Latent Path Models within an IRT Framework

By

Eveline Gebhardt

Dissertation submitted in total fulfillment
of the requirements for the degree of Doctor of Philosophy

February 2016

Melbourne Graduate School of Education
University of Melbourne

Produced on archival quality paper
Abstract

In this thesis a new method is described that includes latent path models within an item response theory (IRT) framework. That is, the latent variables in a path model are estimated in combination with IRT measurement models. Thus far, researchers using IRT modelling had to produce individual person latent trait estimates with IRT software, and subsequently import these into other software packages to link them to other data sources and to perform secondary analysis such as path model analysis. Besides the impracticality of this method, different types of individual person ability estimates often lead to different—and therefore sometimes biased—results in secondary analysis. The approach that is presented in this thesis overcomes both limitations. A multi-step method is introduced in which the IRT model is estimated first. Subsequently, a full sums of squares and cross products (SSCP) matrix is constructed for all the latent variables from the population distribution parameter estimates of the first step (regression coefficients and conditional variance and covariance estimates) and of the observed variables included in the path model. Finally, the SSCP matrix is used to perform two-stage least squares for the estimation of the path model parameters. A simulation study shows that the parameter estimates are practically unbiased and that the standard errors are not worse than standard errors produced by methods employed to estimate latent path models within a structural equation modelling framework, but that their estimation can be improved in future. Another improvement would be to add the computation of fit statistics of the model. The thesis concludes with two examples in which the multi-step method is applied to test hypothetical models using real data.
Declaration

This is to certify that

I. the thesis comprises only my original work towards the PhD

II. due acknowledgement has been made in the text to all other material used

III. the thesis is fewer than 100,000 words in length, exclusive of tables

Candidate’s signature

____________________________________________

Eveline Gebhardt
Dedication

Aan mijn ouders
Acknowledgments

I would like to thank so many people for different reasons. First of all, I would like to thank my beautiful children for giving me such a happy life-work balance that I would have finished my thesis much earlier without them. I would also like to thank my supervisors, Ray Adams and Siek Toon Khoo, for being my teachers and helping me through uncountable moments of despair; the Australian Council for Educational Research for supporting my study; my brother for teaching me maths when we were at school; my friends for gently pushing me out of many lulls; and last but not least my wonderful partner for giving me that final and most important push to finish after supporting me all the way through for 10 years.
# Table of Contents

List of Tables................................................................................................................................................. iv

List of Figures.................................................................................................................................................. v

Chapter 1: Introduction....................................................................................................................................... 1

Chapter 2: IRT Measurement .......................................................................................................................... 10
  A Brief History of IRT ................................................................................................................................... 10
  The MRCML Model..................................................................................................................................... 15
  Estimating IRT Models ................................................................................................................................. 17
  IRT and Factor Analytic Models .................................................................................................................. 22
    Derivation of the Birnbaum Model ........................................................................................................... 23
    The Rasch Model....................................................................................................................................... 25
    The Binary Factor Analytic Model ............................................................................................................ 26
    Transformation from Rasch to Binary Factor Analytic Parameters ....................................................... 28
    Matching Models....................................................................................................................................... 35
    IRT versus Factor Analytic Models in Practice ....................................................................................... 37
  Summary......................................................................................................................................................... 39

Chapter 3: Path Models .................................................................................................................................... 41
  Simultaneous Equations ................................................................................................................................. 41
  Reduced Form Equations ............................................................................................................................... 46
  Matrix Permutation and Model Identification ............................................................................................... 47
    An Illustration............................................................................................................................................. 51
List of Tables

Table 1. Transformations between IRT and Factor Analytic Item Parameters for Four Types of Parameterisations .......................................................... 32
Table 2. Two-Step Transformations between the MRCML Rasch Parameters and Binary Factor Analytic Parameters .......................................................... 33
Table 3. Binary Factor Analytic Parameters Derived from the MRCML Rasch Parameters ................................................................................................................. 36
Table 4. Ordering of Elements of All Matrices in a Path Model ........................................ 53
Table 5. Adding Re-Ordered Elements Together .............................................................. 55
Table 6. Path Model Parameters for Each Version of the Simulation ..................................... 84
Table 7. Results of Simulation Part I, Version 1 ................................................................. 90
Table 8. Results of Simulation Part I, Version 2 ................................................................. 93
Table 9. Results of Simulation Part I, Version 3 ................................................................. 95
Table 10. Comparison of Version 1 Path Analysis Parameters Estimates between ACER

ConQuest, SPSS and Mplus for Simulation Study Part II ............................................. 99
Table 11. Comparison of Version 3 Path Analysis Parameters Estimates between ACER

ConQuest, SPSS and Mplus for Simulation Study Part II ............................................. 100
Table 12. Comparison of Full Model Results from ACER ConQuest and Mplus .......... 103
Table 13. Results of a Latent Path IRT Model Using LSAC Data ........................................ 114
Table 14. Results of a Latent Path IRT Model Using PISA Data .......................................... 120
List of Figures

Figure 1. Generalised latent variable model ................................................................. 4
Figure 2. Example of a general MRCML model ............................................................ 6
Figure 3. Example 1 of a latent path model within an IRT framework ....................... 8
Figure 4. Example 2 of a latent path model within an IRT framework ....................... 8
Figure 5. Scatterplot of parameters estimated by ACER ConQuest and Mplus .......... 36
Figure 6. Example of a path model ........................................................................... 43
Figure 7. Permutation of matrices for each equation m .............................................. 51
Figure 8. Hypothetical latent path model .................................................................. 70
Figure 9. Standard MRCMLM with measurement and multiple latent regression model ... 71
Figure 10. General model for simulation study ......................................................... 82
Figure 11. Scatterplot of item difficulties estimated by ACER ConQuest and by Mplus for simulation Part I ......................................................................................... 89
Figure 11. Parameter estimates of simulation Part I, Version 1 .................................... 91
Figure 12. Standard error estimates of simulation Part I, Version 1 ........................... 92
Figure 13. Parameter estimates of simulation Part I, Version 2 ................................... 94
Figure 14. Standard error estimates of simulation Part I, Version 2 ........................... 94
Figure 15. Parameter estimates of simulation Part I, Version 3 ................................... 96
Figure 16. Standard error estimates of simulation Part I, Version 3 ........................... 97
Figure 17. Scatterplot of item difficulties estimated by ACER ConQuest and by Mplus for simulation Part III ......................................................................................... 102
Figure 18. LSAC latent path model within an IRT framework ......................................... 110

Figure 19. PISA latent path model within an IRT framework........................................ 117
Chapter 1: Introduction

Latent variable modelling is a popular type of analysis used in a broad range of social research. Latent variable modelling is concerned with analysing relationships between variables of which some cannot be observed. These unobserved, or latent, variables are measured by sets of observed variables such as items in a test or questions in a questionnaire. Latent variable modelling combines these two parts, the structural part including the relationships between variables and the measurement part to estimate the latent variables, into one model. More detailed definitions of the structural and measurement parts of latent variable models are given later in the introduction.

Two co-existing traditions have developed to deal with latent variable models: structural equation modelling (SEM) and item response theory (IRT). SEM has focused strongly on methods to combine various structural models with measurement models and determining the fit of structural models. As a result, many types of structural models have been implemented in SEM. The measurement part consisted in the original formulation of factor analytic models that assumed linear relationships between indicators at interval level and a latent trait. Over time, some more flexibility has been introduced in the measurement model to estimate loadings of ordinal and dichotomous indicators of the latent variables.

Within the IRT framework, most emphasis has been placed on measuring latent variables and fitting individual indicators, or items. Consequently, a wide variety of item response models has been implemented. Since IRT is often used for measuring complex constructs such as literacy in academic subjects, methods have been developed to deal with
Chapter 1: Introduction

large item pools and complicated test designs such as rotated booklets, so that individual respondents only respond to a subset of the items, while the full item bank covers all aspects of the constructs. However, the choice of structural models within the IRT framework is limited to latent regression models with latent variables as dependent and observed variables as independent variables.

In summary, while the general goal of the two traditions is the same, the measurement models implemented in SEM have more limitations than in IRT and the structural models are far more restricted in IRT than in SEM. Combining the wide range of existing item response models with a more expansive range of structural models is an important and useful development for social researchers applying latent variable models.

This thesis describes a method to implement latent path models, the most general form of structural models in SEM, to any IRT model so that the IRT measurement parameters and path coefficients can be estimated at the same time. Latent path models are currently not implemented in most item response models. Thus far, researchers working in the IRT tradition needed to create IRT scores for respondents first and then estimate path models using other software packages for secondary analysis of the IRT scores. Before describing the outline of this thesis, the two parts of latent variable models are described in more detail.

Latent variable models generally consist of a structural part and a measurement part. In the structural part, regression coefficients are estimated between (observed or latent) exogenous and (observed or latent) endogenous variables or between endogenous variables. Exogenous variables are variables that are not being explained within the model. An
endogenous variable is being explained by one or more variables in the model and are therefore a dependent variable in one of the equations in the system. If a structural model does not include relationships between endogenous variables, it is a (multiple) regression; if the model includes relationships between endogenous variables, it is a (latent) path model.

The measurement parts of these models describe the relationship between unobserved latent variables and observed manifestations (indicators) of those variables—as such they are used to construct the latent variables.

Figure 1 gives an example of a generalised full latent variable model. Each measurement model is indicated by a factor—the latent variable—represented by an oval shape and a set of observed indicators (in this example, only two per latent variable for simplicity), represented by rectangle shapes. The loadings are denoted as $\Lambda$ (lambda matrix). Each indicator is associated with an error term, or unexplained variance.
Figure 1. Generalised latent variable model

The structural part of the model in this example only includes relationships between latent variables, but it can also include directly observed variables such as respondents’ gender or age. In the structural part of the model, exogenous variables predict endogenous variables and endogenous variables are being explained by exogenous and/or other endogenous variables. In Figure 1, the exogenous variables are $\xi_1$, $\xi_2$ and $\xi_3$ and the endogenous variables are $\eta_1$ and $\eta_2$. Because one relationship is estimated between latent endogenous variables in this example—$\eta_1$ predicts $\eta_2$—the structural model is also known as a latent path model. The path coefficients between exogenous and endogenous variables are denoted as $\mathbf{B}$ (beta matrix). If a structural model only includes a beta matrix, it is a multiple latent regression. In addition, path coefficients between endogenous variables can be estimated, as in Figure 1. The matrix that contains these coefficients is given by $\mathbf{\Gamma}$.
Models that include latent variables and a gamma matrix are examples of latent path models.

Traditional IRT models only include measurement models. However, in specialised IRT models like the multidimensional random coefficients multinomial logit model (MRCMLM; Adams, Wilson, & Wang, 1997) those traditional IRT models were extended to include latent regression models as a structural part (see also Fox [2007, 2010] for the implementation of multilevel regression models). A limitation of the regression model included in MRCMLM is that latent variables can only be endogenous, and observed variables only exogenous. Glöckner-Rist and Hoijtink (2003) also noted that latent regression has been the only structural model that has been studied in IRT. Figure 2 gives an example of such a model.

---

1 Many social scientists are more familiar with what is often known as “the LISREL notation.” The LISREL notation uses gamma for the relationship between exogenous and endogenous variables and beta for the relationships between endogenous variables. The subscripts are also reversed in the LISREL notation.
Chapter 1: Introduction

The aim of this thesis is to introduce a new method of combining logistic IRT measurement models with latent path models. Thus far, researchers working within an IRT framework had to produce individual person latent trait estimates using IRT software. Subsequently, they had to import the ability estimates into other software packages to link the IRT scores with other data and to perform secondary analysis. This method has two major limitations. The first limitation is that it is cumbersome. The second is that different types of estimates for individual IRT scores lead to different results (Adams & Wu, 2007; Mislevy, 1991; Mislevy, Beaton, Kaplan & Sheehan, 1992; Mislevy & Sheehan, 1987). The new method circumvents the need to produce individual person estimates and therefore eliminates both limitations. For this, three model features need to be added to the current MRCML latent regression model:

- linear causal relationships between endogenous variables (\( \Gamma \))
Chapter 1: Introduction

- use of latent variables as exogenous variables
- use of observed variables as endogenous variables.

Figure 3 and Figure 4 present examples of such an IRT model with a latent path model. The results of these examples are described in Chapter 6. These examples were chosen because of their complex latent variables, measured by many items with different formats. Example 1 includes 128 partial credit items, each with three, four or five Likert-type response categories, longitudinally measuring the two latent dimensions social-emotional wellbeing in Year 3 and in Year 7. Example 2 includes 188 cognitive items with a mixture of formats: multiple-choice items and dichotomous and partial credit constructive response items. Thirteen booklets were created so that students only had to respond to a subset of the items and did not have to sit the test for longer than two hours. The second latent dimension in example 2 was measured by five questions with five Likert-type response categories, analysed as partial credit items. The complexities of these item response models cannot be analysed as measurement models within a SEM framework, nor can they be analysed with existing IRT methods.

In addition to the IRT measurement models, the two examples both include a latent path model, in which latent and observed variables are being explained by other latent or observed variables. The latent path model in example 2 is a mediation model. Hitherto, these latent path models have not been included within an IRT framework.
Chapter 1: Introduction

Figure 3. Example 1 of a latent path model within an IRT framework

Figure 4. Example 2 of a latent path model within an IRT framework
Chapter 1: Introduction

To perform latent path model analyses, a full sum of squares and cross products (SSCP) matrix of all included variables needs to be constructed. Currently, the MRCML model is only giving part of these estimates. A multi-step approach has been developed to be able to estimate the new types of relationships. Describing and testing this method is the aim of this thesis.

The thesis starts with a brief history of item response models and a description of the modern MRCMLM and its estimation methods as the basis of the multi-step method. Since factor analytic models have essentially the same purpose as item response models—measuring latent variables and comprising the measurement parts of latent variable models—similarities between the two approaches are described and explored at the end of the next chapter. A method to estimate path models, two stage least squares, is detailed in Chapter 3. Chapter 4 outlines the multi-step approach to include latent path models in the MRCMLM. This approach is tested in a simulation study to examine the empirical properties of the estimates (i.e. bias, mean estimates and standard errors) in Chapter 5 as well as compared with results derived within a SEM framework. In Chapter 6, the two examples of the application of the multi-step method that were briefly presented in this introduction are analysed and reported using real data from the Longitudinal Study of Australian Children and data from the OECD’s Programme for International Student Assessment. The thesis finishes with a conclusion, some limitations of this approach and suggestions for future development.
Chapter 2: IRT Measurement

This chapter starts with a brief history of IRT measurement, followed by a description of the advanced IRT model and its estimation methods that is used as a basis in this thesis. Since factor analytic models have been used for the same purpose in latent variable modeling but within a different framework, the chapter finishes with comparing the simple logistic item response model and the binary factor analytic model. It will be shown that for dichotomous items, the results of the two approaches are identical if the parameterisations of the models are the same.

A Brief History of IRT

In the first half of the twentieth century, the concept of the item characteristic curve (ICC) was developed by Tucker (1946) and the trace line was introduced by Lazarsfeld (1950). An ICC or trace line is the plot of the probability of a correct response on an item against a person’s true ability. The concept of the item characteristic curve was further developed by Frederic Lord (1912–2000) who in 1952 founded a basis for item response theory to estimate an individual’s true abilities simultaneously with item characteristics using the observed responses on test items (Lord, 1952; Suen & French, 2003), which was called the two-parameter normal ogive model. The two item parameters included in the model were the difficulty of the item and the item discrimination (or slope of the ICC). The probability of success on a given item as a function of ability, or the item response function, was chosen to be a cumulative normal distribution function (Van der Linden & Hambleton,
1997). The computations, however, were too demanding and therefore too difficult and expensive to implement, until Alan Birnbaum suggested using a logistic approximation method in 1958 (see Birnbaum, 1968; Suen & French, 2003). Birnbaum also proposed the inclusion of a third parameter in the model, the guessing parameter (Van der Linden & Hambleton, 1997). This guessing parameter was the lowest probability of responding correctly to a multiple-choice item. Despite the simplification by the logistic model, the calculations were still labor intensive and hence unpopular until the 1970s and 1980s when personal computers gave researchers the computing power needed.

In Denmark, a mathematician named Georg Rasch (1901–1980) received his doctorate in pure mathematics in 1930. Despite being considered to be one of the most talented of a new generation of mathematicians, he was unable to find work as a mathematician in the economic climate in Denmark in the 1930s. He chose to become a consultant in applied mathematics, primarily in data analysis and statistics (Andersen, 1982). He started off in the areas of biology and medicine, before he became involved in education and psychology in the 1950s where he made his most original contributions. As a consultant he was asked to work on a method to extract information on individuals from intelligence ability tests. Rasch was working on a statistical method using conditional probabilities to estimate individuals’ latent abilities and item difficulties simultaneously using a logistic function between true ability scores and the odds ratio of a correct response. He worked independently of the Americans, Lord and Birnbaum (Suen & French, 2003). His most simple and elegant model was published in 1960, which became known as the
Rasch model and was soon proven to be mathematically equivalent to the one-parameter logistic IRT model (Rasch, 1960 [1980]). The Rasch model is the most restrictive and most practical IRT model and therefore widely used.

Benjamin Wright (born in 1926) introduced Rasch’s method in the USA and is largely responsible for the widespread adoption of the simple logistic Rasch model. He chaired nearly 70 dissertations, of which the majority was concerned with new Rasch models, estimation methods, fit statistics, or data applications.

Two models for polytomous items were developed by Wright’s students around 1980, in which the difficulty of assigning a particular rating \( k \), rather than the next lower rating \( k-1 \), to a particular item is depicted with item parameter estimates, \( \delta_{ik} \). In one of the models, the distance between category thresholds may be held constant across items. This model, the rating scale model, was published by both Erling Anderson (1977) and David Andrich (1978). In 1982, Benjamin Wright and Geoffrey Masters introduced a more general polytomous rating scale model that allows distances between thresholds to vary across items. This model is known as the partial credit model (Masters, 1982; Wright & Masters, 1982). Linacre built further upon this method to allow for the estimation of raters’ stringency in 1989, known as the many-faceted Rasch model or, in short, the facet model (Linacre, 1989).

In the same period, several psychometricians from the European continent were extending the Rasch model. One such extension is the Linear Logistic Test Model (LLTM) from the Austrian, Fischer, first published in 1972 (Fischer, 1995). In his model, which is a
form of a facet model, the item parameter can be decomposed into the weighted sum of some basic difficulty parameters. That is, the item parameter is a function of certain cognitive operations involved in solving the item or of certain test or experimental conditions such as training or treatment. It is acknowledged that it is actually the ability that changes due to these operations or conditions, not the difficulty of the item, but the ability is assumed to be constant in this conditional model. The core assumption of this model is that differences in item parameters are due to the operations or conditions involved in one item but not another. The implication of this is that each item needs a unique combination of basic parameters.

Another European creator of an extended Rasch model is Verhelst from Belgium and the Netherlands. He introduced the One Parameter Logistic Model (OPLM) in the late 1980s (Verhelst & Glas, 1995), which offers a combination of a Rasch model and its powerful mathematical properties with the flexibility of a two-parameter model. While the Rasch model has powerful mathematical properties, it has rather strict assumptions. If these assumptions are not met and the data for some items do not fit the model, a researcher could decide to eliminate those items or some students from the data. However, these decisions can jeopardise the content validity of the test or the representativeness of the sample. Another option is choosing a less-restrictive model, for example the two-parameter model of Birnbaum, at the cost of losing the theoretical advantages of the Rasch model. While equivalent in form to the two-parameter model, the discrimination parameters in the
OPLM are not unknown quantities. Instead, they are fixed and their values are chosen by the researcher as hypotheses.

A different type of IRT model dealing with polytomous items is the graded response model developed by Samejima (1969). While the probability of interest in the rating scale and partial credit model is the probability of scoring $x$ on an item divided by the same probability plus the probability of scoring $x-1$ for a given individual, the modelled probability is different in the graded response model. Here the probability of interest is the probability of scoring $x$ or higher, minus the probability of scoring $x+1$ or higher. The partial credit model divides by the total probability (of two adjacent categories) and the graded response model is a difference in probabilities. Thissen and Steinberg (1986) have developed a taxonomy of parametric, unidimensional IRT models. According to the authors, most multiple category response models without guessing can be considered either a form of difference models or a version of divided-by-total models. The general IRT model used in this thesis, the MRCMLM, only includes divided-by-total models.

Introducing multiple item parameters and dealing with different response formats have not been the only extension to the simple logistic IRT model by Rasch. In the late 1960s and early 1970s, the first versions of multidimensional IRT (MIRT) models emerged (Lord & Novick, 1968; McDonald, 1967; Samejima, 1974) to measure multiple abilities with single tests. These early models were further developed into practical MIRT models while software became available to estimate the parameters. The unidimensional normal ogive model was extended to a multidimensional model by McDonald (1967, 1997) for
dichotomous response data, the logistic two-parameter model by Reckase (1997, 2009) and the simple logistic model by Fischer (1976).

The MRCML Model

The model used in this thesis is the MRCML model (MRCMLM), which is also a multidimensional extension of the Rasch measurement model (Adams et al., 1997). The model is implemented by the ACER ConQuest software (Adams, Wu and Wilson, 2015). The MRCMLM can be restricted to the partial credit model, the rating scale model and the simple logistic (Rasch) model for dichotomous items. The MRCMLM depicts each person using a profile of multiple correlated latent traits. To enable these features, two matrices are built into the model. The expected relation of items to the underlying dimensions is represented by the design matrix $\mathbf{A}$, and to the underlying item parameters is represented by the scoring matrix $\mathbf{B}$. In the Rasch probability function, the vector of latent traits for each person is multiplied by the design matrix and the vector of item parameters is multiplied by the scoring matrix.

In the case of a unidimensional model, the probability of observing a response in category $j$ of item $i$ with $k$ categories is modelled as

$$\Pr(X_{ij} = 1; \mathbf{A}, \mathbf{b}, \xi | \theta) = \frac{\exp(b_{ij} \theta + a_{ij}' \xi)}{\sum_{k=1}^{k} \exp(b_{ik} \theta + a_{ik}' \xi)}$$
where $b_{ij}$ is the score assigned to category $j$, $\theta$ the ability of a respondent, $a'_{ij}$ a design vector to define the parameter for category $j$ of item $i$, and $\xi$ a vector of item parameters.

The full MRCMLM is a multidimensional extension of this model, with the item response model expressed as

$$
\Pr(X_{ij} = 1; A, B, \xi | \theta) = \frac{\exp(b_{ij} \theta + a'_{ij} \xi)}{\sum_{k=1}^{K} \exp(b_{ik} \theta + a'_{ik} \xi)}.
$$

Here, $\theta$ is a vector of abilities for each respondent with as many elements as there are dimensions, and $b_{ij}$ a vector of scores for category $j$ of item $i$ on each dimension. Since the scoring and design vectors are flexible, many of the before-mentioned models are special cases of the MRCMLM. Examples are the simple logistic Rasch model, the rating scale model, the partial credit model and the linear logistic test model and many more user-defined models or combination of default models when creating and importing design matrices.

Initially, IRT methods were developed for use with standardised achievement tests. IRT methods are now well known among educational researchers and are growing in popularity among psychologists, especially industrial/organisational psychologists, and health scientists. In the early days, very large samples were needed to have a reasonable chance of obtaining unbiased and stable estimates of item parameters. In the last decades, improved statistical algorithms have decreased the sample size requirements to a level similar to factor analysis (Harvey & Hammer, 1999).
Estimating IRT Models

For IRT models within the Rasch framework, some applications of estimation methods have moved from joint and conditional maximum likelihood for specific models to marginal maximum likelihood with an expectation-maximisation algorithm (MML/EM) for all models. Joint maximum likelihood (JML) was the main approach to item parameter estimation since 1968 and is attributed to Birnbaum (Harwell, Baker & Zwartz, 1988). In this approach, the unknown person abilities are estimated alongside the item parameters. A problem with this approach is that the item parameter estimates need not be consistent as sample sizes increase. Conditional maximum likelihood (CML) offered a solution for these inconsistent estimates if sufficient statistics were available, but only for the Rasch model. Marginal maximum likelihood was introduced in 1970 by Bock and Lieberman and could also be applied to two- and three-parameter models provided that the test was very short because of the computational demand by this estimation method. With MML, the input for the analysis, or the sufficient statistics, are the sums of scores for one-parameter models and the original categorical response data for two- and three-parameter models. The abilities are not estimated, they are assumed to be a random sample from a population distribution. Marginalisation occurs over the sufficient statistics and integration is required over the latent factors. The reformulation of the MML approach by Bock and Aitken (1981) using an EM algorithm has made the computations feasible and therefore increased the practicality of the MML estimation method. Marginal maximum likelihood estimation with an EM algorithm (Bock & Aitkin, 1981) is now one of the most commonly used estimation
methods in IRT (Wirth & Edwards, 2007). In some literature, MML is also called full information maximum likelihood or FIML (Forero & Maydeu-Olivares, 2009; Joreskog & Moustaki, 2006; Rijmen, 2009). The estimation method and algorithms are described in more detail below.

In order to avoid estimating all item and ability estimates simultaneously as in joint maximum likelihood, the MML approach integrates over the person-specific parameters (this process is known as marginalisation) and estimates the item parameters in the marginal distribution. The abilities are assumed to be random draws from some, typically normal distribution, similar to missing values completely at random, because all abilities are missing. In this situation, it is possible to integrate over the ability distribution. This essentially means that the abilities are removed from the likelihood by marginalisation. Without being concerned about the abilities, item parameters can be found that maximise this likelihood.

The difference between JML and MML can be illustrated using the simple logistic unidimensional model for dichotomous items. Suppose that a sample of $N$ examinees ($n = 1, ..., N$) responds to a set of $K$ test items ($i = 1, ..., K$) and the response of student $n$ to item $i$ is denoted $x_{ni}$ which takes the value ‘1’ for a correct response and ‘0’ for an incorrect response, then the model can be written as

$$P(x_{ni}; \delta_i, \theta_n) = \frac{\exp[x_{ni}(\theta_n - \delta_i)]}{1 + \exp(\theta_n - \delta_i)},$$

3
where $\theta_n$, $\theta_n$ is referred to as the case parameter, which is the location of case $n$ on the latent continuum, and $\delta_i$ is referred to as the item parameter, which is the location of item $i$ on the latent continuum.

In JML, all item and case parameters are estimated jointly by maximizing the likelihood:

$$
\Lambda(\Theta, \Delta; X) = \prod_{n=1}^{N} \prod_{i=1}^{K} P(x_{ni} ; \delta_i, \theta_n),
$$

with respect to $\Theta$ (all case parameters) and $\Delta$ (all item parameters).

When using MML, the IRT model is usually expressed as a probability conditional on ability $P(x_{ni} ; \delta_i | \theta_n)$, or $P(X = x_n ; \Delta | \theta_n)$ for responses to a set of items by respondent $n$, where the case locations $\theta_n$ are independently distributed according to the probability density function $g$ which has parameters $\alpha$ (for example, the mean and variance). To transform the conditional probability into an unconditional probability, the conditional model is multiplied by the distribution density $g(\theta_n | \alpha)$. In other words, given a particular response pattern for a person sampled randomly from a population with a distribution density $g(\theta_n | \alpha)$, the likelihood of the parameters $\alpha$ and $\Delta$ using $X$ that is maximised in MML is

$$
\Lambda(\alpha, \Delta; X) = \prod_{n} \int_{-\infty}^{\infty} P(X = x_n ; \Delta | \theta) g(\theta | \alpha) d(\theta)
$$
Integration, especially over multiple dimensions, is very complex and highly computer intensive. Therefore, it is common in maximum likelihood estimation to use numerical integration as an approximation of analytical integration. One such method is Gauss-Hermite quadrature. This method approximates the area under the curve by dividing it up in a fixed set of rectangles. Areas of rectangles are easy to compute. By increasing the number of rectangles, the rectangles become narrower and follow more precisely the contour of the curve, which increases the precision of the estimated area under the curve. However, increasing the number of rectangles (also known as nodes) also increases the computational burden, especially for multidimensional models. The number of nodes increases exponentially with each dimension. For example, if 15 nodes are used, the number of nodes for a three-dimensional model is $15^3 = 3375$.

Computing the probability above using Gauss-Hermite quadrature to approximate the integral results in

$$P(X = x_n) = \sum_{k=1}^{d} P(X = x_n | Q_k) A(Q_k),$$

where $Q_k$ is the quadrature point or the location on the latent scale, $A(Q_k)$ the quadrature weight or the width of the rectangle and the summation is over $d$ dimensions.

Despite the efficiency of the numerical approximation, remaining computational difficulties limit the number of items that can be simultaneously analysed. The expectation-maximisation (EM) algorithm (Bock & Aitkin, 1981) is a strategy to overcome this difficulty when obtaining maximum likelihood estimates. The strategy deals with
incomplete data, either the “missing” latent responses in IRT estimation described above or missing response data. The idea underlying the EM algorithm is to assign values to the missing data, estimate parameters, and re-estimate values for missing data until changes between iterations are very small (Rubin, 1991). In case of IRT, the missing data are the latent variables. During the E-step temporary item parameters are regarded as the true item parameters and used to estimate the expected number of correct responses for an expected number of respondents at particular locations on the latent variable scale (the nodes). The M-step involves obtaining new estimates of the item parameters by substituting the E-step probability estimates in the likelihood equations at the same locations on the latent variable scale. The two steps are repeated until convergence.

The EM algorithm deals with the numerical challenges that arise from a model with many items, but it does not solve the computational problem caused by the exponentially increasing number of quadrature points with the number of dimensions. The amount of time needed for the estimation increases linearly with the total number of nodes. A solution is to use an adaptive numerical integration method instead of the Gauss-Hermite quadrature method for multidimensional models. One adaptive numerical integration method is the Monte Carlo EM algorithm (Meng & Schilling, 1996). This approach fits within the MML/EM framework and is applied in ACER ConQuest as well as the Gauss-Hermite quadrature method (Volodin & Adams, 1995).

