ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

DEPARTMENT OF STATISTICS

POSTGRADUATE PROGRAM

Structural Equation Models with Covariate Effects: A Comparison Between the Lisrel and the Item Response Theory Approach

By

Dimitrios {Stylianos} Mauridis

A THESIS

Submitted to the Department of Statistics
of the Athens University of Economics and Business
in partial fulfilment of the requirements for
the degree of Master of Science in Statistics

Athens, Greece
2003

ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

DEPARTMENT OF STATISTICS

POSTGRADUATE PROGRAM

STRUCTURAL EQUATION MODELS WITH COVARIATE EFFECTS: A COMPARISON BETWEEN THE LISREL AND THE ITEM RESPONSE THEORY APPROACH

By

DIMITRIOS {STYLIANOS}. MAVRIDIS

A THESIS
Submitted to the Department of Statistics
of the Athens University of Economics and Business
in partial fulfillment of the requirements for
the degree of Master of Science in Statistics

Athens, Greece
February 2003

ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ

ΥΠΟΔΕΙΓΜΑΤΑ ΔΙΑΡΘΡΩΤΙΚΩΝ ΕΞΙΣΩΣΕΩΝ ΜΕ ΕΡΕΥΝΗΤΙΚΕΣ ΜΕΤΑΒΛΗΤΕΣ: ΜΙΑ ΣΥΓΚΡΙΣΗ ΔΥΟ ΘΕΩΡΗΤΙΚΩΝ ΠΡΟΣΕΓΓΙΣΕΩΝ

ΔΗΜΗΤΡΙΟΣ ΣΤΥΛΙΑΝΟΣ. ΜΑΥΡΙΔΗΣ

ΔΙΑΤΡΙΒΗ

Που υποβλήθηκε στο Τμήμα Στατιστικής
του Οικονομικού Πανεπιστημίου Αθηνών
ως μέρος των απαιτήσεων για την απόκτηση
Μεταπτυχιακού Διπλώματος Ειδίκευσης στη Στατιστική

Αθήνα
Φεβρουάριος 2003

ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

DEPARTMENT OF STATISTICS

A Thesis submitted in partial fulfillment of
the requirements for the degree of
Master of Science

Structural Equation Models with Covariate Effects: A Comparison Between the Lisrel and the Item Response Theory Approach

Dimitrios {Stylianos} Mauridis

Approved by the Graduate Committee

E. Moustaki	S. Psarakis	V. Vasdekis
Assistant Professor	Assistant Professor	Lecturer
Thesis Supervisor	Member of the Committee	

Athens, June 2003

Michael Zazanis, Associate Professor
Director of the Graduate Program

ACKNOWLEDGEMENTS

I would like to thank Professor Irini Moustaki for her co-operation throughout the project and for giving me the opportunity of becoming familiar with factor analysis. Her help and guidance were valuable for the fulfilment of this thesis. I would also like to thank Professor Karl Jöreskog for reading this project and providing me with very useful comments and examples in LISREL

VITA

I was born in Lamia in 1978. I graduated from the Statistics department of the Athens University of Economics and Business in 2000 and I entered the postgraduate Statistics program of the Athens University of Economics and Business in the same year. The second semester was held at the Katholieke Universiteit of Leuven in Belgium.

ABSTRACT

Mavridis Dimitrios

STRUCTURAL EQUATION MODELS WITH COVARIATE EFFECTS: A COMPARISON BETWEEN THE LISREL AND THE ITEM RESPONSE THEORY APPROACH

February 2003

This project reviews latent variable models for analyzing ordinal observed variables. We review two different methodologies, namely the LISREL approach and the Item Response Theory approach. We also investigate the case where there are some covariates affecting the observed variables and some covariates affecting the latent variables. We use two real data sets to compare the two methodologies in terms of model estimation and goodness-of-fit, and we investigate the potentials of each approach.

ΠΕΡΙΛΗΨΗ

Μαυρίδης Δημήτριος

ΥΠΟΔΕΙΓΜΑΤΑ ΔΙΑΡΘΡΩΤΙΚΩΝ ΕΞΙΣΩΣΕΩΝ ΜΕ ΕΡΕΥΝΗΤΙΚΕΣ ΜΕΤΑΒΛΗΤΕΣ: ΜΙΑ ΣΥΓΚΡΙΣΗ ΔΥΟ ΘΕΩΡΗΤΙΚΩΝ ΠΡΟΣΕΓΓΙΣΕΩΝ

Φεβρουάριος 2003

Σε αυτήν την εργασία εξετάζουμε μοντέλα λανθάνουσων μεταβλητών για δεδομένα που βρίσκονται σε κλίμακα διάταξης. Εξετάζουμε δυο μεθοδολογίες, το LISREL μοντέλο και την Item Response Theory προσέγγιση. Εξετάζουμε επίσης την περίπτωση όπου κάποιες μεταβλητές επιδρούν στις παρατηρούμενες μεταβλητές και κάποιες μεταβλητές επιδρούν πάνω στις λανθάνουσες μεταβλητές. Χρησιμοποιούμε δυο αληθινά παραδείγματα για να συγκρίνουμε τις δυο μεθοδολογίες όσον αφορά την εκτίμηση των παραμέτρων, την προσαρμογή του μοντέλου και τις δυνατότητες της κάθε προσέγγισης.

TABLE OF CONTENTS

LIST OF TABLES

Table **Page**

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Multivariate techniques are used in all fields of research. Probably the most famous multivariate technique is factor analysis. It is very common in every field to have numerous variables, many of which are strongly correlated. Our aim with factor analysis is to explain the interrelationships among the observed variables through a set of latent, unobserved variables. The number q of latent variables is much smaller than the number of observed variables p. It is much easier to work with fewer variables and this is a strong reason why someone would conduct factor analysis. Furthermore many variables measure the same thing, probably ability or some other characteristic and their introduction in the analysis adds nothing but complexity. Reduce dimensionality is the most important reason for conducting factor analysis. By doing so we might be able to see relationships among variables that were not obvious before. Another reason that is most common in the social and psychological sciences is to reveal some constructs or characteristics that cannot be measured directly or it is perilous to do so. These characteristics may be unobserved and abstract such as cleverness, sensitiveness, and aggressiveness. Some characteristic may be real like the personal income or the political identification but for some reason we may not consider people's responses as reliable and it would be better to infer this characteristic from a set of indicators of this characteristic. Actually what factor analysis does is to infer these abilities or characteristics from some observed variables which are indicators of these characteristics.

The foundations of factor analysis date back to the beginning of the nineteenth century, but its popularity has increased recently with the improvement of the personal computers, which can now meet with the demanding calculations that factor analysis, and its related topics need. The origins of factor analysis can be found at Spearman (1904), who made the assumption that people who tend to score high on a mental test, also tend to score high on other tests. He assumed that all these mental tests are indicators of a general ability of the student. This general ability that can be represented by a latent variable has different correlations with every item and it is one of the aims of factor analysis to estimate how much every item contributes to this latent variable.

By the name "factor analysis" we usually mean the occasion where all manifest variables are continuous. Actually factor analysis can be conducted when the manifest variables are continuous, nominal, ordinal or a mixture of these. In this project we deal with the case where the observed variables are on an ordinal scale. Only recently, and more specifically in the last decade, people started to think of cases where the observed variables are not continuous or they are a combination of continuous and discrete. When the observed variables are ordinal we should not assign numbers to them with the purpose to treat them as continuous. They are on an ordinal scale and we should not assign them metric properties. The numbers that are usually assigned to such data are in fact labels and the addition or the ratio of these numbers has no actual meaning. Except from the latent variables there might be some covariates that together with the latent variables account for the interrelationships of the observed variables and in addition we might want to investigate the effect of some other covariates on the latent variables.

There are two ways for conducting latent variable analysis for ordinal data. One is the underlying variable approach which assumes that the observed ordinal variables are generated by a set of underlying continuous variables and is supported by commercial software such as LISREL (Jöreskog and Sörbom 1993), EQS (Bentler 1992) and Mplus (Muthén and Muthén 2000). The other approach is the item response theory (IRT) approach where the model specifies the complete p-dimensional response patterns (p denotes the number of observed variables) and makes assumptions about the conditional distribution of the observed variables given the latent variables and the covariates, if there are any.

The underlying variable approach has developed much earlier from the IRT approach and as a consequence it has expanded its potentials to many fields. Hence, there are many programs and many articles in the literature that conduct factor analysis with covariate effects. Jöreskog and Goldberger(1975) discussed a multiple indicators and multiple causes (MIMIC) model for normal manifest variables with a single latent variable that allows for direct effects on the covariates. Muthén(2002) gives an overview of statistical analysis with latent variables and his work focuses on measurement error and hypothetical constructs measured by multiple indicators. IRT methodology has mainly developed theory for measurement models that do not include covariate effects. Sammel, Ryan and Legler (1997) discussed an unidimensional latent trait model for binary and normal outcomes that allow for covariate effects within the IRT framework. Moustaki (2003) has developed a general IRT framework that includes covariate effects both on the manifest variables and the latent variables.

These two methodologies are completely separated and scientists whose interest lies in factor analysis are occupied only with one of the approaches and their knowledge of the other approach, its advantages and its drawbacks, is very limited. There are not many articles in the literature that compare the two approaches. A comparison of the two approaches for measurement models without direct effects can be found in Moustaki(2002) and Moustaki and Jöreskog(2001). In this paper we will present two models, one of each approach, and we will run two examples for measurement models with direct effects to compare their results. From the underlying variable approach we will present the LISREL model, which is probably the most famous and it has contributed to the popularity of the latent variable models. The LISREL program has two steps. Namely PRELIS (preprocessor of LISREL) and LISREL. In PRELIS we compute the covariance or the correlation matrix and in the LISREL step we fit the model to the matrix we have computed from PRELIS. From the item response theory approach we will show a method presented by Moustaki (2003) and we will run the examples with the program GENLAT 1.1(Moustaki 2002).

We will focus mainly on the parameter estimates obtained and the fit of the models. We cannot use the same goodness-of-fit measures for ordinal data as those we use for continuous data. There are not many goodness-of-fit measures or criteria in the literature that test the fit of the model for latent variable models with ordinal data and in this project we present the most widely used. When we have a measurement model with direct effects, the fit of the model gets more difficult to be tested. Through the examples the difficulties in testing the fit of the model will be more clear, as more clear will be the capabilities of each of the two different methods.

In chapter 2 and 3 we will present the Item Response Theory approach and the LISREL model respectively. In chapter 4 we will show the equivalence between the two approaches presented in chapters 2 and 3 and in chapter 5 we will present some tools for measuring the goodness-of-fit of the models. In chapter 6 we will use a real data set, where there is one covariate that affects the ordinal observed variables, to compare the two approaches. In chapter 7 we will use another real data set, where there are some covariates that affect the observed ordinal variables and some other covariates that affect the latent variable, to compare the two approaches. Finally in chapter 8 we will present the main conclusions of our analysis.

CHAPTER 2

ITEM RESPONSE THEORY APPROACH

2.1 Introduction.

In the item response theory approach the whole response pattern is used and as a consequence there is no loss of information. We have to make two major assumptions in order to develop the model. Firstly, we have to formulate a model for the conditional distribution of the ordinal observed variables given the latent variables and secondly, we assume that given the latent variables, the observed variables are independent, the latter assumption is also known as axiom of local or conditional independence. In this chapter we will present these two assumptions and the IRT model with covariate effects.

2.2 MODEL AND ESTIMATION

We start from the first assumption. One of our main tasks is to infer the respondent's ability or skill. The probability of responding a particular response category must depend on her ability. We should assume a probabilistic model that connects the probabilities of responding each response category to the respondent's ability. Thus item response theory starts with a mathematical statement as to how response depends on level of ability. This is achieved by the item response function. McCullagh (1980) discusses a general class of regression models for ordinal data. In his work, there are no latent variables but only an ordinal dependent observed variable and a set of explanatory variables. He showed that these models are multivariate extensions of generalized linear models. Moustaki (2000 and 2003) has extended that regression model for latent variables. The differences are that instead of a single dependent ordinal observed variable we have a set of p ordinal observed variables denoted by the vector $y = (y_1, y_2, \ldots, y_p)$ and instead of a set of explanatory covariates, we have a set of latent variables, denoted by the vector $z = (z_1, z_2 \ldots, z_q)$. The latent variable model considered by Moustaki(2000) has the form :

$$link[\gamma_{is}(z)] = \tau_{is} - \sum_{j=1}^{q} \alpha_{ij} z_j \qquad i=1,2,..,p \qquad (1)$$

where s goes through the m_i response categories of item i and $\gamma_{is}(z)$ is the cumulative probability of a response in category s or lower of item y_i, which is a function of the latent variables, namely $\gamma_{is}(z) = \pi_{i1}(z) + \pi_{i2}(z) + \ldots + \pi_{is}(z)$. Moustaki (2003) extended this model to entail r covariates x, denoted by the vector $x = (x_1, x_2, \ldots, x_r)$, that affect the responses to observed ordinal variables and also k covariates w, denoted by the vector $w = (w_1, w_2, \ldots, w_k)$ that affect the latent variables. The model now takes the form

$$link[\gamma_{is}(z)] = \tau_{is} - \sum_{j=1}^{q} \alpha_{ij} z_j + \sum_{l=1}^{r} \beta_{il} x_l \qquad (2)$$

and the latent variables z_j are related to a set of observed covariates w_h in a simple linear form

$$z_{jn} = \sum_{h=1}^{k} c_{hj} w_h + \delta_n \qquad (3)$$

which can also be written in the form Z=WC+Δ where Z is a (nxq) matrix of latent variables, W is a (nxk) matrix of explanatory variables, C is a (kxq) matrix of regression coefficients and Δ is a (nxq) matrix of error terms.

The parameters τ_{is} of equation (1) are referred as cut-points and they are an increasing function of s $\tau_{io} = -\infty \leq \tau_{i2} \leq \ldots \leq \tau_{im} = +\infty$. The parameters a_{ij} are in fact factor loadings since they measure the effect of the latent variable z_j on some function of the cumulative probability of responding up to a category of the ith item controlling for the effect of the covariates x. The negative sign in front of the slope parameter is used to indicate that as z increases the response on the observed item y_i is more likely to fall at the high end of the scale. The coefficients β_{il} of the covariates x affect only the cut-points and it allows individuals with the same position on the latent variables to have different cumulative and response probabilities, if they have different values on the covariates.

Another issue is the choice of the response function. The link function can be any monotonically increasing function that map (0,1) onto $(-\infty,+\infty)$. Examples of link functions are the logit, the complementary log-log function, the inverse normal function, the inverse Cauchy, the hazard function and the log-log function. McCullagh(1980) claims that the proportional odds model and the proportional hazard model, which use the logit and the complementary log-log functions respectively, are most often used in practice because of the simplicity of their interpretation. Lord(1980) suggests using the logit instead of the probit function and he bases his suggestion on the grounds that individuals on high ability levels should virtually never answer an easy item incorrectly and they won't do many careless mistakes and, since the logistic function approaches its asymptotes less rapidly than the probit, such careless mistakes will do less violence to the logistic than to the normal model. In the examples in sections 6 and 7 of this project we will use the logit as a link function.

In generalized linear models the dependent variable has a distribution from the exponential family, and in our case, we suppose that each of the p ordinal observed variables, conditional on the latent variables and the set of covariates, has a multinomial distribution, which as we will show belongs to the exponential family. The other assumption is that of local or conditional independence. Under this assumption the observed variables are independent if the latent variables and the covariates are held fixed or in other words the latent variables and the covariates account for the interrelationships among the observed variables. In the case where we have ordinal manifest variables the answer an individual gives to an item given the latent variables z and the explanatory variable x is independent of the answers s/he gave to the rest of the items. A mathematically equivalent statement of local independence is that the probability of success on all items is equal to the product of separate probabilities of success.

$$P(y_1 = a_1, y_2 = a_2, \dots, y_p = a_p / z, x) = P(y_1 = a_1 / z, x) P(y_2 = a_2 / z, x) \dots P(y_p = a_p / z, x)$$

where α_i represent response categories or $$g(y/z,x)=\prod_{i=1}^{p} g(y_i/z,x) \qquad (4)$$

We assume that the latent variables are independent with standard normal distributions. This specification identifies the scale of the latent variables, which in turn identifies the scale of the item parameters. Bartholomew (1988) found that any symmetric prior will predict the same first and second order margins and so the effect of the prior is negligible. He suggested the normal because it has rotational advantages when it comes to more than one latent variable. Since the latent variables follow the standard normal distribution their values range from $-\infty$ to $+\infty$.

The joint distribution of the p observed variables is:

$$f(y)=\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} g(y/z,x)h(z/w)dz \qquad (5)$$

Taking into mind the axiom of conditional independence the joint distribution of the p observed variables becomes.

$$f(y)=\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}\prod_{i=1}^{p} g(y_i/z,x)h(z/w)dz \qquad (6)$$

The exponential distribution has the form

$$f_i(y_i,\theta_i,\varphi_i)=\exp\left\{\frac{y_i\theta_i(z,x)-b_i(\theta_i(z,x))}{\varphi_i}+c_i(y_i,\varphi_i)\right\} \qquad (7)$$

Where $\theta_i(z,x)$ is the canonical parameter and $b_i(\theta_i(z,x))$ and $c_i(y_i,\varphi_i)$ are specific functions taking a different form depending on the distribution of the response variables y_i, and ϕ_i is a scale parameter which is one in the ordinal case.

