaging, 343, 350, 364

I
Importance-performance map analysis (IPMA), 122, 210, 330, 367–400
Indirect effects, 11, 23, 32, 175–188, 191, 192, 306, 317, 328, 372
Information criteria, 204, 205, 208
Instrumental variable techniques, 23
Interaction effects, 32, 33, 331–332
Internal consistency
Dijkstra & Henseler's rho, 32
Dillon-Goldstein's rho, 32
ISTAT. *See* Italian National Statistical Institute (ISTAT)
Italian National Statistical Institute (ISTAT)., 82, 266

K
K-means, 198

L
Latent class, 90, 198, 199, 206–209, 211, 212
analysis, 199, 206–208, 211
Latent variable, 10–13, 15, 20, 41, 43, 46, 56, 58, 72, 76, 77, 122, 126, 156, 173, 201, 207, 209, 246–255, 260, 262–264, 269, 285, 300, 303, 305, 306, 308, 325, 328, 339–364, 371, 377, 382, 389
score, 72, 201, 246, 371, 382
Linear structural relations (LISREL), 56, 75, 77
Local optimum, 202, 206, 207

M
Market segmentation, 206
Maximum likelihood (ML), 23, 48, 72, 117, 120, 282–286, 290, 291
ML-SEM, 282, 283, 290, 291
MAXVAR program, 65, 71, 74
Measurement model
formative constructs, 389
full measurement invariance, 224, 235
invariance, 33
measurement invariance assessment, 91
measurement invariance of composite models (MICOM), 201, 209–210, 221, 222, 224, 227, 232–235, 239, 240, 328, 329
partial measurement invariance, 210, 224
reflective constructs, 157
repeated indicators approach, 87
Mediation
analysis, 31, 174–182, 193
competitive partial mediation, 178, 179
complementary mediation, 178–180
inconsistent mediation, 181
Mediation effect, 174–177, 184, 186–188
MIMIC. *See* Multiple indicators, multiple causes (MIMIC)
Minimum description length (MDL), 204
MINVAR program, 65
Missing values
casewise deletion, 203
EM imputation, 203
mean value replacement, 203
missing at random, 203
Mode
mode A, 11, 25, 34, 50, 51, 56, 57, 65–67, 85, 88, 99, 114, 115, 127, 157–160, 165–167, 182, 190, 246, 260–262, 265, 270, 271, 286, 300
mode B, 11, 25, 34, 56, 57, 63–67, 71, 74, 88, 99, 114–116, 157–160, 165–167, 261, 262, 265, 270, 271, 300, 305
mode C, 56, 65, 114, 115
mode Q, 99
new mode A, 261, 262, 265
Model
approximate model fit, 21, 28, 29, 32, 124
assessment, 27, 31, 193, 325
complexity, 90, 199, 203, 255

fit, 20, 21, 27–30, 32, 34, 42, 124, 125, 127, 136, 204, 306, 327
identification, 42
inner, 21, 87, 155, 165, 260, 271
latent factor, 56
outer, 21, 29, 99, 159–160, 260
saturated, 28, 30, 32
Monte Carlo, 27, 28, 49–51, 59, 69, 71–72, 166, 170, 246, 250, 283, 285
simulation, 27, 28, 51, 170, 246, 250
Multicollinearity, 9, 31, 85, 159, 161, 325, 377, 379, 383, 389, 390
Multigroup analysis
parametric approach, 282
partial least squares structural equation modeling multigroup analysis (PLS-MGA), 211, 220, 224–225, 239
permutation test, 90, 211, 224
Multi-method approach, 197–212
Multiple indicators, multiple causes (MIMIC), 34, 115, 325
Multivariate image, 342

N
Nagelkerke R^2, 273, 275, 276
Near Infrared (NIR), 342, 343, 350
spectra, 340
Neural networks, 33
Newton-Raphson method, 206
NILES. *See* Nonlinear iterative least squares (NILES)
Nomological network, 376, 391, 399
Nonlinear iterative least squares (NILES), 10
Nonlinear iterative partial least squares (NIPALS), 10, 299
Nonlinear relationships, 32, 369, 373–375, 393
Non-Metric PLS path modeling (NM-PLSPM), 260, 264, 265, 271, 278
Non-recursive model, 25, 116, 117
Non-symmetrical composite-based path modeling (NSCPM), 155, 165–170
Normed fit index (NFI), 29, 327
NSCPM. *See* Non-symmetrical composite-based path modeling (NSCPM)

O
Omnibus test of group differences, 211
Optimal scaling (OS), 117, 118, 264, 271, 272
Ordinal
consistent PLS (OrdPLSc), 26, 111–113, 120–122, 124–137, 141
PLS (OrdPLS), 110–112, 118, 120–124, 126–137, 140
Ordinary least squares (OLS), 23, 31, 96–98, 110, 112, 115, 116, 118, 126, 159, 260, 261, 340
Orthogonalization, 28, 331, 384, 385, 390
orthogonalizing approach, 383