In the Monte Carlo EM algorithm, the fixed number of nodes is the total number of nodes over all dimensions. For efficiency, the density of the nodes is higher near the
expected values of the multivariate latent distributions and lower near the extremes. The user may have to experiment with the number and range of Monte Carlo nodes to optimise the estimation process and to converge the model. ACER ConQuest can be used to analyse up to 30 dimensions.

The MML/EM estimation method has also become the preferred method for factor analytic models, but other methods are still available.

**IRT and Factor Analytic Models**

Both IRT modelling and factor analytic modelling are methods frequently used to examine whether a set of items can be used to reliably construct measures of an unobserved or latent variable and to estimate scores on the latent variable. While IRT is only used for categorical item variables with nonlinear relationships between item variables and the latent factor, factor analytic models can include either nonlinear relationships with dichotomous or polytomous indicators or linear relationships with continuous indicators. Certain nonlinear variants of factor analytic models are equivalent to IRT models, thus enabling the computation of factor analytic model parameters from IRT parameters and vice versa (Brown, 2006). For example, applying a factor analytic model to the matrix of tetrachoric correlations for dichotomous indicators is equivalent to the two-parameter normal ogive (or probit) IRT model (Birnbaum, 1968). Confirmatory factor analysis with polychoric correlations for polytomous items is equivalent to using the IRT graded response model (Samejima, 1969).
The formal relationship between factor analytic models and IRT was first specified by Lord and Novick in 1968. During the same period, McDonald (1967) also worked on the common foundation of factor analytic models and IRT when dealing with multidimensionality problems by defining a principle of local independence. Most literature, however, refer to Takane and De Leeuw (1987), who provided a systematic series of proofs that show the equivalence of the two-parameter normal ogive model in IRT and the nonlinear confirmatory factor analysis, with dichotomous as well as ordinal item responses.

Before the two approaches are formally written down, a derivation of the Birnbaum two-parameter normal ogive model is given as an introduction from which the Rasch model and the factor analytic model are branched off. The section ends with a discussion of transformations of the parameters between the two models for different parameterisations.

**Derivation of the Birnbaum Model**

In the case of Guttman response patterns, respondents are successful on each item $i$ with a difficulty ($\delta_i$) smaller than the ability of respondent $n$ ($\theta_n$) and unsuccessful on items with a difficulty larger than her or his ability. However, a person’s response is not only dependent on her or his ability or, in other words, the ability does not explain all variance in the response data. This means that the observed response of person $n$ on item $i$ includes an unexplained part or an error term ($\epsilon_i^n$).
Adding this error term to the difficulty of an item can be interpreted as a person-specific item difficulty which would predict exactly the observed response of person \( n \) with ability \( \theta_n \). The formula for the person-specific item difficulty is

\[
\delta_i^n = \delta_i + \varepsilon_i^n
\]

Therefore, \( \theta_n > \delta_i^n \) will lead to success for person \( n \) on item \( i \) \((z_i = 1)\) and the probability of \( \theta_n > \delta_i^n \) is equal to the probability of success.

\[
P(z_i = 1) = P(\theta_n > \delta_i^n) = P(\theta_n > \delta_i + \varepsilon_i^n) = P(\theta_n - \delta_i < \varepsilon_i^n) = P(\varepsilon_i^n < \theta_n - \delta_i)
\]

If \( \varepsilon_i^n \) is assumed to be normally distributed and independent across items and persons, this probability is equal to

\[
P(z_i = 1) = \int_{-\infty}^{\theta_n - \delta_i} \frac{1}{\sqrt{2\pi}\sigma_{in}^2} e^{-t^2/2\sigma_{in}^2} dt = \Phi\left(\frac{\theta_n - \delta_i}{\sigma_{in}}\right)
\]

where \( \sigma_{in}^2 \) is the error variance for each item-person combination. However, it is usually assumed to only vary across items and to be equal for all persons and can therefore be written as \( \sigma_i^2 \).
Chapter 2: IRT Measurement

The normal ogive can be approximated very closely by the logistic ogive (Johnson & Kotz, 1970) which has computational advantages according to Birnbaum.

\[
\Phi\left(\frac{\theta_n - \delta_i}{\sigma_i}\right) \approx \frac{\exp\left\{1.7\alpha_i (\theta_n - \delta_i)\right\}}{1 + \exp\left\{1.7\alpha_i (\theta_n - \delta_i)\right\}}
\]

This logistic ogive is equivalent to the two-parameter IRT model from Birnbaum. The scaling constant 1.7 is used to minimise the difference between the two curves and \(\alpha_i\) is the slope of the item characteristic curve, or the discrimination of the item.

The formula for the two-parameter logistic IRT model for dichotomous items can be rewritten as:

\[
p_i(z_i = 1|\theta_n) = \frac{\exp\left\{1.7\alpha_i (\theta_n - \delta_i)\right\}}{1 + \exp\left\{1.7\alpha_i (\theta_n - \delta_i)\right\}} = \frac{\exp\left\{1.7\alpha_i \theta_n + \beta_i\right\}}{1 + \exp\left\{1.7\alpha_i \theta_n + \beta_i\right\}}
\]

where \(\beta_i = -1.7\alpha_i \delta_i\). Written in this form, \(\alpha_i\) is again the slope parameter and \(\beta_i\) the intercept.

The Rasch Model

The Rasch model is a specific form of the two-parameter logistic IRT model where all slopes are constrained to be equal. Typical parameterisations used in the Rasch model are
\[ E(\theta) = 0 \]
\[ a_i = 1. \]

In addition, the scaling factor of 1.7 is not used in the Rasch model. Linear scaling does not affect the fit of logistic models to the item response data. To relate the IRT parameters to the binary factor analytic parameters, the Rasch model with equal slopes is written in simplified form as a (logistic) function of a linear transformation of the latent trait:

\[ P(z_i = 1) = f(\alpha, \theta_n + \beta_i) \]

where \( f(x) \) is a logistic function.

**The Binary Factor Analytic Model**

While the individuals’ responses to the items are dichotomous, the assumption in binary factor analytic models is that there is a continuous underlying latent response, denoted \( z^*_i \). A one-factor model for the latent response variable can be written as

\[ z^*_i = v_i + \lambda_i \theta_n + \varepsilon_i \]

where \( \theta_n \) is the latent factor score for person \( n \), \( v_i \) is the intercept of item \( i \), \( \lambda_i \) is the factor loading of item \( i \), and \( \varepsilon_i \) is the residual for item \( i \). The residuals \( \varepsilon_i \) are typically assumed to be normally distributed, but a logistic distribution can also be considered.
To accommodate for the dichotomous responses to the items, a threshold model is added to the above linear latent response model:

\[
    z_i = \begin{cases} 
    1 & \text{if } z_i^* \geq \tau_i \\ 
    0 & \text{if } z_i^* < \tau_i 
    \end{cases}
\]

where \(\tau_i\) is the threshold parameter for item \(i\). Since the threshold parameter and the intercept are not jointly identified, either the thresholds or the intercepts are typically assumed to be zero.

The probability of success is equal to \(P(z_i^* \geq \tau_i)\). Following the same logic as above to describe the derivation of the Birnbaum model and fixing the intercept to zero, this probability is equal to

\[
P(z_i^* \geq \tau_i) = P(\lambda_i \theta_n + \epsilon_i \geq \tau_i) = P(\epsilon_i \geq \tau_i - \lambda_i \theta_n) = P(\epsilon_i < - (\tau_i - \lambda_i \theta_n))
\]

The following typical parameterisation can be chosen if slopes are constrained to be equal like in the Rasch model:

\[
\theta_n \sim N(0,1)
\]

\[
\sigma_{\epsilon_i}^2 = 1, \text{ so that } \sigma_{\epsilon_i}^2 = 1 - \lambda^2 \text{ and } \sigma_{\epsilon_i} = \sqrt{1 - \lambda^2}
\]
Using these parameterisations and the normal ogive model, the probability can be computed as

\[
P(z_i^* \geq \tau_i) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_{e_i}} e^{-t^2/2\sigma_{e_i}^2} dt
\]

\[
= \Phi \left( -\frac{(\tau_i - \lambda_i \theta_i)}{\sigma_{e_i}} \right)
\]

In simplified form, the binary factor analytic model is written as a (normal) function of a linear transformation of the latent trait:

\[
P(z_i = 1) = f \left( -\left( \tau_i - \lambda_i \theta_i \right) \right)
\]

\[
= f \left( \lambda_i \theta_i - \tau_i \right)
\]

This model is equivalent to the previously described two-parameter (Birnbaum) model. The next section describes this relationship between the models in detail.

**Transformation from Rasch to Binary Factor Analytic Parameters**

Assuming the common latent factor on the binary factor analytic model has a standard normal distribution, Takane and De Leeuw (1987)—and earlier, in different notation, Lord and Novick (1968)—have shown that the slope parameter of the IRT model (\(\alpha\)) is equal to the slope parameter of the binary factor analytic model (\(\lambda_i\)) divided by the residual standard deviation of the underlying latent response variables (\(\sigma_{e_i}\)). Similarly, the intercept of the IRT model (\(\beta_i\)) is equal to the intercept of the binary factor analytic model (\(\nu_i\)) or the thresholds with the opposite sign (\(-\tau_i\)), depending on which have been fixed to
zero, divided by the residual standard deviation of the underlying latent response variables \( (\sigma_{e_i}) \). In short, these transformations can be written as:

\[
\alpha_i = \frac{\lambda_i}{\sigma_{e_i}} \quad \text{(18)}
\]

\[
\beta_i = \frac{-\tau_i}{\sigma_{e_i}} \quad \text{(19)}
\]

These transformations can be generalised and restricted depending upon the types of parameterisations of the models. Kamata and Bauer (2008) describe transformations for four common parameterisations of a unidimensional factor analytic model with binary indicators (or items). For the parameterisation, two choices need to be made: (1) in constraining (also called scaling) the underlying latent response variables \( (z_i^*) \) of the item and (2) in constraining the common latent factor \( (\theta_n) \), because they are both latent variables without a natural scale.

Possible parameterisations regarding the *underlying latent response variables* (1) are:

a) the variance of the underlying latent item responses, the marginal variance, equals one \( (\sigma_{z_i}^2 = 1) \), which results in \( \sigma_{e_i}^2 = 1 - \lambda_i^2 \sigma_{\theta_n}^2 \) or

---

2 The letter \( V \) is used for variance to simplify notification:
b) the variance of the residuals of the underlying latent responses, the conditional variance, equals one ($\sigma_{\varepsilon_i}^2 = 1$, which results in $\sigma_{z_i^*}^2 = \lambda_i^2 \sigma_{\theta_n}^2 + 1$).

Typical parameterisations of the common latent factor (2) are:

a) fix the parameters (threshold and slope) of one item, the reference indicator (e.g. $\tau_1 = 0$ and $\lambda_1 = 1$) and freely estimate the distribution of the common latent factor or

b) standardise the common latent factor ($E(\theta_n) = 0$ and $\sigma_{\theta_n}^2 = 1$) and freely estimate all item parameters.

The transformations given by Takane and De Leeuw (1987) need to be adjusted for the different parameterisations (see also Kamata & Bauer, 2008). In case the latent common factor $\theta_n$ is standardised (2b), the two options for the standard deviation of the residual of the underlying latent response variable $z_i^*$ are: (1a) $\sigma_{\varepsilon_i} = (1 - \lambda_i^2)^{1/2}$ or (1b) $\sigma_{\varepsilon_i} = 1$. The residual standard deviations in Takane and De Leeuw's transformations are substituted accordingly in the last column of Table 1.

In case the common latent factor is not standardised (2a) and has the distribution $\theta_n \sim N\{E(\theta_n), \sigma_{\theta_n}^2\}$, the two sets of transformations under column 2b need to be adjusted. As previously mentioned, the standard deviation of the residual of the latent responses $z_i^*$ is

\[
V(z_i^*) = V(v_i) + V(\lambda_i \theta_n) + V(\varepsilon_i) = 1 \\
\Rightarrow \lambda_i^2 V(\theta_n) + V(\varepsilon_i) = 1 \\
\Rightarrow V(\varepsilon_i) = 1 - \lambda_i^2 V(\theta_n)
\]
\[ \sigma_{e_i} = (1 - \lambda_i^2 \sigma_{\theta_n}^2)^{1/2} \]. The loading or the slope of the latent response function \( z_i^* \) becomes \( \lambda_i \sigma_{\theta_n} \). The new intercept of the latent response function \( z_i^* \) becomes \( u_i + \lambda_i E(\theta_n) \) or the new threshold becomes \( -(\tau_i - \lambda_i E(\theta_n)) \). The standardised loadings, thresholds and error variances in column 2b of Table 1 are substituted by the rescaled parameters in column 2a, accordingly.

The formulae in Table 1 can be used to transform typical IRT parameters to binary factor analytic parameters. The IRT parameters \( \alpha_i \) and \( \beta_i \) are the typical parameters given by IRT software. However, the parameters estimated in the MRCMLM are somewhat different and first need to be transformed into these typical IRT parameters. When applying a Rasch model in MRCMLM, the slopes of the items are fixed to 1, the scaling factor 1.7 is not used and the population variance is freely estimated. Therefore, the standard deviation of the population as estimated in MRCMLM needs to be divided by 1.7.
Table 1

Transformations between IRT and Factor Analytic Item Parameters for Four Types of Parameterisations

<table>
<thead>
<tr>
<th>Common latent factor (2)</th>
<th>Reference item (a)</th>
<th>Standardised factor (b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marginal (a)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Latent response variables (1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conditional (b)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[
\alpha_i = \frac{\lambda_i \sigma_{\theta_i}}{\sqrt{1 - \lambda_i^2 \sigma_{\theta_i}^2}} \\
\beta_i = -\left(\tau_i - \lambda_i E(\theta_n)\right) \sqrt{1 - \lambda_i^2 \sigma_{\theta_i}^2} \\
\alpha_i = \frac{\lambda_i}{\sqrt{1 - \lambda_i^2}} \\
\beta_i = -\frac{\tau_i}{\sqrt{1 - \lambda_i^2}}
\]

There is also a difference between item difficulty parameters in typical two-parameter IRT software and in MRCMLM. Because two-parameter IRT software typically identifies the model by setting the population variance to one, while it is freely estimated in MRCMLM, the item difficulties estimated in MRCMLM need to be divided by the standard deviation of the population to derive the typical IRT difficulties. The typical IRT intercept for each item is simply the product between the slope and the difficulty with the opposite sign.

These typical IRT parameters can be linked to the matching set of transformations in Table 1 depending on the parameterisation choices of the binary factor analytic model. As mentioned before, a typical parameterisation to apply a binary factor analytic model
with equal loadings is setting the variance of the underlying latent response variable $z^*_i$ to one ($\sigma_{z^*_i}^2 = 1$), which is option (1a), and standardizing the common latent factor ($\theta_n \sim N(0,1)$), which is option (2b).

Table 2

Two-Step Transformations between the MRCML Rasch Parameters and Binary Factor Analytic Parameters

<table>
<thead>
<tr>
<th>MRCML parameters</th>
<th>Typical IRT parameters</th>
<th>Binary factor analytic parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\frac{\alpha_c \sigma_c}{1.7}$</td>
<td>$\alpha_i^*$</td>
<td>$= \frac{\lambda_i}{\sqrt{1 - \lambda_i^2}}$</td>
</tr>
<tr>
<td>$\frac{\delta_i^*}{\sigma_c}$</td>
<td>$\delta_i^*$</td>
<td>$= \frac{-\beta_i^<em>}{\alpha^</em>} = \frac{\tau_i}{\lambda_i}$</td>
</tr>
<tr>
<td>$-\alpha_i^* \cdot \delta_i^*$</td>
<td>$\beta_i^*$</td>
<td>$= \frac{-\tau_i}{\sqrt{1 - \lambda_i^2}}$</td>
</tr>
</tbody>
</table>

From the transformations in Table 2 and by setting $\alpha_c$ to 1, it follows that
Knowledge of the value of the loadings and the item difficulty and population variance estimates from MRCMLM makes it possible to compute the values of the threshold parameters \( \tau_i \) of the binary factor analytic model.

\[
\frac{\delta_i^c}{\sigma_c} \lambda_i = \tau_i \Rightarrow \tau_i = \frac{\delta_i^c \lambda_i}{\sigma_c}
\]

The above shows how to compute the factor loadings from MRCMLM estimates. Similar transformations can be found for multi-category data. The IRT model with multi-category data that is equivalent to a multi-category factor analytic model is the
Samejima model. Since the Samejima model is not a Rasch model and is therefore not a specific model of the MRCMLM, it is outside the scope of this thesis.

Matching Models

To illustrate these transformations, the following small simulation study was carried out. A database with 5000 respondents was created. Responses to 10 dichotomous items were generated using true item difficulties. The item difficulties and population variance were estimated using the MRCMLM as applied by ACER ConQuest. The parameter estimates are presented in Table 3.

The typical IRT parameters are included in the table for completeness, but are not used directly to compute the loadings and thresholds for the binary factor analytic model. Instead, equations 20 and 21 were applied.

Mplus, version 7.11 (Muthén & Muthén, 1998-2015) was used to run a binary confirmatory factor analysis on the data and to check the transformed factor analytic parameter estimates. The variance of the underlying latent response variable $z_i^*$ was equal to 1 (parameterisation option 1a), the common latent factor $\theta_n$ is standardised (parameterisation option 2b) and the loadings are constrained to be equal (Rasch model). Weighted least squares estimation was used. Mplus also produces the typical IRT parameters by default. These are included in the table for comparison. The Mplus results are very close to the transformed parameters from the MRCML model, which can also be seen in Figure 5.
Chapter 2: IRT Measurement

Table 3

Binary Factor Analytic Parameters Derived from the MRCML Rasch Parameters

<table>
<thead>
<tr>
<th>Item</th>
<th>MRCML estimates</th>
<th>Typical IRT</th>
<th>Binary factor analysis</th>
<th>Mplus output (WLSMV)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>δ°</td>
<td>α°</td>
<td>δ*</td>
<td>α*</td>
</tr>
<tr>
<td>1</td>
<td>-1.89</td>
<td>1.00</td>
<td>-1.34</td>
<td>0.82</td>
</tr>
<tr>
<td>2</td>
<td>2.79</td>
<td>1.00</td>
<td>1.99</td>
<td>0.82</td>
</tr>
<tr>
<td>3</td>
<td>-0.53</td>
<td>1.00</td>
<td>-0.37</td>
<td>0.82</td>
</tr>
<tr>
<td>4</td>
<td>-1.42</td>
<td>1.00</td>
<td>-1.01</td>
<td>0.82</td>
</tr>
<tr>
<td>5</td>
<td>2.60</td>
<td>1.00</td>
<td>1.85</td>
<td>0.82</td>
</tr>
<tr>
<td>6</td>
<td>2.97</td>
<td>1.00</td>
<td>2.12</td>
<td>0.82</td>
</tr>
<tr>
<td>7</td>
<td>0.28</td>
<td>1.00</td>
<td>0.20</td>
<td>0.82</td>
</tr>
<tr>
<td>8</td>
<td>0.21</td>
<td>1.00</td>
<td>0.15</td>
<td>0.82</td>
</tr>
<tr>
<td>9</td>
<td>-2.63</td>
<td>1.00</td>
<td>-1.87</td>
<td>0.82</td>
</tr>
<tr>
<td>10</td>
<td>-2.54</td>
<td>1.00</td>
<td>-1.81</td>
<td>0.82</td>
</tr>
<tr>
<td>Variance</td>
<td>1.97</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Figure 5. Scatterplot of parameters estimated by ACER ConQuest and Mplus
IRT versus Factor Analytic Models in Practice

In light of the similarities between IRT and factor analytic models, some authors suggest to use the overarching concept item factor analytic (IFA) models, which was introduced by Bock (1988). In 2010, a small meta study was carried out to evaluate motivations for choosing either factor analysis or IRT in practice (Ten Holt, Van Duijn, & Boomsma, 2010). The researchers selected three journals in the fields of education and psychology that contain many articles on construction or evaluation of a scale as a main topic. Reviewing all articles that were published in 2005, 41 studies were found that reported results from factor analysis, IRT or both.

In this set of articles, far more studies applied confirmatory factor analysis than IRT models, but explicit reasons were not often given. Two studies that applied IRT analysis mentioned a skewed item response distribution as an argument for applying the Rasch model. Another study described factor analysis and IRT as complimentary with IRT better suited for testing equivalence of item parameters and factor analysis more appropriate for multidimensional model testing. Indeed, in all studies that mentioned an interest in the values of item parameters, IRT was used. Furthermore, in all studies applying only factor analysis, the goal of testing whether the data fit a hypothesised structure was mentioned and in almost all studies that applied factor analysis, model fit was formally tested. In contrast, formal tests of model fit were never provided in the studies that applied IRT. The authors stated that this reflects a difference in tradition. It has become standard practice in factor analysis to compare model fit indices to cut-off values, while model fit measures are less
often used in IRT studies. It also seemed that IRT was primarily applied to investigate unidimensional scales, while factor analysis was more often used for multidimensional structures. The authors note that accessibility of software may be another implicit reason for the choice of analysis. When treating the indicator variables as continuous variables, confirmatory factor analysis can be conducted in many general statistical software packages. However, for both categorical factor analytic and IRT, especially multidimensional IRT, modelling more specialised software is needed.

Although the literature study by Ten Holt et al. suggests that many applied researchers do not have explicit motivations for choosing either an IRT or a factor analytic model, for researchers with experience in both methods, the choice depends on the research question of interest (Wirth & Edwards, 2007). IRT models may be more practical for researchers interested in characteristics of individual items and producing scores for individual participants, and factor analysis may provide a more natural framework for studies where the research question focuses on the dimensionality of a set of items. According to the authors, this is due to a historical difference in focus between the two approaches.

Despite the many common features between factor analytic and IRT approaches, some practical differences remain. For example, IRT models cannot deal with continuous indicators, whereas the factor analytic models were originally designed for continuous indicators. Furthermore, model fit is usually not tested in IRT methods, only differences between alternative models can be tested, while factor analysis methods provide general
model fit indices. Instead, the focus in IRT has been on item and person fit, because the main goal of this approach has been the construction of tests, while in factor analytic approaches relatively more attention has been paid to testing of theories on relationships between constructs.

**Summary**

A brief history of IRT was described in this chapter from the very start in the first half of the twentieth century when the concept of the item characteristic curve was developed by Tucker (1946) to modern IRT models such as the MRCMLM, which is used in this thesis.

The MRCMLM as applied in ACER ConQuest, is a very general form of the Rasch model (Adams et al., 1997). It encompasses the simple logistic model (Rasch, 1960 [1980]) for dichotomous indicators, the rating scale model (Andrich, 1978) and the partial credit model (Masters, 1982) for ordered categorical responses, FACETS (Linacre, 1989), among some others. In addition, the MRCMLM may be used to develop many other Rasch models by introducing user-defined design and scoring matrices. As described earlier in this chapter, the item difficulties are multiplied by a design matrix in the MRCMLM, resulting in linear combinations of the item parameters. The design matrix is automatically generated for the above models, but can be adjusted for new models. The abilities are multiplied by a scoring matrix, which describes the performance level of each response category. Hence, the measurement models in the MRCMLM are very flexible within the Rasch framework.
The estimation method for MRCML models is marginal maximum likelihood with an expectation-maximisation algorithm (MML/EM).

This chapter has shown that binary IRT and binary factor analytic measurement models are essentially the same, because the estimated parameters can be transformed from one measurement model to the other or parameterisation can be chosen in such a way that the results are equal.

Despite the similarities between IRT and factor analytic models, software applications from the two traditions, such as Mplus and ACER ConQuest, differ in choice and flexibility of their measurement models. Mplus enables measurement models with continuous indicators, binary indicators and ordered categorical indicators. The model with binary indicators is a two-parameter model, which can be restricted to a one-parameter Rasch model by constraining the slopes to be equal.
Chapter 3: Path Models

The first chapter stated that the goal of this thesis is to implement latent path modelling within an IRT framework, just as structural equation modelling has been developed within a factor analytic framework. Chapter 2 gave an overview of IRT measurement and compared IRT measurement models with factor analytic models. This chapter explains path models, estimation of parameters and identification of equations. Since the method developed for this thesis treats all variables in the path model as observed variables (see Chapter 4), latent path models are simply called path models in this chapter and no distinction is made between latent and observed variables.

**Simultaneous Equations**

The idea behind path models is that there is not one set of dependent and one set of independent variables as in ordinary regression, but that some dependent variables in one equation can be independent variables in another. Variables that are being explained in one of the equations of the path model’s system of structural equations, also called simultaneous equations, are called endogenous \( Y \) variables. All others are exogenous \( X \) variables, including the constant \( X_0 = 1 \) to estimate the intercepts. A system of structural equations has \( K \) exogenous and \( M \) endogenous variables. Therefore, the system has \( M \) equations. The variable that is being explained in equation \( m \) is denoted \( Y_m \). The matrix formulation of the system of equations for all observations \( t \) (where \( t = 1, 2, ..., T \)) is written as follows:
The assumptions for the errors are

\[ E(U) = 0 \]
\[ \text{var}(U) = E(U'U) = \Sigma_{UU} \]
\[ \text{cov}(X, U) = 0 \]

The notation in this chapter is consistent with some literature in econometrics and is different from the LISREL notation. Information used for this chapter comes from the following sources: Johnston (1960), Bollen (1989), Goldberger (1964) and two online course papers by Pollock (2001a, 2001b).

Figure 6 gives an example of a path model that will be used to illustrate the theory in this chapter. Effects between \( X \) and \( Y \) variables are in matrix \( B \) and the effects between the \( Y \) variables are in matrix \( \Gamma \) (which is the other way round in LISREL and Mplus). The first subscript of the coefficients in \( \beta_x \) and \( \gamma_y \) is the tail of the arrow and the second is the head of the arrow (which is also the other way round in LISREL and Mplus). In the example for this chapter, the grey structural parameters \( \beta_{12}, \beta_{31}, \gamma_{12}, \gamma_{13} \) and \( \gamma_{32} \) as well the effects of the \( X \) variables on \( Y_3 \) (\( \beta_{13}, \beta_{23} \) and \( \beta_{33} \), which are omitted in Figure 6 for simplicity) are fixed to zero.
Chapter 3: Path Models

Figure 6. Example of a path model

The estimation of a path model starts with a data set of responses to the $Y$ variables and to the $X$ variables by each person. The data set $Y$ has one row for every person $t$ and one column for each endogenous variable $Y_1, Y_2, \ldots, Y_M$ as shown in equation 24. The top row of $Y$ is called $Y_t^\top$ and contains the observed values of person 1 on all endogenous variables. The first column in $Y$ contains the observed values of $t$ persons on the first endogenous variable, $Y_1$.

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_T \end{bmatrix} = \begin{bmatrix} Y_{11} & Y_{12} & \cdots & Y_{1m} & \cdots & Y_{1M} \\ Y_{21} & Y_{22} & \cdots & Y_{2m} & \cdots & Y_{2M} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ Y_{T1} & Y_{T2} & \cdots & Y_{Tm} & \cdots & Y_{TM} \end{bmatrix} = \begin{bmatrix} Y_1 & Y_2 & \cdots & Y_M \end{bmatrix}$$

24
Matrix $\mathbf{X}$ can be read in a similar way (see equation 25).

$$\mathbf{X} = \begin{bmatrix} X_1 \\ \vdots \\ X_T \end{bmatrix} = \begin{bmatrix} X_{10} & \cdots & X_{1K} \\ \vdots & \ddots & \vdots \\ X_{T0} & \cdots & X_{TK} \end{bmatrix} = \begin{bmatrix} X_0 & \cdots & X_K \end{bmatrix}$$

Three additional matrices contain the parameters of the path model. The matrix $\mathbf{\Gamma}$ of regression coefficients between the endogenous variables is a square matrix, which has a row and a column for each endogenous variable. The $i,j^{th}$ element, $\gamma_{ij}$, of $\mathbf{\Gamma}$ is the coefficient of $Y_j$ regressed on $Y_i$. Typically, we set the diagonal of $\mathbf{\Gamma}$ to -1 ($\gamma_{mm} = -1$), which is called normalisation of the system of equations (see equation 26). Formulated this way, the $m^{th}$ equation explains the observed values of $Y_m$, which will be explained further below.

$$\mathbf{\Gamma} = \begin{bmatrix} -1 & \gamma_{12} & \cdots & \gamma_{1M} \\ \gamma_{21} & -1 & \cdots & \gamma_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{M1} & \gamma_{M2} & \cdots & -1 \end{bmatrix}$$

Matrix $\mathbf{B}$ contains the regression coefficients between exogenous and endogenous variables and has one row for each exogenous variable and one column for each endogenous variable.
Chapter 3: Path Models

The matrix of error terms is denoted $\mathbf{U}$.

$$
\mathbf{U} = \begin{bmatrix}
U_{11} & \cdots & U_{1M} \\
\vdots & \ddots & \vdots \\
U_{T1} & \cdots & U_{TM}
\end{bmatrix} = \begin{bmatrix}
U_1 & \cdots & U_M
\end{bmatrix}
$$

We write the system of equations for a single person in matrix form as

$$
Y_t \Gamma + X_t \mathbf{B} + U_t = 0
$$

and the general system of structural equations for each person $t$ in implicit form as

$$
\gamma_{11} Y_{t1} + \gamma_{21} Y_{t2} + \cdots + \gamma_{M1} Y_{tM} + \beta_{01} X_{t0} + \cdots + \beta_{K1} X_{tk} + U_{t1} = 0 \\
\vdots \\
\gamma_{1M} Y_{t1} + \gamma_{2M} Y_{t2} + \cdots + \gamma_{MM} Y_{tM} + \beta_{0M} X_{t0} + \cdots + \beta_{KM} X_{tk} + U_{tM} = 0
$$

As mentioned above, it is convenient to normalise the system of equations by setting one coefficient in each equation equal to -1, which is typically $\gamma_{mn}$. The following example shows how applying the normalisation rule creates single equations $m$ that explain the observed values of $Y_n$. The example in Figure 6 consists of four exogenous variables ($K=4$), including the constant $X_0=1$ to estimate the intercepts $\beta_{0}$, and three endogenous variables ($M=3$). The system of equations for each single person $t$ when
applying the normalisation rule $\gamma_{11} = -1$, $\gamma_{22} = -1$ and $\gamma_{33} = -1$ becomes equations 31, 32 and 33.