For an observed item y_i the conditional distribution of y_i/z is the multinomial.

$$g(y_i/z,x)=\prod_{s=1}^{m_i}\pi_{is}(z,x)^{y_{is}}=\prod_{s=1}^{m_i}(\gamma_{i,s}-\gamma_{i,s+1})^{y_{is}} \qquad (8)$$

where $y_{i,s} = 1$ if a randomly selected individual responds into category s of the ith item and $y_{i,s} = 0$ otherwise. The above equation can also be written in the form

$$g(y_i / z, x) = \prod_{s=1}^{m_i} \left(\frac{\gamma_{i,s}}{\gamma_{i,s+1}} \right)^{y^*_{i,s}} \left(\frac{\gamma_{i,s+1} - \gamma_{i,s}}{\gamma_{i,s+1}} \right)^{y^*_{i,s+1} - y^*_{i,s}} \quad (9)$$

If we take the logarithm of the above equation we get:

$$\log g(y_i / z, x) = \sum_{s=1}^{m_i - 1} \left[y^*_{i,s} \log \frac{\gamma_{i,s}}{\gamma_{i,s+1} - \gamma_{i,s}} - y^*_{i,s+1} \log \frac{\gamma_{i,s+1}}{\gamma_{i,s+1} - \gamma_{i,s}} \right]$$

$$= \sum_{s=1}^{m_i - 1} \left[y^*_{i,s} \theta_{i,s}(z, x) - y^*_{i,s+1} b(\theta_{i,s}(z, x)) \right] \quad (10)$$

In this way, each component is in the form of the general expression of the exponential family distribution. More specifically:

$$\theta_{i,s}(z, x) = \log \frac{\gamma_{i,s}}{\gamma_{i,s+1} - \gamma_{i,s}}, s = 1,\dots m_i \quad (11)$$

$$b(\theta_{i,s}(z, x)) = \log \frac{\gamma_{i,s+1}}{\gamma_{i,s+1} - \gamma_{i,s}} = \log(1 + \exp(\theta_{i,s}(z, x))) \quad (12)$$

The canonical parameter θ_i is not a linear function of the parameters and as a result of that there is no simple linear function to summarize the information contained in the latent variables as there is for continuous or nominal data for binary cases.

2.3 MODEL IDENTIFICATION

In order for the model to be identified, a necessary condition is that the covariates x that have direct effects on the items should be different from the covariates w that affect the latent variables. We check what would happen to a model with one latent variable and the same covariate (x) affecting both the items and the latent variable. The model would be:

$$link[\gamma_{is}(z, x)] = \tau_{is} - a_{i1} z + \beta_{i1} x$$

$$z = cx + \delta$$

And by substituting the structural part of the model into the measurement model with direct effects we have:

$$link[y_{is}(z,x)] = \tau_{is} - a_{i1}(cx+\delta) + \beta_{i1}x = \tau_{is} - a_{i1}\delta - (a_{i1}c - \beta_{i1})x$$

It is evident from the above equation that the parameters c and β_{i1} cannot be estimated separately and therefore are not identified.

2.4 MODEL ESTIMATION

We aim to estimate all the parameters simultaneously. The model is fitted to the whole response pattern, including both the responses to the p ordinal variables and the values of the r covariates having direct effects on the items and the k covariates affecting the latent variable. The vector y of the observed ordinal items is affected by the latent variables z and the covariates x, whereas the latent variables z are affected only by the covariates w. The covariates x and w are considered fixed. So the joint distribution of the random variables is:

$$f(y,z)=g(y/z,x)h(z/w) \tag{13}$$

The complete log-likelihood for a random sample of size N is written as

$$L = \sum_{n=1}^{N} \log f(y_n, z_n) = \sum_{m=1}^{n} \log g(y_n / z_n, x_n) h(z_n / w_n) = \sum_{m=1}^{n} (\log g(y_n / z_n, x_n) + \log h(z_n / w_n)) \tag{14}$$

And by using the axiom of conditional independence $g(y/z,x) = \prod_{i=1}^{p} g(y_i / z, x)$

we have that

$$L = \sum_{n=1}^{N} \left[\sum_{i=1}^{p} \log g(y_{in} / z_n, x_n) + \log h(z_n / w_n) \right] \tag{15}$$

In the above formulation of the log-likelihood z is unknown and in order to maximize the log-likelihood we will use an EM algorithm. EM algorithm is an iterative technique, which carries out an expectation and a maximization step until convergence is attained, and it has been considered appropriate for models where there is missing information. The latent variables are considered here the missing information. In the expectation step, the expected score function of the model parameters is computed. The score function is the first derivative of the log-likelihood with respect to the parameters.

2.5 ESTIMATION OF C

From the formulation of the log-likelihood (14) we see that the estimation of the parameters in the matrix C that affects the latent variables z does not depend on the first component of the log-likelihood. So, estimation of C can be done separately from the rest of the parameters (τ, α, β).

The latent variables, conditional on w are assumed to follow the normal distribution with mean cw and variance 1 and to be independent so that

$$h(z/w) = \prod_{j=1}^{q} h(z_j/w) \tag{16}$$

The expected score function with respect to the parameter vector c_j is:

$$E(S(c_j)) = \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} S(c_j) h(z/y, x) dz \tag{17}$$

The score function with respect to the parameters c_j is given as:

$$S(c_j) = \frac{\partial L}{\partial c_j} = \frac{\partial \log h(z_j/w, c_j)}{\partial c_j} = \frac{\dfrac{\partial h(z_j/w, c_j)}{\partial c_j}}{h(z_j/w, c_j)}$$

$$= \frac{\partial \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z-c_j w)^2} \Big/ \partial c_j}{\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z-c_j w)^2}} = w(z_j - w c_j) \tag{18}$$

Therefore by substituting (18) into (17) we have:

$$E(S(c_j)) = \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} w(z_j - wc_j) h(z/y, x) dz \tag{19}$$

The fact that the latent variable is unobserved poses a problem to the calculation of the above integral. In order to overcome these difficulties we will approximate the integral by Gauss-Hermite quadrature. We will treat the latent variables as discrete with values $z_{t_1},\dots,z_{t_q}$ and their corresponding probabilities $h(z_{t_1}/w),\dots,h(z_{t_q}/w)$ summing to unity. In this way and by solving $\sum_{m=1}^{n} E(S(c_j)) = 0$ we get an explicit solution for the maximum likelihood estimator of c_j:

$$c_j = \frac{\sum_{n=1}^{N} w_n \sum_{t_1=1}^{v_1} \dots \sum_{t_q=1}^{v_q} z_{t_j} h(z_{t_1},\dots,z_{t_q}/y_n,x_n)}{\sum_{n=1}^{N} w_n w_n^{'}} \tag{20}$$

where

$$h(z_{t_1},\dots,z_{t_q}/y_n,x_n) = \frac{g(y/z_{t_1},\dots,z_{t_q},x_n)\prod_{j=1}^{q} h(z_{t_j}/w_n,c_j)}{f(y_n,x_n)} \tag{21}$$

The maximum likelihood solution for c_j is updated at each step of the EM algorithm, as we will see.

2.6 ESTIMATION OF THE MODEL PARAMETERS α AND β

The estimation of the rest of the parameters, namely τ, α, β depends only on the first component of the complete log-likelihood. Let's denote by d a vector of all the parameters left to be estimated. So d=(τ, α, β).

The expected score function of the parameter vector d when the expectation is taken with respect to h(z/y, x) is:

$$E(S(d_i)) = \int_{-\infty}^{+\infty}\dots\int_{-\infty}^{+\infty} S(d_i) h(z/y,x) dz \quad i=1,2\dots p \tag{22}$$

where

$$S(d_i) = \frac{\partial L}{\partial d_i} = \frac{\partial \log g(y/z,x)}{\partial d_i} \tag{23}$$

Now,

$$\frac{\partial \log g(y/z,x)}{\partial d_i} = \sum_{s=1}^{m_i-1}\left[y^{*}_{i,s,m}\frac{\partial \theta_{i,s,m}}{\partial d_i} - y^{*}_{i,s+1,m}\frac{\partial b(\theta_{i,s,m})}{\partial d_i}\right] \tag{24}$$

Now we replace the above equation into the expected score function in equation (22).

$$E(S_m(d_i)) = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \sum_{s=1}^{m_i-1} \left[y^{*}_{i,s,m} \frac{\partial \theta_{i,s,m}}{\partial d_i} - y^{*}_{i,s+1,m} \frac{\partial b(\theta_{i,s,m})}{\partial d_i} \right] h(z / y_m, x_m) dz \qquad (25)$$

Solving $\sum_{m=1}^{n} E(S_m(d_i)) = 0$ and approximate the integrals with Gauss-Hermite quadrature points we get non-explicit solutions for the parameter vector d.

$$\sum_{t_1=1}^{v_1} \cdots \sum_{t_q=1}^{v_q} \sum_{s=1}^{m_i-1} \left[y^{*}_{i,s,m} \frac{\partial \theta_{i,s,m}}{\partial d_i} - y^{*}_{i,s+1,m} \frac{\partial b(\theta_{i,s,m})}{\partial d_i} \right] h(z_{t_1}, \ldots, z_{t_q} / y \ , x \ ) \qquad (26)$$

The above equation can be written as:

$$\sum_{t_1=1}^{v_1} \cdots \sum_{t_q=1}^{v_q} \sum_{s=1}^{m_i-1} \left[r_{i,s,t_1,\ldots,t_q} - r_{i,s+1,t_1,\ldots,t_q} \right] \qquad (27)$$

Where

$$r_{i,s,t_1,\ldots,t_q} = \sum_{n=1}^{N} h(z_{t_1}, \ldots, z_{t_q} / y_n, x_n) y^{*}_{i,s,m} \frac{\partial \theta_{i,s,m}}{\partial d_i} \qquad (28)$$

And

$$r_{i,s+1,t_1,\ldots,t_q} = \sum_{n=1}^{N} h(z_{t_1}, \ldots, z_{t_q} / y_n, x_n) y^{*}_{i,s+1,m} \frac{\partial b(\theta_{i,s,m})}{\partial d_i} \qquad (29)$$

We need to find the first derivatives of the functions $\theta_{i,s,m}$ and $b(\theta_{i,s,m})$ with respect to the model parameters. These are given in the appendix.

Now we will give a brief description of the EM algorithm that is used to compute the parameters of the model. The steps of the EM algorithm are defined as follows.

Step 1: Choose initial values for the model parameters, namely for all elements of the vector d and also for c.

Step 2: Compute the values $r_{i,s,t_1,\ldots,t_q}$ and $r_{i,s+1,t_1,\ldots,t_q}$. This is the expectation step of the EM algorithm.

Step 3: Obtain improved estimates of the parameters by solving the non-linear maximum likelihood equations for the parameters of the vector d and explicit solutions for the parameters c. At this step a one-step Fisher scoring algorithm, is used to solve the non-linear maximum likelihood equations.

Step 4: Return to step 2 and continue until convergence is attained.

CHAPTER 3

LISREL model

3.1 INTRODUCTION

The name LISREL is an acronym for "linear structural relations" and the LISREL model is actually a set of linear structural relations. Firstly, we consider the simple model without direct effects on the observed variables and without direct effects on the latent variables and then we present the model with covariate effects on the manifest and on the latent variables. The LISREL model consists of two parts. The measurement model that shows how the latent variables are related to the indicator manifest variables. And the structural equation model, also known as SEM that specifies the causal relationships among the latent variables.

3.2 MEASUREMENT MODEL WITHOUT DIRECT EFFECTS

We suppose that we have p observed variables, denoted by y, and q latent variables, denoted by z. The q latent variables $(z_1, z_2, \ldots, z_q)$ are divided into two groups. The independent, denoted by ξ, namely $(\xi_1, \xi_2, \ldots, \xi_{q_1})$ and the dependent, denoted by η, namely $(\eta_1, \eta_2, \ldots, \eta_{q_2})$. Note that $q_1 + q_2 = q$ and $z = \begin{pmatrix} \xi \\ \eta \end{pmatrix}$. We have that E($\xi$)=0 and E($\eta$)=0. The structural equation model has the form

$$\eta = A\eta + \Gamma\xi + \zeta \Leftrightarrow (\mathrm{I} - A)\eta = \Gamma\xi + \zeta \Leftrightarrow \mathrm{B}\eta = \Gamma\xi + \zeta \qquad (30)$$

where

- η is a q_2 x 1 random vector of latent dependent, or endogenous, variables
- ξ is a q_1 x 1 random vector of latent independent, or exogenous, variables

- Γ is a $q_2 \times q_1$ matrix of coefficients of the ξ-variables in the structural relationship
- A is a q_2 x q_2 matrix of coefficients of the η-variables in the structural relationship. (A has zeros in the diagonal, and B=I - A is required to be non-singular)
- ζ is a q_2 x 1 vector of equation errors (random disturbances) in the structural relationship between η and ξ
- ζ is assumed to be uncorrelated with ξ

We now go to the measurement model, which in general has the form

$$y=\mu+\Lambda z+e \tag{31}$$

where $z \sim N_q(0,I)$, $e \sim N_p(0,\Psi)$ and z is independent of e.

LISREL allows us to partition the observed variables y into two groups, namely y_1 and y_2 of dimensions p_1 and p_2 respectively, where p_1+p_2=p, such that the p_1 observed variables of y_1 are indicators of the q_1 latent variables of ξ and the p_2 observed variables of y_2 are indicators of the q_2 latent variables of η. This means that the dependent variables have different indicators from the independent and vice versa. The error term e is also divided into two groups, namely e_1 and e_2 with variances Ψ_{e_1} and Ψ_{e_2} respectively. The measurement model can now be rewritten in the form

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} + \begin{pmatrix} \Lambda_{y_1} & 0 \\ 0 & \Lambda_{y_2} \end{pmatrix} \begin{pmatrix} \xi \\ \eta \end{pmatrix} + \begin{pmatrix} e_1 \\ e_2 \end{pmatrix} \tag{32}$$

- y_1 is a p_1 x 1 vector of observed indicators of the independent latent variables ξ
- y_2 is a p_2 x1 vector of observed indicators of the dependent latent variables η
- e_1 is a p_1 x 1 vector of measurement errors in y_1
- e_2 is a p_2 x 1 vector of measurement errors in y_2

- Λ_{y_1} is a p_1x q_1 matrix of coefficients of the regression of y_1 on ξ
- Λ_{y_2} is a p_2x q_2 matrix of coefficients of the regression of y_2 on η
- e_1 is uncorrelated with ξ
- e_2 is uncorrelated with η
- ζ is uncorrelated with e_1 and e_2

Let's suppose that the variance of ξ and η are Φ_ξ and Φ_η respectively and that the variance of the error term ζ is Θ_ζ.

The covariance matrix of the latent variables $z = \begin{pmatrix} \xi \\ \eta \end{pmatrix}$ is given by

$$\Phi = \begin{pmatrix} \Phi_\xi & \Phi_\xi \Gamma' B'^{-1} \\ \Phi_\xi \Gamma' B'^{-1} & B^{-1}(\Gamma' \Phi_\xi \Gamma + \Theta_\zeta) B'^{-1} \end{pmatrix} \qquad (33)$$

The variance-covariance matrix of the observed variables, computed from the LISREL model (equation 31) is $\Sigma = \Lambda \Phi \Lambda' + \Psi$ (34)

Now, by substituting Φ into (34) we get the covariance matrix of the observed variables.

The covariance matrix of the observed variables under these assumptions is:

$$\Sigma = \begin{pmatrix} \Lambda_{x_1} \Phi_\xi \Lambda'_{x_1} + \Psi_{e_1} & \Lambda_{x_1} \Phi_\xi \Gamma' B'^{-1} \Lambda'_{x_2} \\ \Lambda_{x_2} \Phi_\xi \Gamma' B'^{-1} \Lambda'_{x_1} & \Lambda_{x_2} \left(B^{-1} \Gamma' \Phi_\xi \Gamma B'^{-1} + B^{-1} \Theta_j B'^{-1} \right) \Lambda'_{x_2} + \Psi_{e_2} \end{pmatrix} \qquad (35)$$

The parameters are estimated by minimizing some measure of the distance between the covariance matrix of the observed variables Σ, as it is given from formula (35), or the correlation matrix of Σ as it is preferable to work with the standardized variables, and the sample covariance or correlation matrix S. We should note that if we have p observed variables then the covariance matrix of these variables contains p(p+1)/2 non-duplicated elements, the model won't be identified if the number of parameters to be estimated is more than the non-duplicated elements of the covariance matrix. If the latter happens we will have to fix some parameters of the model in order that the model will be identifiable.

3.3 FACTOR ANALYSIS WITH ORDINAL OBSERVED VARIABLES

LISREL assumes that underlying every ordinal observed variable y there is a continuous variable y*. This continuous variable y* is assumed to follow the standard normal distribution and it is this variable that is used in the structural equation modeling. In this way we have a set of continuous variables. Suppose we have m categories, the relationship between y and y* is

$$y = i \Leftrightarrow \tau_{i-1} < y^* < \tau_i \quad, \text{i=1,2,...m}$$

where

$$-\infty = \tau_0 < \tau_1 < \tau_2 < \ldots < \tau_{m-1} < \tau_m = +\infty$$

The probability of a response in category i is

$$\pi_i = P(y = i) = P(\tau_{i-1} < y^* < \tau_i) = \int_{\tau_{i-1}}^{\tau_i} \phi(u)du = \Phi(\tau_i) - \Phi(\tau_{i-1}) \qquad (36)$$

where Φ is the distribution function of z* and ϕ the corresponding density function. The univariate log-likelihood is $\ln L(\tau_i) = N * \sum_{a=1}^{m_i} p_a^{(i)} * \ln \pi_a^{(i)}$ and the maximum likelihood estimator of τ_i is

$$\tau_i = \Phi^{-1}(\pi_1 + \pi_2 + \ldots \pi_i), \text{ i =1,2,...,m-1.} \qquad (37)$$

The probabilities π_i can be estimated consistently by the corresponding percentage p_i of responses in category i. So the estimates of the thresholds are

$$\hat{\tau}_i = \Phi^{-1}(p_1 + p_2 + \ldots p_i) \qquad (38)$$

There are m-1 parameters τ_i and also m-1 independent sample proportions p_i. The fit is perfect since

$$\pi_i = \Phi(\tau_i) - \Phi(\tau_{i-1}) = p_i \qquad (39)$$

LISREL applies a three-stage estimation method when the manifest variables are ordinal. At the first stage, the thresholds are estimated from the univariate marginal distribution. At the second stage, the correlations among the underlying continuous variables, also know as polychoric correlations, are estimated from the bivariate marginal distributions, for given thresholds. At the third stage, the parameters of the model are estimated by a weighted least squares method, where the weight matrix is an estimate of the asymptotic covariance matrix of the polychoric correlations. The first two steps are done in PRELIS while the third is executed in LISREL. We are not obliged to use weighted least squares at the third stage. We can as well use maximum likelihood or some other method. But by doing so, we assume that the matrix of the polychoric correlations is a sample correlation matrix from a multivariate normal distribution which is not a real assumption since the variables are in an ordinal scale and as a consequence discrete.