P
Partial Least Squares (PLS)
consistent PLS (PLSc), 20, 23, 46, 55, 57, 77, 110–113, 116, 118, 120–122, 124–139, 141–146, 240, 325
discriminant analysis (PLSDA), 360–364
factor-based PLS-SEM (PLSF-SEM), 42, 44, 46–52
finite mixture PLS (FIMIX-PLS), 198, 199, 201–208, 211, 212, 400
genetic algorithm segmentation (PLS-GAS), 33, 99, 199, 207, 208, 212, 240
latent variable path analysis (LVPLS), 08, 299, 300, 305
multigroup analysis (PLS-MGA), 211, 220, 224–239, 327, 329
non-metric PLS (NM-PLS), 117, 118, 260, 264
ordinal consistent PLS (OrdPLSc), 111
partial maximum likelihood PLS (PML-PLS), 117
prediction oriented segmentation (PLS-POS), 33, 199, 207–209, 212, 240
quantile composite-based path modeling (QC-PM), 82, 83, 96–103
response-based procedure for detecting unit segments (REBUS-PLS), 207
semPLS, 286
SmartPLS, 25, 199, 207, 208, 212, 221, 224, 227, 234, 235, 305, 307, 325, 371, 389, 393, 394, 396
SmartPLS 3, 199, 207, 208, 235, 305, 307, 325, 371, 389, 393, 394, 396
structural equation model (PLS-SEM), 41–52, 173–183, 190, 193, 197–212, 219–240, 245–256, 282–293, 299–308, 325, 330, 367–400
typological path modeling, 207
WarpPLS, 42, 49, 52, 246, 256
Path directions incoherence, 155, 165
Penalty-reward contrast analysis (PRCA), 377, 380–381, 390, 392
Performance metrics, 160–164

Prediction
in-sample prediction, 162
out-of-sample prediction, 136, 158–160, 162–164, 393
prediction test, 74
predictive modelling, 154–156, 353
predictive performance, 154, 155, 157, 163, 393
predictive relevance, 155, 160, 163, 168, 327
redundancy-based prediction, 90, 160, 161, 164, 170
Principal component analysis (PCA), 21, 33, 339, 347, 348, 358, 361, 364
Product-indicator approach, 383, 386
Professional accountants, 312, 320, 328, 330, 332
Pseudo factors, 42
Pseudo R^2, 97

Q
Quadratic effects, 32, 382, 383
Quantile composite-based path modeling (QC-PM), 82, 96, 98–103
Quantile correlation (QC), 82, 96, 98, 233
Quantile regression (QR), 82, 96–99, 101, 102

R
Rank condition, 59, 68
rank constraints, 57
Recursive model, 59
Redundancy index, 89, 90, 161, 167, 168
Regression
clusterwise robust regression, 207
diagnostics, 353–359
logistic, 209, 271–273, 276
mixture, 201
multiple linear regression (MLR), 157, 159, 340
multivariate, 114, 339–341
ordinary least squares (OLS), 23, 96, 110, 126, 159, 260, 261, 340
quantile regression (QR), 82, 96–99, 101, 102
robust, 207
trees, 209
Reliability
composite, 30, 34, 46, 111, 126, 305, 323
Cronbach's alpha, 30, 34, 46
Dillon-Goldstein's rho, 30
Jöreskog's rho, 30
omega, 30
rhoA, 21, 30, 34
rho_c, 30
Resource allocation, 13, 373, 393, 400
Robustness, 24, 62, 290
ROC Curve, 274, 276, 277
Root Mean Square (RMS)
root mean square difference (RMSD), 286–290, 292–295
root mean square error correlation (RMStheta), 29

S
Sample
calibration sample, 340, 355
holdout samples, 27, 393
sampling issues, 346–347
saturation rule, 378
size, 21, 26, 27, 31, 42, 46, 49–51, 70, 74, 158, 159, 163, 168, 179, 199, 203, 204, 221, 223, 226, 232, 240, 283–295, 377, 380, 394, 398
split, 377, 393, 398
Satisfaction-profit chain (SPC), 368, 369, 373–376, 378, 384, 399, 400
Segmentation
genetic algorithm, 33, 99, 207, 208, 212, 240
iterative reweighted regressions, 199, 207, 208, 212, 240
maximum number of segments, 204
number of segments to retain, 198, 204, 206, 208
over-segmentation, 204
prediction-oriented, 33, 199, 207–209, 212, 240
profiling of segments, 208
response-based procedure for detecting unit segments (REBUS-PLS), 207
segment retention criteria, 204
under-segmentation, 204
Sign indeterminacy, 21, 34
Soft modeling, 55, 56, 58, 62, 76, 276, 299, 300
Standardized root mean square residual (SRMR), 21, 28, 32, 124, 125, 127, 306, 327
Starting value, 202, 206
Stone-Geisser Q^2
blindfolding, 163
predictive relevance, 163, 327
Stop criterion, 202, 232
Strong concepts, 20, 24

Structural equation model (SEM)
covariance-based SEM (CBSEM), 282, 290, 291
factor-based SEM, 42
structural model, 300
Suppressor effect, 9, 15

T
Three-stage least squares (3SLS), 57, 70, 72
True score theory, 24
Two-stage least squares (2SLS), 23, 56, 60, 69–72, 112, 116

V
Validation
blindfolding, 163, 300, 393
cross-validation, 57, 74, 136, 163, 164, 377, 393
k-fold cross-validation, 136, 377, 393
resampling, 97
sample, 393, 398
split sample, 377, 393, 398
Validity
assessment, 246
content, 324, 379
convergent, 30, 32, 125, 130, 227, 251–253, 305
discriminant validity, 20, 30, 32, 34, 110, 125, 130, 133, 227, 251–253, 306, 324
internal consistency, 32, 125–126, 133, 134, 145
Variance inflation factor (VIF), 31, 253–256, 305, 307, 327, 389
Variation sharing, 48
Visual diagnostics, 360

W
Weighted least squares (WLS), 75
Whistleblowing, 311–333