**Equation 1**

\[
\begin{align*}
\gamma_{11} Y_{t1} + \gamma_{21} Y_{t2} + \gamma_{31} Y_{t3} + \beta_{01} X_{t0} + \beta_{11} X_{t1} + \beta_{21} X_{t2} + \beta_{31} X_{t3} + U_{t1} &= 0 \\
\Rightarrow -Y_{t1} + \gamma_{21} Y_{t2} + \gamma_{31} Y_{t3} + \beta_{01} X_{t0} + \beta_{11} X_{t1} + \beta_{21} X_{t2} + \beta_{31} X_{t3} + U_{t1} &= 0 & \text{31} \\
\Rightarrow Y_{t1} &= \gamma_{21} Y_{t2} + \gamma_{31} Y_{t3} + \beta_{01} X_{t0} + \beta_{11} X_{t1} + \beta_{21} X_{t2} + \beta_{31} X_{t3} + U_{t1}
\end{align*}
\]

**Equation 2**

\[
\begin{align*}
\gamma_{12} Y_{t1} + \gamma_{22} Y_{t2} + \gamma_{32} Y_{t3} + \beta_{02} X_{t0} + \beta_{12} X_{t1} + \beta_{22} X_{t2} + \beta_{32} X_{t3} + U_{t2} &= 0 \\
\Rightarrow \gamma_{12} Y_{t1} - Y_{t2} + \gamma_{32} Y_{t3} + \beta_{02} X_{t0} + \beta_{12} X_{t1} + \beta_{22} X_{t2} + \beta_{32} X_{t3} + U_{t2} &= 0 & \text{32} \\
\Rightarrow Y_{t2} &= \gamma_{12} Y_{t1} + \gamma_{32} Y_{t3} + \beta_{02} X_{t0} + \beta_{12} X_{t1} + \beta_{22} X_{t2} + \beta_{32} X_{t3} + U_{t2}
\end{align*}
\]

**Equation 3**

\[
\begin{align*}
\gamma_{13} Y_{t1} + \gamma_{23} Y_{t2} + \gamma_{33} Y_{t3} + \beta_{03} X_{t0} + \beta_{13} X_{t1} + \beta_{23} X_{t2} + \beta_{33} X_{t3} + U_{t3} &= 0 \\
\Rightarrow \gamma_{13} Y_{t1} + \gamma_{23} Y_{t2} - Y_{t3} + \beta_{03} X_{t0} + \beta_{13} X_{t1} + \beta_{23} X_{t2} + \beta_{33} X_{t3} + U_{t3} &= 0 & \text{33} \\
\Rightarrow Y_{t3} &= \gamma_{13} Y_{t1} + \gamma_{23} Y_{t2} + \beta_{03} X_{t0} + \beta_{13} X_{t1} + \beta_{23} X_{t2} + \beta_{33} X_{t3} + U_{t3}
\end{align*}
\]

**Reduced Form Equations**

For purposes of estimation and consideration of identification, as well as some applications, the reduced form equations of equation 22 are required. If $\mathbf{\Gamma}$ is nonsingular, then the reduced form is derived by post multiplying equation 22 with $\mathbf{\Gamma}^{-1}$.

\[
(\mathbf{Y} \mathbf{\Gamma} + \mathbf{X} \mathbf{B} + \mathbf{U}) \mathbf{\Gamma}^{-1} = 0
\]

\[
\Leftrightarrow \mathbf{Y} + \mathbf{X} \mathbf{B} \mathbf{\Gamma}^{-1} + \mathbf{U} \mathbf{\Gamma}^{-1} = 0 & \text{34}
\]

\[
\Leftrightarrow \mathbf{Y} = -\mathbf{X} \mathbf{B} \mathbf{\Gamma}^{-1} - \mathbf{U} \mathbf{\Gamma}^{-1}
\]

\[
\Leftrightarrow \mathbf{Y} = \mathbf{X} \mathbf{\Pi} + \mathbf{V}
\]

---

46
where

\[ \Pi = -B\Gamma^{-1}, \]

and

\[ V = -U\Gamma^{-1}. \]

The matrix \( \Pi \) is a \( K \times M \) matrix of reduced form coefficients relating each endogenous variable to all exogenous variables. Here, each endogenous variable appears in only one equation. Each column of \( \Pi \) expresses the partial causal relationships between one endogenous and all exogenous variables. It follows that the vector \( V \) of the reduced form disturbances has the properties

\[
E(V) = E(-U\Gamma^{-1}) = 0 \\
\text{var}(V) = \text{var}(-U\Gamma^{-1}) = \Gamma^{-1} \cdot \text{var}(U) \cdot \Gamma^{-1} = \Omega
\]

Matrix Permutation and Model Identification

Different estimation methods have been developed to estimate the coefficients in path models. If there are no identification problems in a model (the model is just identified), alternative estimators will be equivalent. However, in the case of identification problems multiple solutions are valid and the choice of the best solution may differ between estimation methods. If an equation is over-identified, we get more than one estimate for some path coefficients. The estimates differ because of the stochastic nature of our estimates of the relations among variables. All of the different estimation methods for
structural equations amount to alternative ways of combining these several estimates into a better, single estimate. Equations are under-identified if an infinite number of solutions are possible.

Knowing the reduced form parameters $\Pi$ of a particular equation in a simultaneous system, applying the normalisation rule and knowing that $\Gamma$ is nonsingular, are insufficient for estimating only one set of path coefficients $\mathbf{B}$ and $\Gamma$ that yields $\Pi$, because $\Pi = -B\Gamma^{-1}$. Without additional information, we have no way of distinguishing between alternative structural systems that all generate $\Pi$.

Formally, this problem is stated as taking any $M \times M$ matrix $\mathbf{F}$ and $\mathbf{F}'$ and generating alternative structures: $\mathbf{B}' = BF$, $\mathbf{F}' = \mathbf{F} \Gamma$ and $\mathbf{B}'' = BF'$ and $\Gamma'' = \Gamma F'$. The identification problem arises because $B\Gamma^{-1} = B'\Gamma'^{-1} = B''\Gamma''^{-1}$.

The additional information required specifies fixed values for elements of $\mathbf{B}$ and $\Gamma$, which must be maintained in the alternative structures $\mathbf{B}'$, $\mathbf{B}''$, $\mathbf{F}'$ and $\Gamma''$. These restrictions impose conditions on the elements of $\mathbf{F}$. Identification is possible if $\mathbf{B}$ and $\Gamma$ are sufficiently restricted so that only the identity matrix $\mathbf{F} = \mathbf{I}$ maintains the restrictions.

Here, we will deal with identification of single equations within the system of equations with $Y_m$ as the variable being explained. It is useful to divide the model’s endogenous and exogenous variables in two groups: included (A) and excluded (B) variables for each equation $m$, where the number of endogenous variables $M = M^A + M^B$ and the number of exogenous variables $K = K^A + K^B$;

Chapter 3: Path Models
Chapter 3: Path Models

\[ Y^A_m : \text{endogenous variables included in equation } m \text{ (including the variable being explained)}, \]

\[ Y^B_m : \text{all other endogenous variables}, \]

\[ X^A_m : \text{exogenous variables included in equation } m, \]

\[ X^B_m : \text{all other exogenous variables}. \]

Therefore,

\[ \gamma^A_m : \text{column vector of free regression coefficients between } Y^A \text{ and } Y_m \text{ in equation } m \text{ and } \gamma_{mn}, \]

\[ \gamma^B_m : \text{column vector of parameters fixed to zero between endogenous variables and } Y_m, \]

\[ \beta^A_m : \text{column vector of free regression coefficients between } X^A \text{ and } Y_m \text{ in equation } m, \]

\[ \beta^B_m : \text{column vector of parameters fixed to zero between exogenous variables and } Y_m. \]

The parts of the matrices that are of interest for a single equation are ordered in such a way that the included variables and corresponding parameters that are not fixed to zero appear first, and the excluded variables and parameters that are fixed to zero appear last. This permutation is schematically represented in Figure 7 for a single equation \( m \).
Chapter 3: Path Models

\[ Y = \begin{bmatrix} Y_{11} & \cdots & Y_{14} & \cdots & Y_{1M} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ Y_{T1} & \cdots & Y_{Ta} & \cdots & Y_{TM} \end{bmatrix} \quad \gamma_m = \begin{bmatrix} \gamma_m^A \\ \gamma_m^B \end{bmatrix} \]

\[ X = \begin{bmatrix} X_{10} & \cdots & X_{14} & \cdots & X_{1K} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ X_{T0} & \cdots & X_{Ta} & \cdots & X_{T_K} \end{bmatrix} \quad \beta_m = \begin{bmatrix} \beta_m^A \\ \beta_m^B \end{bmatrix} \]

\[ U = \begin{bmatrix} U_{11} & \cdots & U_{14} & \cdots & U_{1M} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ U_{Ta} & \cdots & U_{Ta} & \cdots & U_{TM} \end{bmatrix} \]
Chapter 3: Path Models

Figure 7. Permutation of matrices for each equation $m$

Consequently, the $\Pi$ matrix must be reordered and partitioned in four parts: $\Pi^{AA}$, $\Pi^{AB}$, $\Pi^{BA}$ and $\Pi^{BB}$, where $\Pi^{AA}$ is a $K^A \times M^A$ matrix, $\Pi^{AB}$ is a $K^A \times M^B$ matrix, $\Pi^{BA}$ is a $K^B \times M^A$ matrix, and $\Pi^{BB}$ is a $K^B \times M^B$ matrix.

Once the matrices are partitioned and permuted, identification type of each equation can be determined. This is easiest to explain with an example.

An Illustration

The model in Figure 6 is used here to illustrate the permutation of the matrices and the issues of identification for each equation. The model is presented by a system of three equations (including parameters that are fixed to zero).

$$\begin{align*}
Y_1 &= \gamma_{21} Y_2 + \gamma_{31} Y_3 + \beta_{01} X_0 + \beta_{11} X_1 + \beta_{21} X_2 + \beta_{31} X_3 + U_1 \\
Y_2 &= \gamma_{12} Y_1 + \gamma_{32} Y_3 + \beta_{02} X_0 + \beta_{12} X_1 + \beta_{22} X_2 + \beta_{32} X_3 + U_2 \\
Y_3 &= \gamma_{13} Y_1 + \gamma_{23} Y_2 + \beta_{03} X_0 + \beta_{13} X_1 + \beta_{23} X_2 + \beta_{33} X_3 + U_3
\end{align*}$$

This model has 3 endogenous and 4 exogenous variables ($M = 3$ and $K = 4$). The original path and reduced matrices are:

$$\Gamma = \begin{bmatrix} -1 & 0 & 0 \\ \gamma_{21} & -1 & \gamma_{23} \\ \gamma_{31} & 0 & -1 \end{bmatrix}$$
Chapter 3: Path Models

\[ B = \begin{bmatrix} \beta_{01} & \beta_{02} & \beta_{03} \\ \beta_{11} & 0 & 0 \\ \beta_{21} & \beta_{22} & 0 \\ 0 & \beta_{32} & 0 \end{bmatrix} \]

\[ \Pi = \begin{bmatrix} \pi_{01} & \pi_{02} & \pi_{03} \\ \pi_{11} & \pi_{12} & \pi_{13} \\ \pi_{21} & \pi_{22} & \pi_{23} \\ \pi_{31} & \pi_{32} & \pi_{33} \end{bmatrix} \]

Table 4 shows the ordering of elements in each matrix of \( Y\Gamma + XB + U = 0 \) and of \( \Pi \) by equation.
Table 4

Ordering of Elements of All Matrices in a Path Model

<table>
<thead>
<tr>
<th>Equation 1 ($Y_m = Y_1$)</th>
<th>Equation 2 ($Y_m = Y_2$)</th>
<th>Equation 3 ($Y_m = Y_3$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Y_1^A = [Y_1 \ Y_2 \ Y_3]$</td>
<td>$Y_2^A = Y_2$</td>
<td>$Y_3^A = [Y_2 \ Y_3]$</td>
</tr>
<tr>
<td>$Y_1^B = empty$</td>
<td>$Y_2^B = [Y_1 \ Y_3]$</td>
<td>$Y_3^B = Y_4$</td>
</tr>
<tr>
<td>$X_1^A = [X_0 \ X_1 \ X_2]$</td>
<td>$X_2^A = [X_0 \ X_2 \ X_3]$</td>
<td>$X_3^A = X_0$</td>
</tr>
<tr>
<td>$X_1^B = X_3$</td>
<td>$X_2^B = X_1$</td>
<td>$X_3^B = [X_1 \ X_2 \ X_3]$</td>
</tr>
<tr>
<td>$\gamma_1^A = \begin{bmatrix} \gamma_{11} \ \gamma_{21} \ \gamma_{31} \end{bmatrix}$</td>
<td>$\gamma_2^A = \begin{bmatrix} \gamma_{22} = -1 \ \gamma_{21} \ \gamma_{31} \end{bmatrix}$</td>
<td>$\gamma_3^A = \begin{bmatrix} \gamma_{23} \ \gamma_{33} \end{bmatrix}$</td>
</tr>
<tr>
<td>$\gamma_1^B = empty$</td>
<td>$\gamma_2^B = \begin{bmatrix} \gamma_{12} \ \gamma_{32} \end{bmatrix}$</td>
<td>$\gamma_3^B = \gamma_{13} = 0$</td>
</tr>
<tr>
<td>$\beta_1^A = \begin{bmatrix} \beta_{01} \ \beta_{11} \ \beta_{21} \end{bmatrix}$</td>
<td>$\beta_2^A = \begin{bmatrix} \beta_{02} \ \beta_{22} \ \beta_{32} \end{bmatrix}$</td>
<td>$\beta_3^A = \begin{bmatrix} \beta_{03} \ \beta_{13} \ \beta_{23} \end{bmatrix}$</td>
</tr>
<tr>
<td>$\beta_1^B = \beta_{31} = 0$</td>
<td>$\beta_2^B = \beta_{12} = 0$</td>
<td>$\beta_3^B = \begin{bmatrix} \beta_{03} \ \beta_{23} \end{bmatrix}$</td>
</tr>
</tbody>
</table>
### Chapter 3: Path Models

<table>
<thead>
<tr>
<th>Equation 1 ( Y_m = Y_1 )</th>
<th>Equation 2 ( Y_m = Y_2 )</th>
<th>Equation 3 ( Y_m = Y_3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \Pi_A^{IA} = \begin{bmatrix} \pi_{01} &amp; \pi_{02} &amp; \pi_{03} \ \pi_{11} &amp; \pi_{12} &amp; \pi_{13} \ \pi_{21} &amp; \pi_{22} &amp; \pi_{23} \end{bmatrix} )</td>
<td>( \Pi_A^{IA} = \begin{bmatrix} \pi_{02} \ \pi_{22} \ \pi_{32} \end{bmatrix} )</td>
<td>( \Pi_A^{IA} = \begin{bmatrix} \pi_{02} &amp; \pi_{03} \end{bmatrix} )</td>
</tr>
<tr>
<td>( \Pi_B^{IB} = empty )</td>
<td>( \Pi_B^{IB} = empty )</td>
<td>( \Pi_B^{IB} = \pi_{01} )</td>
</tr>
<tr>
<td>( \Pi_B^{IB} = \begin{bmatrix} \pi_{31} &amp; \pi_{32} &amp; \pi_{33} \end{bmatrix} )</td>
<td>( \Pi_B^{IB} = \pi_{12} )</td>
<td>( \Pi_B^{IB} = \begin{bmatrix} \pi_{11} \end{bmatrix} )</td>
</tr>
<tr>
<td>( \Pi_B^{IB} = empty )</td>
<td>( \Pi_B^{IB} = \pi_{12} )</td>
<td>( \Pi_B^{IB} = \begin{bmatrix} \pi_{11} \end{bmatrix} )</td>
</tr>
</tbody>
</table>
Chapter 3: Path Models

Table 5 shows the resulting permuted matrices by equation \( m \) after putting the ordered sub matrices back together.

Table 5

Adding Re-Ordered Elements Together

<table>
<thead>
<tr>
<th>Equation 1</th>
<th>Equation 2</th>
<th>Equation 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \mathbf{Y}_1 = [Y_1 \ Y_2 \ Y_5] )</td>
<td>( \mathbf{Y}_2 = [Y_2 \ Y_1 \ Y_3] )</td>
<td>( \mathbf{Y}_3 = [Y_2 \ Y_3 \ Y_1] )</td>
</tr>
<tr>
<td>( \mathbf{X}_1 = [X_0 \ X_1 \ X_2 \ X_3] )</td>
<td>( \mathbf{X}_2 = [X_0 \ X_2 \ X_3 \ X_1] )</td>
<td>( \mathbf{X}_3 = [X_0 \ X_1 \ X_2 \ X_3] )</td>
</tr>
<tr>
<td>( \mathbf{\gamma}<em>1 = \begin{bmatrix} -1 \ \gamma</em>{21} \ \gamma_{31} \end{bmatrix} )</td>
<td>( \mathbf{\gamma}_2 = \begin{bmatrix} -1 \ 0 \ 0 \end{bmatrix} )</td>
<td>( \mathbf{\gamma}<em>3 = \begin{bmatrix} \gamma</em>{23} \ -1 \ 0 \end{bmatrix} )</td>
</tr>
<tr>
<td>( \mathbf{\beta}<em>1 = \begin{bmatrix} \beta</em>{01} \ \beta_{11} \ \beta_{21} \ 0 \end{bmatrix} )</td>
<td>( \mathbf{\beta}<em>2 = \begin{bmatrix} \beta</em>{02} \ \beta_{22} \ \beta_{32} \ 0 \end{bmatrix} )</td>
<td>( \mathbf{\beta}<em>3 = \begin{bmatrix} \beta</em>{03} \ 0 \ 0 \ 0 \end{bmatrix} )</td>
</tr>
<tr>
<td>( \mathbf{\Pi}<em>1 = \begin{bmatrix} \pi</em>{01} &amp; \pi_{02} &amp; \pi_{03} \ \pi_{11} &amp; \pi_{12} &amp; \pi_{13} \ \pi_{21} &amp; \pi_{22} &amp; \pi_{23} \ \pi_{31} &amp; \pi_{32} &amp; \pi_{33} \end{bmatrix} )</td>
<td>( \mathbf{\Pi}<em>2 = \begin{bmatrix} \pi</em>{02} &amp; \pi_{01} &amp; \pi_{03} \ \pi_{22} &amp; \pi_{21} &amp; \pi_{23} \ \pi_{32} &amp; \pi_{31} &amp; \pi_{33} \ \pi_{12} &amp; \pi_{11} &amp; \pi_{13} \end{bmatrix} )</td>
<td>( \mathbf{\Pi}<em>3 = \begin{bmatrix} \pi</em>{02} &amp; \pi_{03} &amp; \pi_{01} \ \pi_{12} &amp; \pi_{13} &amp; \pi_{11} \ \pi_{22} &amp; \pi_{23} &amp; \pi_{21} \ \pi_{32} &amp; \pi_{33} &amp; \pi_{31} \end{bmatrix} )</td>
</tr>
</tbody>
</table>

The following sections expand the permuted matrices of \( \mathbf{Y}_m \mathbf{\gamma}_m + \mathbf{X}_m \mathbf{\beta}_m + \mathbf{U}_m = \mathbf{0} \) and \( \mathbf{\Pi}_m \mathbf{\gamma}_m = -\mathbf{\beta}_m \) and tests identification for each equation.

Expanding permuted matrices and identification of Equation 1

First, the permuted matrices of each equation and its reduced form are expanded.
Chapter 3: Path Models

\[
Y_1, \gamma_1 + X_1, \beta_i + U_i = 0
\]

\[
\Rightarrow \begin{bmatrix} Y_1 & Y_2 & Y_3 \end{bmatrix} \begin{bmatrix} -1 \\ \gamma_{21} \\ \gamma_{31} \end{bmatrix} + \begin{bmatrix} X_0 & X_1 & X_2 & X_3 \end{bmatrix} \begin{bmatrix} \beta_{01} \\ \beta_{11} \\ \beta_{21} \\ 0 \end{bmatrix} + U_i = 0
\]

\[
\Pi_1 \gamma_1 = -\beta_i
\]

\[
\Rightarrow \begin{bmatrix} \pi_{01} & \pi_{02} & \pi_{03} \\ \pi_{11} & \pi_{12} & \pi_{13} \\ \pi_{21} & \pi_{22} & \pi_{23} \\ \pi_{31} & \pi_{32} & \pi_{33} \end{bmatrix} \begin{bmatrix} -1 \\ \gamma_{21} \\ \gamma_{31} \end{bmatrix} = \begin{bmatrix} \beta_{01} \\ \beta_{11} \\ \beta_{21} \\ 0 \end{bmatrix}
\]

\[
\Rightarrow \begin{bmatrix} -\pi_{01} + \pi_{02}\gamma_{21} + \pi_{03}\gamma_{31} \\ -\pi_{11} + \pi_{12}\gamma_{21} + \pi_{13}\gamma_{31} \\ -\pi_{21} + \pi_{22}\gamma_{21} + \pi_{23}\gamma_{31} \end{bmatrix} = \begin{bmatrix} \beta_{01} \\ \beta_{11} \\ \beta_{21} \end{bmatrix}
\]

\[
\Rightarrow -\pi_{31} + \pi_{32}\gamma_{21} + \pi_{33}\gamma_{31} = 0
\]

The key to identifying the first equation in the model is equation 47. If we could solve this equation after substituting the reduced parameters by their estimated values (see section on estimation), we know \( \gamma_{21} \) and \( \gamma_{31} \) (or \( \gamma_{1}^4 \)). Then, equation 46 will yield estimates of \( \beta_{01}, \beta_{11} \) and \( \beta_{21} \) (or \( \beta_1^4 \)).

It turns out that equation 47 is a set of \( K^4 \) linear equations with \( M_m^4 - 1 \) unknown values (because \( \gamma_{mm} \) is fixed to -1 for normalisation). After applying the normalisation rule, equation 47 can be solved for a unique solution if

\[
\text{rank } \Pi_m^{\beta_i} = M_m^4 - 1
\]
This condition is called the *rank condition* and is only valid if $M_m^A > 1$, otherwise the equation is a simple regression. In the first equation $M_i^A - 1 = 2$ and the rank of $\Pi_i^{BA} = 1$, because it has only one row. Therefore, the first equation of the example is *under-identified*, which is clear from equation 47, because one equation with two unknowns has an infinite number of solutions.

Calculating the rank of matrices with more rows can be difficult. Therefore, the identification criterion is changed into the form of necessary but not sufficient conditions that relate to the sizes of the submatrices of $\Pi_m^*$. 

$\Pi_m^{BA}$ is a $K^B_m \times M^A_m$ matrix. If its rank has to be $M_m^A - 1$, the number of rows (the number of excluded exogenous variables, $K^B$), must be at least the number of columns (the number of included endogenous variables) minus one ($M_m^A - 1$). 

$$K^B_m \geq M^A_m - 1.$$  

This is called the *order condition*. This is not sufficient, because satisfying the order condition does not mean that $\text{rank } \Pi_m^{BA} = M_m^A - 1$.

The rank and order conditions can be related to the type of identification for each equation in a system. There are three cases of identification: the *just identified* case, the *over-identified* case and the *under-identified* case. Equations are

- *under-identified* if $\text{rank } (\Pi_m^{BA}) < M_m^A - 1$,
- *just identified* if $\text{rank } (\Pi_m^{BA}) = M_m^A - 1$ and $K^B_m = M_m^A - 1$, or
Chapter 3: Path Models

- over-identified if \( \text{rank}\left(\Pi^B_m\right) = M^A_m - 1 \) and \( K^B_m > M^A_m - 1 \).

Two hypothetical cases illustrate satisfying the order condition but not the rank condition of the just identified case. The first case is when one exogenous variable is left out of all equations. In that case, its reduced form coefficients are zero and the rank of \( \Pi^B_m \) is reduced by 1. The second case is when two equations have exactly the same set of excluded variables, because it results in \( \Pi^B_m \) having at least two rows that are linearly dependent once the multiplication \( -\bm{B}\Gamma^{-1} \) is done\(^3\).

\(^3\) Say we have the system of equations

\[
\begin{align*}
Y_1 &= \gamma_{21}Y_2 + \beta_{01}X_0 + U_1 \\
Y_2 &= \beta_{02}X_0 + \beta_{12}X_1 + \beta_{22}X_2 + U_2, \\
Y_3 &= \gamma_{23}Y_2 + \beta_{03}X_0 + U_3
\end{align*}
\]

so that the first and third equation exclude the same set of variables \( X_1 \) and \( X_2 \). The structural parameters in simplified notation are as follows:

\[
\begin{align*}
\bm{B} &= \begin{bmatrix} a & b & c \\ 0 & e & 0 \\ 0 & h & 0 \end{bmatrix}, & \Gamma &= \begin{bmatrix} -1 & 0 & 0 \\ 0 & l & -1 \\ 0 & 0 & -1 \end{bmatrix}, \\
-\bm{B}\Gamma^{-1} &= \begin{bmatrix} a & b & c \\ 0 & e & 0 \\ 0 & h & 0 \end{bmatrix} \begin{bmatrix} -1 & 0 & 0 \\ -m & -1 & -o \\ 0 & 0 & -1 \end{bmatrix} = \begin{bmatrix} a + bm & b & bo + c \\ em & e & eo \\ hm & h & ho \end{bmatrix}, \\
\Pi^B_m &= \begin{bmatrix} \pi_{12} & \pi_{13} \\ \pi_{22} & \pi_{23} \end{bmatrix} = \begin{bmatrix} e & eo \\ h & ho \end{bmatrix}, \text{which has two linear dependent rows}
\end{align*}
\]
Equation 1 satisfies condition (a) \( \text{rank}(\Pi_m^A) < M_m^A - 1 \). It is therefore under-identified.

Expanding permuted matrices and identification of Equation 2

The second equation is an ordinary regression and therefore the rank and order conditions are not valid and identification is not an issue. Expanding the matrices has the following results.

\[
Y_2' + X_2 \beta_2 + U_2 = 0
\]

\[
\Rightarrow [Y_2' Y_1' Y_3'] [-1] + [X_0' X_2' X_3' X_1'] \begin{bmatrix} \beta_{02} \\ \beta_{22} \\ \beta_{32} \\ 0 \end{bmatrix} + U_2 = 0
\]

\[
\Pi_2' \gamma_2 = -\beta_2
\]

\[
\Rightarrow \begin{bmatrix} \pi_{02} & \pi_{01} & \pi_{03} \\ \pi_{22} & \pi_{21} & \pi_{23} \\ \pi_{32} & \pi_{31} & \pi_{33} \\ \pi_{12} & \pi_{11} & \pi_{13} \end{bmatrix} \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \beta_{02} \\ \beta_{22} \\ \beta_{32} \\ 0 \end{bmatrix}
\]

\[
\Rightarrow \begin{bmatrix} -\pi_{02} \\ -\pi_{22} \\ -\pi_{32} \end{bmatrix} = \begin{bmatrix} \beta_{02} \\ \beta_{22} \\ \beta_{32} \end{bmatrix}
\]

\[
\Rightarrow -\pi_{12} = 0
\]

Expanding permuted matrices and identification of Equation 3

The third equation has \( K_3^B = 3 \) and \( M_3^A - 1 = 1 \), which satisfies the necessary but insufficient order condition. These values also cancel out condition (b) in which
Chapter 3: Path Models

$K_m^B > M_m^A - 1$. In addition, while the exact rank of $\Pi_m^{BA}$ is not known, it cannot be less than 1 because it has three rows, which cancels out condition (a). Therefore, the third equation is likely to be overidentified.

\[ Y_3 \gamma_3 + X_3 \beta_3 + U_3 = 0 \]

\[ \Rightarrow [Y_2, Y_3, Y_4] \begin{bmatrix} \gamma_{23} \\ -1 \\ 0 \end{bmatrix} + [X_0, X_1, X_2, X_3] \begin{bmatrix} \beta_{01} \\ 0 \\ 0 \\ 0 \end{bmatrix} + U_3 = 0 \]  

\[ \Pi_3 \gamma_3 = -\beta_3 \]

\[ \Rightarrow \begin{bmatrix} \pi_{02} & \pi_{03} & \pi_{01} \\ \pi_{12} & \pi_{13} & \pi_{11} \\ \pi_{22} & \pi_{23} & \pi_{21} \\ \pi_{32} & \pi_{33} & \pi_{31} \end{bmatrix} \begin{bmatrix} \gamma_{23} \\ -1 \\ 0 \end{bmatrix} = - \begin{bmatrix} \beta_{03} \\ 0 \\ 0 \\ 0 \end{bmatrix} \]  

\[ \Rightarrow \pi_{02} \gamma_{23} - \pi_{03} = -\beta_{03} \]  

\[ \Rightarrow \begin{bmatrix} \pi_{12} \gamma_{23} - \pi_{13} \\ \pi_{22} \gamma_{23} - \pi_{23} \\ \pi_{32} \gamma_{23} - \pi_{33} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \]  

Equation 57 shows that $\gamma_{23}$ has up to three different solutions and is therefore indeed over-identified.