Now we will estimate the polychoric correlations from the bivariate margins. Suppose we have y_1^* and y_2^* which follow the standard normal distribution. This doesn't mean that we can assume that they follow the bivariate normal distribution. The polychoric correlation ρ is the correlation in the bivariate normal distribution of the underlying variables y_1^* and y_2^*. The polychoric correlation is robust to violations of the bivariate normal distribution (Qiuroga 1992).

The polychoric correlation can be estimated by maximizing the log-likelihood of the multinomial distribution,

$$\ln L = \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} n_{ij} * \log \pi_{ij}(\theta)$$

where θ is a parameter vector that contains the thresholds and the polychoric correlations

Where

$$\pi_{ij}(\theta) = P(y_1 = a_1, y_2 = a_2) = \int_{\tau^1_{a_1-1}}^{\tau^1_{a_1}} \int_{\tau^2_{a_2-1}}^{\tau^2_{a_2}} \phi_2(u,v) dudv \qquad (40)$$

Once the matrix of polychoric correlations has been calculated, we consider this matrix to be a correlation matrix from continuous data and we proceed to compute from this matrix the parameters of the model. It is wiser to use weighted least squares, where the weight matrix W is the inverse of the asymptotic covariance matrix of the polychoric correlations, instead of maximum likelihood because the latter method assumes that our data stem from a multivariate normal distribution, which is not a real assumption in our case. This matrix W is symmetric. Any positive definite matrix can give consistent parameter estimates but the asymptotic covariance matrix of the polychoric correlations gives correct standard errors and chi-squares in large samples. We should note that for the calculations of the asymptotic covariance matrix a large sample is needed. There is no specific rule about the sample size and its relation with the number of items but a rule of thumb that we should bear in mind is that if there are many zero cells or low frequencies in the bivariate contingency tables then our sample should be considered small. If this is the case, it is better to use maximum likelihood or some other technique. If this is not the case, the loadings are fitted to the estimated polychoric correlations by weighted least squares by minimizing the fit function

$$F(\theta)=\sum_{i=2}^{p}\sum_{j=1}^{i-1}\sum_{g=2}^{p}\sum_{h=1}^{g-1} w^{ij,gh}\left(\hat{\rho}_{ij}-\sum_{l=1}^{q}\lambda_{il}\lambda_{jl}\right)\left(\hat{\rho}_{gh}-\sum_{l=1}^{q}\lambda_{gl}\lambda_{hl}\right) \qquad (41)$$

where $w^{ij.gh}$ is an element of the inverse of W and λ_{il} is an element of the matrix Λ denoting the factor loading for the i item and the l latent variable.

3.4 LISREL MODEL FOR ORDINAL DATA WITH COVARIATES BOTH ON THE MANIFEST AND ON THE LATENT VARIABLES

Suppose that there are r covariates x that account together with the latent variables z for the correlations of the p ordinal observed items and there also k covariates w, whose effect on the latent variables we want to investigate. The covariates x have direct effects on the latent variables, whereas the covariates w have indirect effects through the latent variables z. So the latent variables z induce a spurious correlation between the observed items and the covariates w.

The measurement part of the model is changed from equation (31) to

$$y=\mu+\Lambda z+Bx+e \qquad (42)$$

The structural part of the model is

$$z=Vw+\omega \qquad (43)$$

We can apply two methods in LISREL to deal with such situations. The first method is to find either the polychoric correlations for the case where all variables (manifests & covariates) are on an ordinal scale or the polyserial correlation among the covariates that are continuous with the variables (manifests & covariates) that are on an ordinal scale and the Pearson correlation among the continuous covariates.

Then we get a correlation matrix with dimensions (p+k+r)x(p+k+r) and we fit the model to this matrix to get the factor loadings and the coefficients.

The other method that we can apply in LISREL to deal with situations where we have ordinal variables with covariates is presented by K.G.Jöreskog in the article "Analysis of Ordinal Variables 5: Covariates" (2002) that can be found in the following web address www.ssicentral.com/LISREL/corner.htm. We will give a short outline of Jöreskog's method.

Let us denote by b all the covariates, both those that have direct effects on the observed variables (x), and those that affect the latent variables (w). So b=(x,w). We should formulate a model that explains the relationship between the underlying variables y* and the covariates b. Consider the regression of y_i^* on b. $y_i^* = \alpha_i + \gamma_i' b + e_i$. e_i is assumed to follow the normal distribution with mean 0 and variance ψ_i^2. We can fit either the probit model that uses the normal distribution or the logit model that makes use of the logistic distribution. Let's suppose that

$$y_i^* / b \sim N(\alpha_i + \gamma_i' b, \psi_i^2) \tag{44}$$

In the model without the covariates we assumed that each of the underlying continuous variables followed the standard normal distribution. Now we consider that $y_i^* / b \sim N(\alpha_i + \gamma_i' b, \psi_i^2)$and the probability of a response in category k is defined as

$$P(y_i = k / b) = \pi_k(b) - \pi_{k-1}(b) = \Phi\left(\frac{\tau_{ik} - \alpha_i - \gamma_i' b}{\psi_i}\right) - \Phi\left(\frac{\tau_{i,k-1} - \alpha_i - \gamma_i' b}{\psi_i}\right) \tag{45}$$

The likelihood of the sample is $L = \prod_{n=1}^{N} \left\{ \prod_{a=1}^{m} \left[\pi_{na}(b_n)^{k_{ia}} \right] p(b_n) \right\}$ where p(b) is the density function of b which is unspecified and does not contain any parameters of interest and $k_{ia} = 1$ if $y_i = a$ and $k_{ia} = 0$ otherwise. The parameter vector γ as well as the parameters a, τ and ψ are estimated by maximizing the above likelihood.

Now suppose we have two underlying variables y_g^* and y_h^* with equations $y_g^* = a_g + \gamma_g' b + e_g$ and $y_h^* = a_h + \gamma_h' b + e_h$ respectively. e_g and e_h are assumed to follow the bivariate normal distribution with mean 0 and covariance matrix $\begin{pmatrix} \psi_g^2 & \\ \psi_{gh} & \psi_h^2 \end{pmatrix}$ then the probability that an individual i answers k on y_g and l on y_h is

$$\pi_{igh,kl} = \int_{\tau^*_{ig,k-1}}^{\tau^*_{ig,k}} \int_{\tau^*_{ih,l-1}}^{\tau^*_{ih,l}} \phi^{(2)}(u, v, \rho_{gh}) du dv \tag{46}$$

Where $\tau^*_{ig,k} = \dfrac{\tau_{gk} - a_g - \gamma_g b_i}{\psi_g}$ and $\phi^{(2)}(u,v,\rho)$ is the density function of the standardized bivariate normal distribution with correlation ρ.

The likelihood function is $L = \prod_{n=1}^{N} \prod_{k=1}^{m_g} \prod_{l=1}^{m_h} \pi_{ngh,kl}^{k_{ngh,kl}} p(b_n)$ where $k_{ngh,kl} = 1$ if the individual n responds in category k on y_g and l on y_h and $k_{ngh,kl} = 0$ otherwise. By maximizing the above likelihood we estimate ρ_{gh}.

Now if we consider a vector y* (px1), it is assumed that y*/b~N(α+Γb,Ψ). The rows of α and Γ and the diagonal estimates of Ψ are estimated from the univariate margins as we show below, whereas the off-diagonal estimates of Ψ are estimated from the bivariate margins.

From the above formulation of the model it is evident that the estimated conditional covariance matrix of y* for given b is Ψ. The estimated unconditional covariance matrix of y* is

$$\hat{\Gamma} VAR(b) \hat{\Gamma}' + \hat{\Psi} \tag{47}$$

and the covariance of y* and b is: cov(y*,b) = cov(α+Γb,b) = cov(Γb,b) = cov(b,b)Γ′=VAR(b)Γ′.

Hence, the estimated joint unconditional covariance matrix is

$$\hat{\Sigma} = \begin{pmatrix} \hat{\Gamma} VAR(b) \hat{\Gamma}' + \hat{\Psi} & \\ VAR(b) \hat{\Gamma}' & VAR(b) \end{pmatrix} \tag{48}$$

The above matrix can be used for modeling in LISREL just as sample covariance matrix for continuous variables.

CHAPTER 4

EQUIVALENCE OF THE ITEM RESPONSE THEORY AND THE UNDERLYING VARIABLE APPROACH

Let's denote by $y=(y_1,y_2,\ldots,y_p)$ the p ordinal observed variables. As we have seen the IRT model is given by the formula

$$link[\gamma_{is}(z,x)]=\tau_{is}-\sum_{j=1}^{q}a_{ij}z_j+\sum_{l=1}^{r}\beta_{il}x_l$$
$$i=1,2,\ldots,p$$
$$s=1,\ldots,c_i$$

where $\gamma_{is}(z,x)$ is the cumulative probability of a response in category s or lower of item y_i for given z and x written as $\gamma_{is}(z,x)=\pi_{i1}(z,x)+\pi_{i2}(z,x)+\ldots+\pi_{is}(z,x)$.

The Underlying variable approach model is given by the formula $y^*=\mu+\Lambda z+\mathrm{B}x+\varepsilon$ where Λ is a pxq matrix of factor loadings λ_{ij}, z is a qx1 matrix of the latent variables $(z_1,z_2,\ldots,z_q)$, B is a pxl matrix of the parameters of the covariates b_{il}, x is a lx1 matrix of the covariates $(x_1,x_2,\ldots,x_l)$ and ϵ is a px1 vector of measurement errors that is uncorrelated with z and x.

A necessary condition for the two approaches to be equivalent is that their joint probability distribution must be the same.

In the IRT approach because of the axiom of local or conditional independence the joint probability distribution of the ordinal observed variables is:

$$P\left(\bigcap_{i=1}^{p}(y_i=a_i)\right)=\int_{-\infty}^{+\infty}\ldots\int_{-\infty}^{+\infty}\prod_{i=1}^{p}P(y_i=a_i/z,x)h(z/w)dz \qquad (49)$$

In the underlying variable approach, we know that the measurements errors are independent with each other and uncorrelated with z and x. So for given z and x the y*'s must be independent. We have that $P(y_i\le a_i)=P(y_i^*\le\tau_{a_i})$. So

$$P\left(\bigcap_{i=1}^{p}(y_i = a_i)\right) = \int_{-\infty}^{+\infty}\dots\int_{-\infty}^{+\infty}\prod_{i=1}^{p} P\left(\tau_{a_i-1} \le y_i^{*} \le \tau_{a_i} / z, x\right)h(z / w)dz \qquad (50)$$

So the two approaches lead to similar joint probability distributions. We have that in the IRT approach

$$P(y_i \le a_{is}) = \Phi\left(\tau_{is} - \sum_{j=1}^{q} a_{ij} z_j + \sum_{l=1}^{r} b_{il} x_l\right) \qquad (51)$$

In the structural equation modeling approach we have assumed that the measurement errors $\varepsilon_i \sim N(0,\psi_i) \Leftrightarrow \frac{\varepsilon_i}{\sqrt{\psi_i}} \sim N(0,1)$.

In the SEM framework we have that

$$P(y_i \le a\ ) = P(y_i^{*} \le \tau_a\ {*}) = P\left(\mu_i + \sum_{j=1}^{q}\lambda_{ij} z_j + \sum_{l=1}^{r}\beta_{il} x_l + \varepsilon_i \le \tau_a {*}\right)$$

$$= P\left(\frac{\varepsilon_i}{\sqrt{\psi_i}} \le \frac{\tau_a {*} - \mu_i - \sum_{j=1}^{q}\lambda_{ij} z_j - \sum_{l=1}^{r}\beta_{il} x_l}{\sqrt{\psi_i}}\right) = R\left(\frac{\tau_a {*} - \mu_i - \sum_{j=1}^{q}\lambda_{ij} z_j - \sum_{l=1}^{r}\beta_{il} x_l}{\sqrt{\psi_i}}\right) \qquad (52)$$

R is the standard normal distribution or it can be any other distribution for $\frac{\varepsilon_i}{\sqrt{\psi_i}}$. In order for the two approaches (51) and (52) to be equivalent we must have $\Phi \equiv \Psi$, and also we must have

$$\tau_{is} = \frac{\tau_i^{*} - \mu_i}{\sqrt{\psi_i}},\ a_{ij} = \frac{\lambda_{ij}}{\sqrt{\psi_i}} \text{ and } b_{il} = -\frac{\beta_{il}}{\sqrt{\psi_i}} \qquad (53)$$

We should note that we have assumed that the variance of y* in the SEM approach is one. Hence

$$\psi_i = \operatorname{var}(\varepsilon) = 1 - \sum_{j=1}^{q}\lambda_{ij}^2 \operatorname{var}(z_j) - \sum_{l=1}^{r} b_{il}^2 \operatorname{var}(x_l) \qquad (54)$$

if there weren't any covariates affecting the latent variable then the variance of the latter would be one and as a consequence equation (52) would be:

$$\psi_i = \operatorname{var}(\varepsilon) = 1 - \sum_{j=1}^{q}\lambda_{ij}^2 - \sum_{l=1}^{r} b_{il}^2 \operatorname{var}(x_l) \qquad (55)$$

Now we know the equivalence between the parameters in the two approaches and we will see how we can standardize the parameters of the IRT in such a way that we can compare them to the parameters from the underlying variable approach. We assume that there are covariates that affect the latent variable. The covariance between y* and z is $\text{cov}(y^*,z)=\text{cov}(\mu+\lambda z+bx+\epsilon,z)=\lambda\text{cov}(z,z)=\lambda\text{var}(z)$ and the corresponding correlation is

$$\frac{\text{cov}(y^*,z)}{\sqrt{\text{var}(y^*)\text{var}(z)}}=\frac{\lambda\,\text{var}(z)}{\sqrt{\left(\lambda^2\,\text{var}(z)+b^2\,\text{var}(x)+\psi\right)\text{var}(z)}} \quad (56)$$

and by making use of the equivalence equations (53) we get

$$\frac{\alpha\sqrt{\psi}\sqrt{\text{var}(z)}}{\sqrt{\left(a^2\psi\,\text{var}(z)+\beta^2\psi\,\text{var}(x)+\psi\right)}}=\frac{a\sqrt{c^2\,\text{var}(w)}}{\sqrt{\left(1+\alpha^2c^2\,\text{var}(w)+\beta^2\,\text{var}(x)\right)}} \quad (57)$$

Similarly, in order to standardize the regression parameters β we find the correlation of y* with x. $\text{Cov}(y^*,x)=\text{cov}(\mu+\lambda z+bx+\epsilon,x)=b\text{cov}(x,x)=b\text{var}(x)$ and the corresponding correlation is :

$$\begin{aligned}\frac{\text{cov}(y^*,x)}{\sqrt{\text{var}(y^*)\text{var}(x)}}&=\frac{b\,\text{var}(x)}{\sqrt{\left(\lambda^2\,\text{var}(z)+b^2\,\text{var}(x)+\psi\right)\text{var}(x)}}\\&=\frac{\beta\sqrt{\psi}\sqrt{\text{var}(x)}}{\sqrt{\left(a^2\psi\,\text{var}(z)+\beta^2\psi\,\text{var}(x)+\psi\right)}}=\frac{\beta\sqrt{\text{var}(x)}}{\sqrt{\left(1+a^2c^2\,\text{var}(w)+\beta^2\,\text{var}(x)\right)}}\end{aligned} \quad (58)$$

Formulas (57) and (58) are referring to vectors .If we use subscripts to get the explicit standardization for α_{ij} and β_{il} equation (57) can be written in the following form

$$\frac{a_{ij}\sqrt{\sum_{h=1}^{k}c_{jh}^2\,\text{var}(w_h)}}{\sqrt{\left(1+\sum_{j=1}^{q}\alpha_{ij}^2\sum_{h=1}^{k}c_{jh}^2\,\text{var}(w_h)+\sum_{l=1}^{r}\beta_{il}^2\,\text{var}(x_l)\right)}} \quad (59)$$

And equation (58) becomes

$$\frac{\beta_{il}\sqrt{\text{var}(x_l)}}{\sqrt{\left(1+\sum_{j=1}^{q}\alpha_{ij}^2\sum_{h=1}^{k}c_{jh}^2\,\text{var}(w_h)+\sum_{l=1}^{r}\beta_{il}^2\,\text{var}(x_l)\right)}} \quad (60)$$

CHAPTER 5

GOODNESS-OF-FIT

5.1 INTRODUCTION

In this chapter we will show some methods for testing the goodness-of-fit of the models presented in chapters 2 and 3. In the subsequent chapters we will run two examples sing both models (IRT and LISREL) and we will infer our conclusions regarding which model fits best the data using the goodness-of-fit measures and the model-selection criteria presented in this chapter.

5.2 OVERALL GOODNESS-OF -FIT

Suppose we have $y_1, y_2, \ldots y_p$ observed ordinal variables with m_i denoting the number of response categories of variable i. There are $\prod_{i=1}^{p} m_i$ possible response patterns. If the sample size N is large compared to the number of likely response patterns, it is very possible that there will be much occurrences in every response pattern and its expected frequency will be large enough so that we will be able to carry out a valid chi-squared or log-likelihood ratio test to compare the observed and the expected frequencies and to check how good the model fits. The chi-squared test statistic, also known as the goodness-of-fit(GF) test statistic is given by the formula

$$X_{GF}^2 = \sum_{r=1} \left[\frac{\left(n_r - N\hat{\pi}_r\right)^2}{N\hat{\pi}_r} \right] = N\sum_{r=1} \frac{\left(p_r - \hat{\pi}_r\right)^2}{\hat{\pi}_r} \tag{61}$$

Where r denotes a response pattern and n_r the size of a response pattern r.

The likelihood ratio test statistic is given by the formula:

$$X_{LR}^2 = 2\sum_{r=1} n_r \ln\left(\frac{p_r}{\hat{\pi}_r}\right) = 2N\sum_{r=1} p_r \ln\left(\frac{p_r}{\hat{\pi}_r}\right) = 2NF\left(\hat{\theta}\right) \quad (62)$$

Hence, the likelihood ratio test statistic is 2N times the minimum value of the fit function

$$F\left(\hat{\theta}\right) = \sum_{r=1} p_r \ln\left(\frac{p_r}{\hat{\pi}_r(\theta)}\right) \quad (63)$$

If the model holds, both statistics have the same asymptotic distribution under H_0 and this is X^2 with $\sum_{i=1}^{p} m_i - p + pk - \frac{k(k-1)}{2}$ degrees of freedom. Usually we compute these statistics for every response pattern to check how much each response pattern contributes to the overall test. The LR test statistic has some advantages over the GF. The LR can be either positive or negative. A positive contribution means that the model underestimates the observed proportions, whereas a negative contribution means that the model overestimates the observed proportions. This does not happen with the GF where all the contributions are positive. This is obvious from its formula where the difference between the observed and the expected frequencies is squared. Also the GF derives most of its value from patterns with small probabilities, whereas the LR derives most of its value from response patterns with large probabilities. So the fit is not influenced so much by response patterns that occur rarely. On the other hand, the LR-statistic is not defined for a response pattern that does not occur. Generally if the model holds the two statistics will have similar values, whereas when the model does not hold they can be quite different. Both tests are not valid when the number N of individuals in the sample is not large, when the expected frequencies of some response patterns are small and when the contingency table of some pair of items is sparse. Joreskog and Moustaki(2001) consider two alternative ways of reducing the distorting effects of chi-square.