Two-Stage Least Squares (2SLS) Estimation

For nonrecursive models (i.e. models in which endogenous variables do not affect each other), two-stage least squares (2SLS) is the most common estimation method (Bollen,
1989). Just identified and over-identified equations can both be solved using the 2SLS method. In the over-identified case, there is no unique solution for the structural parameters due to the deviations of the estimated reduced form parameters from their true values. The third equation of the system of equations of the model in Figure 6 is an example of the over-identified case. Equation 57 shows that substitution of the reduced parameters could yield three different estimates for $\gamma_{23}$.

A method needs to be found for combining the information contained in all possible estimates for each coefficient. The most common procedure is two-stage least squares. This method estimates the parameters equation by equation. For each equation, the values of included endogenous variables are replaced by their predicted values (also called instruments), obtained from the reduced parameter estimates. In this first stage, each of the included (explained and explanatory) endogenous variables is regressed on the complete set of exogenous variables using the ordinary least squares (OLS) method. The resulting predicted values of these endogenous variables are subsequently used to estimate the path model parameters in the second stage. This combined method is called indirect least squares.

However, the computations do not need to be performed in two stages. A computationally more efficient way to estimate the two-stage least squares coefficients for a single equation can be derived. For each equation, the dependent endogenous variable is denoted by $y_m$, the included endogenous variables that are not being explained ($Y_m^{(d-1)}$) are
substituted by \( \hat{Y}_m^{(A-1)} \) (A-I is used to indicate that the dependent variable is not included in this set of endogenous variables). Now, the \( m^{th} \) structural equation is written as

\[
\hat{Y}_m^{(A-1)} + X_m A^m + U_m = Y_m. \tag{58}
\]

Premultiplying the left and the right side of equation 58 by \( A_m' \) gives

\[
(Y_m^{(A-1)})'X_m' \hat{Y}_m^{(A-1)} + X_m' A_m^m + X_m' U_m = (Y_m^{(A-1)})' Y_m, \tag{59}
\]

and premultiplying the left and the right side of equation 58 by \( X_m' \) gives

\[
X_m' X_m' \hat{Y}_m^{(A-1)} + X_m' A_m^m + X_m' U_m = X_m' Y_m. \tag{60}
\]

Since the expected values of \( \hat{Y}_m^{(A-1)} U_m \) and \( X_m' U_m \) are zero, equation 59 and 60 can be written in matrix form;

\[
\begin{bmatrix}
Y_m^{(A-1)} & X_m' A_m^m
\end{bmatrix}
\begin{bmatrix}
\gamma_m^{(A-1)} \\
\beta_m
\end{bmatrix}
= \begin{bmatrix}
\hat{Y}_m^{(A-1)} \\
X_m' Y_m
\end{bmatrix}. \tag{61}
\]

In order to avoid calculating the path model parameters in two steps, the sums of squares and cross products that include \( \hat{Y}_m^{(A-1)} \) have to be substituted. Two formulae will be used for the substitution:

\[
\hat{Y}_m^{(A-1)} = Y_m^{(A-1)} - V^*_m \quad \text{and} \quad \hat{Y}_m^{(A-1)} = Y_m^{(A-1)} - V^*_m \tag{62}
\]
Chapter 3: Path Models

\[ \hat{Y}_m^{(A-1)} = X(X'X)^{-1}XY_m^{(A-1)} \]

which is the reduced form in the first step of 2SLS. Using both equation 62 and 63 for substitution of \( \hat{Y}_m^{(A-1)} \hat{Y}_m^{(A-1)} \) in equation 61 gives

\[
\begin{align*}
\hat{Y}_m^{(A-1)'} \hat{Y}_m^{(A-1)} &= \hat{Y}_m^{(A-1)'} \left( Y_m^{(A-1)} - V_m \right) \\
 &= \hat{Y}_m^{(A-1)'} Y_m^{(A-1)} - \hat{Y}_m^{(A-1)'} V_m \\
 &= \hat{Y}_m^{(A-1)'} Y_m^{(A-1)} \\
 &= \left\{ X(X'X)^{-1}X'Y_m^{(A-1)} \right\} Y_m^{(A-1)} \\
 &= Y_m^{(A-1)'} X(X'X)^{-1}X'Y_m^{(A-1)}
\end{align*}
\]

Similarly, \( \hat{Y}_m^{(A-1)'} X_m^A \) and \( \hat{Y}_m^{(A-1)'} Y_m \) can respectively be written as equations 65 and 66.

\[
\begin{align*}
\hat{Y}_m^{(A-1)'} X_m^A &= \left( Y_m^{(A-1)} - V_m \right)' X_m^A \\
 &= Y_m^{(A-1)'} X_m^A - V_m' X_m^A \\
 &= Y_m^{(A-1)'} X_m^A
\end{align*}
\]

\[
\begin{align*}
\hat{Y}_m^{(A-1)'} Y_m &= \left( X(X'X)^{-1} X'Y_m^{(A-1)} \right)' Y_m \\
 &= Y_m^{(A-1)'} X(X'X)^{-1} X'Y_m
\end{align*}
\]

Using equations 65 and 66 to substitute the indicated sums of squares and cross products in equation 61 leads to equation 67. This is called the \textit{limited information method}.  

\footnote{Using the OLS estimator \( B = (X'X)^{-1} X'Y \) in \( \hat{Y} = XB \) (see also equation 74 in Chapter 4).}
Chapter 3: Path Models

\[
\begin{bmatrix}
Y_m^{(d-1)'} X (X'X)^{-1} X' Y_m^{(d-1)} & Y_m^{(d-1)'} X_m^{d}\beta_m^{(d-1)} \\
X_m^{d'} Y_m^{(d-1)} & X_m^{d'} X_m^{d}\end{bmatrix}
\begin{bmatrix}
\gamma_m^{(d-1)} \\
\beta_m^{d'}
\end{bmatrix}
= \begin{bmatrix}
Y_m^{(d-1)'} X (X'X)^{-1} X' Y_m \\
X_m^{d'} Y_m
\end{bmatrix}
\]

\[
\begin{bmatrix}
Y_m^{(d-1)'} X (X'X)^{-1} X' Y_m^{(d-1)} & Y_m^{(d-1)'} X_m^{d} \\
X_m^{d'} Y_m^{(d-1)} & X_m^{d'} X_m^{d}\end{bmatrix}^{-1}
\begin{bmatrix}
Y_m^{(d-1)'} X (X'X)^{-1} X' Y_m \\
X_m^{d'} Y_m
\end{bmatrix}
\]

Standard Errors

The asymptotic covariance matrix of the parameters may be written as

\[
S_m = \frac{1}{s_m^2}
\begin{bmatrix}
Y_m^{(d-1)'} X (X'X)^{-1} X' Y_m^{(d-1)} & Y_m^{(d-1)'} X_m^{d} \\
X_m^{d'} Y_m^{(d-1)} & X_m^{d'} X_m^{d}\end{bmatrix}^{-1},
\]

where \(s_m^2\) is the estimated variance of the error term or the residual variance of the explained variable (Goldberger, 1964).
\[ s_m^2 = \frac{1}{N} \left( Y_m - Y_m^{(d-1)} Y_m^{(d-1)} - X_m^d \beta_m^d \right) \left( Y_m - Y_m^{(d-1)} Y_m^{(d-1)} - X_m^d \beta_m^d \right)^T \]

\[ = \frac{1}{N} \left( \gamma_m m Y_m^{(d-1)} Y_m^{(d-1)} - Y_m^d X_m^d \beta_m^d - \gamma_m m Y_m^{(d-1)} Y_m^{(d-1)} + \gamma_m m Y_m^{(d-1)} Y_m^{(d-1)} \right) \]

\[ + \gamma_m m Y_m^{(d-1)} X_m^d \beta_m^d - \beta_m^d X_m^d Y_m^{(d-1)} + \beta_m^d X_m^d X_m^d \beta_m^d \]

\[ = \frac{1}{N} \left( \gamma_m m Y_m^{(d-1)} Y_m^{(d-1)} - Y_m^d Y_m^{(d-1)} - \gamma_m m Y_m^{(d-1)} Y_m^{(d-1)} + \gamma_m m Y_m^{(d-1)} Y_m^{(d-1)} \right) \]

\[ + \gamma_m m Y_m^{(d-1)} X_m^d Y_m^{(d-1)} + \beta_m^d X_m^d Y_m^{(d-1)} + \beta_m^d X_m^d X_m^d \beta_m^d \]

\[ = \frac{1}{N} \left( \gamma_m m Y_m^{(d-1)} Y_m^{(d-1)} + \gamma_m m Y_m^{(d-1)} Y_m^{(d-1)} \right) \]

\[ - \gamma_m m Y_m^{(d-1)} Y_m^{(d-1)} \]

\[ + 2 \gamma_m m Y_m^{(d-1)} X_m^d Y_m^{(d-1)} + \beta_m^d X_m^d Y_m^{(d-1)} + \beta_m^d X_m^d X_m^d \beta_m^d \]

\[ = \frac{1}{N} \left( \gamma_m m Y_m^{(d-1)} Y_m^{(d-1)} \right) \]

\[ - \gamma_m m Y_m^{(d-1)} Y_m^{(d-1)} \]

\[ + 2 \gamma_m m Y_m^{(d-1)} X_m^d Y_m^{(d-1)} + \beta_m^d X_m^d Y_m^{(d-1)} + \beta_m^d X_m^d X_m^d \beta_m^d \]

\[ = \frac{1}{N} \left( \gamma_m m Y_m^{(d-1)} Y_m^{(d-1)} \right) \]

\[ - \gamma_m m Y_m^{(d-1)} Y_m^{(d-1)} \]

\[ + 2 \gamma_m m Y_m^{(d-1)} X_m^d Y_m^{(d-1)} + \beta_m^d X_m^d Y_m^{(d-1)} + \beta_m^d X_m^d X_m^d \beta_m^d \]

\[ = \frac{1}{N} \left( \gamma_m m Y_m^{(d-1)} Y_m^{(d-1)} - \gamma_m m Y_m^{(d-1)} Y_m^{(d-1)} + \gamma_m m Y_m^{(d-1)} Y_m^{(d-1)} \right) \]

\[ - 2 \gamma_m m Y_m^{(d-1)} Y_m^{(d-1)} \]

\[ + 2 \gamma_m m Y_m^{(d-1)} X_m^d Y_m^{(d-1)} + \beta_m^d X_m^d Y_m^{(d-1)} + \beta_m^d X_m^d X_m^d \beta_m^d \]

\[ = \frac{1}{N} \left( \gamma_m m Y_m^{(d-1)} Y_m^{(d-1)} \right) \]

\[ - \gamma_m m Y_m^{(d-1)} Y_m^{(d-1)} \]

\[ + 2 \gamma_m m Y_m^{(d-1)} X_m^d Y_m^{(d-1)} + \beta_m^d X_m^d Y_m^{(d-1)} + \beta_m^d X_m^d X_m^d \beta_m^d \]

\[ = \frac{1}{N} \left( \gamma_m m Y_m^{(d-1)} Y_m^{(d-1)} - \gamma_m m Y_m^{(d-1)} Y_m^{(d-1)} + \gamma_m m Y_m^{(d-1)} Y_m^{(d-1)} \right) \]

\[ - 2 \gamma_m m Y_m^{(d-1)} Y_m^{(d-1)} \]

\[ + 2 \gamma_m m Y_m^{(d-1)} X_m^d Y_m^{(d-1)} + \beta_m^d X_m^d Y_m^{(d-1)} + \beta_m^d X_m^d X_m^d \beta_m^d \]

\[ = \frac{1}{N} \left( \gamma_m m Y_m^{(d-1)} Y_m^{(d-1)} - \gamma_m m Y_m^{(d-1)} Y_m^{(d-1)} + \gamma_m m Y_m^{(d-1)} Y_m^{(d-1)} \right) \]

\[ - 2 \gamma_m m Y_m^{(d-1)} Y_m^{(d-1)} \]

\[ + 2 \gamma_m m Y_m^{(d-1)} X_m^d Y_m^{(d-1)} + \beta_m^d X_m^d Y_m^{(d-1)} + \beta_m^d X_m^d X_m^d \beta_m^d \]

\[ = \frac{1}{N} \left( \gamma_m m Y_m^{(d-1)} Y_m^{(d-1)} - \gamma_m m Y_m^{(d-1)} Y_m^{(d-1)} + \gamma_m m Y_m^{(d-1)} Y_m^{(d-1)} \right) \]

\[ - 2 \gamma_m m Y_m^{(d-1)} Y_m^{(d-1)} \]

\[ + 2 \gamma_m m Y_m^{(d-1)} X_m^d Y_m^{(d-1)} + \beta_m^d X_m^d Y_m^{(d-1)} + \beta_m^d X_m^d X_m^d \beta_m^d \]

where \( X_m^d Y_m - \gamma_m m Y_m^{(d-1)} Y_m^{(d-1)} X_m^d \) in the fifth line is equal to \( X_m^d X_m^d \beta_m^d \) in the sixth line, because the second row of the first part of equation 67 shows that

\[ \beta_m^d = \left( X_m^d X_m^d \right)^{-1} \left( X_m^d Y_m - X_m^d Y_m^{(d-1)} Y_m \right) \]
Explained Variance

Multiplying the conditional variance \( s_m^2 \) with the degrees of freedom results in the conditional sums of squares

\[
U_m'U_m = s_m^2 \left( T - (K_m^4 + M_m^4) \right).
\]

The conditional sums of squares can be used to compute the explained variance \( R_m^2 \).
Chapter 3: Path Models

\[ R_m^2 = 1 - \left( \frac{U'_m U_m}{\sum (Y_m - \bar{Y}_m)^2} \right) \]

\[ = 1 - \left( \frac{U'_m U_m}{\sum (Y_m - \bar{Y}_m)(Y_m - \bar{Y}_m)} \right) \]

\[ = 1 - \left( \frac{U'_m U_m}{\sum Y_m^2 - 2 \sum Y_m \bar{Y}_m + \bar{Y}_m^2} \right) \]

\[ = 1 - \left( \frac{U'_m U_m}{\sum Y_m^2 - 2 \bar{Y}_m \sum Y_m + N \bar{Y}_m^2} \right) \]

\[ = 1 - \left( \frac{U'_m U_m}{\sum Y_m^2 - 2 N \bar{Y}_m^2 + N \bar{Y}_m^2} \right) \]

\[ = 1 - \left( \frac{U'_m U_m}{\sum Y_m^2 - N \bar{Y}_m^2} \right) \]

\[ = 1 - \left( \frac{U'_m U_m}{\sum Y_m^2 - N \left( N^{-1} \sum Y_m \right)^2} \right) \]

\[ = 1 - \left( \frac{U'_m U_m}{\sum Y_m^2 - N \left( N^{-2} \sum Y_m \right) \sum Y_m} \right) \]

\[ = 1 - \left( \frac{U'_m U_m}{\sum Y_m^2 - N^{-1} \left( \sum Y_m \right)^2} \right) \]

All elements in this equation have known values as will be shown in the next chapter.

One problem that can occur when using 2SLS is a negative or very low explained variance. This is caused by the substitution of \( Y_m^{(A-1)} \) by \( \hat{Y}_m^{(A-1)} \) in equation 58. Because
Chapter 3: Path Models

\( \hat{Y}_m^{(A-1)} = \tilde{Y}_m^{(A-1)} + V_m \), the error term \( U_m \) includes the error term of the reduced form.

Depending on the predictive power of the instrumental variables, the error term may become very large and therefore the explained variance very low or even negative. This would make the explained variance statistically not meaningful. Many software packages, such as Stata, do not print negative R-squares or fix them to zero. Fortunately, a very low or negative R-squared does not imply that the model parameter estimates are no good (Sribney, Wiggins, Drukker, & StataCorp, 1999).
Chapter 4: Multi-Step Method

This chapter will bring together the IRT model that was described in Chapter 2 and the 2SLS method from Chapter 3 to estimate latent path models. The current MRCML model (see Figure 2) does not include latent path models; it can only include multiple latent regression models. A further limitation of the MRCML regression model is that latent variables can only be explained by observed variables. This chapter will describe a new method that expands the regression model of the MRCMLM so that latent path models can be estimated in conjunction with its IRT measurement models.

As will be shown, constructing a sums of squares and cross products (SSCP) matrix of all observed and latent variables in the latent path model—also referred to as structural variables—enables us to estimate the wider range of regression coefficients required in path modelling. The SSCP is not a direct product of an MRCML model, but it will be shown here how an estimate can be constructed from the parameter estimates of the MRCML model.

Suppose we are interested in estimating the parameters of the model in Figure 8. This model includes both a measurement and a latent path model. The full model has two latent variables ($L_1$ and $L_2$), measured by $k + m$ indicators ($Z_{11}$ to $Z_{1k}$ and $Z_{21}$ to $Z_{2m}$) and five observed variables ($O_1$ to $O_5$).
Figure 8. Hypothetical latent path model

The latent path model consists of three endogenous variables (Y), and four exogenous variables (X). The endogenous variables are the latent variables $L_1$ and $L_2$ re-labelled as $Y_1$ and $Y_2$, respectively, and the observed variable $O_5$, re-labelled as $Y_3$. The exogenous variables are $O_1$ to $O_4$, re-labelled as $X_1$ to $X_4$, respectively. Consistent with previous chapters, the path coefficients between endogenous and exogenous variables are called $\beta_{ij}$ and between endogenous variables $\gamma_{ij}$ with $i$ indicating the explanatory variable and $j$ the explained variable. The labels of the parameters in the measurement model are omitted from the figures in this chapter, because they do not play a role in the method to estimate the path model parameters.

To estimate the parameters of this path model, three steps are necessary:
Chapter 4: Multi-Step Method

1. estimate the standard MRCML model
2. estimate the SSCP matrix for all path model variables
3. perform two-stage least squares (2SLS).

These three steps are described in detail in this chapter.

Step 1: Estimate the MRCML Model

In the first step, a standard MRCML model with a measurement model and a multiple latent regression model with all observed variables as independent and latent variables as dependent variables is estimated. This model is shown in Figure 9.

![Figure 9. Standard MRCMLM with measurement and multiple latent regression model](image)
Chapter 4: Multi-Step Method

The regression coefficients between the observed and latent variables are indicated by \( \beta_{ij} \), with \( i \) referring to the explanatory variable and \( j \) to the explained variable. The conditional variances (\( \sigma_{u_1}^2 \) and \( \sigma_{u_2}^2 \)) and the covariance between the residuals \( u_1 \) and \( u_2 \) are estimated as well. Both sets of parameter estimates are needed in step 2. The matrix \( \hat{B}^c \) denotes the matrix of regression coefficients and \( \hat{\Sigma}_{c\theta} \) is the conditional variance-covariance matrix of the latent variables as produced by estimation of the MRCML model in ACER ConQuest in this first step.

**Step 2: Estimate the SSCP Matrix for All Path Model Variables**

From the resulting MRCML regression and variance estimates the full SSCP can be obtained using ordinary least squares (OLS) equations. Suppose we have the two sets of variables of the multiple regression model \( L=(L_1, L_2, \ldots, L_M) \) and \( O=(O_1, O_2, \ldots, O_K) \), where \( L \) are latent variables and \( O \) are observed variables. If the \( L \) set is regressed on the \( O \) set, as in Figure 9, and if the variables are ordered appropriately, then the SSCP of all variables can be partitioned into four parts:

\[
SSCP = \begin{bmatrix}
L'L & L'O \\
O'L & O'O
\end{bmatrix}
\]

The regression model can be formally written as

\[
L = OB + E
\]

72
where \( \mathbf{L} \) is the vector of \( M \) latent dependent variables, \( \mathbf{O} \) is the vector of \( K \) observed independent variables, \( \mathbf{B} \) is a \( K \times M \) matrix of regression coefficients, and \( \mathbf{E} \) is a vector of \( M \) residuals. The variance of the residuals is \( \text{VAR}(\mathbf{E}) \)—the variance of \( \mathbf{L} \) that is not explained by \( \mathbf{O} \). The OLS estimators (Kendall, Stuart, & Ord, 1987) can be used to obtain the full SSCP. The OLS estimator of the regression parameters is given in equation 74.

\[
\mathbf{B} = (\mathbf{O}'\mathbf{O})^{-1} \mathbf{O}'\mathbf{L}
\]

(74)

The OLS estimator of the residual variance is the total variance of \( \mathbf{L} \) minus the explained variance of \( \mathbf{L} \) by \( \mathbf{O} \). When using sums of squares and cross products instead of variances and covariances, the computation is

\[
\mathbf{E}'\mathbf{E} = \mathbf{L}'\mathbf{L} - \mathbf{L}'\mathbf{O}(\mathbf{O}'\mathbf{O})^{-1}\mathbf{O}'\mathbf{L},
\]

(75)

where \( \mathbf{E}'\mathbf{E} = \hat{\boldsymbol{\Sigma}}_{\ell\ell}^{e} (N-K) \) and \( N \) is the number of respondents. This information can be used to construct the full marginal SSCP of the latent and observed variables in the latent path model.

After fitting the model in Figure 9 in step 1 we have estimates of the conditional variances and covariances of the latent variables, denoted by \( \hat{\boldsymbol{\Sigma}}_{\ell\ell}^{e} \), and a matrix of regression coefficients between the latent variables and the observed variables, denoted by \( \hat{\mathbf{B}}^{c} \). Note that the superscript \( c \) is used to indicate that the parameter estimates are obtained from the MRCML model in step 1.
Substituting this information into equations 74 and 75 will enable us to estimate all parts of the marginal SSCP: $L'O$, $L'L$, $O'L$ and $O'O$. Note that the SSCP of $O$, the observed variables, can be estimated directly from the data.

\[
L'O = \hat{B}'(O'O) \tag{76}
\]

\[
L'L = E'E + L'O(O'O)^{-1} O'L
= E'E + \hat{B}'(O'O)(O'O)^{-1} O'L
= E'E + \hat{B}'(O'O)\hat{B} \tag{77}
\]

\[
O'L = (O'O)\hat{B} \tag{78}
\]

\[
SSCP = \begin{bmatrix} L'L & L'O \\ O'L & O'O \end{bmatrix} = \begin{bmatrix} E'E + \hat{B}'(O'O)\hat{B} & \hat{B}'(O'O) \\ (O'O)\hat{B} & O'O \end{bmatrix} \tag{79}
\]

---

6 Equation 79 is a simplified form of the full covariance matrix of full structural equation models. In this simplified form, each variable of the structural part is only measured by one indicator in the measurement parts on both endogenous and exogenous sides and relationships between endogenous structural variables are fixed to zero. In its full form, using SEM notation, the covariance matrix is as follows.

\[
\Sigma = \begin{bmatrix} \Sigma_y & \Sigma_{yx} \\ \Sigma_{xy} & \Sigma_x \end{bmatrix} = \begin{bmatrix} \Lambda_y\Lambda(y\Gamma \Gamma' + \Psi)\Lambda_y' + \Theta \varepsilon & \Lambda_y\Lambda\Gamma\Phi\Lambda_x' \\ \Lambda_x\Phi\Gamma'\Lambda_y' & \Lambda_x\Phi\Lambda_x' + \Theta \delta \end{bmatrix}
\]

where $\Lambda_y$ is a matrix of factor loadings on the endogenous side of the model, $\Lambda_x$ is a matrix of factor loadings on the exogenous side of the model, $\Gamma$ is a matrix of regression coefficients between...
Chapter 4: Multi-Step Method

From the MRCML parameter estimates we have been able to construct an estimate of the full marginal SSCP matrix of all variables in the path model (the latent and the observed). Access to this matrix allows a suite of latent path models for these variables, for example treating the latent background variable, such as SES, as an exogenous variable and estimating nonrecursive effects between endogenous variables. Step 3 explains how the 2SLS estimation method as described in Chapter 3 can be used to produce estimates for the latent path model parameters.

\[ \Sigma = \begin{bmatrix} \Sigma_{YY} & \Sigma_{YX} \\ \Sigma_{XY} & \Sigma_{XX} \end{bmatrix} = \begin{bmatrix} A(\Phi\Gamma' + \Psi)A' & A\Gamma\Phi \\ \Phi\Gamma'A' & \Phi \end{bmatrix} \]

When no relationships between the endogenous variables are hypothesised, as in the examples of this proposal, \( A \) is the identity matrix. In this case, the equations can be further simplified to equation 79.
Step 3: Perform Two-Stage Least Squares (2SLS)

In the first two steps, the path model variables are classified as latent or observed. When applying 2SLS for estimating path model parameters, the variables need to be re-labelled as endogenous and exogenous variables. As explained in Chapter 3, endogenous variables are dependent variables in one of the equations of the system of equations and exogenous variables are the variables that are not being explained by the model (they have no arrows pointing toward them). In this third step the set of endogenous variables are denoted by $Y$ and the set of exogenous variables by $X$. In the example in this chapter there are three $Y$ variables ($L_1, L_2$ and $O_5$) and four $X$ variables ($O_1$ to $O_4$).

After re-labelling the variables as endogenous and exogenous, a permutation of the SSCP matrix is necessary. The first part of the permutation consists of reordering the variables so that the full SSCP matrix starts with the endogenous variables followed by all the exogenous variables (of which the first one is a constant of 1 to estimate the intercept):

$$SSCP^* = \begin{bmatrix} Y'Y & Y'X \\ X'Y & X'X \end{bmatrix}$$

This permuted SSCP matrix with four partitions is used as a starting point to create the submatrices that are used for each equation of the system of equations as described in Chapter 3. The next step is to further permute each partition of $SSCP^*$ so that the order of the variables for each equation in the path model is consistent with the descriptions in Chapter 3; that is:
Chapter 4: Multi-Step Method

1. endogenous variables included in equation $m$ ($Y^A_m$)

2. endogenous variables excluded from equation $m$ ($Y^B_m$)

3. exogenous variables included in equation $m$ ($X^A_m$)

4. exogenous variables excluded from equation $m$ ($X^B_m$)

Chapter 3 also explains how to apply 2SLS to estimate the parameters of the path model and provides formulae for estimating the standard errors of the parameters and the explained variance of the model. These three formulae are repeated here.

The path model parameters are calculated as follows (see also equation 67 in Chapter 3):

$$
\begin{bmatrix}
Y^{(A-1)}_m \\
\beta^A_m
\end{bmatrix}
= 
\begin{bmatrix}
Y^{(A-1)y}_m (X'X)^{-1} X'Y^{(A-1)}_m & Y^{(A-1)y}_m X^A_m \\
X^A_m (X'X)^{-1} X'Y^{(A-1)}_m & X^A_m X^A_m
\end{bmatrix}^{-1}
\begin{bmatrix}
Y^{(A-1)y}_m (X'X)^{-1} X'Y_m \\
X^A_m Y_m
\end{bmatrix}
$$

The superscript $(A-I)$ is used for the endogenous variables included in equation $m$, minus the explained variable. The next formulae are used to compute the standard errors, with $s^2_m$ representing the residual variance of the explained variable:

$$
S_m = s^2_m 
\begin{bmatrix}
Y^{(A-1)y}_m (X'X)^{-1} X'Y^{(A-1)}_m & Y^{(A-1)y}_m X^A_m \\
X^A_m (X'X)^{-1} X'Y^{(A-1)}_m & X^A_m X^A_m
\end{bmatrix}
$$

$$
s^2_m = \frac{1}{N}
\left(Y_m'Y_m + Y^{(A-1)y}_m Y^{(A-1)y}_m Y^{(A-1)y}_m Y^{(A-1)y}_m - 2Y^{(A-1)y}_m Y^{(A-1)y}_m - \beta^A_m X^A_m X^A_m \beta^A_m \right)
$$
Chapter 4: Multi-Step Method

If $s^2$ is the residual variance, then the conditional sums of squares can be computed by multiplying this variance with its degrees of freedom:

$$U_m'U_m = s_m^2 \left( T - (K_m^A + M_m^A) \right),$$

where $T$ is the number of observations, $K_m^A$ the number of included exogenous variables and $M_m^A$ the number of included endogenous variables. The conditional sums of squares can be used to estimate the explained variance by each equation:

$$R_m^2 = 1 - \left( \frac{U_m'U_m}{\sum Y_m^2 - N^{-1} \left( \sum Y_m \right)^2} \right)$$

The parts of the denominators are cells from the full sums of squares and cross products (SSCP) matrix. $\sum Y_m^2$ is the SSCP of $Y_m$, $N$ is the SSCP of the constant, $\sum Y_m$ is the SSCP of $Y_m$ with the constant.

Summary

The multi-step model that has been developed and described in this chapter enables researchers to combine the full range of MRCML models with nonrecursive latent path models. In the standard MRCML model, latent variables can only be regressed on observed variables in a latent multiple regression model. The new feature in ACER Conquest enriches the MRCML model so that any latent or observed variables can explain and/or be explained by any of the other variables in the model.
Chapter 4: Multi-Step Method

The method comprises three steps. First a standard MRCML model is estimated where all latent variables are regressed on all observed variables—a latent multiple regression model. In step 2, the estimated regression coefficients and the residual variances and covariances from step 1 are used in OLS equations to obtain an estimate of the marginal SSCP matrix of all latent and observed variables in the path model of interest. In step 3, the variables are re-labelled to exogenous and endogenous variables and the SSCP matrix is permuted accordingly. Before applying 2SLS for the estimation of path coefficients, standard errors of the parameters and explained variance of the model, the partitions of the permuted SSCP matrix are further permuted for each equation of the path model. An interesting advantage of this multi-step model is that step two and three of this method can also be applied to other models than the item response models included in the MRCMLM.