Alternative 1

Suppose we calculate the sum of the LR and the GF contributions over those response patterns whose expected frequency exceeds a value ν and use these as test statistics. The expected frequencies are multiplied by a constant such that their sum equals the sum of observed frequencies. These are not real chi-square statistics since the model has not been fitted on the subset of observations.

Alternative 2

Combine categories in such a way that all retained categories have expected frequencies exceeding a value ν and proceed to calculate GF and LR from this reduced set of categories. Although the distorting effects of chi-square have been reduced and the response patterns that cause the most lack of fit have been revealed, we have no clue about what should be done to improve the fit of the model. Two ways of improving the fit are:

1) Reduce the number of categories.
2) Eliminate the most offending variables thereby obtaining more homogeneity for the retained variables.

The fit of the latent trait model for binary response models is treated in Bartholomew and Tzamourani(1999). The main conclusion of this work is that the best way to make a global test of fit is to generate the empirical sampling distribution of the statistic (either GF or LR) using the parametric bootstrap method. This proceeds as follows:

1) Fit the desired model
2) Generate a random sample of the same size from the population in which the parameter values are equal to those estimated from the actual sample.
3) Fit the model in each case and compute the chosen test of fit.
4) Compare the actual value of the statistic with the bootstrap sampling distribution.

As we have seen we need a large sample for the chi-square statistics to be valid. In the polytomous latent trait models it is very likely that there will be numerous response patterns and some of them won't happen or they would have low occurrences. Sparseness in contingency tables distorts the chi-square statistics. Hence, it is possible that chi-square statistics won't give us any information about how well the model fits the data. We should also note that chi-square statistics always decrease as we add parameters in the model. So a two-factor model will have smaller chi-square than a one-factor model. By adding parameters we get a better fit. Overfitting is considered as a more serious problem than underfitting. Jöreskog and Rayment (1996) claim that if your model overfits, it means that you are "capitalizing on chance" and some of your factor loadings may not have a real meaning. They propose a method for understanding when nonsense parameters are considered in the model. You start with a single model with one factor. You compute the chi-square and the degrees of freedom d. Then you proceed in more complicated models (for example two factors). So you have computed X_k^2 and d_k for a number of different factors k=1,2,3... Then you start from the simplest case where k=1 and you compute the difference $X_k^2 - X_{k+1}^2$. This difference follows the chi-square distribution with $d_k - d_{k+1}$ degrees of freedom. If this difference is significant we continue to compute the same difference for larger k until it is not significant. This would be the proper value for k.

5.3 MODEL SELECTION CRITERIA

Instead of computing chi-square statistics one can use a criterion that not only take the value of the likelihood at the maximum likelihood solution but also the number of parameters estimated. One such criterion is AIC given by the formula:

$$AIC = -2\left[\log l\left(\hat{a}\right)\right] + 2m \qquad (64)$$

where $l(\hat{a})$ is the value of the likelihood at the maximum likelihood solution and m is the number of model parameters. The smaller the AIC the better the model. AIC tries to prevent overfitting by assigning a cost to the introduction of each additional parameter. This is a model-selection criterion that shows which model gives the best fit. It doesn't give any information about how well the fitted model fits the data.

5.3 RESIDUAL ANALYSIS

The computation of the GF-statistic and the LR-statistic involves the estimation of the probability of each response pattern. Another way to check if the model fits the data is by examining the GF-statistic or the LR-statistic for pairs of responses. For every possible combination of categories for every pair of items one has to compute the predicted probabilities $\hat{\pi}_{ij}$.

Under the assumption of conditional independence, the conditional probability of a pair of responses in two items, for given z and without any covariates w that affect the latent variable, is

$$P\left(y_i = k \bigcap y_j = l / z, x\right) = \int\int P\left(y_i = k / z, x\right) h(z / w) P\left(y_j = l / z, x\right) h(z / w) dz$$

We have that $P\left(y_i \leq k / z, x\right) = \Phi\left(\tau_{ik} - \sum_{j=1}^{q} a_{ij} z_j + \sum_{l=1}^{r} \beta_{il} x_l\right)$. where h(z/w) is the density function of z, conditional on w, and Φ is the response function, e.g. the probit or the logit. The latent variables are assumed to be independent with standard normal distributions, so that $h(z) = \prod_{j=1}^{q} \phi\left(z_j\right)$. The above integral is not in a closed form. The only thing we know about the z is its probability distribution that we have assumed to be the standard normal. We can approximate the integral using Gauss-Hermite quadrature.

So

$$\begin{aligned} P\left(y_i = k \bigcap y_j = l / z, x\right) &= \int \ldots \int P\left(y_i = k / z, x\right) P\left(y_j = l / z, x\right) h\left(z_1 / w\right) \ldots h\left(z_q / w\right) dz_1 \ldots dz_q \\ &= \sum_{t_1}^{v_1} \ldots \sum_{t_q}^{v_q} P\left(y_i = k / z, x\right) P\left(y_j = l / z, x\right) h\left(z_1 / w\right) \ldots h\left(z_q / w\right) \end{aligned} \tag{65}$$

The points for the integral approximations are the Gauss-Hermite quadrature points given in Straud and Sechrest(1966). In this way we can compute the probability of pairs of responses for all pairs of items or even for triplets of responses or for the whole response pattern. Since we have computed the probabilities of every pair of items we continue to calculate the

LR chi-square statistic or the GF chi-square statistic. The smaller they are the better the model. If a chi-square residual or a LR residual is large, for example bigger than 4, then probably the model does not fit well this pair of responses.

CHAPTER 6

FIRST EXAMPLE

6.1 INTRODUCTION

The first data set consists of five ordinal variables $(y_1, y_2, y_3, y_4, y_5)$ given below.

On the whole do you think it should or not be the government's responsibility to...

- provide a job for everyone who wants one. [JobEvery]
- keep prices under control [PriCon]
- provide a decent standard of living for the unemployed [LivUnem]
- reduce income differences between the rich and the poor [IncDiff]
- provide decent housing for those who can't afford it

The response alternatives given to the respondents are: definitely should be, probably should be, probably not be and definitely should not be. Missing values seems not to be a serious problem in this example. Item nonresponse varies from 2% to 6% among the items and we proceed by doing listwise deletion, meaning that all missing values will be excluded from the analysis. After we excluded the missing values we were left with 822 respondents.

A covariate x that is constructed to measure left to right political identification is used, after it has been standardized, as a continuous explanatory variable for the manifest ordinal variables. This variable is available in the 1996 British Social Attitudes (BSA) survey and it has been constructed from a set of five items that are related to redistribution and equality. It is usually used for distinguishing party identification (Heath, Jowell, Curtice and Witherspoon 1986). The covariate is a latent variable constructed from the same sample in a different survey. So when we have a model with one factor and the covariate, we could consider it as a two-factor model.

6.1 PRELIS STEP

We start the analysis by fitting the measurement model with no direct effects and we investigate whether the five items measure a unidimensional latent trait. Each item has four categories, thus there are $4^5 = 1024$ possible response patterns. Since we have only 822 respondents it is reasonable to presume that not every response pattern can be present in the data. In fact we notice that there are only 252 different response patterns. The most common response pattern is that of answering definitely should be on all five items occurring in 88 cases, followed by that of answering probably should be, which occurred in 41 cases. We should note that 131 out of the 252 different response patterns were observed only once and that 51 response patterns only twice. Probably this is a clue that our sample is small. The 15 most common response patterns are given in table 1 and it is interesting to observe that there is no response 'definitely should not be' in any of them. It is possible that this form of questions may deter respondents from taking a very negative stand.

Table 1:Example 1: 15 most common response patterns.

88	1	1	1	1	1
41	2	2	2	2	2
23	2	1	2	1	2
22	2	2	2	1	2
22	3	2	2	2	2
19	2	2	2	3	2
18	1	1	2	1	1
17	3	2	2	3	2
15	2	1	2	2	2
14	1	1	2	2	2
13	2	1	1	1	1
11	2	2	1	1	1
11	2	1	2	1	1
10	1	2	1	1	1
10	2	2	2	2	1

The percentage for each category of every item are given in table 2.

Table 2: Example 1: Frequency distributions for the ordinal observed items.

	JobEvery	PriCon	LivUnem	IncDiff	Housing
Definitely should be	30	43.3	29.3	36.4	37.6
Probably should be	38.8	41.7	49	31.8	50.9
Probably not be	19.3	10.2	15.1	21.5	9.2
Definitely should not be	11.8	4.7	6.6	10.3	2.3

We see that the bulk of the answers are in the first two categories. Especially for the PriCon and the Housing item. This is an indication that we can combine the last two categories either for these two items or even for all items and perform the analysis again.

With 5 variables there are $\frac{5!}{2!*(5-2)!} = 10$ pairs of variables. The contingency tables for every pair of variables are given in table 3:

Table 3: Example 1: Bivariate distributions for the ordinal observed items.

	PriCon				LivUnem				IncDiff			
JobEvery	AS	A	D	DS	AS	A	D	DS	AS	A	D	DS
AS	187	48	9	3	148	72	16	11	158	67	13	9
A	118	175	22	4	65	201	44	9	108	123	76	12
D	30	88	29	12	18	96	35	10	21	56	56	26
DS	21	32	24	20	10	34	29	24	12	15	32	38

	LivUnem				IncDiff				Housing			
PriCon	AS	A	D	DS	AS	A	D	DS	AS	A	D	DS
AS	162	139	35	20	199	98	43	16	190	139	23	4
A	56	209	60	18	77	137	90	39	87	216	35	5
D	13	39	21	11	14	17	38	15	20	44	15	5
DS	10	16	8	5	9	9	6	15	12	19	3	5

JobEvery	Housing AS	A	D	DS
AS	159	79	7	2
A	101	192	22	4
D	28	100	27	4
DS	21	47	20	9

LivUnem	IncDiff AS	A	D	DS	Housing AS	A	D	DS
AS	161	58	15	7	191	50	0	0
A	117	154	104	28	101	278	22	2
D	13	37	48	26	8	76	37	3
DS	8	12	10	24	9	14	17	14

IncDiff	Housing AS	A	D	DS
AS	195	97	5	2
A	71	172	15	3
D	31	109	36	1
DS	12	40	20	13

There are two zero cells in the table of LivUnem and Housing. The fit function is not defined if $p_{ij} = 0$. Prelis skips such zero cells in order to compute the matrix of the polychoric correlations but too many zero cells can be a problem. The fact that there are some zero cells in the bivariate contingency tables is another indication that our sample is small.

As we have seen the assumption of underlying bivariate normality is needed for the calculation of the polychoric coefficients. LR-chi-square is very sensitive to violations of the underlying bivariate normality and for this reason Jöreskog(2001) has developed a measure of population discrepancy for structural equation models named Root Mean Square Error of Approximation(RMSEA). He has found that there are no serious effects of non-normality unless RMSEA is larger than 0.1. The last column of the table 4 is the P-Value for the test of the hypothesis that the population value of

RMSEA is less than 0.1. None of the pairs of variables has RMSEA less than 0.1, hence there are no distortions due to non-normality.

Table 4: Example1: Correlations and test statistics.

				Test of Model			Test of Close Fit	
Variable	vs.	Variable	Correlation	Chi-Squ.	D.F.	P-Value	RMSEA	P-Value
PriCon	vs.	JobEvery	0.558 (PC)	35.670	8	0.000	0.065	0.996
LivUnem	vs.	JobEvery	0.505 (PC)	51.062	8	0.000	0.081	0.923
LivUnem	vs.	PriCon	0.314 (PC)	40.389	8	0.000	0.070	0.988
IncDiff	vs.	JobEvery	0.573 (PC)	23.943	8	0.002	0.049	1.000
IncDiff	vs.	PriCon	0.451 (PC)	30.638	8	0.000	0.059	0.999
IncDiff	vs.	LivUnem	0.552 (PC)	22.298	8	0.004	0.047	1.000
Housing	vs.	JobEvery	0.462 (PC)	20.539	8	0.008	0.044	1.000
Housing	vs.	PriCon	0.328 (PC)	27.854	8	0.001	0.055	1.000
Housing	vs.	LivUnem	0.712 (PC)	51.742	8	0.000	0.082	0.916
Housing	vs.	IncDiff	0.566 (PC)	28.798	8	0.000	0.056	0.999

The matrix of polychoric correlation that is going to be used for further modeling in LISREL is given in table 5. All the correlations are highly significant.

Table 5:Example 1: Matrix of polychoric correlations.

	JobEvery	PriCon	LivUnem	IncDiff	Housing
JobEvery	1.000				
PriCon	0.558	1.000			
LivUnem	0.505	0.314	1.000		
IncDiff	0.573	0.451	0.552	1.000	
Housing	0.462	0.328	0.712	0.566	1.000

All the results of the first five tables, as well as the thresholds for the LISREL model (table 6) were computed from PRELIS using input 1.

PRELIS/LISREL input 1: Example 1:

Computing Polychoric Correlations and Asymptotic Covariance Matrix
Data Ninputvars = 5
Labels
JobEvery PriCon LivUnem IncDiff Housing PolIden
Rawdata = goverst.dat
Clabels JobEvery- Housing 1=AS 2=A 3=D 4=DS
Output BT MA=PM PM=goverst.PM AC=goverst.ACP

The estimated thresholds for the one-factor model with no direct effects under the LISREL & IRT are given in table 6.

Table 6: Example 1: Estimated thresholds for the one-factor model from the IRT and the LISREL model.

		IRT	*PRELIS*
Item	*Category*	*thresholds* $\left(\hat{t}_{is}\right)_2$	*thresholds* $\left(\hat{t}_{is}\right)$
JobEvery	*1*	*-1.26*	*-0.523*
	2	*1.24*	*0.492*
	3	*2.95*	*1.185*
PriCon	*1*	*-0.32*	*-0.169*
	2	*2.16*	*1.038*
	3	*3.57*	*1.67*
LivUnem	*1*	*-1.49*	*-0.544*
	2	*2.26*	*0.784*
	3	*4.28*	*1.509*
IncDiff	*1*	*-0.9*	*-0.348*
	2	*1.3*	*0.471*
	3	*3.44*	*1.262*
Housing	*1*	*-0.89*	*-0.316*
	2	*3.41*	*1.197*
	3	*5.81*	*1.993*

6.3 FITTING THE MODEL TO THE MATRIX OF POLYCHORIC CORRELATIONS

We cannot directly compare the factor loadings of the IRT model with those of LISREL because they are measured on a different scale. In order to make them comparable we have to standardize those of IRT. In order to standardize the loadings of IRT we find the correlation of the underlying continuous variable with the latent variable and we use the equivalence equations from equation (53). If we consider that there are no direct effects x and there are no covariates w that affect the latent variable and as a consequence the variance of the latent variable is unity, then equation (56) changes to $\frac{\operatorname{cov}(y^*, z)}{\sqrt{\operatorname{var}(y^*)\operatorname{var}(z)}} = \frac{\lambda}{\sqrt{\lambda^2 + \psi}}$. So equation (59) becomes $\frac{a_{ij}}{\sqrt{1 + \sum_{j=1}^{q} a_{ij}^2}}$ and we use this equation to standardize the loadings. This equation gives the correlation between y* and z under the equivalence of IRT and SEM. The standard errors are not comparable too and we cannot make them comparable by using the same way as for the coefficients.

In the structural equation modeling approach we estimate the loadings by using two different methods. Namely weighted least squares and maximum likelihood. Although the loadings in the second way are computed by maximum likelihood, the standard errors and the chi-squares are computed by weighted least squares using the asymptotic covariance matrix, hence, we get the correct chi-squares and standard errors in large samples. The LISREL input 2 is used to fit the model to the matrix of polychoric correlation estimated in PRELIS input 1.

PRELIS/LISREL input 2: Example 1:

Testing Measurement Model
Observed Variables: JobEvery PriCon LivUnem IncDiff Housing
Correlation Matrix from File goverst.PM
Asymptotic Covariance Matrix from File goverst.ACP
Sample Size: 822
Latent Variable: z

Relationships:

JobEvery PriCon LivUnem IncDiff Housing = z

Path Diagram

End of Problem

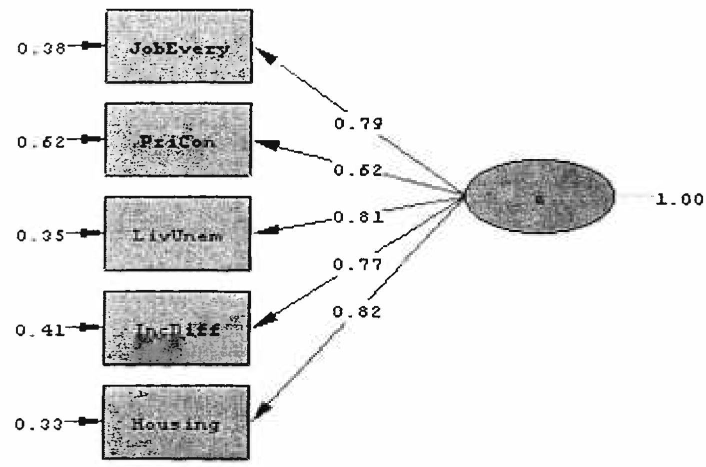
The loadings of LISREL, IRT and the standardized loadings of IRT as well as their standard errors are given in table 7.

Table 7: example 1: Estimated loadings for the one-factor model from the IRT and the LISREL model.