The multi-step method is tested in a simulation study in Chapter 5 and applied to real data in Chapter 6.
Chapter 5: Simulation Study

A simulation study was designed to evaluate results from the multi-step approach to estimate latent path models as an extension of the MRCML model as described in Chapter 4. Response data was generated knowing the true parameters for the path model and the explained variance of the endogenous variables. Indicator variables for measuring the latent variables were dichotomous so that equivalent measurement models could be analysed in ACER ConQuest and in Mplus. Three versions were created for the same model, but with different values for the parameters and different numbers of items in the measurement models. For each version, 100 replications were created of this data set to take random variation into account. The simulation study consisted of three parts.

The purpose of the first part of this simulation study was to detect bias in the path model estimates from the multi-step method. In this part, the response data sets were analysed in ACER ConQuest using the multi-step approach combining the simple logistic IRT model with a latent path model. The same model was applied to the 100 data set and the parameter estimates were averaged over the 100 replications. The average estimates were compared with the true parameters to find possible bias in the path model parameter estimates.

The purpose of the second part of the simulation study was to evaluate the implementation of 2SLS in the multi-step method by comparing the path model estimates with estimates from other software packages. Only the path model was estimated in this part by using the true values for the latent variables, thus ignoring the measurement part of
the model. Two data sets were analysed again with different true parameters using the 2SLS procedure in ACER ConQuest and in SPSS and using full information maximum likelihood in Mplus. Since the same variables were used using true values for latent variables, no replications were necessary. The path coefficients were compared between the three applications of the model.

The third part of the simulation study was designed to compare two approaches for combining measurement models with latent path models. Results from the full model, including the measurement part, were compared between the multi-step approach in ACER ConQuest and a structural equation model including a measurement part for dichotomous items in Mplus.

**Generating Data**

The theoretical model in Figure 10 was used for all simulations in this study. The variables called $L$ in the ovals are the latent variables. The latent variables are measured by the dichotomous items, $Z$. The first latent variable is measured by $k$ items, the second latent variable by $m$ items. The observed variables, $O$, are included in the latent path model. The endogenous variables are $L_1, L_2$ and $O_5$ (re-labelled for the purpose of the path model to $Y_1$, $Y_2$ and $Y_3$, respectively); the exogenous variables are $O_1$ to $O_4$ (re-labelled to $X_1$ to $X_4$, respectively).
Latent Path Model

Three equations describe the path model that is illustrated in Figure 10:

Equation 1 \[ y_1 = \beta_{01} + \beta_{11}x_1 + \beta_{21}x_2 + \mu_i \]

Equation 2 \[ y_2 = \beta_{02} + \gamma_{12}y_1 + \mu_2 \]

Equation 3 \[ y_3 = \beta_{03} + \gamma_{13}y_1 + \beta_{31}x_3 + \beta_{43}x_4 + \mu_3 \]

where \( \beta_{ij} \) are path coefficients between endogenous and exogenous variables, \( \gamma_{ij} \) are path coefficients between pairs of endogenous variables, and \( \mu_j \) are the residuals for each of the equations, with \( \text{var}(\mu_j) = \sigma_j^2 \).

For generating data, the distributions of the variables were first defined as follows:
$x_1 \sim N(0,1)$

$x_2 \sim \text{Bernoulli}(p_2)$

$x_3 \sim N(0,1)$

$x_4 \sim \text{Bernoulli}(p_4)$

$y_1 \sim N(0,1); \mu_1 \sim N(0,\sigma_1^2)$

$y_2 \sim N(0,1); \mu_2 \sim N(0,\sigma_2^2)$

$y_3 \sim N(0,1); \mu_3 \sim N(0,\sigma_3^2)$

Subsequently, realistic true values were specified for some parameters. Then, a second version was created where the specified true parameters were more extreme. The third version was identical to the second version, except for an increased number of items that measured the latent variables. The specified values and the constraints of the distributions listed below were used to derive the values of the remaining parameters. The three values of each parameter of both the chosen and the derived coefficients are included in Table 6, followed by the formulae used to derive values for the remaining parameters. The specified values are shaded, the derived are not shaded.
Table 6

Path Model Parameters for Each Version of the Simulation

<table>
<thead>
<tr>
<th></th>
<th>Version 1</th>
<th>Version 2</th>
<th>Version 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Equation 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$p_2$</td>
<td>0.50</td>
<td>0.10</td>
<td>0.10</td>
</tr>
<tr>
<td>$\beta_{01}$</td>
<td>-0.15</td>
<td>-0.01</td>
<td>-0.01</td>
</tr>
<tr>
<td>$\beta_{11}$</td>
<td>0.85</td>
<td>0.31</td>
<td>0.31</td>
</tr>
<tr>
<td>$\beta_{21}$</td>
<td>0.30</td>
<td>0.10</td>
<td>0.10</td>
</tr>
<tr>
<td>$\sigma_1^2$</td>
<td>0.25</td>
<td>0.90</td>
<td>0.90</td>
</tr>
<tr>
<td>Equation 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_{02}$</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>$\gamma_{12}$</td>
<td>0.60</td>
<td>0.14</td>
<td>0.14</td>
</tr>
<tr>
<td>$\sigma_2^2$</td>
<td>0.64</td>
<td>0.98</td>
<td>0.98</td>
</tr>
<tr>
<td>Equation 3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$p_4$</td>
<td>0.20</td>
<td>0.10</td>
<td>0.10</td>
</tr>
<tr>
<td>$\beta_{03}$</td>
<td>-0.12</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>$\gamma_{13}$</td>
<td>0.30</td>
<td>0.04</td>
<td>0.04</td>
</tr>
<tr>
<td>$\beta_{33}$</td>
<td>0.20</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>$\beta_{43}$</td>
<td>0.60</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>$\sigma_3^2$</td>
<td>0.81</td>
<td>0.996</td>
<td>0.996</td>
</tr>
<tr>
<td>Number of items</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$l_k$</td>
<td>10</td>
<td>30</td>
<td>50</td>
</tr>
<tr>
<td>$l_m$</td>
<td>30</td>
<td>10</td>
<td>50</td>
</tr>
</tbody>
</table>

The remaining parameters were derived in the following ways.
Chapter 5: Simulation Study

**Equation 1:** \[ y_1 = \beta_{01} + \beta_{11}x_1 + \beta_{21}x_2 + \mu_i \]

\[ \text{var}(y_1) = \beta_{11}^2 \cdot \text{var}(x_1) + \beta_{21}^2 \cdot \text{var}(x_2) + \sigma_i^2 \]

\[ 1 = \beta_{11}^2 \cdot \text{var}(x_1) + \beta_{21}^2 \cdot \text{var}(x_2) + \sigma_i^2 \]

\[ 1 = \beta_{11}^2 + \beta_{21}^2 \cdot \text{var}(x_2) + \sigma_i^2 \]

\[ \beta_{11}^2 = 1 - \beta_{21}^2 \cdot \text{var}(x_2) - \sigma_i^2 \]

\[ \beta_{11}^2 = 1 - \beta_{21}^2 \cdot p_2 (1 - p_2) - \sigma_i^2 \]

\[ \beta_{11} = \sqrt{1 - \beta_{21}^2 \cdot p_2 (1 - p_2) - \sigma_i^2} \]

\[ E(y_1) = \beta_{01} + \beta_{11} \cdot 0 + \beta_{21} \cdot p_2 + 0 \]

\[ 0 = \beta_{01} + \beta_{11} \cdot 0 + \beta_{21} \cdot p_2 + 0 \]

\[ \beta_{01} = -\beta_{21} \cdot p_2 \]

After generating random responses to the variables \( x_1 \) and \( x_2 \) using their distribution parameters, the true path coefficients were used to generate \( y_1 \) responses (estimated values plus residuals).

**Equation 2:** \[ y_2 = \beta_{02} + \gamma_{12}y_1 + \mu_2 \]

\[ \text{var}(y_2) = \beta_{12}^2 \cdot \text{var}(y_1) + \sigma_2^2 \]

\[ 1 = \beta_{12}^2 \cdot \text{var}(y_1) + \sigma_2^2 \]

\[ \beta_{12}^2 = 1 - \sigma_2^2 \]

\[ \beta_{12} = \sqrt{1 - \sigma_2^2} \]

\[ E(y_2) = \beta_{02} + \gamma_{12} \cdot 0 + 0 = 0 \]

\[ \beta_{02} = 0 \]
When the parameters were fixed and responses to \( y_1 \) were generated in the previous step, values for \( y_2 \) could be created using \textit{Equation 2}.

\textit{Equation 3:} \[ y_3 = \beta_{03} + \gamma_{13} y_1 + \beta_{33} x_3 + \beta_{43} x_4 + \mu_3 \]

\[ \text{var}(y_3) = \gamma_{13}^2 \cdot \text{var}(y_1) + \beta_{33}^2 \cdot \text{var}(x_3) + \beta_{43}^2 \cdot \text{var}(x_4) + \sigma_3^2 \]

\[ 1 = \gamma_{13}^2 \cdot \text{var}(y_1) + \beta_{33}^2 \cdot \text{var}(x_3) + \beta_{43}^2 \cdot \text{var}(x_4) + \sigma_3^2 \]

\[ \gamma_{13}^2 \cdot \text{var}(y_1) = 1 - \beta_{33}^2 \cdot \text{var}(x_3) - \beta_{43}^2 \cdot \text{var}(x_4) - \sigma_3^2 \]

\[ \gamma_{13}^2 = 1 - \beta_{33}^2 - \beta_{43}^2 \cdot p_4 (1 - p_4) - \sigma_3^2 \]

\[ \gamma_{13} = \sqrt{1 - \beta_{33}^2 - \beta_{43}^2 \cdot p_4 (1 - p_4) - \sigma_3^2} \]

\[ E(y_3) = \beta_{03} + \gamma_{13} \cdot 0 + \beta_{33} \cdot 0 + \beta_{43} p_4 + 0 \]

\[ 0 = \beta_{03} + \beta_{43} p_4 \]

\[ \beta_{03} = -\beta_{43} p_4 \]

After generating random responses to \( x_3 \) and \( x_4 \) from their known distributions, values could be generated for \( y_3 \) using the chosen and derived path model parameters and the responses to \( y_1 \) from \textit{Equation 1}.

\textbf{Measurement Model}

The values for \( y_1 \) and \( y_2 \) (equations 86 and 87) were used to generate responses to dichotomous items that measure these dimensions. First item difficulties (\( \delta_k \) measuring \( y_1 \) and \( \delta_m \) measuring \( y_2 \)) were randomly drawn from a uniform distribution between \(-3\) and \(3\). The sum of the item difficulties within each dimension were constrained to be zero for
simplicity, because that is the default option in ACER ConQuest to identify the model.

Stated formally:

\[ \delta \sim \text{uniform}(-3,3), \]

\[ \sum_{i=1}^{k} \delta_i = 0 \text{ and } \sum_{j=1}^{m} \delta_j = 0. \]

Subsequently, the item response probabilities \( p_k \) and \( p_m \) were computed for each respondent, by assuming that the item response data conformed to the Rasch IRT. Under this assumption, for each respondent,

\[ p_i(z_i = 1) = \frac{e^{(y_i - \delta_i)}}{1 + e^{(y_i - \delta_i)}} \]

\[ p_j(z_j = 1) = \frac{e^{(y_j - \delta_j)}}{1 + e^{(y_j - \delta_j)}} \]

The item responses could then be generated by drawing Bernoulli random variables as follows:

\[ z_i \sim \text{Bernoulli}(p_i) \]

\[ z_j \sim \text{Bernoulli}(p_j) \]
Replication

Each data set contained 5,000 respondents. While keeping the path model parameters and the item difficulties constant for each of the three versions of the same model, 100 replications where undertaken for the first part of the simulation study. This was needed to account for random noise in the response data. The average of the 100 estimated parameter values should be very close to the true values of the parameters. In addition, the standard deviation of the 100 estimates is an estimate of the sampling variation. For simplicity, this will be referred to as the true standard error. It is used for testing the computation of the standard errors of the parameters by the multi-step approach.

The SPSS syntax to generate the 100 data sets is included in Appendix A.

Analysis of the Data

Part I: Latent Path Coefficients versus True Parameters

The path model parameters that were estimated with the multi-step approach were compared with the true path model parameters and their standard errors. The true path model parameters were known and the true standard errors were computed by taking the standard deviation of the 100 estimated parameters.

The analyses were carried out three times, once for each version of the parameter values (see Table 6).
Part I, Version 1

Prior to evaluating the parameter recovery of the path model, the average of the 100 sets of item parameter estimates were compared with the true item parameters used to generate the item responses. While this test is not the main purpose of this part of the simulation, the accuracy of the path model estimates depends on the accuracy of the item response model estimates. Figure 11 shows that the item parameter recovery was successful. The same was true for version 2 and 3 of the generated data. The scatter plots will not be repeated in those sections.

Figure 11. Scatterplot of item difficulties estimated by ACER ConQuest and by Mplus for simulation Part I

The results of Version 1 where reasonable values were chosen for the latent path parameters are presented in Table 7 and in Figure 12 and Figure 13. The bias of the mean estimate was tested for significance via comparison with a standard normal distribution (i.e.
Chapter 5: Simulation Study

a Z-test) using a significance level of .05 (two-tailed). The ratio of the squared mean standard error and the squared true standard error was tested for significance using a chi-square test of the variance. Under the null hypothesis of no bias in the variance estimate, the ratio of the observed variance to the true variance should follow scaled chi-square.7

Table 7

Results of Simulation Part I, Version 1

<table>
<thead>
<tr>
<th>Parameter</th>
<th>True parameter</th>
<th>Mean estimate</th>
<th>Bias*</th>
<th>True SE</th>
<th>Mean SE</th>
<th>Ratio*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Equation 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_{01}$</td>
<td>-0.150</td>
<td>-0.140</td>
<td>0.010</td>
<td>0.019</td>
<td>0.010</td>
<td>0.526</td>
</tr>
<tr>
<td>$\beta_{11}$</td>
<td>0.853</td>
<td>0.855</td>
<td>0.002</td>
<td>0.013</td>
<td>0.007</td>
<td>0.538</td>
</tr>
<tr>
<td>$\beta_{21}$</td>
<td>0.300</td>
<td>0.299</td>
<td>-0.001</td>
<td>0.031</td>
<td>0.014</td>
<td>0.452</td>
</tr>
<tr>
<td>$R^2_1$</td>
<td>0.750</td>
<td>0.746</td>
<td>-0.004</td>
<td>0.015</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equation 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_{02}$</td>
<td>0.000</td>
<td>-0.004</td>
<td>-0.004</td>
<td>0.014</td>
<td>0.011</td>
<td>0.786</td>
</tr>
<tr>
<td>$\gamma_{12}$</td>
<td>0.600</td>
<td>0.602</td>
<td>0.002</td>
<td>0.019</td>
<td>0.013</td>
<td>0.684</td>
</tr>
<tr>
<td>$R^2_2$</td>
<td>0.360</td>
<td>0.360</td>
<td>0.000</td>
<td>0.017</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equation 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_{03}$</td>
<td>-0.120</td>
<td>-0.121</td>
<td>-0.001</td>
<td>0.014</td>
<td>0.014</td>
<td>1.000</td>
</tr>
<tr>
<td>$\gamma_{13}$</td>
<td>0.304</td>
<td>0.303</td>
<td>-0.001</td>
<td>0.014</td>
<td>0.015</td>
<td>1.071</td>
</tr>
<tr>
<td>$\beta_{33}$</td>
<td>0.200</td>
<td>0.201</td>
<td>0.001</td>
<td>0.012</td>
<td>0.013</td>
<td>1.083</td>
</tr>
<tr>
<td>$\beta_{43}$</td>
<td>0.600</td>
<td>0.604</td>
<td>0.004</td>
<td>0.032</td>
<td>0.032</td>
<td>1.000</td>
</tr>
<tr>
<td>$R^2_3$</td>
<td>0.190</td>
<td>0.191</td>
<td>0.001</td>
<td>0.011</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Bold if significant (p < 0.05, two-tailed)

7 $Z(\text{bias}) = (\text{mean estimate} - \text{true estimate}) / \text{True SE}; \alpha = 0.05; \text{two-tailed.}$

$X^2(\text{ratio}) = (\text{No of replications} - 1) \times (\text{mean SE} / \text{true SE})^2; \alpha = 0.05; \text{two-tailed.}$
The parameter estimates were, after 100 replications, very close to their true values and did not show a bias (see also Figure 12). The estimated standard errors, as shown in Figure 13, were underestimated in Equation 1 and, to a lesser extent, in Equation 2 and fairly correct in Equation 3. In Equation 1, the explained variable was a latent variable, measured by only 10 items. The bias was larger for a dichotomous predictor (\( X_2 \) with \( p_2 = 0.5 \)) than for a normally distributed, continuous predictor (\( X_1 \)). In Equation 2, on the other hand, the explained variable was a latent variable measured by 30 items. The dependent variable in Equation 3 was an observed variable. The latent variable with only 10 indicators (\( Y_1 \)) was an independent variable in both Equation 2 and Equation 3. These results suggested that low reliability of a dependent variable caused a bias in the estimated standard errors. This hypothesis was tested in versions 2 and 3 of simulation Part I.

Figure 12. Parameter estimates of simulation Part I, Version 1
Chapter 5: Simulation Study

Part I, Version 2

In Version 2 of simulation Part I, the numbers of indicators for the latent variables were reversed. The explained variable in Equation 1 was now measured by 30 items, while the explained variable in Equation 2 was measured by only 10 items. In addition, the explained variance for each equation was chosen to be low, to test if the parameter estimates were still unbiased. The results are shown in Table 8, Figure 14 and Figure 15.

Figure 13. Standard error estimates of simulation Part I, Version 1
Table 8

Results of Simulation Part I, Version 2

<table>
<thead>
<tr>
<th>Parameter</th>
<th>True parameter</th>
<th>Mean estimate</th>
<th>Bias*</th>
<th>True SE</th>
<th>Mean SE</th>
<th>Ratio*</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Equation 1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_{01}$</td>
<td>-0.010</td>
<td>-0.004</td>
<td>0.006</td>
<td>0.016</td>
<td>0.014</td>
<td>0.875</td>
</tr>
<tr>
<td>$\beta_{11}$</td>
<td>0.311</td>
<td>0.312</td>
<td>0.001</td>
<td>0.015</td>
<td>0.013</td>
<td>0.867</td>
</tr>
<tr>
<td>$\beta_{21}$</td>
<td>0.100</td>
<td>0.107</td>
<td>0.007</td>
<td>0.052</td>
<td>0.045</td>
<td>0.865</td>
</tr>
<tr>
<td>$R^2_1$</td>
<td>0.098</td>
<td>0.098</td>
<td>0.000</td>
<td>0.000</td>
<td>0.009</td>
<td></td>
</tr>
<tr>
<td><strong>Equation 2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_{02}$</td>
<td>0.000</td>
<td>-0.001</td>
<td>-0.001</td>
<td>0.017</td>
<td>0.014</td>
<td>0.824</td>
</tr>
<tr>
<td>$\gamma_{12}$</td>
<td>0.141</td>
<td>0.144</td>
<td>0.003</td>
<td>0.061</td>
<td>0.044</td>
<td>0.721</td>
</tr>
<tr>
<td>$R^2_1$</td>
<td>0.020</td>
<td>0.017</td>
<td>0.003</td>
<td>0.006</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Equation 3</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_{03}$</td>
<td>-0.001</td>
<td>0.001</td>
<td>0.002</td>
<td>0.015</td>
<td>0.014</td>
<td>0.933</td>
</tr>
<tr>
<td>$\gamma_{13}$</td>
<td>0.039</td>
<td>0.037</td>
<td>-0.002</td>
<td>0.041</td>
<td>0.045</td>
<td>1.098</td>
</tr>
<tr>
<td>$\beta_{33}$</td>
<td>0.050</td>
<td>0.052</td>
<td>0.002</td>
<td>0.012</td>
<td>0.014</td>
<td>1.167</td>
</tr>
<tr>
<td>$\beta_{43}$</td>
<td>0.010</td>
<td>0.010</td>
<td>0.000</td>
<td>0.045</td>
<td>0.047</td>
<td>1.044</td>
</tr>
<tr>
<td>$R^2_1$</td>
<td>0.004</td>
<td>0.004</td>
<td>0.000</td>
<td>0.002</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Bold if significant

Again, the parameter estimates were unbiased and after 100 replications very close to their true values (see Figure 14). Therefore, the level of explanatory power of the model did not affect the quality of the parameter estimates. Regarding the standard errors (see Figure 15), they were fairly correct in Equation 1, but significantly underestimated in Equation 2. The explained variable in Equation 2 was the latent variable with only 10 indicators. In addition, the standard error of the path coefficient between $x_3$ and $y_3$ in Equation 3 was significantly overestimated.
Chapter 5: Simulation Study

Figure 14. Parameter estimates of simulation Part I, Version 2

Figure 15. Standard error estimates of simulation Part I, Version 2

Overall, the results of Version 2 supported the hypothesis that the reliability of the latent variable, especially when it is a dependent variable, is not taken into account when estimating the standard errors. Version 3 of simulation Part I was designed as a conclusive
test by increasing the number of indicators for the latent variables and keeping all other parameters equal to Version 2 of the model.

**Part I, Version 3**

The number of items was increased in Version 3 of the model parameters so that the latent variables were more reliable and therefore the true standard errors smaller. Fifty items were used to measure each of the two latent variables. The remaining parameters were identical to Version 2. The results are presented in Table 9, Figure 16 and Figure 17.

*Table 9*

**Results of Simulation Part I, Version 3**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>True parameter</th>
<th>Mean estimate</th>
<th>Bias*</th>
<th>True SE</th>
<th>Mean SE</th>
<th>Ratio*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Equation 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_{01}$</td>
<td>-0.010</td>
<td>-0.013</td>
<td>-0.003</td>
<td>0.016</td>
<td>0.014</td>
<td>0.875</td>
</tr>
<tr>
<td>$\beta_{11}$</td>
<td>0.311</td>
<td>0.311</td>
<td>0.000</td>
<td>0.013</td>
<td>0.013</td>
<td>1.000</td>
</tr>
<tr>
<td>$\beta_{21}$</td>
<td>0.100</td>
<td>0.107</td>
<td>0.007</td>
<td>0.049</td>
<td>0.045</td>
<td>0.918</td>
</tr>
<tr>
<td>$R_1^2$</td>
<td>0.098</td>
<td>0.097</td>
<td>-0.001</td>
<td>0.008</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equation 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_{02}$</td>
<td>0.000</td>
<td>-0.008</td>
<td>-0.008</td>
<td>0.015</td>
<td>0.014</td>
<td>0.933</td>
</tr>
<tr>
<td>$\gamma_{12}$</td>
<td>0.141</td>
<td>0.148</td>
<td>0.007</td>
<td>0.052</td>
<td>0.044</td>
<td>0.846</td>
</tr>
<tr>
<td>$R_2^2$</td>
<td>0.020</td>
<td>0.018</td>
<td>-0.002</td>
<td>0.005</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equation 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_{03}$</td>
<td>-0.001</td>
<td>0.002</td>
<td>0.003</td>
<td>0.015</td>
<td>0.014</td>
<td>0.933</td>
</tr>
<tr>
<td>$\gamma_{13}$</td>
<td>0.039</td>
<td>0.038</td>
<td>-0.001</td>
<td>0.041</td>
<td>0.045</td>
<td>1.098</td>
</tr>
<tr>
<td>$\beta_{33}$</td>
<td>0.050</td>
<td>0.052</td>
<td>0.002</td>
<td>0.012</td>
<td>0.014</td>
<td>1.167</td>
</tr>
<tr>
<td>$\beta_{33}$</td>
<td>0.010</td>
<td>0.010</td>
<td>0.000</td>
<td>0.045</td>
<td>0.047</td>
<td>1.044</td>
</tr>
<tr>
<td>$R_3^2$</td>
<td>0.004</td>
<td>0.003</td>
<td>-0.001</td>
<td>0.002</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Bold if significant
As expected, the path model parameters were estimated as well as in Version 2 and the estimated standard errors on the parameters were identical to the ones from Version 2. However, the true standard errors reflected the increase in reliability of the latent variables resulting in a decrease of the true standard errors. The true standard errors of the parameters in Equation 1 and Equation 2, each predicting a latent variable, decreased and became closer to the estimated standard errors. Only the standard error on $\gamma_{12}$ was still significantly underestimated. The true standard errors in Equation 3, where an observed variable was predicted by one latent and two other observed variables did not change by increasing the number of indicators of the latent variable.

\begin{figure}
\centering
\includegraphics[width=\textwidth]{sim1.png}
\caption{Parameter estimates of simulation Part I, Version 3}
\end{figure}
In summary, path coefficients and explained variances of latent path models were well estimated by the multi-step method. Estimation of the standard errors seemed correct when the latent variables are measured by a large number of indicators, which increases the reliability of the scales. This is caused by the fact that the measurement model of the full model is disregarded when estimating the latent path model.

Part II: Path Coefficients by ACER ConQuest versus SPSS and Mplus

In the second part of the simulation study the path coefficients and explained variances as estimated by 2SLS in ACER ConQuest were compared with path model estimates from 2SLS in SPSS and full information maximum likelihood in Mplus. The aim of this analysis is to test if the estimation of the path coefficients and the explained variance is consistent across methods. Therefore, the true values for the latent variables were used in
the path analysis and the measurement models were ignored. The first data set of Version 1 and Version 3 of the model parameters was chosen for the analysis.

In ACER ConQuest, SPSS and Mplus, the three generated endogenous variables and the four generated exogenous were read in and directly used as variables in the path model, thus circumventing the measurement model. In Table 10 and Table 11 the different versions of the parameters estimates are compared between 2SLS in ACER ConQuest and in SPSS and with full information maximum likelihood in Mplus.
### Table 10

**Comparison of Version 1 Path Analysis Parameters Estimates between ACER ConQuest, SPSS and Mplus for Simulation Study Part II**

<table>
<thead>
<tr>
<th></th>
<th>ACER ConQuest</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>SPSS (2SLS)</th>
<th></th>
<th></th>
<th></th>
<th>Mplus (ML)</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
<td>Std. Error</td>
<td>Estimate</td>
<td>Std. Error</td>
<td>Estimate</td>
<td>Std. Error</td>
<td>Estimate</td>
<td>Std. Error</td>
<td>Estimate</td>
<td>Std. Error</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equation 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_{01}$</td>
<td>-0.153</td>
<td>(0.010)</td>
<td>-0.153</td>
<td>(0.010)</td>
<td>-0.153</td>
<td>(0.010)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_{11}$</td>
<td>0.864</td>
<td>(0.007)</td>
<td>0.864</td>
<td>(0.007)</td>
<td>0.864</td>
<td>(0.007)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_{21}$</td>
<td>0.309</td>
<td>(0.014)</td>
<td>0.309</td>
<td>(0.014)</td>
<td>0.309</td>
<td>(0.014)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$R^2_1$</td>
<td>0.758</td>
<td></td>
<td>0.758</td>
<td></td>
<td>0.752</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equation 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_{02}$</td>
<td>-0.012</td>
<td>(0.011)</td>
<td>-0.012</td>
<td>(0.011)</td>
<td>-0.012</td>
<td>(0.011)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\gamma_{12}$</td>
<td>0.603</td>
<td>(0.013)</td>
<td>0.603</td>
<td>(0.013)</td>
<td>0.606</td>
<td>(0.011)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$R^2_2$</td>
<td>0.376</td>
<td></td>
<td></td>
<td></td>
<td>0.311</td>
<td>0.374</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equation 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_{03}$</td>
<td>-0.119</td>
<td>(0.014)</td>
<td>-0.119</td>
<td>(0.014)</td>
<td>-0.119</td>
<td>(0.014)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\gamma_{13}$</td>
<td>0.328</td>
<td>(0.014)</td>
<td>0.328</td>
<td>(0.014)</td>
<td>0.327</td>
<td>(0.012)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_{33}$</td>
<td>0.193</td>
<td>(0.013)</td>
<td>0.193</td>
<td>(0.013)</td>
<td>0.193</td>
<td>(0.013)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_{43}$</td>
<td>0.544</td>
<td>(0.031)</td>
<td>0.544</td>
<td>(0.031)</td>
<td>0.544</td>
<td>(0.031)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$R^2_3$</td>
<td>0.200</td>
<td></td>
<td></td>
<td></td>
<td>0.178</td>
<td>0.214</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Table 11

*Comparison of Version 3 Path Analysis Parameters Estimates between ACER ConQuest, SPSS and Mplus for Simulation Study Part II*

<table>
<thead>
<tr>
<th></th>
<th>ACER ConQuest</th>
<th>SPSS (2SLS)</th>
<th>Mplus (ML)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
<td>Std. Error</td>
<td>Estimate</td>
</tr>
<tr>
<td><strong>Equation 1</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_{01}$</td>
<td>-0.014</td>
<td>(0.014)</td>
<td>-0.014</td>
</tr>
<tr>
<td>$\beta_{11}$</td>
<td>0.333</td>
<td>(0.013)</td>
<td>0.333</td>
</tr>
<tr>
<td>$\beta_{21}$</td>
<td>0.157</td>
<td>(0.045)</td>
<td>0.157</td>
</tr>
<tr>
<td>$R^2_{1}$</td>
<td>0.111</td>
<td></td>
<td>0.111</td>
</tr>
<tr>
<td><strong>Equation 2</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_{02}$</td>
<td>-0.015</td>
<td>(0.014)</td>
<td>-0.015</td>
</tr>
<tr>
<td>$\gamma_{12}$</td>
<td>0.144</td>
<td>(0.041)</td>
<td>0.144</td>
</tr>
<tr>
<td>$R^2_{1}$</td>
<td>0.023</td>
<td></td>
<td>0.020</td>
</tr>
<tr>
<td><strong>Equation 3</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_{03}$</td>
<td>-0.007</td>
<td>(0.015)</td>
<td>-0.007</td>
</tr>
<tr>
<td>$\gamma_{13}$</td>
<td>0.118</td>
<td>(0.042)</td>
<td>0.118</td>
</tr>
<tr>
<td>$\beta_{33}$</td>
<td>0.043</td>
<td>(0.014)</td>
<td>0.043</td>
</tr>
<tr>
<td>$\beta_{43}$</td>
<td>-0.053</td>
<td>(0.047)</td>
<td>-0.053</td>
</tr>
<tr>
<td>$R^2_{2}$</td>
<td>0.002</td>
<td></td>
<td>0.040</td>
</tr>
</tbody>
</table>

The path coefficients are very similar across software packages and estimation methods. The explained variance ($R^2$) estimated by SPSS appeared somewhat lower than estimated by ACER ConQuest and Mplus.

In Version 3 (Table 11), one estimated path coefficient and its standard error (the regression of $Y_3$ on $Y_1$) and one standard error (the regression of $Y_2$ on $Y_1$) were smaller from Mplus than from ACER ConQuest and SPSS. These are the two relationships in the
model between endogenous variables. The difference is likely caused by the different estimation method used in Mplus (full information using maximum likelihood instead of limited information using 2SLS).

Part III: Full Model in ACER ConQuest and Mplus

In Part III of the simulation study, a model was evaluated that included both a latent path model and measurement models. The full model was estimated in both ACER ConQuest and Mplus using matching parameterisations. Version 1 of the parameters was chosen and the first data set was analysed. Replication was not needed because the goal was not to find a bias in the estimates—as in Part I of the simulation study—but to test if similar results were obtained from both methods.

By default in ACER ConQuest, the loadings of the items on the latent variables are fixed to one and the sum of the item difficulties within each dimension is constrained to be zero. In Mplus, the loading can either be constrained to be equal or the first loading can be fixed to one while the others are freely estimated. Hence, the loadings were constrained to be equal within each dimension. In addition, the residual variance and the intercept needed to be fixed to identify the model. In order to mimic the parameterisation of ACER ConQuest, the residual variance and the intercept of the latent variables were fixed to the values that were estimated by ACER ConQuest.

Maximum likelihood estimation with robust standard errors and a logit link was chosen in Mplus for these comparisons.
Using the described parametrisation and estimation method, the loadings as estimated by Mplus were very close to one. The loadings of all 10 items measuring dimension 1 was 1.007 and the loadings of the 30 items measuring dimension 2 were 0.943. In addition, the estimated item difficulties lined up perfectly between the two methods as shown in Figure 18.

![Figure 18](scatterplot.png)

*Figure 18. Scatterplot of item difficulties estimated by ACER ConQuest and by Mplus for simulation Part III*

As reported in Table 12, all path coefficients and explained variance estimates were very close between the two methods. The standard errors for equations with a latent variable as dependent variable (*Equation 1 and Equation 2*) appear larger from Mplus than from ACER ConQuest. While these standard errors were underestimated in ACER ConQuest as shown in Part I of this simulation study, the standard errors from Mplus were in some cases larger than the standard deviation of the 100 replicates (the so-called “true standard error”). In comparison, the standard errors estimated by ACER ConQuest were
closer to the standard deviation of the replicates than the standard errors estimated by Mplus.

Table 12

Comparison of Full Model Results from ACER ConQuest and Mplus

<table>
<thead>
<tr>
<th>Equation</th>
<th>ACER ConQuest</th>
<th>Mplus (MLR)</th>
<th>True values</th>
<th>SD of replicates</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Equation 1 (Y1 with 10 items)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_{01}$</td>
<td>-0.149 (0.010)</td>
<td>-0.149 (0.010)</td>
<td>-0.150 (0.019)</td>
<td></td>
</tr>
<tr>
<td>$\beta_{11}$</td>
<td>0.878 (0.007)</td>
<td>0.872 (0.033)</td>
<td>0.853 (0.013)</td>
<td></td>
</tr>
<tr>
<td>$\beta_{21}$</td>
<td>0.327 (0.014)</td>
<td>0.304 (0.027)</td>
<td>0.300 (0.031)</td>
<td></td>
</tr>
<tr>
<td>$R^2_1$</td>
<td>0.772</td>
<td>0.775</td>
<td>0.750 (0.015)</td>
<td></td>
</tr>
<tr>
<td><strong>Equation 2 (Y2 with 30 items)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_{02}$</td>
<td>-0.017 (0.011)</td>
<td>-0.017 (0.011)</td>
<td>0.000 (0.014)</td>
<td></td>
</tr>
<tr>
<td>$\gamma_{12}$</td>
<td>0.600 (0.013)</td>
<td>0.641 (0.031)</td>
<td>0.600 (0.019)</td>
<td></td>
</tr>
<tr>
<td>$R^2_2$</td>
<td>0.368</td>
<td>0.374</td>
<td>0.360 (0.017)</td>
<td></td>
</tr>
<tr>
<td><strong>Equation 3</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_{03}$</td>
<td>-0.122 (0.014)</td>
<td>-0.119 (0.015)</td>
<td>-0.120 (0.014)</td>
<td></td>
</tr>
<tr>
<td>$\gamma_{13}$</td>
<td>0.322 (0.014)</td>
<td>0.327 (0.017)</td>
<td>0.304 (0.014)</td>
<td></td>
</tr>
<tr>
<td>$\beta_{33}$</td>
<td>0.191 (0.013)</td>
<td>0.190 (0.013)</td>
<td>0.200 (0.012)</td>
<td></td>
</tr>
<tr>
<td>$\beta_{43}$</td>
<td>0.536 (0.031)</td>
<td>0.541 (0.032)</td>
<td>0.600 (0.032)</td>
<td></td>
</tr>
<tr>
<td>$R^2_3$</td>
<td>0.202</td>
<td>0.214 (0.017)</td>
<td>0.190 (0.011)</td>
<td></td>
</tr>
<tr>
<td><strong>Variance (residual)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$y_1$</td>
<td>0.229 (0.005)</td>
<td>0.229 (0.005)</td>
<td>0.250</td>
<td></td>
</tr>
<tr>
<td>$y_2$</td>
<td>0.700 (0.014)</td>
<td>0.700 (0.014)</td>
<td>0.640</td>
<td></td>
</tr>
<tr>
<td><strong>Variance (marginal)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$y_1$</td>
<td>1.044 (0.023)</td>
<td>1.019 (0.062)</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>$y_2$</td>
<td>0.996 (0.022)</td>
<td>1.119 (0.028)</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td><strong>Meansa</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$y_1$</td>
<td>0.013 (0.016)</td>
<td>0.001 (0.019)</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>$y_2$</td>
<td>-0.007 (0.016)</td>
<td>-0.016 (0.012)</td>
<td>0.000</td>
<td></td>
</tr>
</tbody>
</table>

a The means are estimated after drawing plausible values in ACER ConQuest

Values in green italics are fixed in Mplus
The only other noticeable difference between the estimates by the different methods was the marginal variance of the second latent variable ($Y_2$). The true variance was one, the variance estimated by ACER ConQuest was 0.996 and the variance estimated by Mplus was 1.119.

The command files for running this model in ACER ConQuest and Mplus are included in Appendix B.

Summary

The simulation study showed that the multi-step method to extend the MRCML model with latent path models produced unbiased estimates of latent path coefficients and explained variance estimates. The standard errors were somewhat underestimated if the explained variable was a latent variable with low reliability, because the computation of the standard error did not take the reliability of the latent variable into account. However, if the measurement model was ignored, the standard errors were equal to standard errors estimated by 2SLS in SPSS and full information maximum likelihood in Mplus. Relationships between endogenous variables appeared somewhat different when applying a full information method as in Mplus than when using a limited information method such as the multi-step method in ACER ConQuest. Comparison of the results of a latent path model within an IRT framework between ACER ConQuest and Mplus revealed that all parameter

---

8 Because the mean of $Y_2$ was zero, the variance could be computed by dividing the sums of squares of $Y_2$—as estimated by the multi-step method—by its degrees of freedom.
Chapter 5: Simulation Study

estimates were very similar. However, for equations where the dependent variable was a latent variable, the standard errors estimated in ACER ConQuest appeared smaller than the standard errors estimated in Mplus.
After testing the multi-step method for latent path models within an IRT framework in a simulation study, it was applied to real data testing a meaningful hypothetical model. Data from two surveys were chosen for this purpose. The surveys were the Longitudinal Study of Australian Children (LSAC) and the Programme for International Student Assessment (PISA). Both models include multiple complex latent variables, measured by many mixed items types combined with a latent path model. The PISA survey also uses a rotated booklet design so that students only respond to part of the items. Currently it is not possible to run such a model in other software packages, especially not with data from surveys with complex test designs like PISA.

Longitudinal Study of Australian Children

The data chosen for this study were prepared by the Australian Institute of Family Studies (AIFS) and the Australian Bureau of Statistics (ABS). The study of interest is Growing Up in Australia: The Longitudinal Study of Australian Children (LSAC). A personal licence had to be obtained to access the data. Data from this study was chosen because of its high quality and the detailed documentation (Australian Institute of Family Studies, 2013). The longitudinal nature of the study and diversity of the data collection instruments enabled comparisons over time in both academic and social-emotional areas. Compared to large scale educational surveys where schools are sampled at the first stage
and students at the second stage, the inflation of the standard errors caused by the sample design was much smaller for the LSAC study. This was also a desirable feature of the study because the MRCMLM assumes simple random samples when estimating the standard errors.

The LSAC study follows the development of 10,000 children and families from all parts of Australia. *Growing Up in Australia* is investigating the contribution of children’s social, economic and cultural environments to their adjustment and wellbeing. The study commenced in 2004 with two cohorts: families with 4–5 year old children and families with 0–1 year old infants. Every second year new data are collected from the same sample and in 2014, Wave 5 was released.

The LSAC sample has been linked to several other sources of data that were not collected for the purpose of this study. One source is the National Assessment Program – Literacy and Numeracy (NAPLAN). This is an annual assessment for students in years 3, 5, 7 and 9. NAPLAN claims to test the sorts of skills that are essential for children to progress through school and life, such as reading, writing, spelling and numeracy. The assessments are undertaken every year nationwide.

**Sample**

The analysis undertaken for this thesis used Wave 3 and Wave 5 data from the older LSAC cohort. The ages of the students were 8 to 9 years in Wave 3 (2008) and 12 to 13 years in Wave 5 (2012). The model grade of this cohort was Year 3 in 2008 and Year 7 in
Chapter 6: Application on Real Data

2012. The purpose of the analysis was to predict student social-emotional wellbeing and numeracy achievement at the start of high school, in Year 7, with information of the students that was gathered earlier in their lives.

The full sample that were assessed in both waves and that had a student sample weight larger than zero consisted of 3,682 students. Of these students, 2,214 (60%) were linked to Year 3 and Year 7 non-missing NAPLAN numeracy scores.

To examine the sample for bias caused by list-wise deletion, the distributions of the NAPLAN scores and the socio-economic position were compared with the national distributions. In 2008, the national mean and standard deviation of Year 3 numeracy was 397 and 70, respectively. In comparison, the average numeracy performance of the selected sample was 420 and the standard deviation was 70. In 2012, the national average performance of Year 7 was 538 and the standard deviation 74, while the Year 7 sample included in this thesis had a mean of 554 and a standard deviation of 75. Therefore, the results of the analysis need to be interpreted with caution because the students in the sample achieved higher than the national average and is therefore not representative of the full population. The national distribution of socio-economic position had a mean of zero and a standard deviation of 1. The current sample’s mean position was -0.08 and the standard deviation 0.99.
Chapter 6: Application on Real Data

Method

The hypothetical model included two latent variables: social-emotional wellbeing in Year 3 and in Year 7. A parent and a teacher rated the Year 3 student’s wellbeing at home and at school. The parent questionnaire included 25 questions about the study child and the teacher questionnaire had 38 questions. Four questions from the parent and three questions from the teacher questionnaire were removed from the scale because they correlated less well with the other questions and therefore did not fit the Rasch model. The self-report questionnaire for Year 7 students included 66 questions of which six were removed to form the wellbeing scale. Appendix C lists the questions that were included and excluded from the scales and their response category options. All questions were answered on a three-, four- or five-point rating scale. Negatively worded questions were reverse coded before analysis. A score of 0 was used for the lowest level of social-emotional wellbeing for each of the questions in the questionnaires. The partial credit IRT model was applied for estimating the measurement models.

Figure 19 shows the latent path model with IRT measurement parts that was analysed using the extension of the MRCML model. The latent variables are indicated with an oval shape, the observed variables with a rectangle shape. In theory, the NAPLAN numeracy scores are latent scores, but only the student IRT scores were released without the original item responses. They were therefore treated as observed variables or, more precisely, variables that were read directly from the input data file. The original NAPLAN scores were rescaled so that the national standard deviation was equal to 1 in Year 7.
Measurement variance was expected for some of the items measuring social-emotional wellbeing between male and female children. In IRT terminology this is called differential item functioning (DIF) and it violates one of the assumptions of the Rasch model. Rasch models assume that the probability of giving a certain response to an item only depends on a respondent’s latent score and not on any membership of a group such as males or females. In order to test this assumption, a DIF model was estimated to test the difference in item difficulty between males and females for each of the questionnaire items (the ACER ConQuest command file for this model is included in Appendix D). Items with significant differences in difficulty of more than one logit between males and females were
regarded as different items for males than for females. All other items were regarded as common items for males and females.

It was further decided to estimate the model separately for males and females to explore if the path model revealed different relationships between variables for the two gender groups. Item parameters were freely estimated for each group while applying the multi-step method described in this thesis. In order to equate the two latent variables onto the same scale, the common item shift method was applied. The shift was calculated by taking the difference between the mean difficulties of the common items (the items that did not show DIF) for male and female children. This shift was then applied to the intercepts of the path model for male children (0.063 for Year 3 and 0.065 for Year 7).

Four of the variables in the hypothesised path model were endogenous. They are shaded in blue. The variables explained by the model were Year 3 and Year 7 social-emotional wellbeing and Year 3 and Year 7 numeracy performance. The remaining variables, shaded in green, were exogenous variables; they were not explained by the model. The exogenous variables were socio-economic position and attendance of a kinder or preschool program four years prior to Grade 3 (the year before starting primary school). The index for socio-economic position was derived from three components: parents’ educational attainment, their income and their occupational prestige (Blakemore, Gibbings, & Strazdins, 2006). Kinder attendance was dummy coded so that the reference group was the students that did not attend a kinder or preschool program; one dummy variable indicated that the student attended a program for less than 20 hours a week, one that the
student attended a program for 20 hours or more. A third variable was an indicator for missing information (18% of the included students). Appendix E includes the command codes that were run in ACER ConQuest to estimates the parameters in this latent path IRT model.

**Results**

The results of the model are recorded in Table 13. The results show that social-emotional wellbeing in Year 3—as reported by a parent and a teacher—was positively associated with socio-economic position ($\beta=0.13$ for girls and $\beta=0.20$ for boys), but was not related to kinder program attendance for either males or females. The strength of the association between wellbeing and socio-economic position was small for both girls and boys (the standard deviations of the latent variables are reported between brackets in Table 13). While the intercepts for girls and boys cannot be compared directly because the standard errors are dependent and we do not know the covariance between them, the size of the difference and the small standard errors suggest that girls are reported to have a higher social-emotional wellbeing in Year 3 than boys after taking into account differences in socio-economic position and kinder program attendance. The exogenous variables explained three per cent of the variance in wellbeing of girls and five per cent of boys.

Numeracy performance in Year 3 was also positively related to socio-economic position for both girls and boys ($\beta=0.26$ and $\beta=0.28$, respectively). The effect was small to moderate (0.28 for girls and 0.30 for boys). Attending a kinder program less than 20 hours
a week was positively associated with numeracy performance in Year 3 ($\beta=0.19$ for girls and $\beta=0.18$ for boys). Students that attended a kinder program less than 20 hours a week performed better than students that had not attended a kinder program (the reference group) by about one-fifth of a standard deviation. A similar effect was found for boys attending a kinder program for more than 20 hours a week ($\beta=0.21$), but not for girls. The exogenous variables explained nine per cent of the variance in numeracy performance in Year 3 for both girls and boys.

Social-emotional wellbeing and numeracy performance in Year 3 were used to explain the same constructs in Year 7. Self-reported wellbeing in Year 7 was not predicted by parent and teacher reported wellbeing or numeracy performance in Year 3. Numeracy performance in Year 7, on the other hand, was predicted by numeracy performance in Year 3, but not by social-emotional wellbeing in Year 3. The effect was very large (more than a standard deviation).
Table 13

Results of a Latent Path IRT Model Using LSAC Data

<table>
<thead>
<tr>
<th></th>
<th>Females</th>
<th>Males</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(SD=0.85)</td>
<td>(SD=0.87)</td>
</tr>
<tr>
<td>Year 3 wellbeing on</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>1.87 (0.05)</td>
<td>1.55 (0.05)</td>
</tr>
<tr>
<td>Socio-economic position</td>
<td><strong>0.13</strong> (0.03)</td>
<td><strong>0.20</strong> (0.03)</td>
</tr>
<tr>
<td>Kinder program (≥ 20 hrs)</td>
<td>-0.13 (0.10)</td>
<td>-0.15 (0.10)</td>
</tr>
<tr>
<td>Kinder program (&lt; 20 hrs)</td>
<td>0.08 (0.06)</td>
<td>0.01 (0.07)</td>
</tr>
<tr>
<td>Kinder program (missing)</td>
<td>0.10 (0.08)</td>
<td>-0.02 (0.08)</td>
</tr>
<tr>
<td><strong>R-squared</strong></td>
<td>0.03</td>
<td>0.05</td>
</tr>
<tr>
<td>Year 3 numeracy on</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>0.19 (0.05)</td>
<td>0.20 (0.06)</td>
</tr>
<tr>
<td>Socio-economic position</td>
<td><strong>0.26</strong> (0.03)</td>
<td><strong>0.28</strong> (0.03)</td>
</tr>
<tr>
<td>Kinder program (≥ 20 hrs)</td>
<td>0.13 (0.10)</td>
<td><strong>0.21</strong> (0.10)</td>
</tr>
<tr>
<td>Kinder program (&lt; 20 hrs)</td>
<td><strong>0.19</strong> (0.06)</td>
<td><strong>0.18</strong> (0.07)</td>
</tr>
<tr>
<td>Kinder program (missing)</td>
<td>0.07 (0.08)</td>
<td>0.12 (0.08)</td>
</tr>
<tr>
<td><strong>R-squared</strong></td>
<td>0.09</td>
<td>0.09</td>
</tr>
<tr>
<td>Year 7 wellbeing on</td>
<td>(SD=0.95)</td>
<td>(SD=0.85)</td>
</tr>
<tr>
<td>Intercept</td>
<td>0.00 (0.92)</td>
<td>1.22 (0.42)</td>
</tr>
<tr>
<td>Year 3 wellbeing</td>
<td>1.01 (0.52)</td>
<td>0.25 (0.33)</td>
</tr>
<tr>
<td>Year 3 numeracy</td>
<td>-0.29 (0.26)</td>
<td>0.34 (0.22)</td>
</tr>
<tr>
<td><strong>R-squared</strong></td>
<td>0.00</td>
<td>0.04</td>
</tr>
<tr>
<td>Year 7 numeracy on</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>2.04 (0.70)</td>
<td>1.46 (0.46)</td>
</tr>
<tr>
<td>Year 3 wellbeing</td>
<td>-0.16 (0.40)</td>
<td>0.23 (0.36)</td>
</tr>
<tr>
<td>Year 3 numeracy</td>
<td><strong>1.15</strong> (0.20)</td>
<td><strong>1.23</strong> (0.24)</td>
</tr>
<tr>
<td><strong>R-squared</strong></td>
<td>0.28</td>
<td>0.25</td>
</tr>
</tbody>
</table>

*Significant path coefficients in bold.*

**Programme for International Student Assessment**

The OECD’s Programme for International Student Assessment (PISA) is a triennial international survey that aims to evaluate education systems worldwide by testing the skills
and knowledge of 15-year-old students. Since the year 2000, every three years, 15-year-old students from randomly selected schools worldwide take tests in the subjects reading, mathematics and science, with a main focus on one subject in each year of assessment. Data from PISA 2009 were used in this thesis. Reading was the main domain in this PISA cycle.

Students take a test that lasts two hours. The tests are a mixture of open-ended and multiple-choice questions that are organised in units based on a passage setting out a real-life situation. A total of about 390 minutes of test items are covered. Students take different combinations of items. The PISA test is based on a balanced, incomplete block design. Booklets are randomly assigned to students. Each student responds to items in the main domain, but not every student responds to items in minor domains. However, using a multidimensional MRCML model, population parameter estimates are based on all students in the sample. For example, a country’s mean achievement in mathematics is based on all students in the national sample, not only on students that responded to mathematics items.

The students and their school principals also answer questionnaires to provide information about the students’ backgrounds, schools and learning experiences and about the broader school system and learning environment.

**Sample**

The Australian national sample was selected for the current analysis. The total sample consisted of 14,251 15-year-old students. Of these students, 4.4 per cent did not
have valid data for the background variables parental education or family wealth. Those students were removed for the analysis, which resulted in a final sample size of 13,626.

**Method**

The IRT model that was analysed for this chapter included all cognitive items measuring the latent variables reading, mathematics and science. In addition, questionnaire items were included that measure another latent variable, control strategies for learning. PISA focuses on three kinds of learning strategies: memorisation, elaboration and control strategies. Memorisation strategies include rote learning of facts and rehearsal of examples, while elaboration strategies involve relating new material to something the student already knows. Control strategies involve determining what one has already learned in order to determine what one still needs to learn. In Australia, control strategies were found to have a moderate relationship with achievement: students with greater awareness of control strategies scored slightly higher in reading, mathematical and scientific literacy. Hence, it was decided to delve deeper in the relationship between control strategies and performance.

Control strategies were measured by five questionnaire items on a four-point Likert scale (almost never, sometimes, often, almost always). Items of the index of control strategies were:

- *When I study, I start by figuring out what exactly I need to learn*
- *When I study, I check if I understand what I have read*
- *When I study, I try to figure out which concepts I still haven’t really understood*
• When I study, I make sure that I remember the most important points in the text
• When I study and I don’t understand something, I look for additional information to clarify this.

The full analytic model is presented in Figure 20. Student achievement in reading was regressed on gender (1 = girl, 0 = boy), student’s use of control strategies and school type (1 = government school, 0 = independent or Catholic). In turn, both gender and school type were regressed on control strategies. This is an example of a mediational model, because the effect of gender and school type on achievement is measured both directly and indirectly via control strategies. Furthermore, control strategies were predicted by highest level of parental education and school type was predicted by family wealth.

Figure 20. PISA latent path model within an IRT framework
An index for parental education was constructed by selecting the highest level for each parent and then assigning them to the following categories: (0) none, (1) primary education, (2) lower secondary, (3) vocational/pre-vocational upper secondary, (4) upper secondary and/or non-tertiary post-secondary, (5) vocational tertiary and (6) theoretically oriented tertiary and post-graduate. The highest level of two parents was used for the analysis. The index of family wealth was based on the students’ responses on whether they had the following at home: a room of their own, a link to the internet, a dishwasher (treated as a country-specific item), a DVD player, and three other country-specific items; and their responses on the number of cellular phones, televisions, computers, cars and rooms with a bath or shower. Both these indices were provided in the public database for PISA 2009. School type was also predicted by geographic location, because a lower percentage of students attend government schools in metropolitan areas.

In the 2009 cycle of PISA, all students responded to items measuring their reading ability. Therefore, it was decided to use reading achievement as the measure for achievement. However, items for mathematics and science were also included in the IRT model (to improve the estimation of the reading ability scores in the multi-dimensional IRT model) and could be used as alternatives.

Since PISA is a long test, not all students reach the end of their test booklet within two hours. If not reached items were scored as incorrect responses instead of not attempted, the difficulties of the items at the end of a booklet would be underestimated. On the other hand, if the not-reached items were scored as not attempted instead of incorrect, the
abilities of students not reaching the end of their booklets would be overestimated. Because of this conflict, the IRT model is estimated in two steps for PISA. First, the item parameters are estimated and the not-reached items are treated as not administered. Second, the student abilities are estimated treating the not-reached items as incorrect responses while anchoring the item parameters to their values from the first step. The latent path model is estimated during the second step. The command files are included in Appendix F.

The 2SLS method does not take the complex two-stage clustered sample design of the PISA survey into account. The computation of the standard errors assumes a simple random sample, which results in an underestimation of the standard error for most PISA analyses. To correct for this bias, the design effect on the mean of each of the student-level endogenous variables was estimated using the intra-class correlation and the average cluster size (number of students sampled in each school).\(^9\) The standard error of the path coefficients were multiplied by the square root of the design effect for each equation in which a student-level variable was explained.\(^10\)

**Results**

Two versions of the latent path model were estimated: one simple multiple regression analysis of reading performance on gender, use of control strategies and school

\(^9\) Design effect = 1 + (cluster size – 1) * intra-class correlation

\(^10\) Design effects were 9.6 for reading and 4.3 for use of control strategies.
type, and one full path model as presented in Figure 20. The results are included in Table 14.

Table 14

*Results of a Latent Path IRT Model Using PISA Data*

<table>
<thead>
<tr>
<th></th>
<th>Regression Estimate (SD=1.29)</th>
<th>Path model Estimate (SD=1.29)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reading achievement on</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>0.20 (0.06)</td>
<td>-0.15 (0.13)</td>
</tr>
<tr>
<td>Control strategies</td>
<td>0.31 (0.02)</td>
<td>1.05 (0.21)</td>
</tr>
<tr>
<td>Government school</td>
<td>-0.38 (0.06)</td>
<td>0.45 (0.58)</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.21</td>
<td>0.00</td>
</tr>
<tr>
<td>Control strategies on</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>0.49 (0.06)</td>
<td></td>
</tr>
<tr>
<td>Government school</td>
<td>-1.25 (0.29)</td>
<td></td>
</tr>
<tr>
<td>Parental education</td>
<td>0.19 (0.04)</td>
<td></td>
</tr>
<tr>
<td>R-squared</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>Government school on</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Family wealth</td>
<td>-0.12 (0.01)</td>
<td></td>
</tr>
<tr>
<td>Metropolitan</td>
<td>-0.12 (0.01)</td>
<td></td>
</tr>
<tr>
<td>R-squared</td>
<td>0.06</td>
<td></td>
</tr>
</tbody>
</table>

In a simple multiple regression, all three predictors had a significant effect on reading achievement. That is, girls performed better than boys, students that use more control strategies scored higher and students in government schools did less well on the reading test than students in independent or Catholic schools. However, when indirect effects of gender and school type through control strategies were added to the model, the direct effect of gender and school type on reading achievement was not significant
anymore. Only the use of control strategies had a significant direct effect on reading performance.

As for predicting the use of control strategies, gender, school type and highest level of parental education all had a significant effect: girls used more control strategies than boys, students in government schools used less control strategies than students in independent and Catholic schools, and parental education was positively associated with the use of control strategies. Finally, students with higher levels of family wealth and students in metropolitan areas were less likely to be enrolled in government schools.

According to the results of this analysis model, the difference in reading performance between girls and boys and between different school types was explained by the use of control strategies. That is, the use of control strategies was related to better reading performance; and girls used more control strategies than boys; and students in government schools used less control strategies than students in independent and Catholic schools.

The explained variance was low for the three equations. Given that the three independent variables explained over 20 per cent of the variance in reading when applying a simple multiple regression, the exogenous variables in this model—which are the instrumental variables in the 2SLS estimation method—may not have been the best choice. It is not easy to solve this problem without intimate knowledge of the research area. Since this analysis was a mere illustration of an application of the latent path model within an IRT
framework and not primarily driven by theory, it was deemed sufficiently suitable for its purpose. However, the results should be interpreted with caution.

**Summary**

This chapter has shown how the multi-step method for estimating latent path models within an IRT framework can be applied to real data. The IRT measurement models in both these examples were partial credit models, but any of the many Rasch models that are special cases of the MRCMLM—and any other IRT measurement models that may be implemented in future—can be combined with a latent path model.

The aim of this chapter was to apply the method to a model that currently cannot be estimated in other methods with the IRT or SEM frameworks. Other IRT software packages have not implemented latent path modelling and SEM software packages, such as Mplus and GLLAMM (Rabe-Hesketh, Skrondal, & Pickles, 2004), have currently implemented only a limited number of categorical measurement within the Rasch family of divide-by-total models and are also limited in the number of variables.
Chapter 7: Discussion, Conclusions, and Recommendations

The multi-step method to estimate latent path models is an important addition to the IRT framework because thus far researchers using IRT modelling had to produce individual person latent trait estimates with IRT software, and subsequently import these into other software packages to link them to other data sources and perform secondary analysis such as path analysis. Besides the impracticality of this method, different types of individual person ability estimates often lead to different results in secondary analysis. The approach that is presented in this thesis overcomes both limitations. The current approach estimates the path coefficients directly from the latent trait distribution parameters, hence avoids producing latent trait estimates for individual respondents.

The presented approach is a multi-step method in which the IRT model is estimated before the latent path model parameters are estimated. A great advantage of this approach is that any type of IRT model that is currently a special case of the MRCML model can be combined with a latent path model. The MRCML model includes many different predefined unidimensional and multidimensional Rasch models, but also user-defined models by creating specific design matrices that can be imported. In addition, a two-parameter model that estimates the slopes of each item has also been implemented. All these IRT models, and the IRT models that will be implemented in future, can be combined with the latent path model presented in this thesis. The software that is built to estimate the MRCML model, ACER ConQuest, is not limited by sample size, can estimate up to 10,000 items and
can handle holes in response data files caused by test designs with multiple booklets (sometimes called missing by design).