	IRT			*LISREL*			
				WLS		ML	
Item	unstandardized loadings α	s.e	standardized loadings stα	loadings λ	s.e	loadings λ	s.e
JobEvery	1.78	0.13	0.87	0.79	0.024	0.69	0.029
PriCon	1.18	0.11	0.76	0.62	0.034	0.52	0.039
LivUnem	2.24	0.16	0.91	0.81	0.025	0.79	0.027
IncDiff	2.1	0.16	0.9	0.77	0.024	0.75	0.027
Housing	2.31	0.21	0.92	0.82	0.024	0.78	0.026

We see from table 7 that the loadings estimated from the item response theory approach using a logit function, as a link function, are in all cases larger that those of the underlying response variable approach. The two approaches use different distributional assumptions. In the IRT case we make use of the logistic distribution. If we accept that the results for the IRT model using a logit function are 1.79 times the results of the IRT model using a probit function, and as a result of this assumption divide all IRT loadings with 1.79 and then standardize them using equation (59), we will see that the new loadings are smaller than the corresponding LISREL models. In the LISREL model there are larger differences among the loadings that in the IRT case. It is interesting to see that in LISREL, the results we get when we use weighted least square are different from the results of maximum likelihood. The path diagram for the LISREL model with weighted least squares is given in figure 1:

Figure 1: Example 1: Path diagram for the LISREL model with one latent variable when we use weighted least squares.



6.3.1 GOODNESS-OF-FIT OF THE MODEL

One way to check the fit of the model is by looking at the chi-squared residuals for the two-way margins. The GF cell contributions in row i and column j also known as chi-squared residual is $\frac{(o_{ij} - e_{ij})^2}{e_{ij}}$ where o_{ij} is the observed frequency for a pair of items which are given in Table 3 and e_{ij} is the expected frequency for a pair of items and is estimated using equation (64). The sum of all these contributions is the GF-statistic. We will use mainly the GF-statistic instead of the LR-statistic because the latter is not defined if the observed frequency for a bivariate margin is zero, something that occurs in our data and especially for different values of the covariate there are many zero observed bivariate frequencies that make the calculation of the LR chi-square statistic not possible. We could assume that every chi-squared residual follows the chi-square distribution with one degree of freedom and as a consequence any chi-Squared residual greater than four is an indication of poor fit. For the first pair of variables, namely JobEvery and PriCon the expected frequencies under the two approaches are given in tables 8 and 9 respectively. The expected frequencies come out if we multiply equation (65) by the sample size N. We can get the corresponding observed frequencies from Table 3.

Table 8: Example 1: Expected frequencies for the first pair of items for the one-factor LISREL(WLS) model

	PriCon			
JobEvery	1	2	3	4
1	164.94	72.59	7.72	1.60
2	136.43	145.13	28.60	9.19
3	41.40	80.77	25.2	11.6
4	13.04	44.79	22.42	16.59

Table 9: Example1 : Expected frequencies for the first pair of items for the one-factor IRT model

	PriCon			
JobEvery	1	2	3	4
1	158.75	76.77	10.88	3.99
2	137.07	139.95	28.76	11.7
3	44.99	76.89	23.22	10.98
4	18.76	45.93	20.61	12.76

The residuals for the first pair of variables for the LISREL and the IRT model are given in tables 10 and 11 respectively.

Table 10 : Example 1: Residuals for the first pair of items for the one-factor LISREL(WLS) model

	PriCon			
JobEvery	1	2	3	4
1	2.95	8.33	0.21	1.21
2	2.49	6.15	1.52	2.93
3	3.14	0.65	0.57	0.02
4	4.85	3.65	0.11	0.7

Table 11 : Example 1: Residuals for the first pair of items for the one-factor IRT model

	PriCon			
JobEvery	1	2	3	4
1	5.03	10.78	0.32	0.25
2	2.65	8.78	1.59	5.06
3	4.99	1.60	1.44	0.09
4	0.27	4.22	0.56	4.11

As we see using the LISREL approach there are 3 Chi-Squared residuals greater than four, namely the residuals for pairs (1,2),(2,2) and (4,1). The sum of the GF cell contributions is 39.48. For the IRT approach there are 7 Chi-Squared residuals greater than four, namely the residuals for pairs (1,2), (1,2), (2,2), (2,4), (3,1), (4,2) and (4,4). The sum of the GF cell contributions is 51.75. We extend this analysis to the rest of the pairs and we see that there are chi-squared residuals exceeding four in all pairs of items and in most cases there are many. These are indications that the model does not fit. With the exception of the first pair of variables the GF is much larger for the LISREL model. Based on these findings it is obvious that none of the models give a good fit. In the case of the LISREL model we see that the model does not give a good fit for response categories 1 and 4 for all pairs of items. Although, using both approaches the results are not satisfactory, we see that the total GF contribution of the IRT model is less than half of that of the LISREL model. Every pair of variables has 16 possible combinations of response categories and if the GF contribution for a pair of items is larger than 16*4=64 then the fit is bad. We see from tables (12), (13) and (14) that in the IRT case this does not happen for any pair of variables, whereas, it happens for some pairs in the LISREL model. The average GF-contribution is less than four in the IRT model and if we had used this as a measure of goodness-of-fit we would have accepted the model. Table 12 shows the chi-squared residuals that exceed four for the one-factor IRT model, whereas tables 13 and 14 show the chi-squared residuals that exceed four for the one-factor LISREL model using weighted least squares and maximum likelihood respectively.

Table 12: Example 1: Chi-squared residuals greater than four, for two-way margins, from the IRT model with one latent variable and four response categories

item	2	3	4	5
1	*(1,2), (2,1), (2,2), (2,4),(3,1),(4,2),(4,4)*	*(1,2), (1,4)*	*(2,4)*	*(4,1)*
SUM(GF)	*51.75*	*56.24*	*18.71*	*29.52*
SUM(LR)	54.92	42.46	19.7	22.68
2		*(1,4), (4,1)*	*(3,3)*	*(1,3), (4,1)*
SUM(GF)		*50.69*	*24.28*	*31.83*
SUM(LR)		41.47	24.27	29.17
3			*(4,1), (4,3)*	*(1,2), (1,4), (2,2), (3,3), (4,1), (4,2)*
SUM(GF)			*21.19*	*59.06*
SUM(LR)			17.7	
4				*(3,4), (4,1)*
SUM(GF)				*24.45*
SUM(LR)				23.65
TOTAL(GF)=367.72				
TOTAL(LR)=276				

Table 13: Example 1: Chi-squared residuals greater than four, for two-way margins, from the LISREL model(WLS) with one latent variable and four response categories)

item	2	3	4	5
1	(1,2), (2,2), (4,1)	*(1,2), (1,3), (1,4), (4,1)*	*(1,4)*	*(1,3), (1,4), (4,1)*
SUM(GF)	39.48	200.14	36.06	144.59
SUM(LR)	40.79	78.45	25.69	65.68
2		*(1,4), (4,1)*	*(4,1)*	*(1,3), (1,4), (3,1), (4,1), (4,3)*
SUM(GF)		119.36	36.53	81.08
SUM(LR)		71.18	31.3	54.86
3			*(1,4), (4,1))*	*(1,2), (2,1), (2,2) (4,1)*
SUM(GF)			49.62	89.58
SUM(LR)			29.83	
4				*(1,4), (4,1)*
SUM(GF)				51.87
SUM(LR)				34.67
TOTAL(GF)=848.32				
TOTAL(LR)=432.46				

Table 14: Example 1: Chi-squared residuals greater than four, for two-way margins, from the LISREL model(ML) with one latent variable and four response categories)

item	2	3	4	5
1	(1,1), (1,2), (2,1), (3,1), (4,1), (4,3), (4,4)	*(1,1), (1,3), (1,4), (2,1), (4,1)*	*(1,4), (2,4), (4,1), (4,4)*	*(1,3), (1,4), (2,1), (4,1)*
SUM(GF)	94.64	157.75	55.75	125.36
SUM(LR)	87.17	86.41	47.83	66.21
2		*(1,2), (1,4), (2,1), (4,1)*	*(2,1), (3,3), (4,1), (4,4)*	*(1,3), (1,4), (3,1), (4,1)*
SUM(GF)		95.25	45.17	72.06
SUM(LR)		59.78	38.95	49.21
3			*(1,4), (4,1))*	*(1,1), (1,2), (2,1), (2,2), (3,3), (4,1), (4,2), (4,4)*
SUM(GF)			49	98.27
SUM(LR)			29.62	
4				*(1,4), (3,3), (4,1), (4,4)*
SUM(GF)				61.34
SUM(LR)				38.99
TOTAL(GF)=854.60695				
TOTAL(LR)=504.16495				

Although in Table 12 there are some large chi-squared residuals none of the GF contributions for any pair of items is larger than 64. Tables 13 and 14 show that Lisrel does not give such a good fit. We have less large chi-squared residuals when we use weighted least squares but the total GF almost the same as when we use maximum likelihood.

6.4 LATENT VARIABLE MODEL WITH ONE FACTOR, FOUR RESPONSE CATEGORIES AND ONE COVARIATE

Now we consider the one-factor model allowing for the political identification of the individuals to affect directly the 5 items. When we fit the model we consider that the covariate together with the latent variable accounts for the interrelationships among the items. In LISREL we can think of two ways to fit the model. The first way is to construct the correlation matrix among all 6 variables including the covariate. Since the observed variables are ordinal and the one covariate continuous, the polychoric correlation is estimated among the manifest ordinal variables and the polyserial correlation is estimated between the continuous covariate and each observed ordinal variable. This way doesn't allow us to use the asymptotic covariance matrix and we will fit the model by using maximum likelihood. If we try to use the asymptotic covariance matrix we will get the following warning "The number of variables in the matrix to be analyzed exceeds the number of variables used to compute the Asymtotic Covariance Matrix". The other way is to compute the joint unconditional covariance matrix and its' asymptotic covariance matrix as it is shown in the presentation of the LISREL model in section (3.3) and to fit the model using this covariance matrix.

The covariate takes numerous values, to check the fit of the model we take three of its values with many occurrences and we check how good the model predicts the observed frequencies of the bivariate margins for these values. We select the values such that the first one comes from the left tail of the distribution, the second from the middle and the third from the right tail. We select the values –1.239,–0.126 and 0.987 with frequencies 44,103 and 53 respectively.

6.4.1 IRT MODEL WITH DIRECT EFFECTS

Tables 15 and 16 give the thresholds and the factor loadings respectively of the one-factor IRT model with one covariate affecting the five indicators.

Table 15: Example 1: Estimated thresholds from the IRT model.

		IRT	
Item	*Category*	*thresholds* $\left(\hat{t}_{is}\right)$	*s.e.*
JobEvery	*1*	*-1.24*	*0.22*
	2	*1.18*	*0.12*
	3	*2.88*	*0.18*
PriCon	*1*	*-0.33*	*0.16*
	2	*2.11*	*0.11*
	3	*3.53*	*0.2*
LivUnem	*1*	*-1.64*	*0.3*
	2	*2.47*	*0.17*
	3	*4.65*	*0.29*
IncDiff	*1*	*-0.93*	*0.18*
	2	*1.26*	*0.14*
	3	*3.45*	*0.21*
Housing	*1*	*-0.99*	*0.28*
	2	*3.83*	*0.3*
	3	*6.51*	*0.62*

Table 16: Example 1: Estimated loadings and coefficients for the IRT model.

		IRT		
Item	α	s.e.	β	s.e.
JobEvery	1.11	0.1	-1.27	0.1
PriCon	0.68	0.09	-0.86	0.08
LivUnem	2.21	0.18	-1.26	0.12
IncDiff	1.34	0.12	-1.58	0.11
Housing	2.39	0.27	-1.34	0.17

The standardized loadings and the regression coefficients of the model using equations (59) and (60) respectively are given in Table 17.

Table 17: Example 1: Standardized loadings and coefficients from the IRT model.

item		
	standardized α	standardized β
JobEvery	0.57	0.65
PriCon	0.46	0.58
LivUnem	0.81	0.46
IncDiff	0.58	0.69
Housing	0.82	0.46

It is evident from Table 17 that the covariate has a bigger effect on the items 1, 2 and 4 where β is large and the corresponding loadings have decreased by 34%, 26% and 35% respectively.

Tables 18,19 and 20 give the chi-squared residuals for the three values of the covariate we have chosen a-priori.

Table 18: Example 1: Chi-squared residuals greater than four, for two-way margins, for the one-factor IRT model with covariates, when the covariate takes the value –1.239

item	*2*	*3*	*4*	*5*
1		*(1,4)*	*(1,4)*	
sum(GF)	10.53	41.28	13.12	6.96
2		*(1,4)*		
sum(GF)		19.43	7.07	3.16
3			*(4,2)*	*(4,2)*
sum(GF)			34.3	14.15
4				*(4,1)*
sum(GF)				12.41
TOTAL(GF)=162.42				

Table 19: Example 1: Chi-squared residuals greater than four, for two-way margins, for the one-factor IRT model with covariates, when the covariate takes the value –0.126

item	*2*	*3*	*4*	*5*
1	*(2,1), (2,2), (3,4), (4,4)*	*(1,2), (1,3)*		*(1,3), (1,4)*
sum(GF)	29.48	21.64	8.72	26.59
2		*(4,1)*	*(2,2), (4,1)*	*(4,1)*
sum(GF)		57.29	30.6	33.95
3			*(4,1)*	
sum(GF)			18.74	11.2
4				*(1,4)*
sum(GF)				28.67
TOTAL(GF)=266.88				

Table 20: Example 1: Chi-squared residuals greater than four, for two-way margins, for the one-factor IRT model with covariates, when the covariate takes the value 0.987

item	*2*	*3*	*4*	*5*
1		(4,1)		(4,1)
sum(GF)	15.38	12.55	13.16	26.54
2		(3,1)	(1,1)	(3,1)
sum(GF)		19.38	18.78	14.53
3				(4,1)
sum(GF)			11.92	11.82
4				(3,1)
sum(GF)				10.76
TOTAL(GF)=154.84				

It is evident from Tables 18,19 and 20 that the fit in the IRT model has improved in the three cases we examined. The covariate together with the latent variable gives a very good fit. The "problematic" chi-square residuals are only a few and the total GF has decreased in comparison with the one-factor model without the covariate in all three cases. The reduction in the total GF for the values −1.239,0.126 and 0.789 is 56%,27% and 58% respectively. We should note that although we give the chi-squared residuals and the total GF for three values of the covariate we have checked the fit for many other values of the covariate and they all give similar results. Also we should note that most of the "problematic" chi-squared residual involve the response categories 1 and 4.

6.4.2 LISREL MODEL WITH DIRECT EFFECTS WITH THE USE OF THE MATRIX OF POLYSERIAL AND POLYCHORIC CORRELATIONS

The PRELIS input 3 is used for the estimation of the matrix of polychoric and polyserial correlations.

PRELIS/LISREL input 3: Example 1:

Computing Correlation Matrix

Data Ninputvars = 6

Labels

JobEvery PriCon LivUnem IncDiff Housing PolIden

Rawdata = goverst.ls8

Clabels JobEvery-Housing 1=AS 2=A 3=D 4=DS

Output MA=PM PM=goverst.PM AC=goverst.ACC WP

Table 21 gives the correlation matrix that is going to be used for finding the parameters of the model with the covariate. This is a matrix of polychoric correlations with the exception of the last row where it gives the polyserial correlations of the covariate with the rest of the items. We will use this as a correlation matrix to fit the model by considering that the covariate PolIden together with the latent variable accounts for the interrelationships among the items.

Table 21: Example 1: Matrix of Polychoric and Polyserial correlations

	JobEvery	PriCon	LivUnem	IncDiff	Housing	PolIden
JobEvery	1.000					
PriCon	0.558	1.000				
LivUnem	0.505	0.314	1.000			
IncDiff	0.573	0.451	0.552	1.000		
Housing	0.462	0.328	0.712	0.566	1.000	
PolIden	0.531	0.412	0.399	0.582	0.411	1.000

As we see from Table 21 the covariate has significant correlations with all items and it is more correlated with items 1 and 4. So the introduction of the covariate in the model is expected to decrease the magnitude of all loadings and especially of the loadings of the items 1 and 4.

LISREL input 4 gives the code used in LISREL to fit the one-factor model with the covariate.

PRELIS/LISREL input 4: Example 1:

```
MIMIC Model
Observed Variables: JobEvery PriCon LivUnem IncDiff Housing PolIden
Covariance Matrix from File goverst.PM
!Asymptotic Covariance Matrix from File goverst.ACC
Sample Size: 822
Latent Variables: JOBEVERY PRICON LIVUNEM INCDIFF HOUSING
POLIDEN ETA
Relationships:
JobEvery=1*JOBEVERY
PriCon=1*PRICON
LivUnem=1*LIVUNEM
IncDiff=1*INCDIFF
Housing=1*HOUSING
PolIden=1*POLIDEN
Set the error variances of JobEvery-PolIden to 0
JOBEVERY=0.5*ETA POLIDEN
PRICON LIVUNEM INCDIFF HOUSING= ETA POLIDEN
Path Diagram
Set the correlations of  POLIDEN-ETA to 0
Method: Maximum Likelihood
End of Problem
```

The path diagram for the LISREL model is shown in figure 2. For the calculation of the model we had to construct six latent variables, identical to the observed variables, and we also had to set a value for a loading in order to identify the scale for the latent variable, which in turn identifies the scale for the item parameters, in our case we chose 0.5 for the first loading. Once we have found our estimates we standardize the latent variable and find the new estimates in order to make them comparable with the estimates from the IRT approach. All loadings and coefficients are significant.

Figure 2: Example 1: Path diagram for the LISREL model with covariate effects, using the matrix of polyserial and polychoric correlations.

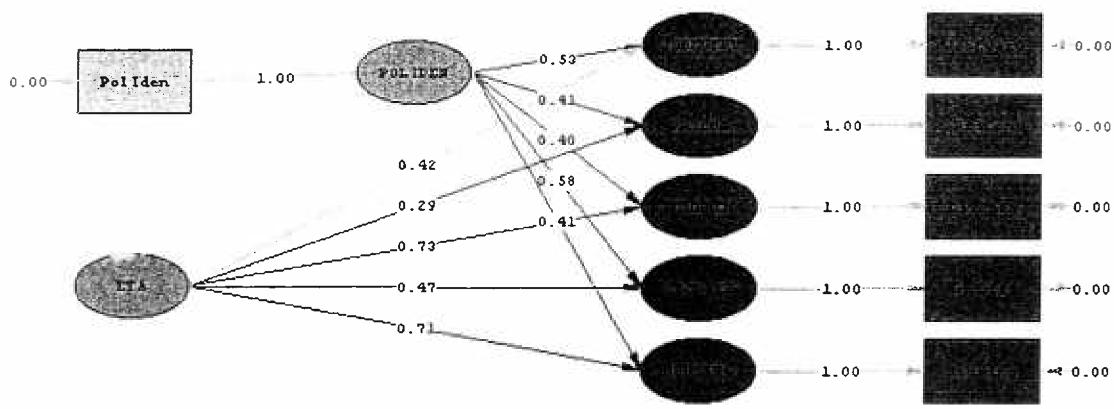


Table 22 gives the standardized loadings and coefficients of the IRT method and that of the LISREL method.