The multi-step method for estimating latent path models can be further developed and improved. Results from the simulation study suggested that the reliability of a latent variable is not taken into account when estimating standard errors of path coefficients. That is, path models that include latent variables with low reliabilities produce the same path coefficients and standard errors as path models that include latent variables with high reliabilities. In theory, the standard errors are expected to be larger when the latent variables have low reliabilities and thus are measured less well.

In relation to this, the accuracy of the matrices with regression coefficients and conditional co-variances derived from the first step of the presented multi-step method is a function of the sample size and the reliability of the latent variable estimation (see A20 on page 71 and A22 on page 72 in Adams, Wilson & Wu, 1997). Therefore, the accuracy of the results depends on the sample size and the reliability of the latent constructs. A further evaluation of the conditions under which the estimates are sufficiently accurate is needed in a future comprehensive suite of simulation studies.

The simulation study also showed that some of the standard errors produced by the method employed within a SEM framework (Mplus) were larger than the standard errors estimated by the method developed for this thesis for one model, but that these standard errors were sometimes larger and less close to the true standard errors (the standard errors of the 100 replicated path coefficients) than the standard errors produced by ACER
ConQuest. When developing the latent path model as part of the MRCML model further, attempts will be made to find a solution for taking into account the reliability of the latent variables when estimating the standard errors of the path coefficients.

Another future contribution will be the addition of a fit statistic for the full model, including the measurement and path models. Fit statistics for latent path models within the SEM framework have received great attention and many indices are available. SEM software packages produce multiple model fit statistics (Kline, 2011). Fit indices that are often reported are the model chi-square, the root mean square error of approximation (RMSEA), the comparative fit index (CFI), standardized root mean square residual (SRMSR) and Akaike information criterion (AIC). Only for the RMSEA a confidence interval is estimated. A criterion value is used to evaluate values of the CFI and the SRMSR. The model chi-square and the AIC are generally used to compare the fit of different models. A recent promising development within the IRT framework is the Bayesian method called posterior predictive model checking (PPMC) to test the fit of measurement models. It is expected that the PPMC can be further developed to include latent path models.

The method that was implemented for estimating path models was 2SLS. This is a popular method, but not the only method. It is a limited information method, because each equation in the path model is estimated independent of the other equations and it cannot estimate recursive relationships. Further exploration of other methods, such as full information methods in which all equations of the latent path model are estimated in one
Chapter 7: Discussion, Conclusions, and Recommendations

step, is a desirable goal for future development of the latent path model extension of the MRCML model.

Despite these limitations, the presented method allows researchers to combine latent path models with a wide range of item response models, including complex latent traits measured by very large numbers of items in mixed formats. In addition, another advantage of this multi-step model is that step two and three can also be implemented in other models than the item response models included in the MRCMLM.
References


References


References


References


estimation of a structural equation.

http://www.le.ac.uk/users/dsgp1/COURSES/TOPICS/2sls.pdf (course notes).


132
References


Appendix A: SPSS Syntax to Generate 100 Response Data Sets

*** EQUATIONS ***.
/* y1 = beta01 + beta11*x1 + beta21*x2 + u1.
/* y2 = beta02 + gamma21*y1 + u2.
/* y3 = beta03 + gamma21*y1 + beta33*x3 + beta43*x4 + u3.

define !p1() 0.5 !enddefine.
define !beta21() 0.3 !enddefine.
define !sigma1() 0.5 !enddefine.
define !sigma2() 0.8 !enddefine.
define !p4() 0.2 !enddefine.
define !beta33() 0.2 !enddefine.
define !beta43() 0.6 !enddefine.
define !sigma3() 0.9 !enddefine.
/* R-squared = 1 - sigma^2 */

define !struc () .
input program.
loop id=1 to 5000.
end case.
end loop.
end file.
end input program.
exe.

d!do !a = 1 !to 100.
*-----------------------------------------------------------------.
*** Equation 1: y1 = beta01 + beta11*x1 + beta21*x2 + u1 ***.
/* y1 ~ N(0,1)
/* x1 ~ N(0,1)
/* x2 ~ binary (p1)
/* u1 ~ N(0,sigma1^2).
set seed = !a.
compute x1=rv.normal(0,1).
compute x2=rv.bernoulli(!p1).
compute u1=rv.normal(0,!sigma1).
compute beta01=beta11*!p1.
compute beta11=sqrt(1-!sigma1*!sigma1-!beta21*!beta21*!p1*(1-!p1)).
compute y1=beta01+beta11*x1+beta21*x2+u1.
exe.

*fre x2.
*means x1 y1 u1.
*REGRESSION /DEPENDENT y1 /METHOD=ENTER x1 x2.
**** Equation 2: y2 = beta02 + gamma12*y1 + u2 ***.
/* y2 ~ N(0,1) */
/* u2 ~ N(0,sigma2^2). */
compute gamma12=sqrt(1-!sigma2*!sigma2).
compute u2=RV.NORMAL(0,!sigma2).
compute y2=gamma12*y1+u2.
*means y2 u2.
*REGRESSION /DEPENDENT y2 /METHOD=ENTER y1.

**** Equation 3: y3 = beta03 + gamma13*y1 + beta33*x3 + beta43*x4 + u3 ***.
/* y3 ~ N(0,1) */
/* u3 ~ N(0,sigma3^2). */
/* x3 ~ N(0,1) */
/* x4 ~ binary (p4). */
compute x3 = rv.normal(0,1).
compute x4 = rv.bernoulli(!p4).
compute u3=RV.NORMAL(0,!sigma3).
compute gamma13=sqrt(1-!beta33*!beta33-!beta43*!beta43*!p4*(1-!p4)-!sigma3*!sigma3).
compute beta03=-!beta43*!p4.
compute y3=beta03+gamma13*y1+!beta33*x3+!beta43*x4+u3.
*fre x4.
*means y3 x3 u3 /cells=mean variance.
*REGRESSION /DEPENDENT y3 /METHOD=ENTER y1 x3 x4.

compute cons=1.
save outfile=!quote(!concat('c:\temp\','!a','.sav')).
!doend.
!enddefine.
set mprint=no.
!struc.

*** Create item responses for latent variables y1 (10 items) and y2 (30 items)***.
define !resp (total=!tokens(1)/dim1=!tokens(1)/sim=!tokens(1)).
!do !a=1 !to 100.
input program.
loop id=1 to !total.
end case.
end loop.
end file.
end input program.
formats ID (f3.0).
exe.

!if (!a=1) !then
set seed=1.
compute DELTA=rv.uniform(-3,3).
compute DIM=1.
if ($casenum>!dim1) DIM=2.
aggregate /break=DIM /MN=mean(Delta).
compute D=DELTA-MN.
string ITEM (a4).
compute ITEM=concat('d',Ltrim(string(ID,f3.0))).
exe.
delete var ID DELTA DIM MN.
FLIP VARIABLES=D /NEWNAMES=ITEM.
delete var CASE_LVL.
compute cons=1.
save outfile='c:\temp\deltas.sav'.
!ifend.

match file file=!quote(!concat('c:\temp\',!a,'.sav'))
/table='c:\temp\deltas.sav' /by cons.
exe.

!do !i=1 !to !Dim1.
!let !s=!length(!concat(!blanks(!i),!blanks(!a)))).
set seed=!s.
compute !concat('p',!i)=exp(y1 -!concat('d',!i))
/(1+exp(y1-!concat('d',!i))).
compute !concat('z',!i)=rv.bernoulli(!concat('p',!i)).
!doend.

!do !i=!length(!concat(!blanks(!dim1),!blanks(1))) !to !total.
!let !s=!length(!concat(!blanks(!i),!blanks(1)))).
set seed=!s.
compute !concat('p',!i)=exp(y2 -!concat('d',!i))
/(1+exp(y2-!concat('d',!i))).
compute !concat('z',!i)=rv.bernoulli(!concat('p',!i)).
!doend.

/* reorder variables and output response data */.
save outfile=!quote(!concat('c:\temp\qqq',!a,'.sav'))
/keep= x1 x2 x3 x4 y1 y2 y3 !do !i=1 !to !total !concat('z',!i) !doend
u1 u2 u3 beta01 beta11 gamma12 gamma13 beta03
!do !d=1 !to !total !concat('d',!d) !doend !do !p=1 !to !total !concat('p',!p) !doend.
Appendix A

get file=!quote(!concat('c:\temp\qqq',!a,'.sav')).
formats z1 to !concat('z',!total) (f1.0).
write
outfile=!quote(!concat('D:\WorkGebhardt\Thesis\Simulation\StructuralModel\2SLS\SIM','!sim','\Data\responses_SIM','!sim','_rep',!a,'.dat'))
   /z1 to !concat('z',!total) y1 y2 y3 x1 x2 x3 x4.
save
outfile=!quote(!concat('D:\WorkGebhardt\Thesis\Simulation\StructuralModel\2SLS\SIM','!sim','\Data\responses_SIM','!sim','_rep',!a,'.sav'))
   /keep=z1 to !concat('z',!total) y1 y2 y3 x1 x2 x3 x4.
exe.
/* store true parameters */.
get file=!quote(!concat('c:\temp\qqq',!a,'.sav')) /keep=d1 to d40.
select if $casenum=1.
flip.
rename var ( CASE_LBL var001=ITEM DELTA).
compute substr(item,1,1)='z'.
save translate
outfile=!quote(!concat('D:\WorkGebhardt\Thesis\Simulation\StructuralModel\2SLS\SIM','!sim','\True\TrueDeltas_SIM','!sim','_rep',!a,'.csv'))
   /type=csv /fieldnames /replace.
exe.
!if (!a=1) !then
get file=!quote(!concat('c:\temp\qqq',!a,'.sav')) /keep=beta01 beta11 gamma12 gamma13 beta03.
select if $casenum=1.
compute beta21=!beta21.
compute beta33=!beta33.
compute beta43=!beta43.
compute rsq1 =1-!sigma1*!sigma1.
compute rsq2 =1-!sigma2*!sigma2.
compute rsq3 =1-!sigma3*!sigma3.
exe.
flip.
rename var ( CASE_LBL var001=NAME PARAMETER).
SORT CASES BY name.
formats PARAMETER (f5.3).
save translate outfile=
!quote(!concat('D:\WorkGebhardt\Thesis\Simulation\StructuralModel\2SLS\SIM','!sim','\True\TrueStructural_SIM','!sim','.csv'))
   /type=csv /fieldnames /replace.
exe.
!ifend.
!doend.
!enddefine.
Appendix A

set mprint=no.
!resp total=40 dim1=10 sim=1.
Appendix B: Commands in ACER ConQuest and Mplus for Simulation Part III

In ACER ConQuest

let path=C:\Work_Gebhardt\My Work\Thesis\Simulation\Part 3;
set warnings=no, constraints=items;
export logfile >> %path%\ConQuest\rep1.log;
data %path%\data\responses_SIM1_rep1.sav !filetype=SPSS, responses=z1 to z40, keeps=x1 x2 x3 x4 y3;
score (0,1)(0,1)() !items(1-10);
score (0,1)() (0,1) !items(11-40);
model item;
regression x1 x2 x3 x4 y3;
export parameter >> %path%\ConQuest\rep1.par;
estimate !stderr=quick, fit=no;
show !estimates=latent >> %path%\ConQuest\rep1.shw;
show cases !estimates=latent, filetype=spss >> %path%\ConQuest\rep1.sav;

structural
/dimension_1 on x1 x2
/dimension_2 on dimension_1
/y3 on dimension_1 x3 x4
!matrixout=p3;
descriptives !estimates=pv >> %path%\ConQuest\rep1.des;
print p3_fullsscp !filetype=xlsx >>%path%\ConQuest\fullsscp.xlsx;
print p3_osscp !filetype=xlsx >>%path%\ConQuest\osscp.xlsx;
print p3_lossscp !filetype=xlsx >>%path%\ConQuest\lossscp.xlsx;
print p3_lsscp !filetype=xlsx >>%path%\ConQuest\lsscp.xlsx;

In Mplus

data:
file is C:\Work_Gebhardt\My Work\Thesis\Simulation\Part 1\Version 1\Mplus\rep1.csv;

variable:
names are z1 z2 z3 z4 z5 z6 z7 z8 z9 z10 z11 z12 z13 z14 z15 z16 z17 z18 z19 z20 z21 z22 z23 z24 z25 z26 z27 z28 z29 z30 z31 z32 z33 z34 z35 z36 z37 z38 z39 z40 Y1 Y2 Y3 X1 X2 X3 X4;
usevariables are z1 z2 z3 z4 z5 z6 z7 z8 z9 z10 z11 z12 z13 z14 z15 z16 z17 z18 z19 z20 z21 z22 z23 z24 z25 z26 z27 z28 z29 z30 z31 z32 z33 z34 z35 z36 z37 z38 z39 z40 Y3 X1 X2 X3 X4;
categorical are z1 z2 z3 z4 z5 z6 z7 z8 z9 z10 z11 z12 z13 z14 z15 z16 z17 z18 z19 z20 z21 z22 z23 z24 z25 z26 z27 z28 z29 z30 z31 z32 z33 z34 z35 z36 z37 z38 z39 z40;

model:
[y1@-0.149];
y2@-0.017;  
Y1 by z1-z10*(1);  
Y2 by z11-z40*(2);  
y1 @0.229;  
y2 @0.700;  
Y1 on x1 x2;  
Y2 on Y1;  
Y3 on Y1 x3 x4;

analysis: estimator=mlr;  
link=logit;
output: sampstat tech4 tech8;
## Appendix C: Indicators of Wellbeing in LSAC

<table>
<thead>
<tr>
<th>Item number*</th>
<th>VARIABLE</th>
<th>Difference (Males-Females)</th>
<th>CONSTRUCT</th>
<th>MEASURE</th>
<th>QUESTION</th>
<th>RESPONSES</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>esse03a1a</td>
<td>0.17 Social development</td>
<td>SDQ Pro-social scale</td>
<td>Considerate of other people’s feelings</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
<td>POS</td>
</tr>
<tr>
<td>2</td>
<td>esse03a1b</td>
<td>-0.21 Social development</td>
<td>SDQ Pro-social scale</td>
<td>Shares readily with other children (treats, toys, pencils, etc.)</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
<td>POS</td>
</tr>
<tr>
<td>3</td>
<td>esse03a1c</td>
<td>0.24 Social development</td>
<td>SDQ Pro-social scale</td>
<td>Helpful if someone is hurt, upset or feeling ill</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
<td>POS</td>
</tr>
<tr>
<td>4</td>
<td>esse03a1d</td>
<td>0.01 Social development</td>
<td>SDQ Pro-social scale</td>
<td>Kind to younger children</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
<td>POS</td>
</tr>
<tr>
<td>5</td>
<td>esse03a1e</td>
<td>0.34 Social development</td>
<td>SDQ Pro-social scale</td>
<td>Often volunteers to help others (parents, teachers, other children)</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
<td>POS</td>
</tr>
<tr>
<td>6</td>
<td>esse03a2a</td>
<td>0.23 Behaviour</td>
<td>SDQ Hyperactivity scale</td>
<td>Restless, overactive, cannot stay still for long</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
<td>NEG</td>
</tr>
<tr>
<td>7</td>
<td>esse03a2b</td>
<td>0.13 Behaviour</td>
<td>SDQ Hyperactivity scale</td>
<td>Constantly fidgeting or squirming</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
<td>NEG</td>
</tr>
<tr>
<td>8</td>
<td>esse03a2c</td>
<td>0.27 Behaviour</td>
<td>SDQ Hyperactivity scale</td>
<td>Easily distracted, concentration wanders</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
<td>NEG</td>
</tr>
<tr>
<td>9</td>
<td>esse03a2d</td>
<td>0.22 Behaviour</td>
<td>SDQ Hyperactivity scale</td>
<td>Thinks things out before acting</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
<td>POS</td>
</tr>
<tr>
<td>10</td>
<td>esse03a2e</td>
<td>0.21 Behaviour</td>
<td>SDQ Hyperactivity scale</td>
<td>Good attention span, sees chores or homework through to the end</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
<td>POS</td>
</tr>
<tr>
<td>11</td>
<td>esse03a3a</td>
<td>Males -0.91 Emotional development</td>
<td>SDQ Emotional problems scale</td>
<td>Often complains of headaches, stomachaches or sickness</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
<td>NEG</td>
</tr>
<tr>
<td>12</td>
<td>esse03a3b</td>
<td>-0.38 Emotional development</td>
<td>SDQ Emotional problems scale</td>
<td>Many worries, often seems worried</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
<td>NEG</td>
</tr>
<tr>
<td>13</td>
<td>esse03a3c</td>
<td>-0.44 Emotional development</td>
<td>SDQ Emotional problems scale</td>
<td>Often unhappy, depressed or tearful</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
<td>NEG</td>
</tr>
<tr>
<td>Item number</td>
<td>VARIABLE</td>
<td>CONSTRUCT</td>
<td>MEASURE</td>
<td>QUESTION</td>
<td>RESPONSES</td>
<td>WRODED</td>
</tr>
<tr>
<td>-------------</td>
<td>----------</td>
<td>-----------</td>
<td>---------</td>
<td>----------</td>
<td>-----------</td>
<td>--------</td>
</tr>
<tr>
<td>14</td>
<td>ese03a3d</td>
<td>Males</td>
<td>-0.76</td>
<td>Emotional development</td>
<td>SDQ Emotional problems scale</td>
<td>Nervous or clingy in new situations, easily loses confidence</td>
</tr>
<tr>
<td>15</td>
<td>ese03a3e</td>
<td>Males</td>
<td>-0.53</td>
<td>Emotional development</td>
<td>SDQ Emotional problems scale</td>
<td>Many fears, easily scared</td>
</tr>
<tr>
<td>16</td>
<td>ese03a4a</td>
<td>Behaviour</td>
<td>-0.07</td>
<td>Behaviour</td>
<td>SDQ Conduct problems scale</td>
<td>Often loses temper</td>
</tr>
<tr>
<td>17</td>
<td>ese03a4b</td>
<td>Behaviour</td>
<td>0.04</td>
<td>Behaviour</td>
<td>SDQ Conduct problems scale</td>
<td>Generally well behaved, usually does what adults request</td>
</tr>
<tr>
<td>18</td>
<td>ese03a4c</td>
<td>Behaviour</td>
<td>0.05</td>
<td>Behaviour</td>
<td>SDQ Conduct problems scale</td>
<td>Often fights with other children or bullies them</td>
</tr>
<tr>
<td>19</td>
<td>ese03a4f</td>
<td>Behaviour</td>
<td>-0.11</td>
<td>Behaviour</td>
<td>SDQ Conduct problems scale</td>
<td>Often lies or cheats</td>
</tr>
<tr>
<td>20</td>
<td>ese03a4g</td>
<td>Behaviour</td>
<td>-0.27</td>
<td>Behaviour</td>
<td>SDQ Conduct problems scale</td>
<td>Steals from home, school or elsewhere</td>
</tr>
<tr>
<td>21</td>
<td>ese03a5a</td>
<td>Social development</td>
<td>-0.41</td>
<td>Social development</td>
<td>SDQ Peer problems scale</td>
<td>Rather solitary, tends to play alone</td>
</tr>
<tr>
<td>22</td>
<td>ese03a5b</td>
<td>Social development</td>
<td>-0.15</td>
<td>Social development</td>
<td>SDQ Peer problems scale</td>
<td>Has at least one good friend</td>
</tr>
<tr>
<td>23</td>
<td>ese03a5c</td>
<td>Social development</td>
<td>-0.14</td>
<td>Social development</td>
<td>SDQ Peer problems scale</td>
<td>Generally liked by other children</td>
</tr>
<tr>
<td>24</td>
<td>ese03a5d</td>
<td>Social development</td>
<td>-0.25</td>
<td>Social development</td>
<td>SDQ Peer problems scale</td>
<td>Picked on or bullied by other children</td>
</tr>
<tr>
<td>25</td>
<td>ese03a5e</td>
<td>Social development</td>
<td>-0.41</td>
<td>Social development</td>
<td>SDQ Peer problems scale</td>
<td>Gets on better with adults than with other children</td>
</tr>
<tr>
<td>26</td>
<td>ese06a</td>
<td>Temperament</td>
<td>-0.07</td>
<td>Temperament</td>
<td>General temperament</td>
<td>Overall, compared with other children of the same age, do you think this child ...?</td>
</tr>
<tr>
<td>27</td>
<td>ese13a1</td>
<td>Emotional development</td>
<td>-0.20</td>
<td>Emotional development</td>
<td>Temperament</td>
<td>Reacts strongly (cries or complains loudly) to a disappointment or failure</td>
</tr>
<tr>
<td>28</td>
<td>ese13a2</td>
<td>Emotional development</td>
<td>-0.30</td>
<td>Emotional development</td>
<td>Temperament</td>
<td>When angry, yells or snaps at others</td>
</tr>
<tr>
<td>29</td>
<td>ese13a3</td>
<td>Emotional development</td>
<td>-0.33</td>
<td>Emotional development</td>
<td>Temperament</td>
<td>Moody when corrected for misbehaviour</td>
</tr>
<tr>
<td>30</td>
<td>ese13a4</td>
<td>Emotional development</td>
<td>-0.36</td>
<td>Emotional development</td>
<td>Temperament</td>
<td>Responds intensely to disapproval (shouts, cries, etc.)</td>
</tr>
</tbody>
</table>
## Appendix C

<table>
<thead>
<tr>
<th>Item number</th>
<th>VARIABLE</th>
<th>Easier to have a relatively high socio-emotional score for Difference (Males-Females)</th>
<th>CONSTRUCT</th>
<th>MEASURE</th>
<th>QUESTION</th>
<th>RESPONSES</th>
<th>WORDED</th>
</tr>
</thead>
<tbody>
<tr>
<td>31</td>
<td>ese13b1</td>
<td>0.11</td>
<td>Emotional development</td>
<td>Temperament</td>
<td>Does not complete homework unless reminders are given</td>
<td>1 Never; 2 Rarely; 3 Half the time; 4 Frequently; 5 Always</td>
<td>NEG</td>
</tr>
<tr>
<td>32</td>
<td>ese13b2</td>
<td>0.14</td>
<td>Emotional development</td>
<td>Temperament</td>
<td>Remembers to do homework without being reminded</td>
<td>1 Never; 2 Rarely; 3 Half the time; 4 Frequently; 5 Always</td>
<td>POS</td>
</tr>
<tr>
<td>33</td>
<td>ese13b3</td>
<td>-0.05</td>
<td>Emotional development</td>
<td>Temperament</td>
<td>Goes back to the task at hand (chores, housework, etc.) after an interruption</td>
<td>1 Never; 2 Rarely; 3 Half the time; 4 Frequently; 5 Always</td>
<td>POS</td>
</tr>
<tr>
<td>34</td>
<td>ese13b4</td>
<td>-0.06</td>
<td>Emotional development</td>
<td>Temperament</td>
<td>Has difficulty completing assignments (homework, chores, etc.)</td>
<td>1 Never; 2 Rarely; 3 Half the time; 4 Frequently; 5 Always</td>
<td>NEG</td>
</tr>
<tr>
<td>X</td>
<td>ese13c1</td>
<td></td>
<td>Emotional development</td>
<td>Temperament</td>
<td>Approaches children his/her age even when he/she doesn’t know them</td>
<td>1 Never; 2 Rarely; 3 Half the time; 4 Frequently; 5 Always</td>
<td>POS</td>
</tr>
<tr>
<td>X</td>
<td>ese13c2</td>
<td></td>
<td>Emotional development</td>
<td>Temperament</td>
<td>Is shy with adults he/she doesn’t know</td>
<td>1 Never; 2 Rarely; 3 Half the time; 4 Frequently; 5 Always</td>
<td>NEG</td>
</tr>
<tr>
<td>X</td>
<td>ese13c3</td>
<td></td>
<td>Emotional development</td>
<td>Temperament</td>
<td>When meeting new children acts bashful</td>
<td>1 Never; 2 Rarely; 3 Half the time; 4 Frequently; 5 Always</td>
<td>NEG</td>
</tr>
<tr>
<td>X</td>
<td>ese13c4</td>
<td></td>
<td>Emotional development</td>
<td>Temperament</td>
<td>Seems uncomfortable when at someone’s house for the first time</td>
<td>1 Never; 2 Rarely; 3 Half the time; 4 Frequently; 5 Always</td>
<td>NEG</td>
</tr>
</tbody>
</table>

### YEAR 3 – TEACHER REPORT

<table>
<thead>
<tr>
<th>Item number</th>
<th>VARIABLE</th>
<th>SOCIAL DEVELOPMENT</th>
<th>MEASURE</th>
<th>QUESTION</th>
<th>RESPONSES</th>
<th>WORDED</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ese03t1a</td>
<td>0.43</td>
<td>SDQ Pro-social scale</td>
<td>Considerate of other people’s feelings</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
<td>POS</td>
</tr>
<tr>
<td>2</td>
<td>ese03t1b</td>
<td>0.26</td>
<td>SDQ Pro-social scale</td>
<td>Shares readily with other children (treats, toys, pencils etc.)</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
<td>POS</td>
</tr>
<tr>
<td>3</td>
<td>ese03t1c</td>
<td>Females 0.75</td>
<td>SDQ Pro-social scale</td>
<td>Helpful if someone is hurt, upset or feeling ill</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
<td>POS</td>
</tr>
<tr>
<td>4</td>
<td>ese03t1d</td>
<td>Females 0.55</td>
<td>SDQ Pro-social scale</td>
<td>Kind to younger children</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
<td>POS</td>
</tr>
<tr>
<td>5</td>
<td>ese03t1e</td>
<td>Females 0.56</td>
<td>SDQ Pro-social scale</td>
<td>Often volunteers to help others (parents, teachers, other children)</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
<td>POS</td>
</tr>
<tr>
<td>6</td>
<td>ese03t2a</td>
<td>Females 0.79</td>
<td>Behaviour</td>
<td>SDQ Hyperactivity scale</td>
<td>Restless, overactive, cannot stay still for long</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
</tr>
<tr>
<td>7</td>
<td>ese03t2b</td>
<td>Females 0.66</td>
<td>Behaviour</td>
<td>SDQ Hyperactivity scale</td>
<td>Constantly fidgeting or squirming</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
</tr>
<tr>
<td>8</td>
<td>ese03t2c</td>
<td>0.45</td>
<td>Behaviour</td>
<td>SDQ Hyperactivity scale</td>
<td>Easily distracted, concentration wanders</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
</tr>
</tbody>
</table>
## Appendix C