Table 22 : Example 1: Estimated loadings and coefficients of the one-factor model with covariates

item	*Loadings*		*Coefficients*	
	IRT α	LISREL λ	IRT β	LISREL b
JobEvery	0.57	0.42	0.65	0.53
PriCon	0.46	0.29	0.58	0.41
LivUnem	0.81	0.73	0.46	0.4
IncDiff	0.58	0.47	0.69	0.58
Housing	0.82	0.71	0.46	0.41

Both the loadings and the coefficients are larger in the IRT model. The different signs for the coefficients in Table 22 are caused by the different model assumptions and it is not he sign that we compare but the magnitude of the coefficients. Now we test the fit of the model for the same values of the covariate as in the IRT case.

Table 23: Example 1: Chi-squared residuals greater than four for two-way margins for the one-factor LISREL model with covariates, when the covariate takes the value –1.239

item	2	3	4	5
1	(1,1), (2,1), (2,2), (3,2), (3,3), (4,1)	(1,1), (1,2), (1,4) (2,2), (2,4), (3,1), (3,3), (4,2)	(1,1), (1,3), (1,4), (2,2), (2,3), (3,2), (4,1)	(1,1), (1,3), (2,1), (2,2), (3,2), (4,1)
sum(GF)	179.82	208.72	472.66	127.11
2		(1,4), (2,2), (3,3)	(1,1), (1,2), (1,3), (1,4), (2,2), (2,3), (3,2)	(1,1), (1,3), (2,2), (3,2)
sum(GF)		115.61	306.37	21.67
3			1,1), (1,2), (1,3), (2,2), (2,3), (2,4), (4,2)	(1,1), (4,2), (4,3)
sum(GF)			518.11	95.84
4				(1,1), (2,1), (2,2), (2,3), (3,1), (3,2), (4,1)
sum(GF)				736.59
TOTAL(GF)=2782.49				

Table 24: Example 1: Chi-squared residuals greater than four, for two-way margins, for the one-factor LISREL model with covariates, when the covariate takes the value –0.126

item	2	3	4	5
1	(2,1), (2,2), (3,4)	(1,3)		(1,3), (1,4), (3,3)
sum(GF)	38.41	14.19	9.54	28.89
2		(4,1)	(2,2), (4,1)	(1,2), (2,2), (4,1)
sum(GF)		75.25	36.98	44.03
3			(2,2), (4,1)	(4,1)
sum(GF)			21.04	17.94
4				(1,4), (2,2), (3,3)
sum(GF)				36.06
TOTAL(GF)=322.32				

Table 25: Example 1: Chi-squared residuals greater than four, for two-way margins, for the one-factor LISREL model with covariates, when the covariate takes the value 0.987

item	*2*	*3*	*4*	*5*
1	(1,1), (2,2), (4,2)	(1,1), (2,2), (3,2), (4,4)	(1,1), (2,1), (3,3)	(1,1), (2,1), (4,1), (4,3)
sum(GF)	45.46	34.36	119.58	39.08
2		(1,1), (2,2), (3,1)	(1,1), (2,2), (2,3), (2,4)	(1,1), (2,1), (3,1)
sum(GF)		38.2	137.52	34.8
3			(1,1), (2,1), (2,3), (4,4)	(1,1), (2,1), (4,1)
sum(GF)			69.23	40.53
4				(1,1), (1,2), (3,1), (4,2), (4,3)
sum(GF)				60.35
TOTAL(GF)=619.11				

With the exception of the value –0.126, where the fit has improved , in all other cases the fit has deteriorated considerably. For the value –0.126 as we see in table 24 there are only a few chi-squared residuals exceeding four and the total GF has decreased by 62%. For the value 0.987, although there are more chi-square residuals exceeding four than in the latent variable model with no covariates, the average GF-contribution has decreased by 27%.

6.4.3 LISREL MODEL WITH DIRECT EFFECTS WITH THE USE OF THE JOINT UNCONDITIONAL COVARIANCE MATRIX

Now we will fit the model using the joint unconditional covariance matrix as it was presented in section 3. PRELIS input 6 gives the conditional covariance matrix of the 5 items on the covariate and it also gives the joint unconditional covariance matrix.

PRELIS/LISREL input 6: Example 1:

```
Computing Covariance Matrix
Data Ninputvars = 6
Labels
JobEvery PriCon LivUnem IncDiff Housing PolIden
Rawdata = goverst.dat
Clabels JobEvery-Housing 1=AS 2=A 3=D 4=DS
Fixedvariables: PolIden
Output MA=CM CM=goverst.CM AC=goverst.ACC WP
```

Although, we have taken the covariate into account, we see that all correlations remain highly significant, as we see from the conditional covariance matrix (table 28), meaning that the covariate alone does not account for the correlations of the variables underlying the ordinal variables. Probably the introduction of a latent variable along with the covariate will account for the correlations among the observed ordinal variables.

Table 26 : Example 1 : Conditional Covariance Matrix of the items on the covariate PolIden

	JobEvery	PriCon	LivUnem	IncDiff	Housing
JobEvery	1.000				
PriCon	0.434 (0.037) 11.857	1.000			
LivUnem	0.375 (0.037) 10.012	0.169 (0.043) 3.964	1.000		
IncDiff	0.376 (0.038) 10.023	0.278 (0.041) 6.763	0.428 (0.036) 11.799	1.000	
Housing	0.314 (0.040) 7.776	0.184 (0.044) 4.189	0.657 (0.028) 23.190	0.441 (0.037) 11.839	1.000

The joint unconditional covariance matrix that is going to be used for finding the loadings and the coefficients in the second way is given in Table 27.

Table 27:Example 1: Unconditional covariance matrix

	JobEvery	PriCon	LivUnem	IncDiff	Housing	PolIden
JobEvery	1.388					
PriCon	0.714	1.202				
LivUnem	0.645	0.364	1.188			
IncDiff	0.818	0.598	0.736	1.504		
Housing	0.593	0.386	0.851	0.760	1.201	
PolIden	0.623	0.450	0.434	0.710	0.449	1.001

We see from table 27 that as in the matrix of polychoric and polyserial correlations (table 10) the covariate is more related to items 1 and 4

Now we will present the LISREL results that come out from using the joint unconditional covariance matrix as we showed in part 3. LISREL input 6 fits the model to the covariance matrix estimated in PRELIS (input 5)

PRELIS/LISREL input 6: Example 1:

MIMIC Model
Observed Variables: JobEvery PriCon LivUnem IncDiff Housing PolIden
Covariance Matrix from File goverst.CM
Asymptotic Covariance Matrix from File goverst.ACC
Sample Size: 822
Latent Variables: JOBEVERY PRICON LIVUNEM INCDIFF HOUSING POLIDEN ETA
Relationships:
JobEvery=1*JOBEVERY
PriCon=1*PRICON
LivUnem=1*LIVUNEM
IncDiff=1*INCDIFF
Housing=1*HOUSING
PolIden=1*POLIDEN
Set the error variances of JobEvery-PolIden to 0

JOBEVERY=0.5*ETA POLIDEN

PRICON LIVUNEM INCDIFF HOUSING= ETA POLIDEN

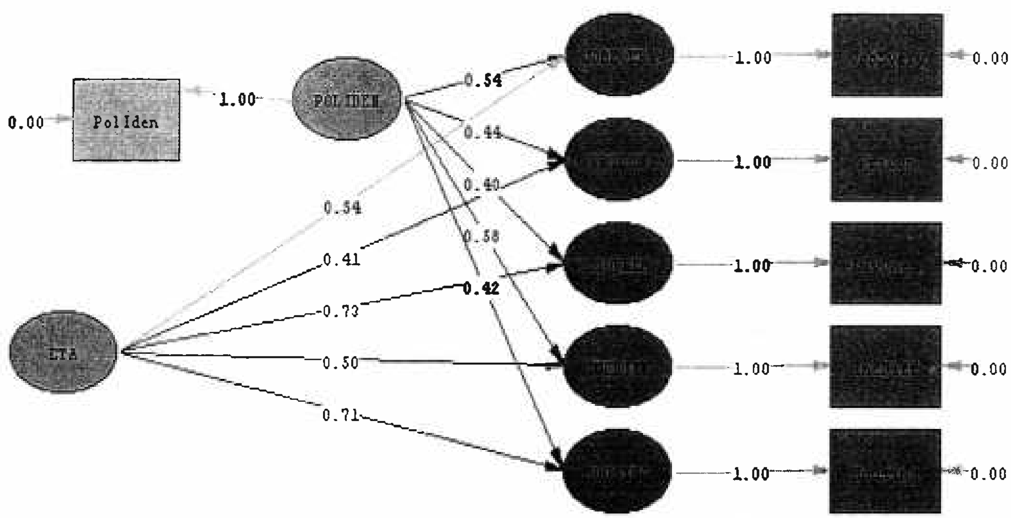
Path Diagram

Set the correlations of POLIDEN-ETA to 0

End of Problem

Using the joint unconditional covariance matrix and weighted least squares the results we get are shown in the path diagram in figure 3.

Figure 3: Example 1: Path diagram for the one-factor LISREL (WLS) model with covariates.



All the loadings and the coefficients are significant. We will find the "problematic" chi-squared residuals for the same values of the covariate as in section 6.3.2. Tables 28, 29 and 30 give these chi-squared residuals for the values –1.239, -0.126 and 0.987 respectively.

Table 28: Example 1: Chi-squared residuals greater than four, for two-way margins, for the one-factor LISREL model with covariates, when the covariate takes the value −1.239

item	*2*	*3*	*4*	*5*
1	(1,1), (2,1), (2,2), (3,2), (3,3), (4,1)	(1,1), (1,4), (2,2), (2,4),(3,1), (3,3), (4,2)	(1,1),(1,3), (1,4), (2,2),(2,3), (3,2), (4,1)	(1,1),(1,3), (2,1), (2,2), (3,1),(3,2), (4,1)
sum(GF)	157.56	265.71	539.85	214.3
2		(1,1), (2,4), (3,3)	(1,1), (1,2), (1,3), (1,4), (2,2), (2,3), (3,2)	(1,1), (1,3), (3,2)
sum(GF)		117.97	335.57	24.82
3			1,1), (1,2), (1,3), (2,2), (2,3), (2,4), (4,2),	(1,1), (4,2), (4,3)
sum(GF)			518.42	96.82
4				(1,1), (2,1), (2,2), (2,3), (3,1), (3,2), (4,1)
sum(GF)				870.34
TOTAL(GF)=3141.36				

Table 29: Example 1: Chi-squared residuals greater than four, for two-way margins, for the one-factor LISREL model with covariates, when the covariate takes the value −0.126

item	*2*	*3*	*4*	*5*
1	(2,1), (2,2), (3,4), (4,4)	(1,3), (1,4)		(1,3), (1,4)
sum(GF)	29.33	27.62	11.16	68.7
2		(4,1)	(2,2), (4,1), (4,2)	(1,4), (4,1)
sum(GF)		141.9	49.23	69.52
3			(2,2), (4,1)	(4,1)
sum(GF)			22.22	18.34
4				(1,4), (2,2), (3,3)
sum(GF)				40.5
TOTAL(GF)=377.55				

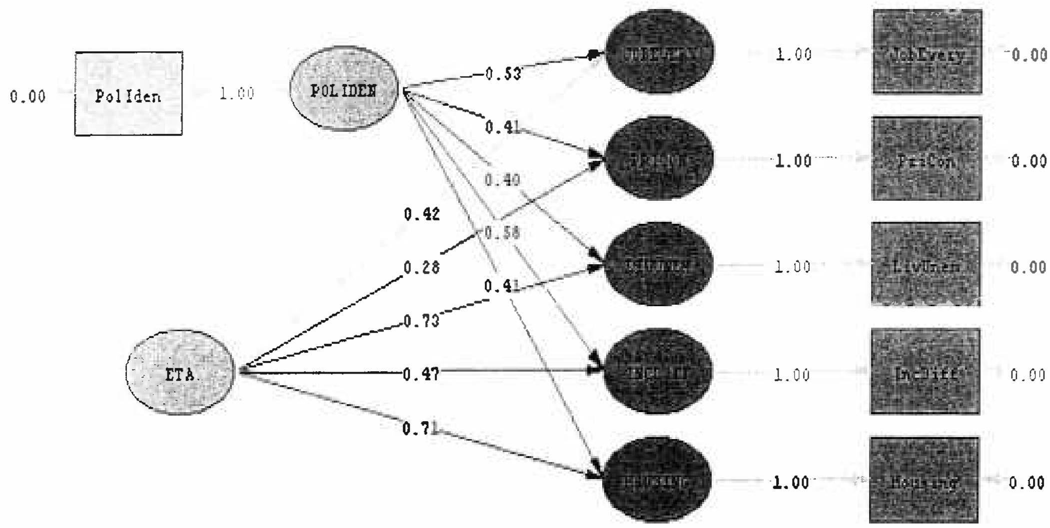
Table 30: Example 1: Chi-squared residuals greater than four, for two-way margins, for the one-factor LISREL model with covariates, when the covariate takes the value 0.987

item	*2*	*3*	*4*	*5*
1	(1,1), (2,2)	(1,1), (1,3), (2,2), (4,1)	(1,1), (2,1), (3,3), (4,4)	(4,1), (4,3)
sum(GF)	36.37	37.47	95.73	48.79
2		(1,1), (2,2), (3,1), (4,4)	(1,1), (2,2), (2,4)	(1,1), (2,1), (3,1)
sum(GF)		39.22	111.47	37.44
3			(1,1), (2,1), (2,3) (4,2), (4,4)	(1,1), (2,1), (4,1)
sum(GF)			67.02	42.86
4				(1,1), (1,2), (3,1), (4,2), (4,3)
sum(GF)				60.64
TOTAL(GF)=577				

From Tables 28, 29 and 30 we see that the results are similar to those obtained from section 6.3.2 where we fit the model to the matrix of polychoric and polyserial correlation. Table 28 shows that the fit has deteriorated considerably for the value –1.239. Table 29 shows a considerable increase in the fit since there are only a few chi-squared residuals exceeding four and the total GF has decreased by 65%. In table 30 we see that although there are more "problematic" chi-squared residuals than in the one-factor model without the covariate(table 13) the total GF has decreased by 32%.

Now we consider the LISREL model, where we use the joint unconditional covariance matrix but with maximum likelihood instead of weighted least squares. The coefficients and the loadings are identical to those of the LISREL model we fitted to the matrix of polychoric and polyserial correlations, and as a consequence the fit is similar as in section 6.3.2. Figure 4 gives the corresponding path diagram.

Figure 4: Example 1: Path diagram for the one-factor LISREL model (ML) with covariates.



6.4.4 CONCLUSIONS REGARDING THE INTRODUCTION OF THE COVARIATE

In the LISREL model, for values of the covariate near the mean value 0 the fit has improved, whereas for values at the tails of its distribution the fit has deteriorated. The two different methods of fitting the model in LISREL give similar fit, though they have big differences in the loadings as we see in table 34. When we use maximum likelihood using both ways the results are exactly the same. This is not a general rule since in other examples different results have come out, but the results are not very different and our data happen to give exactly the same results. So we have actually two different types of results, one from maximum likelihood and one from weighted least squares. From the two methods we infer similar results regarding the fit of the model. Although the fit from Tables (28) and (30) seem to have deteriorated in comparison with the model without the covariate, the AIC of the model with the covariates has decreased to 119.48 from 434.18, which is the AIC for the model with no covariates. There is some evidence that the model with four response categories, one factor and the covariate might fit well if we increase our sample. For the value –0.126 where we had 103 observations the fit is very good. This happens also for the value 0.156 with 87 observations as we see from Table 31 where there are almost no "problematic" chi-squared residuals at all and the total GF has decreased by 82%. Perhaps if we had a larger sample we would have been able to infer more reliable results.

Table 31: Example 1: Chi-squared residuals greater than four, for two-way margins, for the one-factor LISREL model(WLS) with covariates, when the covariate takes the value 0.156

item	*2*	*3*	*4*	*5*
1				
sum(GF)	16.93	20.78	9.39	12.96
2			(3,3)	
sum(GF)		17.0165 6	13.87561	9.92
3			(2,1)	(4,1)
sum(GF)			15.69	16
4				(4,1)
sum(GF)				21.46
TOTAL(GF)= 154.02				

In general, the results from the LISREL model with the covariate are contradictory and it is difficult to infer conclusions.

Table 32: Example 1: Estimated factor loadings for the LISREL model with direct effects

	ML	***WLS***
JobEvery	0.42	**0.54**
PriCon	0.29	0.41
LivUnem	0.73	0.73
IncDiff	0.47	0.5
Housing	0.71	0.71

6.5 LATENT VARIABLE MODEL WITH TWO FACTORS

Now we consider the model without covariates and with two factors. We investigate the chi-square residuals of the model to check if they give good predictions of the frequencies of the observed bivariate margins. In order for the LISREL model to be identified, we had to fix the values of some parameters. We chose to fix the dependence between the latent variables to 0 and also to fix the loadings of item 3 and 5 to 0. In all LISREL models we consider the same assumptions. Tables 33 and 34 give the chi-squared residuals exceeding four for the two-factor LISREL model and the two-factor IRT model respectively.

Table 33: Example 1: Chi-squared residuals greater than four, for two-way margins, for the two-factor LISREL model

item	*2*	*3*	*4*	*5*
1	(1,3), (1,4), (4,1)	(1,2), (1,3), (1,4), (4,1)	(1,1), (1,2), (1,3), (1,4), (2,2), (2,3), (2,4), (3,1), (3,3), (4,1), (4,2), (4,4)	(1,3), (1,4), (4,1), (4,3)
sum(GF)	250.65	162.17	1607.1	102.54
sum(LR)	91.15	72.81	233.82	56.56
2		(1,2), (2,1), (2,2), (4,1)	(1,4), (3,1), (4,1)	(2,1), (2,2), (4,1), (4,3)
sum(GF)		48.99	91.02	32.59
sum(LR)		43.74	53.17	31.21
3			(1,3), (1,4), (3,1), (4,1), (4,2), (4,3)	(2,3), (2,4), (3,1), (4,1), (4,3)
sum(GF)			415.71	2865.7
sum(LR)			102.43	
4				(1,3), (1,4), (2,4), (3,1), (4,1), (4,3)
sum(GF)				370.57
sum(LR)				108.09
TOTAL(GF)=5947.04				
TOTAL(LR)=792.96				

Table 33 shows that the fit has not improved. We should note that the big GF-contributions that occur in some pairs of variables are mainly caused by some combinations of some response categories, mainly the combination of (1,4) and (4,1). For example, when we investigate the case of items 1 and 4 where the GF is 1607.097, the chi-squared residual for response categories (1,4) is 1064.53, and the corresponding residual for response categories (4,1) is 409.78. So 2 out of 16 combinations of response categories account for the 92% of the total GF. This happens with all pairs of variables that have a large

GF. This is a bad property of the GF statistic. One way to overcome this problem is by collapsing categories. It is also interesting to see the big differences between GF-statistic and the LR-statistic. It is true that when the model does not hold the LR-statistic may be much smaller than the GF-statistic.

Table 34: Example 1: Chi-squared residuals greater than four, for two-way margins, for the two-factor IRT model

item	2	3	4	5
1	(1,1), (1,2), (2,2), (2,4), (3,1), (4,2), (4,4)	(1,2), (1,4)	(2,4)	(4,1)
sum(GF)	51.76	56.25	18.71	29.53
2		(1,4), (4,1)	(3,3)	(1,3), (4,1), (4,3)
sum(GF)		50.69	24.28	31.83
3			(4,1), (4,3)	(1,2), (2,1), (2,2), (3,3),(4,1), (4,2)
sum(GF)			21.2	59.03
4				(3,4), (4,1)
sum(GF)				24.45
TOTAL(GF)=367.74				

In the IRT case (table 34) the results are almost identical to those of the one-factor model. Hence the second factor has not contributed anything to the fit of the model but complexity and on these grounds we should reject the second factor. One of the most widely used model selection criteria is the AIC (equation 64). Table 35, which gives the AIC of the one-factor and two-factor IRT & LISREL models show some contradictory results. In LISREL, in contrast with the IRT model, the model with two factors gives a better fit than the model with one factor. This doesn't seem to be in agreement with the average GF-contribution and the examination of the chi-square residuals that suggest using the model with one factor. We should note that we cannot compare the AIC of the LISREL model with that of the IRT model because the latter uses the whole response pattern and therefore the likelihood is larger than the likelihood in the LISREL model which use only the bivariate frequencies.

Table 35: Example 1: Model selection criteria (AIC)

	1-FACTOR	*2-FACTORS*
LISREL	84.11	28.7
IRT	8636.19	8646.09

6.6 LATENT VARIABLE MODEL WITH THREE RESPONSE CATEGORIES(ONE-FACTOR AND TWO FACTOR)

Most of the "problematic" chi-squared residuals, as we see from tables 12 and 13, involve either the first or the last response categories. Also from tables 1 and 2 we see that the last response category is rarely answered. By collapsing the last two response categories into one we might get a better fit. By doing so we lose information that might be valuable and since the fit of the one-factor model is not very bad it is perilous to combine response categories. Probably the collapsing of categories should have been done before allocating the questionnaire to the individuals. Grier (1975) investigated the problem of the optimal number of choices per item and found that three-choice items are best when the total number of alternatives is fixed. Two-choice items are next best. We will fit the model with two response categories in section 6.6.

Table 36 gives the loadings and their standard errors for the LISREL and the IRT model as well as the standardized IRT loadings.

Table 36: Example 1: Estimated loadings and coefficients for the one-factor model with three response categories.

	IRT			*LISREL(WLS)*	
item	loadings	s.e.	standardized loadings	loadings	s.e.
JobEvery	1.817	0.152	0.876	0.79	0.053
PriCon	1.195	0.108	0.7669	0.63	0.055
LivUnem	2.263	0.17	0.9147	0.82	0.051
IncDiff	2.102	0.19	0.903	0.78	0.051
Housing	2.287	0.209	0.9163	0.84	0.051

As in the one-factor model with four response categories Table 36 shows that the loadings from the IRT approach are larger. Tables 37 and 38 give the chi-squared residuals exceeding four for the one-factor LISREL model with three response categories and the one-factor IRT model with three response categories.

Table 37: Example 1: Chi-squared residuals greater than four, for two-way margins, for the one-factor LISREL model with three response categories

item	*2*	*3*	*4*	*5*
1	(1,1), (1,3), (2,1), (2,2), (3,1), (3,3)	(1,1), (1,3), (2,2), (3,1), (3,3)	(1,2) (1,3), (2,2), (3,1), (3,3)	(1,2), (1,3), (2,1), (2,2), (2,3), (3,1), (3,3)
sum(GF)	116.16	471.96	93.19	657.95
sum(LR)	95.28	200.12	203.03	203.03
2		(1,1), (1,2), (1,3), (2,2), (3,1), (3,2), (3,3)	(1,1), (1,2), (1,3), (2,2), (3,1), (3,3)	(1,3), (2,2), (2,3), (3,1), (3,3)
sum(GF)		353.57	118.68	418.92
sum(LR)		182.01	86.16	186.61
3			(1,1), (1,3), (2,2), (3,1), (3,2), (3,3)	(1,1), (2,1), (2,2), (2,3), (3,1), (3,3)
sum(GF)			164.91	184.05
sum(LR)			122.29	
4				(1,1), (1,3), (2,1), (2,2), (2,3), (3,1), (3,3)
sum(GF)				234.11
sum(LR)				135.23
TOTAL(GF)=2813.51				
TOTAL(LR)=1413.76				

Table 38: Example 1: Chi-squared residuals greater than four, for two-way margins, for the one-factor IRT model with three response categories

item	*2*	*3*	*4*	*5*
1	(1,1), (1,2), (2,2), (2,3)	(1,2), (1,3)		(3,1)
sum(GF)	38.49	31.66	1.33	12.76
2		(1,3), (3,1)		(1,3), (3,1)
sum(GF)		33.22	11.2	25.44
3				(1,2), (2,1), (2,2)
sum(GF)			4.29	29.92
4				
sum(GF)				9.11
TOTAL(GF)=197.4				

The results we get from the IRT approach (Table 38) are satisfactory, though the fit is not as good as in the case with four response categories and the covariates. The GF as a criterion of goodness of fit has the drawback that is highly influenced by the bivariate frequencies that are not large and this may have inflated the average GF in this case. The LISREL results (Table 37) show a very bad fit. We will now fit the LISREL model with two factors and three response categories. Tables 39 and 40 give the chi-squared residuals exceeding four for the two-factor LISREL model with three response categories and the two-factor IRT model with three response categories respectively.

Table 39: Example 1: Chi-squared residuals greater than four, for two-way margins, for the two-factor LISREL model with three response categories

item	*2*	*3*	*4*	*5*
1	(2,2), (2,3)	(1,2), (1,3), (2,1), (2,2)	(1,2), (1,3)	
sum(GF)	34.7	31.63	16.25	8.7
sum(LR)	36.58	31.59	16.71	8.4
2		(1,2), (2,1), (2,2)	(1,3), (2,2), (3,1), (3,2)	(2,1)
sum(GF)		27.12	31.47	18.9
sum(LR)		27.54	31.86	18.58
3			(1,3), (3,2), (3,3)	(3,1)
sum(GF)			43.74	25.96
sum(LR)			37.93	
4				
sum(GF)				10.76
sum(LR)				10.81
TOTAL(GF)=249.21				
TOTAL(LR)=219.99				

The fit (Table 39) is satisfactory, though some improvement seems feasible.

Table 40: Example 1: Chi-squared residuals greater than four, for two-way margins, for the two-factor IRT model with three response categories

item	*2*	*3*	*4*	*5*
1	(2,2)	(1,2), (1,3), (2,1), (2,2)		
sum(GF)	16.11	29.26	0.9	7.14
2		(1,2), (2,1), (2,2)	(3,1)	(2,1), (2,2)
sum(GF)		25.88	11.57	17.18
3				(3,1)
sum(GF)			2.88	16.87
4				
sum(GF)				8.88
TOTAL(GF)=136.7				

The introduction of the second factor improves the fit when we have three response categories (Table 40). In the case with four response categories we saw that the introduction of the second factor added nothing to the overall fit (Table 34).

Table 41: Example 1: Model selection criteria (AIC)

	1-FACTOR	*2-FACTORS*
LISREL	83.45	27.19
IRT	7550.5126	7459.331

The results from the AIC (table 43) are similar to the case with four response categories. We should note that we cannot compare the AIC between models with different number of response categories.

6.7 LATENT VARIABLE MODEL WITH TWO RESPONSE CATEGORIES(ONE-FACTOR AND TWO FACTOR)

Now we consider the model with two response categories. The loadings of the IRT and the LISREL model as well as their standard errors are given in table 42.

Table 42:Example 1: Estimated factor loadings for the model with two response categories.

	IRT		*LISREL(WLS)*		
item	loadings	s.e.	standardized loadings	loadings	s.e.
JobEvery	1.761	0.199	0.8696	0.79	0.065
PriCon	1.197	0.158	0.7675	0.63	0.073
LivUnem	1.84	0.193	0.8787	0.78	0.068
IncDiff	2.127	0.238	0.9049	0.79	0.064
Housing	2.398	0.284	0.923	0.89	0.077

Tables 43 and 44 give the chi-squared residuals exceeding four for the one-factor LISREL model with two response categories and one-factor IRT model with two response categories respectively.

Table 43: Example 1: Chi-squared residuals greater than four, for two-way margins, for the one-factor LISREL model with two response categories

item	*2*	*3*	*4*	*5*
1	(1,1), (2,2)	(1,1), (1,2), (2,1), (2,2)	(1,1), (2,2)	(1,1), (1,2), (2,2)
sum(GF)	41.02	52.72	19.25	169.76
sum(LR)	34.16	46.61	18.64	92.42
2		(1,1), (1,2), (2,1), (2,2)	(1,1), (2,1), (2,2)	(1,1), (1,2), (2,1), (2,2)
sum(GF)		66.26	38.26	119.41
sum(LR)		58.3	33.81	88.86
3			(1,1), (2,1), (2,2)	(1,1), (1,2), (2,2)
sum(GF)			40.2	83.45
sum(LR)			36.39	62.24
4				(1,1), (1,2), (2,2)
sum(GF)				107.68
sum(LR)				69.78
TOTAL(GF)=738.01				
TOTAL(LR)=541.2				

Table 44: Example 1: Chi-squared residuals greater than four, for two-way margins, for the one-factor IRT model with two response categories

item	*2*	*3*	*4*	*5*
1	(1,2), (2,2)			
sum(GF)	11.43	1.93	0.89	3.12
2				
sum(GF)		3.08	0.18	3.79
3				(1,2)
sum(GF)			0.6	11.75
4				
sum(GF)				0.15
TOTAL(GF)=36.93				

It is obvious that the fit in the LISREL model (Table 43) has deteriorated in contrast with the IRT method (Table 44) where big improvement has been made from the model with one factor and three response categories, which, in addition was a good model. In the LISREL case we see that the models with one factor are not satisfactory. If we try to put the covariate in the case with three or two response factors we will get insignificant coefficients on the probit regressions of the underlying continuous variables on PolIden suggesting that we should not add the covariate in the model. Adding a second factor in the model improved the fit for the model with three response categories. We should do the same for the model with two response categories. Table 45 give the chi-squared residuals exceeding four for the two-factor LISREL model with two response categories.

Table 45: Example 1: Chi-squared residuals greater than four, for two-way margins, for the two-factor LISREL model with two response categories

item	*2*	*3*	*4*	*5*
1				
sum(GF)	0.000742272	0.06	0.001146644	0.31
sum(LR)	0.0007427274	0.06	0.001147111	0.31
2				
sum(GF)		0.02	0.001818066	0.04
sum(LR)		0.02	0.001818793	0.04
3				
sum(GF)			0.03	1.41
sum(LR)			0.03	1.35
4				
sum(GF)				0.22
sum(LR)				0.22
TOTAL(GF)=2.1				

The fit from Table 45 is perfect. It gives an excellent fit for all pairs of items and the total GF is negligible.

6.8 CONCLUSION

For the IRT model, one can get a good fit in all cases with the exception of the models with one factor and four response categories and the model with two factors and four response categories, though, even in these cases the fit isn't very bad. By collapsing categories we lose some information that might be valuable. The model with the covariate gives a very satisfactory fit. In the LISREL case, we see that in order to get a good model, we have to reduce the response categories and to add a second factor. If we consider the GF-statistic as a measure of fit, it is interesting to see in Table 46 the differences between the two approaches for the same model.

Table 46 : Example 1: Total GF per model

MODELS	TOTAL GF CONTRIBUTION	
	IRT	LISREL
1 FACTOR-4 RESPONSE CATEGORIES	367.72	848.32
2 FACTORS-4 RESPONSE CATEGORIES	367.74	5947.04
1 FACTOR-3 RESPONSE CATEGORIES	197.4	2813.5
2 FACTORS-3 RESPONSE CATEGORIES	136.7	249.21
1 FACTOR-2 RESPONSE CATEGORIES	36.93	738.01
2 FACTORS-2 RESPONSE CATEGORIES	41.32	2,1

From the Table (46), and if we take as a measure of goodness of fit the total GF, we see that the IRT models fit considerably well in all cases, whereas only two of the LISREL models fit well the data. The one-factor model with the covariate improves the fit in the IRT model but in the LISREL model things are more obscure. Probably if we increased the sample, we would get a better fit. We should note that no matter which model we choose the loadings are always positive and of similar magnitude. For example the second item (PriCon) has always the lowest loading. The introduction of the covariate brought a considerable decrease in the loadings of item one and four in all cases. It is also evident that the measures of goodness-of-fit such as the total GF or the chi-square residuals that exceed four, and the model selection criteria give sometimes different results.

CHAPTER 7

SECOND EXAMPLE

7.1 INTRODUCTION

The second application is also from the 1996 British Social Attitudes(BSA) Survey. Five ordinal manifest variables were selected for the analysis. The items measure satisfaction with the National Health Service in respondent's area. The items asked are whether the National Health Service in your area is, on the whole, satisfactory or in need of improvement.

- GP's appointment systems [Appointment]
- Amount of time GP gives to each patient [AmountTime]
- Being able to choose which GP to see [ChooseGP]
- Quality of medical treatment by GPs [Quality]
- Waiting areas at GP's surgeries [WaitingArea]

The response alternatives given to the respondents are: in need of a lot of improvement, in need of some improvement, satisfactory, and very good. Item nonresponse varies between 1.5%-2.5%. After we excluded the missing values we were left with 841 respondents. First we fit the measurement model with no direct effects to see if the five ordinal manifest variables measure one unidimensional latent trait. Next we consider several models of different combinations of factors and response categories and by letting the covariate political identification having direct effects on the five ordinal manifest variables and the covariates age and gender having direct effects in the latent variable.

7.2 PRELIS STEP

Each item has four categories, thus there are $4^5 = 1024$ possible response patterns. Since we have only 841 respondents it is reasonable to presume that not every response pattern can be present in the data. In fact we notice that there are only 205 different response patterns. The most common response pattern is that of answering satisfactory on all five items occurring in 149 cases. We should note that 99 out of the 205 different response patterns were observed only once and that 37 response patterns only twice. Probably this is a clue that our sample is small. The 20 most common response patterns are given in Table 47 and it is interesting to observe that there is no response 'in need of a lot of improvement' in any of them.

Table 47: Example 2: The 20 most common response patterns

cases					
149	3	3	3	3	3
41	2	3	3	3	3
30	4	4	4	4	4
23	2	2	3	3	3
22	3	3	3	4	3
18	3	3	3	3	2
16	3	3	3	3	4
15	2	3	2	3	3
14	2	2	2	2	2
13	3	2	3	3	3
13	2	2	2	3	3
12	3	3	2	3	3
12	3	3	3	2	3
11	4	4	4	4	3
11	3	4	3	4	4
10	2	3	3	4	3
10	3	2	2	3	3
9	3	4	4	4	4
9	2	2	3	2	3
9	3	3	4	4	3

The percentages for each category for each item are shown in table 48.

Table 48: Example 2: Frequency distribution for the observed ordinal items

	Appointment	*AmountTime*	*ChooseGP*	*Quality*	*WaitingArea*
in need of a lot of improvement	*11.4*	*6.5*	*6.7*	*3.8*	*3.6*
in need of some improvement	*29.4*	*22.8*	*20.9*	*19*	*16.1*
satisfactory	*47.2*	*57.9*	*58.3*	*53.9*	*63.3*
very good	*12*	*12.7*	*14.1*	*23.3*	*17.1*

We see that the majority of the responses fall in the two middle categories. For the first category, with the exception of the first item and especially for the last two items the percentages are negligible.

With 5 variables there are $\frac{5!}{2!*(5-2)!}=10$ pairs of variables. The contingency tables for every pair of variables are given in table 49:

Table 49: Example 2: Bivariate distributions for the observed ordinal items.

	AmountTi				ChooseGP				Quality			
Appointm	LI	SI	S	VR	LI	SI	S	VR	LI	SI	S	VR
LI	34	31	27	4	29	34	30	3	19	30	35	12
SI	13	100	128	6	14	80	140	13	8	74	133	32
S	8	56	296	37	13	59	283	42	5	50	260	82
VR	0	5	36	60	0	3	37	61	0	6	25	70

	WaitingA			
Appointm	LI	SI	S	VR
LI	15	30	45	6
SI	9	60	158	20
S	6	43	284	64
VR	0	2	45	54

	ChooseGP				Quality				WaitingA			
AmountTi	LI	SI	S	VR	LI	SI	S	VR	LI	SI	S	VR
LI	23	20	11	1	19	27	8	1	13	15	24	3
SI	20	77	92	3	10	69	98	15	8	55	118	11
S	13	77	347	50	3	59	333	92	8	60	353	66
VR	0	2	40	65	0	5	14	88	1	5	37	64

	Quality				WaitingA			
ChooseGP	LI	SI	S	VR	LI	SI	S	VR
LI	15	22	18	1	13	11	27	5
SI	12	72	80	12	6	47	112	11
S	5	60	327	98	10	69	340	71
VR	0	6	28	85	1	8	53	57

	WaitingA			
Quality	LI	SI	S	VR
LI	9	11	12	0
SI	12	52	82	14
S	6	60	342	45
VR	3	12	96	85

We can see that it is not very common that an individual will answer "in need of a lot of improvement" in an item and "very good" on another item and there are many zero cells in the contingency tables for this combination of responses. Too many zero cells can be problematic and give estimates that are imprecise and unreliable.