<table>
<thead>
<tr>
<th>Item number</th>
<th>VARIABLE</th>
<th>Easier to have a relatively high socio-emotional score for</th>
<th>Difference (Males-Females)</th>
<th>CONSTRUCT</th>
<th>MEASURE</th>
<th>QUESTION</th>
<th>RESPONSES</th>
<th>WORDED</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>ese03t2d</td>
<td>Females</td>
<td>0.59</td>
<td>Behaviour</td>
<td>SDQ Hyperactivity scale</td>
<td>Thinks things out before acting</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
<td>POS</td>
</tr>
<tr>
<td>10</td>
<td>ese03t2e</td>
<td>Females</td>
<td>0.53</td>
<td>Behaviour</td>
<td>SDQ Hyperactivity scale</td>
<td>Good attention span, sees chores or homework through to the end</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
<td>POS</td>
</tr>
<tr>
<td>X</td>
<td>ese03t3a</td>
<td>Emotional development</td>
<td>-0.41</td>
<td>Emotional development</td>
<td>SDQ Emotional problems scale</td>
<td>Often complains of headaches, stomach aches or sickness</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
<td>NEG</td>
</tr>
<tr>
<td>12</td>
<td>ese03t3b</td>
<td>Emotional development</td>
<td>-0.41</td>
<td>Emotional development</td>
<td>SDQ Emotional problems scale</td>
<td>Many worries, often seems worried</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
<td>NEG</td>
</tr>
<tr>
<td>13</td>
<td>ese03t3c</td>
<td>Emotional development</td>
<td>-0.47</td>
<td>Emotional development</td>
<td>SDQ Emotional problems scale</td>
<td>Often unhappy, depressed or tearful</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
<td>NEG</td>
</tr>
<tr>
<td>14</td>
<td>ese03t3d</td>
<td>Emotional development</td>
<td>-0.47</td>
<td>Emotional development</td>
<td>SDQ Emotional problems scale</td>
<td>Nervous or clingy in new situations, easily loses confidence</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
<td>NEG</td>
</tr>
<tr>
<td>15</td>
<td>ese03t3e</td>
<td>Males</td>
<td>-0.62</td>
<td>Emotional development</td>
<td>SDQ Emotional problems scale</td>
<td>Many fears, easily scared</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
<td>NEG</td>
</tr>
<tr>
<td>16</td>
<td>ese03t4a</td>
<td>Females</td>
<td>0.77</td>
<td>Behaviour</td>
<td>SDQ Conduct problems scale</td>
<td>Often loses temper</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
<td>NEG</td>
</tr>
<tr>
<td>17</td>
<td>ese03t4b</td>
<td>Females</td>
<td>0.68</td>
<td>Behaviour</td>
<td>SDQ Conduct problems scale</td>
<td>Generally well behaved, usually does what adults request</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
<td>POS</td>
</tr>
<tr>
<td>18</td>
<td>ese03t4c</td>
<td>Behaviour</td>
<td>-0.14</td>
<td>Behaviour</td>
<td>SDQ Conduct problems scale</td>
<td>Often fights with other children or bullies them</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
<td>NEG</td>
</tr>
<tr>
<td>19</td>
<td>ese03t4d</td>
<td>Behaviour</td>
<td>-0.05</td>
<td>Behaviour</td>
<td>SDQ Conduct problems scale</td>
<td>Often lies or cheats</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
<td>NEG</td>
</tr>
<tr>
<td>20</td>
<td>ese03t4g</td>
<td>Behaviour</td>
<td>0.13</td>
<td>Behaviour</td>
<td>SDQ Conduct problems scale</td>
<td>Steals from home, school or elsewhere</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
<td>NEG</td>
</tr>
<tr>
<td>X</td>
<td>ese03t5a</td>
<td>Social development</td>
<td>0.13</td>
<td>Social development</td>
<td>SDQ Peer problems scale</td>
<td>Rather solitary, tends to play alone</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
<td>NEG</td>
</tr>
<tr>
<td>22</td>
<td>ese03t5b</td>
<td>Social development</td>
<td>-0.23</td>
<td>Social development</td>
<td>SDQ Peer problems scale</td>
<td>Has at least one good friend</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
<td>POS</td>
</tr>
<tr>
<td>23</td>
<td>ese03t5c</td>
<td>Social development</td>
<td>-0.25</td>
<td>Social development</td>
<td>SDQ Peer problems scale</td>
<td>Generally liked by other children</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
<td>POS</td>
</tr>
<tr>
<td>24</td>
<td>ese03t5d</td>
<td>Social development</td>
<td>-0.26</td>
<td>Social development</td>
<td>SDQ Peer problems scale</td>
<td>Picked on or bullied by other children</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
<td>NEG</td>
</tr>
<tr>
<td>X</td>
<td>ese03t5e</td>
<td>Social development</td>
<td>0.13</td>
<td>Social development</td>
<td>SDQ Peer problems scale</td>
<td>Gets on better with adults than with other children</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
<td>NEG</td>
</tr>
<tr>
<td>Item number</td>
<td>VARIABLE</td>
<td>CONSTRUCT</td>
<td>MEASURE</td>
<td>QUESTION</td>
<td>RESPONSES</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>----------</td>
<td>-----------</td>
<td>---------</td>
<td>----------</td>
<td>-----------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>YEAR 7 – STUDENT SELF-REPORT</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>gse03c1a</td>
<td>Females</td>
<td>0.99</td>
<td>Social development</td>
<td>SDQ Pro-social scale</td>
<td>I try to be nice to other people. I care about their feelings.</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
<td></td>
</tr>
<tr>
<td>X</td>
<td>gse03c1b</td>
<td>Social development</td>
<td>SDQ Pro-social scale</td>
<td>I usually share with others, for example CDs, games, food.</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>gse03c1c</td>
<td>Females</td>
<td>0.91</td>
<td>Social development</td>
<td>SDQ Pro-social scale</td>
<td>I am helpful if someone is hurt, upset or feeling ill.</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>gse03c1d</td>
<td>Females</td>
<td>0.82</td>
<td>Social development</td>
<td>SDQ Pro-social scale</td>
<td>I am kind to younger children.</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>gse03c1e</td>
<td>Females</td>
<td>0.38</td>
<td>Social development</td>
<td>SDQ Pro-social scale</td>
<td>I often volunteer to help others (parents, teachers, children).</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>gse03c2a</td>
<td>Behaviour</td>
<td>0.10</td>
<td>SDQ Hyperactivity scale</td>
<td>I am restless, I cannot stay still for long.</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>gse03c2b</td>
<td>Behaviour</td>
<td>0.03</td>
<td>SDQ Hyperactivity scale</td>
<td>I am constantly fidgeting or squirming.</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>gse03c2c</td>
<td>Behaviour</td>
<td>0.05</td>
<td>SDQ Hyperactivity scale</td>
<td>I am easily distracted, I find it difficult to concentrate.</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>gse03c2d</td>
<td>Behaviour</td>
<td>0.18</td>
<td>SDQ Hyperactivity scale</td>
<td>I think before I do things.</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>gse03c2e</td>
<td>Behaviour</td>
<td>0.03</td>
<td>SDQ Hyperactivity scale</td>
<td>I finish the work I am doing, my attention is good.</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>gse03c3a</td>
<td>Emotional development</td>
<td>-0.18</td>
<td>SDQ Emotional problems scale</td>
<td>I get a lot of headaches, stomach-aches or sickness.</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>gse03c3b</td>
<td>Emotional development</td>
<td>-0.34</td>
<td>SDQ Emotional problems scale</td>
<td>I worry a lot.</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>gse03c3c</td>
<td>Emotional development</td>
<td>-0.26</td>
<td>SDQ Emotional problems scale</td>
<td>I am often unhappy, depressed or tearful.</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>gse03c3d</td>
<td>Males</td>
<td>-0.70</td>
<td>SDQ Emotional problems scale</td>
<td>I am nervous in new situations. I easily lose confidence.</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>gse03c3e</td>
<td>Males</td>
<td>-0.80</td>
<td>SDQ Emotional problems scale</td>
<td>I have many fears, I am easily scared.</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>gse03c4a</td>
<td>Behaviour</td>
<td>0.19</td>
<td>SDQ Conduct problems scale</td>
<td>I get very angry and lose my temper.</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item number</td>
<td>VARIABLE</td>
<td>CONSTRUCT</td>
<td>CONSTRUCT MEASURE</td>
<td>QUESTION</td>
<td>RESPONSES</td>
<td>WORDED</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>----------</td>
<td>-----------</td>
<td>-------------------</td>
<td>----------</td>
<td>-----------</td>
<td>--------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>gse03c4b</td>
<td>Behaviour</td>
<td>SDQ Conduct problems scale</td>
<td>I usually do as I am told.</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
<td>POS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>gse03c4c</td>
<td>Behaviour</td>
<td>SDQ Conduct problems scale</td>
<td>I fight a lot. I can make other people do what I want.</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
<td>NEG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>gse03c4f</td>
<td>Females</td>
<td>SDQ Conduct problems scale</td>
<td>I am often accused of lying or cheating.</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
<td>NEG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>gse03c4g</td>
<td>Behaviour</td>
<td>SDQ Conduct problems scale</td>
<td>I take things that are not mine from home, school or elsewhere.</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
<td>NEG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>gse03c5a</td>
<td>Social development</td>
<td>SDQ Peer problems scale</td>
<td>I would rather be alone than with people my age.</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
<td>NEG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>gse03c5b</td>
<td>Social development</td>
<td>SDQ Peer problems scale</td>
<td>I have one good friend or more.</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
<td>POS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>gse03c5c</td>
<td>Social development</td>
<td>SDQ Peer problems scale</td>
<td>Other people my own age generally like me.</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
<td>POS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>gse03c5d</td>
<td>Social development</td>
<td>SDQ Peer problems scale</td>
<td>Other children or young people pick on me or bully me.</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
<td>NEG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>gse03c5e</td>
<td>Social development</td>
<td>SDQ Peer problems scale</td>
<td>I get on better with adults than people my age.</td>
<td>1 Not true; 2 Somewhat true; 3 Certainly true</td>
<td>NEG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>gse14a</td>
<td>Females</td>
<td>Social Skills Rating Scale</td>
<td>I feel sorry for others when bad things happen to them.</td>
<td>1 Rarely or never; 2 Sometimes; 3 Very often</td>
<td>POS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>gse14b</td>
<td>Females</td>
<td>Social Skills Rating Scale</td>
<td>I listen to my friends when they talk about the problems they are having.</td>
<td>1 Rarely or never; 2 Sometimes; 3 Very often</td>
<td>POS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>gse14c</td>
<td>Females</td>
<td>Social Skills Rating Scale</td>
<td>I try to understand how my friends feel when they are angry, upset or sad.</td>
<td>1 Rarely or never; 2 Sometimes; 3 Very often</td>
<td>POS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>gse14d</td>
<td>Females</td>
<td>Social Skills Rating Scale</td>
<td>I accept people who are different.</td>
<td>1 Rarely or never; 2 Sometimes; 3 Very often</td>
<td>POS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>gse14e</td>
<td>Females</td>
<td>Social Skills Rating Scale</td>
<td>I say nice things to others when they have done something well.</td>
<td>1 Rarely or never; 2 Sometimes; 3 Very often</td>
<td>POS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>X</td>
<td>gse16a1</td>
<td>Worries</td>
<td>Child worries and concerns</td>
<td>the environment (climate change, drought, pollution)</td>
<td>1 Not at all worried; 2 A little worried; 3 Fairly worried; 4 Very worried</td>
<td>NEG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>X</td>
<td>gse16a10</td>
<td>Worries</td>
<td>Child worries and concerns</td>
<td>changing schools</td>
<td>1 Not at all worried; 2 A little worried; 3 Fairly worried; 4 Very worried</td>
<td>NEG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item number</td>
<td>VARIABLE</td>
<td>CONSTRUCT</td>
<td>MEASURE</td>
<td>QUESTION</td>
<td>RESPONSES</td>
<td>WORDED</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>----------</td>
<td>-----------</td>
<td>---------</td>
<td>----------</td>
<td>-----------</td>
<td>--------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>X gse16a2</td>
<td>Worries</td>
<td>Child worries and concerns</td>
<td>terrorism or war</td>
<td>1 Not at all worried; 2 A little worried; 3 Fairly worried; 4 Very worried</td>
<td>NEG</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X gse16a3</td>
<td>Worries</td>
<td>Child worries and concerns</td>
<td>use of alcohol and other drugs by children or teenagers</td>
<td>1 Not at all worried; 2 A little worried; 3 Fairly worried; 4 Very worried</td>
<td>NEG</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X gse16a4</td>
<td>Worries</td>
<td>Child worries and concerns</td>
<td>someone in your family becoming seriously ill or injured</td>
<td>1 Not at all worried; 2 A little worried; 3 Fairly worried; 4 Very worried</td>
<td>NEG</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X gse16a5</td>
<td>Worries</td>
<td>Child worries and concerns</td>
<td>people in your family fighting</td>
<td>1 Not at all worried; 2 A little worried; 3 Fairly worried; 4 Very worried</td>
<td>NEG</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>37 gse16a6</td>
<td>Males</td>
<td>-0.89 Worries</td>
<td>Child worries and concerns</td>
<td>the way you look</td>
<td>1 Not at all worried; 2 A little worried; 3 Fairly worried; 4 Very worried</td>
<td>NEG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>38 gse16a7</td>
<td>-0.49 Worries</td>
<td>Child worries and concerns</td>
<td>not fitting in with your friends</td>
<td>1 Not at all worried; 2 A little worried; 3 Fairly worried; 4 Very worried</td>
<td>NEG</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X gse16a8</td>
<td>Worries</td>
<td>Child worries and concerns</td>
<td>not doing well at school</td>
<td>1 Not at all worried; 2 A little worried; 3 Fairly worried; 4 Very worried</td>
<td>NEG</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40 gse16b1</td>
<td>Males</td>
<td>-0.60 Worries</td>
<td>Child worries and concerns</td>
<td>I worry about things.</td>
<td>1 Never; 2 Sometimes; 3 Often; 4 Always</td>
<td>NEG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>41 gse16b2</td>
<td>Males</td>
<td>-0.61 Worries</td>
<td>Child worries and concerns</td>
<td>I feel afraid.</td>
<td>1 Never; 2 Sometimes; 3 Often; 4 Always</td>
<td>NEG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>42 gse16b3</td>
<td>-0.44 Worries</td>
<td>Child worries and concerns</td>
<td>I feel afraid that I will make a fool of myself in front of people.</td>
<td>1 Never; 2 Sometimes; 3 Often; 4 Always</td>
<td>NEG</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>43 gse16b4</td>
<td>-0.30 Worries</td>
<td>Child worries and concerns</td>
<td>I worry that something bad will happen to me.</td>
<td>1 Never; 2 Sometimes; 3 Often; 4 Always</td>
<td>NEG</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>44 gse16b5</td>
<td>Males</td>
<td>-0.60 Worries</td>
<td>Child worries and concerns</td>
<td>I feel nervous.</td>
<td>1 Never; 2 Sometimes; 3 Often; 4 Always</td>
<td>NEG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>45 gse16b6</td>
<td>-0.18 Worries</td>
<td>Child worries and concerns</td>
<td>I wake up feeling scared.</td>
<td>1 Never; 2 Sometimes; 3 Often; 4 Always</td>
<td>NEG</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>46 gse16b7</td>
<td>Males</td>
<td>-0.63 Worries</td>
<td>Child worries and concerns</td>
<td>I worry what other people think of me.</td>
<td>1 Never; 2 Sometimes; 3 Often; 4 Always</td>
<td>NEG</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Appendix C

<table>
<thead>
<tr>
<th>Item number</th>
<th>VARIABLE</th>
<th>CONSTRUCT</th>
<th>MEASURE</th>
<th>QUESTION</th>
<th>RESPONSES</th>
<th>WORDED</th>
</tr>
</thead>
<tbody>
<tr>
<td>47</td>
<td>gse16b8</td>
<td>Worries</td>
<td>Child worries and concerns</td>
<td>All of a sudden I feel really scared for no reason at all.</td>
<td>1 Never; 2 Sometimes; 3 Often; 4 Always</td>
<td>NEG</td>
</tr>
<tr>
<td>48</td>
<td>gse21a1</td>
<td>Child self-perceptions</td>
<td>Self-efficacy</td>
<td>Overall, I have a lot to be proud of.</td>
<td>1 False; 2 Mostly false; 3 Sometimes false sometimes true; 4 Mostly true; 5 True</td>
<td>POS</td>
</tr>
<tr>
<td>49</td>
<td>gse21a2</td>
<td>Child self-perceptions</td>
<td>Self-efficacy</td>
<td>Most things I do, I do well.</td>
<td>1 False; 2 Mostly false; 3 Sometimes false sometimes true; 4 Mostly true; 5 True</td>
<td>POS</td>
</tr>
<tr>
<td>50</td>
<td>gse21a3</td>
<td>Child self-perceptions</td>
<td>Self-efficacy</td>
<td>Overall, most things I do turn out well.</td>
<td>1 False; 2 Mostly false; 3 Sometimes false sometimes true; 4 Mostly true; 5 True</td>
<td>POS</td>
</tr>
<tr>
<td>51</td>
<td>gse21a4</td>
<td>Child self-perceptions</td>
<td>Self-efficacy</td>
<td>I can do things as well as most people.</td>
<td>1 False; 2 Mostly false; 3 Sometimes false sometimes true; 4 Mostly true; 5 True</td>
<td>POS</td>
</tr>
<tr>
<td>52</td>
<td>gse21a5</td>
<td>Child self-perceptions</td>
<td>Self-efficacy</td>
<td>If I really try, I can do almost anything I want to.</td>
<td>1 False; 2 Mostly false; 3 Sometimes false sometimes true; 4 Mostly true; 5 True</td>
<td>POS</td>
</tr>
<tr>
<td>X</td>
<td>gse21b1</td>
<td>Child self-perceptions</td>
<td>Overall happiness</td>
<td>In general, I am happy with how things are for me in my life right now.</td>
<td>1 Strongly disagree; 2 Disagree; 3 Neither agree nor disagree; 4 Agree; 5 Strongly agree</td>
<td>POS</td>
</tr>
<tr>
<td>54</td>
<td>gse21c1</td>
<td>Internalising behaviours</td>
<td>Depressed feelings</td>
<td>I felt miserable or unhappy.</td>
<td>1 True; 2 Sometimes; 3 Not true</td>
<td>POS</td>
</tr>
<tr>
<td>55</td>
<td>gse21c10</td>
<td>Internalising behaviours</td>
<td>Depressed feelings</td>
<td>I felt lonely.</td>
<td>1 True; 2 Sometimes; 3 Not true</td>
<td>POS</td>
</tr>
<tr>
<td>56</td>
<td>gse21c11</td>
<td>Internalising behaviours</td>
<td>Depressed feelings</td>
<td>I thought nobody really loved me.</td>
<td>1 True; 2 Sometimes; 3 Not true</td>
<td>POS</td>
</tr>
<tr>
<td>57</td>
<td>gse21c12</td>
<td>Internalising behaviours</td>
<td>Depressed feelings</td>
<td>I thought I could never be as good as other kids.</td>
<td>1 True; 2 Sometimes; 3 Not true</td>
<td>POS</td>
</tr>
<tr>
<td>58</td>
<td>gse21c13</td>
<td>Internalising behaviours</td>
<td>Depressed feelings</td>
<td>I did everything wrong.</td>
<td>1 True; 2 Sometimes; 3 Not true</td>
<td>POS</td>
</tr>
<tr>
<td>59</td>
<td>gse21c2</td>
<td>Internalising behaviours</td>
<td>Depressed feelings</td>
<td>I didn't enjoy anything at all.</td>
<td>1 True; 2 Sometimes; 3 Not true</td>
<td>POS</td>
</tr>
<tr>
<td>60</td>
<td>gse21c3</td>
<td>Internalising behaviours</td>
<td>Depressed feelings</td>
<td>I felt so tired I just sat around and did nothing.</td>
<td>1 True; 2 Sometimes; 3 Not true</td>
<td>POS</td>
</tr>
<tr>
<td>61</td>
<td>gse21c4</td>
<td>Internalising behaviours</td>
<td>Depressed feelings</td>
<td>I was very restless.</td>
<td>1 True; 2 Sometimes; 3 Not true</td>
<td>POS</td>
</tr>
<tr>
<td>62</td>
<td>gse21c5</td>
<td>Internalising behaviours</td>
<td>Depressed feelings</td>
<td>I felt I was no good anymore.</td>
<td>1 True; 2 Sometimes; 3 Not true</td>
<td>POS</td>
</tr>
<tr>
<td>63</td>
<td>gse21c6</td>
<td>Males</td>
<td>Depressed feelings</td>
<td>I cried a lot.</td>
<td>1 True; 2 Sometimes; 3 Not true</td>
<td>POS</td>
</tr>
</tbody>
</table>
### Appendix C

<table>
<thead>
<tr>
<th>Item number</th>
<th>VARIABLE</th>
<th>Construct</th>
<th>Measure</th>
<th>Question</th>
<th>Responses</th>
<th>Worded</th>
</tr>
</thead>
<tbody>
<tr>
<td>64 gse21c7</td>
<td>Internalising behaviours</td>
<td>Depressed feelings</td>
<td>I found it hard to think properly or concentrate.</td>
<td>1 True; 2 Sometimes; 3 Not true</td>
<td>POS</td>
<td></td>
</tr>
<tr>
<td>65 gse21c8</td>
<td>Internalising behaviours</td>
<td>Depressed feelings</td>
<td>I hated myself.</td>
<td>1 True; 2 Sometimes; 3 Not true</td>
<td>POS</td>
<td></td>
</tr>
<tr>
<td>66 gse21c9</td>
<td>Internalising behaviours</td>
<td>Depressed feelings</td>
<td>I was a bad person.</td>
<td>1 True; 2 Sometimes; 3 Not true</td>
<td>POS</td>
<td></td>
</tr>
</tbody>
</table>

* Items with X as item number are removed from the scale due to poor discrimination.  
  Item in green font are positively worded; items in red font are negatively worded.

---

150
Appendix D: ACER ConQuest Command File to Estimate Gender DIF in the LSAC Example

let name=DIF3;
doif %name%==DIF3;
let items=ese03a1a_R ese03a1b_R ese03a1c_R ese03a1d_R ese03a1e_R ese03a2a_R ese03a2b_R ese03a2c_R ese03a2d_R ese03a2e_R ese03a3a_R ese03a3b_R ese03a3c_R ese03a3d_R ese03a3e_R ese03a4a_R ese03a4b_R ese03a4c_R ese03a4d_R ese03a4e_R ese03a5a_R ese03a5b_R ese03a5c_R ese03a5d_R ese03a5e_R ese06a_R ese13a1_R ese13a2_R ese13a3_R ese13a4_R ese13b1_R ese13b2_R ese13b3_R ese13b4_R ese03t1a_R ese03t1b_R ese03t1c_R ese03t1d_R ese03t1e_R ese03t2a_R ese03t2b_R ese03t2c_R ese03t2d_R ese03t2e_R ese03t3a_R ese03t3b_R ese03t3c_R ese03t3d_R ese03t3e_R ese03t4a_R ese03t4b_R ese03t4c_R ese03t4d_R ese03t4e_R ese03t5a_R ese03t5b_R ese03t5c_R ese03t5d_R;
endif;
doif %name%==DIF7;
let items=gse03c1a_R gse03c1b_R gse03c1c_R gse03c1d_R gse03c1e_R gse03c2a_R gse03c2b_R gse03c2c_R gse03c2d_R gse03c2e_R gse03c3a_R gse03c3b_R gse03c3c_R gse03c3d_R gse03c3e_R gse03c4a_R gse03c4b_R gse03c4c_R gse03c4d_R gse03c4e_R gse03c5a_R gse03c5b_R gse03c5c_R gse03c5d_R gse03c5e_R gse14a_R gse14b_R gse14c_R gse14d_R gse14e_R gse16a_R gse16b1_R gse16b2_R gse16b3_R gse16b4_R gse16b5_R gse16b6_R gse16b7_R gse16b8_R gse21a1_R gse21a2_R gse21a3_R gse21a4_R gse21a5_R gse21c1_R gse21c10_R gse21c11_R gse21c12_R gse21c13_R gse21c2_R gse21c3_R gse21c4_R gse21c5_R gse21c6_R gse21c7_R gse21c8_R gse21c9_R;
endif;
let path= C:\Work_Gebhardt\_My Work\Thesis\LSAC \Scaling;

datafile %path%\input\d3.For Scaling.sav !filetype=spss, responses=%items%, keep= KINDER_G20 KINDER_L20 KINDER_M esos esep FEMALE y3num y7num, weight=defgwt, pid=hicid;
labels << %path%\input\i57_wb7.lab;
export logfile >> %path%\output\dif\%name%.log;
model item + item*step -female +item*female;
export parameters >> %path%\output\dif\%name%.par;
estimate! nodes=25,iter=5000, deviancechange=0.000001, converge=0.00001, stderr=empirical, minnode=-10, maxnode=6, fit=yes;
show !estimates=latent,expand=no, table=1:2:3:4:5 >> %path%\output\dif\%name%.shw;
descriptives !estimates=pv, filetype=excel >> %path%\output\dif\%name%_des.xls;
Appendix E: ACER ConQuest Command File to Estimate a Full Latent Path Model within an IRT Framework for the LSAC Example

let name=male;
let comp=C: \Work_Gebhardt\My Work\Thesis\LSAC;
let path=%comp%\Scaling;
set seed=3, kepplastest=yes, warnings=no, iterlimit=500, constraints=items;
datafile %path%\input\d3.For Scaling.sav !filetype=spss, responses=ese03a1a_R to ese03t5d_R gse03c1a_R to gse21c9_R, keep=KINDER_G20 KINDER_L20 KINDER_M esos esep FEMALE INDIG y3num y7num ENJOY TROUBLE BULLY NOTWELL HIGHSCH CHANGED HOMEWORK FRIENDS INDIVIDUAL SELF DIRECT EXTRACLASS COMPUTER LEVELS MISSTCH MISS HW Y7_PINTEREST, weight=defgwt2, pid=hicid;
labels << %path%\input\i128_2d.lab;
export logfile >> %path%\output\%name%.log;
score (0,1,2,3,4)(0,1,2,3,4)() !items (1-62);
score (0,1,2,3,4)() (0,1,2,3,4) !items (63-128);
model item + item*step;
delete !item (35-38,49,59,64,93-98,101,115);
doif %name%==male;
keepcases 0 !female;
endif;
doif %name%==female;
keepcases 1 !female;
endif;
regression KINDER_G20 KINDER_L20 KINDER_M esep y3num y7num;
export parameters >> %path%\output\%name%.par;
estimate! method=quadrature, nodes=25, iter=5000, deviancechange=0.000001, converge=0.00001, stderr=quick, minnode=-10, maxnode=6, fit=yes;
show !estimates=latent, expand=no, table=1:2:3:4:5 > %path%\output\%name%.shw;
itanal !estimates=latent >> %path%\output\%name%.ita;
itanal! estimates=latent, format=summary, filetype=excel > %path%\output\%name%_summary.xls;
descriptives !estimates=pv, filetype=excel > %path%\output\%name%_des.xls;
show cases !estimates=latent, filetype=spss > %path%\output\%name%_PVS.sav;
structural
\/wb3 on esep KINDER_G20 KINDER_L20 KINDER_M
\/y3num on esep KINDER_G20 KINDER_L20 KINDER_M
\/wb7 on wb3 y3num
\/y7num on y3num wb3
!matrixout=m >> %path%\output\%name%.str;
print m_fullsscp !filetype=xlsx >> %path%\output\%name%_fullsscp.xlsx;
print m_osscp !filetype=xlsx >> %path%\output\%name%_osscp.xlsx;
print m_lossscp !filetype=xlsx >> %path%\output\%name%_lossscp.xlsx;
print m_lsscp !filetype=xlsx >> %path%\output\%name%_lsscp.xlsx;
Appendix F: ACER ConQuest Command File to Estimate a Full Latent Path Model within an IRT Framework for the PISA Example

/* (1) Estimate item parameters (excluding not reached) */
/*===================================================================*/
reset all;
let path=C:\Work_Gebhardt\_My Work\Thesis\PISA;
let name=learning_item;
set keeplast=yes, logestimates=yes, warnings=no, iterlimit=200,
constraints=items, nodefilter=0.01;
datafile %path%\data\AUS09_input.sav !filetype=spss, responses=M033Q01 to S527Q04T con1 con2 con3 con4 con5, pid=fullstd, weight=W_FSTUWT,
keep=female hisced hiscedm wealth wealthm government conm;
labels << %path%\learning.lab;
export logfile >> %path%\%name%.log;
codes     0,1,2,3,4,5,6,8,9,A,B,C,D,E,F,G,H,I,O,P,Q;
key
411GG4211414411114G11414234313111141213111114132511131112311
1113211122111132111111C1311413121123411144114114114111131GG11G
411234134321423211234134G132323323111112123231111 !1;
key
xxxHHx3xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx6xxxxxxxxxxxxx
xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxHHxxH
xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx !1;
key
xxxIIxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx !1;
key
xxxOOxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx !2;
key
xxxPxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx !2;
key
xxxQxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx !2;
key
xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx !3;
score (0,1,2) (0,1,2) () () !items (1-35);
Appendix F

score (0,1,2) ()(0,1,2)()() !items (36-135);
score (0,1,2) ()()() (0,1,2)() !items (136-188);
score (0,1,2,3,4)()()()() (0,1,2,3,4) !items(189-193);

model item+item*step;
keepcases 0 !hiscedm;
keepcases 0 !wealthm;
import init_parameters >> %path%\%name%.par;
export parameters >> %path%\%name%.par;
estimate ! method=gauss, fit=yes, iter=1000, nodes=15, conv=0.0001,
stderr=quick;
show !estimates=latent, expand=no, table=1:3:4:5 >> %path%\%name\.shw;
itanal !estimates=latent >> %path%\%name\.ita;
itanal! estimates=latent, format=summary, filetype=excel >>
%path%\%name\_summary.xls;

/* (2) Estimate population parameters (including not reached) */
/*--------------------------------------------------------------------------------*/
reset all;
let path=C:\Work_Gebhardt\_My Work\Thesis\PISA;
let name=learning_stud;
set keeplast=yes, logestimates=yes, warnings=no, iterlimit=200,
constraints=items, nodefilter=0;
datafile %path\data\AUS09_input.sav !filetype=spss, responses=M033Q01 to
S527Q04T con1 con2 con3 con4 con5, pid=fullid, weight=W_FSTUWT,
keep=female hisced hiscedm wealth wealthm government metro conm;
labels << %path\learning.lab;
export logfile >> %path%\%name%.log;
codes 0,1,2,3,4,5,6,8,9,A,B,C,D,E,F,G,H,I,O,P,Q,r;
key
411GG42111414411114G11414234313111141213111114111143325111131112311
111321112211113111111111C131141312112341114114111141111131GG11G
411234134321423211234134G1323233211111121223231111 !1;
key
xxxHHx3xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx6xxxxxxxxxxx
xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxHxxH
xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx !1;
key
xxxIIxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxHxxxxxH
xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx !1;
key
xxxOxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx22xxxxxxxxxx222222222222xxxxxxxxxxxxxxxxxxxx
xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx !2;
key
xxxPxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx !2;
key

score (0,1,2) (0,1,2)()()() !items (1-35);
score (0,1,2) ()(0,1,2)()() !items (36-135);
score (0,1,2) ()()0,1,2)() !items (136-188);
score (0,1,2,3,4)()()()0,1,2,3,4) !items(189-193);

model item+item*step;
regression  female hisced government wealth metro;
keepcases 0 !hiscedm;
keepcases 0 !wealthm;
import anchor_parameters << %path%\learning_item.par;
export parameters >> %path%\%name%.par;
export reg_coefficients >> %path%\%name%.reg;
export covariance >> %path%\%name%.cov;
estimate ! method=montecarlo, fit=no, nodes=2000, conv=0.0001,
stderr=quick;
show !estimates=latent, expand=no, table=1:2:3:4:5 >> %path%\%name%.shw;
itanal !estimates=latent >> %path%\%name%.ita;
itanal !estimates=latent, format=summary, filetype=excel >>
%path%\%name%_summary.xls;
descriptives !estimates=pv, filetype=excel >> %path%\%name%_des.xls;
show cases !estimates=latent, filetype=spss >> %path%\%name%_PVS.sav;
structural
/road on female government control
!matrixout=r1 >> %path\r1.str;
structural
/road on female government control
/control on female hisced government
/government on wealth metro
!matrixout=r2 >> %path\r2.str;
print r2_fullsscp !filetype=xlsx >> %path%\name\_fullsscp.xlsx;
Curriculum Vitae

2007–current  Senior Research Fellow (Australian Council for Educational Research)


1997–2001  Evaluation of two educational preschool programs in socio-economic disadvantaged areas (University of Amsterdam)

1990–1996  Masters in clinical psychology (University of Amsterdam)
Author/s: Gebhardt, Eveline

Title: Latent path models within an IRT framework

Date: 2016

Persistent Link: http://hdl.handle.net/11343/90848

File Description: Latent path models within an IRT framework