From Table 50 we see that there are no serious distortions due to non normality.

Table 50: Example 2: Correlations and test statistics

			Test of Model			Test of Close Fit	
Variable vs. Variable	Correlation		Chi-Squ.	D.F.	P-Value	RMSEA	P-Value
AmountTime vs. Appointment	0.632	(PC)	44.376	8	0.000	0.074	0.978
ChooseGP vs. Appointment	0.581	(PC)	31.757	8	0.000	0.059	0.999
ChooseGP vs. AmountTime	0.669	(PC)	14.377	8	0.072	0.031	1.000
Quality vs. Appointment	0.490	(PC)	43.737	8	0.000	0.073	0.981
Quality vs. AmountTime	0.676	(PC)	46.219	8	0.000	0.075	0.970
Quality vs. ChooseGP	0.632	(PC)	23.957	8	0.002	0.049	1.000
WaitingArea vs. Appointment	0.493	(PC)	14.068	8	0.080	0.030	1.000
WaitingArea vs. AmountTime	0.513	(PC)	30.755	8	0.000	0.058	0.999
WaitingArea vs. ChooseGP	0.420	(PC)	25.974	8	0.001	0.052	1.000
WaitingArea vs. Quality	0.507	(PC)	33.386	8	0.000	0.061	0.998

The matrix of polychoric correlation that is going to be used for further modeling in LISREL is given in Table 51. All the correlations are highly significant.

Table 51: Example 2: Matrix of polychoric correlations.

	Appointment	AmountTime	ChooseGP	Quality	WaitingArea
Appointment	1.000				
AmountTime	0.632	1.000			
ChooseGP	0.581	0.669	1.000		
Quality	0.490	0.676	0.632	1.000	
WaitingArea	0.493	0.513	0.420	0.507	1.000

Table 52 gives the estimated thresholds for the IRT model and those estimated from PRELIS.

Table 52: Example 2: Estimated thresholds for the IRT and for the LISREL model.

		IRT	*PRELIS*
Item	*Category*	*thresholds* $\left(\hat{t}_{is}\right)$	*thresholds* $\left(\hat{t}_{is}\right)$
Appointment	*1*	*-3.058*	*-1.205*
	2	*-0.636*	*-0.233*
	3	*3.024*	*1.175*
AmountTime	*1*	*-5.392*	*-1.511*
	2	*-1.949*	*-0.543*
	3	*4.189*	*1.140*
ChooseGP	*1*	*-4.268*	*-1.502*
	2	*-1.691*	*-0.595*
	3	*3.095*	*1.074*
Quality	*1*	*-5.142*	*-1.774*
	2	*-2.078*	*-0.744*
	3	*2.042*	*0.729*
WaitingArea	*1*	*-4.125*	*-1.803*
	2	*-1.873*	*-0.855*
	3	*2.116*	*0.949*

All the results in Tables 47,48,49,50,51 and the thresholds for the LISREL model in Table 52 came out by the PRELIS input 7.

PRELIS/LISREL input 7: Example 2:

Computing Polychoric Correlations and Asymptotic Covariance Matrix
Data Ninputvars = 5
Labels
Appointment AmountTime ChooseGP Quality WaitingArea
Rawdata = GP.dat
Clabels Appointment-WaitingArea 1=LI 2=SI 3=S 4=VR
Output BT MA=PM PM=GP.PM AC=GP.ACP

7.3 FITTING THE LATENT VARIABLE MODEL WITH ONE FACTOR TO THE MATRIX OF POLYCHORIC CORRELATIONS ESTIMATED IN 7.2

LISREL input 8 will fit the model to the matrix of polychoric correlations (Table 53) estimated by PRELIS input 7.

PRELIS/LISREL input 8: Example 2:

Testing Measurement Model
Observed Variables: Appointment AmountTime ChooseGP Quality
WaitingArea
Correlation Matrix from File GP.PM
Asymptotic Covariance Matrix from File GP.ACP
Sample Size: 841
Latent Variable: z
Relationships:
Appointment AmountTime ChooseGP Quality WaitingArea = z
Path Diagram
End of Problem

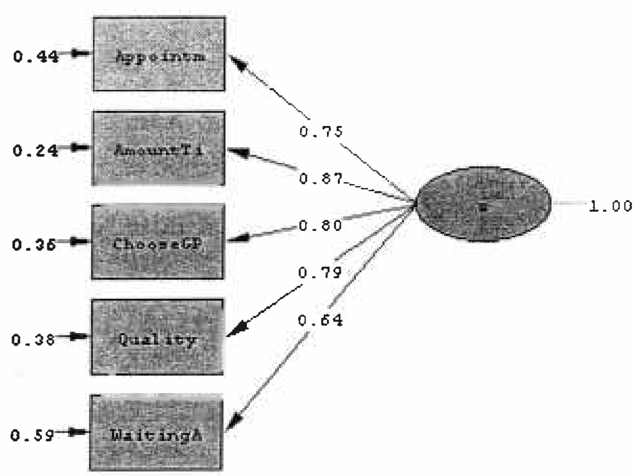
The results we get from fitting a one-factor model with the two approaches are shown in table 53.

Table 53: Example 2: Estimated factor loadings for the one-factor IRT model and for the one-factor LISREL model

	IRT			*LISREL*			
				WLS		*ML*	
Item	*unstandardized loadings α*	*s.e*	*standardized loadings stα*	*loadings λ*	*s.e*	*loadings λ*	*s.e*
Appointment	1.913	0.155	0.8863	*0.75*	*0.052*	*0.72*	*0.053*
AmountTime	3.224	0.229	0.9551	*0.87*	*0.052*	*0.86*	*0.053*
ChooseGP	2.293	0.163	0.9167	*0.8*	*0.051*	*0.78*	*0.053*
Quality	2.310	0.149	0.9177	*0.79*	*0.055*	*0.78*	*0.055*
WaitingArea	1.443	0.104	0.8219	*0.64*	*0.055*	*0.61*	*0.055*

LISREL loadings from every different method are always smaller that those of the IRT approach. The two different methods of estimating the parameters in LISREL give similar results. The path diagram for the LISREL model with weighted least squares is shown in Figure 5:

Figure 5: Example 2: Path diagram for the LISREL model with one latent factor and the use of weighted least squares.



We will check the models by computing the chi-squared. Tables 54 and 55 give the chi-squared residuals greater than four for the two-way margins for the one-factor IRT model and for the one-factor LISREL model respectively.

Table 54: Example 2: Chi-squared residuals greater than four for the two-way margins of the one-factor IRT model with four response categories

item	*2*	*3*	*4*	*5*
1	*(1,4), (3,4), (4,3)*	*(3,4), (4,3), (4,4)*	*(1,2), (1,4), (2,4), (3,4)*	
SUM(GF)	41.18	27.9	54.67	13.94
2			*(2,4), (3,4), (4,2), (4,3)*	*(1,2), (1,4)*
SUM(GF)		12.24	42.7	26.36
3				*(1,2), (1,4)*
SUM(GF)			19.28	28.45
4				*(2,4), (3,3), (3,4)*
SUM(GF)				28.45
TOTAL(GF)=295.18				

Table 55: Example 2: Chi-squared residuals greater than four for the two-way margins of the one-factor LISREL model with four response categories

item	*2*	*3*	*4*	*5*
1	*(1,4), (3,4),(4,2), (4,3)*	*(1,4), (3,1), (3,4), (4,3), (4,4)*	*(1,4), (2,4), (3,4), (4,2)*	*(1,4)*
SUM(GF)	93.96655	37.15153	91.22707	15.47415
2		*(1,4)*	*(2,4), (3,4), (4,2), (4,3)*	*(1,4), (3,4), (4,1), (4,2), (4,3)*
SUM(GF)		39.71579	61.16554	48.06529
3			*(4,2), (4,3)*	*(1,2), (1,4), (4,2)*
SUM(GF)			27.22329	49.73762
4				*(2,4), (3,4), (4,1)*
SUM(GF)				40.51032
TOTAL(GF)=504.23715	AVERAGE(GF)=3.15			

There are some **chi-squared residuals that exceed four in** Table 54 but all pairs of items have **a GF contribution less than 64. Almost the** same holds for Table 55. The LISREL model has almost double GF-statistic in comparison with IRT. We will fit various models with different combinations of number of factors and response categories to check how they fit the data. We start with the IRT case.

7.4 FIT OF VARIOUS IRT MODELS

In all possible models we will check the chi-squared residuals that exceed four and the total GF. We start with the IRT model with four response categories and two factors (Table 56).

Table 56: Example 2: Chi-squared residuals greater than four for the two-way margins of the two-factor IRT model with four response categories

item	*2*	*3*	*4*	*5*
1	*(1,1), (1,2), (2,2) ,(2,4), (3,1), (4,2), (4,4)*	*(1,2), (1,4)*		*(4,1)*
SUM(GF)	51.76	56.25	18.71	29.53
2		*(1,4), (4,1)*	*(3,3)*	*(1,3), (4,1), (4,3)*
SUM(GF)		50.69	24.28	31.83
3			*(4,1), (4,3)*	*(1,2), (2,1), (2,2), (3,3), (4,1), (4,2)*
SUM(GF)			21.2	59.03
4				*(3,4), (4,1)*
SUM(GF)				24.45
TOTAL(GF)=367.74				

The fit of the model (Table 56) has deteriorated, hence there is no reason to add a second factor. We stay in the IRT model and we check if we can get a better fit by collapsing the first two categories of each question into one (Table 57).

Table 57: Example 2: Chi-squared residuals greater than four for the two-way margins of the one-factor IRT model with three response categories

item	*2*	*3*	*4*	*5*
1	*(3,2)*	*(3,2), (3,3)*	*(1,3), (2,3)*	
SUM(GF)	17.43	21.7	34.19	9.98
2			*(1,3), (2,3), (3,1), (3,2)*	*(3,2)*
SUM(GF)		7.48	35.81	17.99
3			*(3,2)*	
SUM(GF)			15.77	14.02
4				*(2,3)*
SUM(GF)				21.82
TOTAL(GF)=196.19				

Most pairs of items in Table 57 give a satisfactory fit. We add a second latent variable to see if there is any considerable improvement in the fit (Table 58).

Table 58: Example 2: Chi-squared residuals greater than four for the two-way margins of the two-factor IRT model with three response categories

item	*2*	*3*	*4*	*5*
1	*(2,3), (3,2)*	*(3,2), (3,3)*	*(1,3), (2,3)*	
SUM(GF)	17.43	21.7	34.19	9.98
2			*(1,3), (2,3), (3,1), (3,2)*	*(3,2)*
SUM(GF)		7.48	35.81	17.99
3			*(3,2)*	
SUM(GF)			15.77	14.02
4				*(2,2), (2,3)*
SUM(GF)				21.82
TOTAL(GF)=196.18				

As we see from Table 58 the fit has not improved and it is nearly the same as with the case with three response categories and one factor. So we have added nothing but complexity. Now we collapse the last two categories of each item and we calculate the chi-squared residuals for the new model (Table 59).

Table 59: Example 2: Chi-squared residuals greater than four for the two-way margins of the one-factor IRT model with two response categories

item	2	3	4	5
1				
SUM(GF)	0.35	0.2	2.84	1.16
2				
SUM(GF)		0.05	0.27	0.4
3				
SUM(GF)			0.71	3.09
4				
SUM(GF)				1.15
TOTAL(GF)=10.22				

The fit in Table 59 is perfect.

7.5 FIT OF VARIOUS LISREL MODELS

Now we will try to find the LISREL model that gives the best fit to the data by examining the chi-squared residuals that exceed four and the total GF. We try the model with all response categories and two factors (Table 60). In order for our model to be identified we set the correlation between the latent variables to 0 and the values of the loadings of items 1 and 5 to 0 also.

Table 60: Example 2: Chi-squared residuals greater than four for the two-way margins of the two-factor LISREL model with four response categories

item	2	3	4	5
1	*(1,4), (3,4), (4,3), (4,4)*	*(1,4), (3,4), (4,3), (4,4)*	*(1,4), (3,3), (3,4), (4,3)*	*(1,4)*
SUM(GF)	78.49	34	60.53	18.47
2		*(1,4)*	*(1,4), (2,4), (4,2), (4,3)*	*(1,4), (4,1), (4,3), (4,4)*
SUM(GF)		32.45	64.61	41.53
3			*(2,4), (4,2)*	*(1,2), (1,4), (4,2)*
SUM(GF)			34.27	39.47
4				*(2,4), (3,3), (3,4), (4,1), (4,4)*
SUM(GF)				38.91
TOTAL(GF)=442.73				

We see from Table 60 that the fit is not satisfactory, though only one GF contribution for the pairs of items 2 and 4 barely exceeds the value 64. We continue by checking the chi-square residuals for every combination of response categories and number of factors and we see that the best model is that with two response categories and two factors.

Table 61: Example 2: Chi-squared residuals greater than four for the two-way margins of the one-factor LISREL model with three response categories

item	*2*	*3*	*4*	*5*
1	*(1,1), (1,3), (2,1), (2,2), (3,1), (3,3)*	*(1,1), (1,3), (2,1), (2,2), (3,1), (3,3)*	*(1,1), (1,2), (1,3), (2,1), (2,2), (3,1), (3,3)*	*(1,1), (1,3), (2,1), (2,2), (3,2), (3,3)*
SUM(GF)	329.07	243.32	362.17	142.68
SUM(LR)	152.02	148.69	177.68	111.34
2		*(1,1), (1,2), (1,3), (2,1), (2,2), (3,1), (3,2) (3,3)*	*(1,1), (1,3), (2,2), (3,1), (3,3)*	*(1,1), (1,3), (2,1), (2,2), (3,1), (3,3)*
SUM(GF)		280.95	706.85	308.47
SUM(LR)		149.63	190.13	159.47
3			*(1,1), (1,3), (2,2), (3,1), (3,3)*	*(1,1), (1,2), (1,3), (2,1), (2,2), (3,1), (3,2) (3,3)*
SUM(GF)			335.7	288.54
SUM(LR)			154.03	159.85
4				*(1,1), (1,3), (2,2), (3,1), (3,3)*
SUM(GF)				230.74
SUM(LR)				145.68
TOTAL(GF)=3228.49				
TOTAL(LR)=1548.53				

The fit in Table 61 is very bad. Now we consider the model with three response categories and two factors (Table 62).

Table 62: Example 2: Chi-squared residuals greater than four for the two-way margins of the two-factor LISREL model with three response categories

item	*2*	*3*	*4*	*5*
1	*(2,3)*	*(2,2), (2,3)*	*(1,2), (2,2), (2,3)*	
SUM(GF)	24.67	35.1	64.78	8.96
SUM(LR)	26.3	39.37	68.55	9.16
2		*(2,3)*	*(1,3), (2,2), (2,3), (3,1)*	*(2,3), (3,1), (3,2)*
SUM(GF)		23.63	54.72	22.99
SUM(LR)		26.21	63.55	21.65
3			*(2,2), (2,3), (3,1)*	*(3,1)*
SUM(GF)			41.96	15.06
SUM(LR)			35.09	13.88
4				*(2,2), (2,3)*
SUM(GF)				28.78
SUM(LR)				29
TOTAL(GF)=320.64				
TOTAL(LR)=332.77				

The fit as we see from Table 62 has improved.

We consider the model where the first two categories have collapsed into one and the last two categories have collapsed into another (Table 63).

Table 63: Example 2: Chi-squared residuals greater than four for the two-way margins of the one-factor LISREL model with two response categories

item	*2*	*3*	*4*	*5*
1	*(1,1), (2,1)*	*(1,1), (2,1), (2,2)*	*(1,1), (2,1), (2,2)*	*(1,1), (2,1)*
SUM(GF)	18.98	20.27	45.4	24.8
SUM(LR)	17.93	18.98	37.93	22.35
2		*(1,1), (2,2)*	*(1,1), (2,1), (2,2)*	*(1,1), (1,2), (2,1), (2,2)*
SUM(GF)		27.06	39.07	35.58
SUM(LR)		25.61	35.96	33.2
3			*(1,1), (2,1), (2,2)*	*(1,1), (1,2), (2,1), (2,2)*
SUM(GF)			37.97	40.83
SUM(LR)			34.17	38.06
4				*(1,1), (1,2), (2,2)*
SUM(GF)				44.56
SUM(LR)				39.47
TOTAL(GF)=334.52				
TOTAL(LR)=303.65				

The fit as we see from Table 63 is not satisfactory and we add a second latent variable. We see from Table 64 that this model gives a perfect fit. We also see that GF and LR have similar values when the model holds.

Table 64: Example 2: Chi-squared residuals greater than four for the two-way margins of the two-factor LISREL model with two response categories

item	*2*	*3*	*4*	*5*
1				
SUM(GF)	0.19	0.26	2.14	0.18
SUM(LR)	0.19	0.26	2.09	0.18
2				
SUM(GF)		0.09	0.17	0.91
SUM(LR)		0.09	0.17	0.91
3				
SUM(GF)			0.37	1.48
SUM(LR)			0.37	1.47
4				
SUM(GF)				1.19
SUM(LR)				1.19
TOTAL(GF)=6.97				

7.6 LATENT VARIABLE MODEL WITH ONE FACTOR AND COVARIATES

Now we assume that in the analysis we are interested in measuring overall satisfaction with GP's controlling for respondents political identification (measured by an observed covariate with four categories: conservative, labour, liberal democrat and other). We also want to measure the effect of gender and age on the latent variable satisfaction (Age is given in four categories: 18-25, 26-44, 45-64, 65+). The AIC of the IRT model without covariates is 7966.5, whereas for the model with the covariates it has decreased to7890.5. So the AIC proposes the introduction of the covariates in the IRT case.

PRELIS input 9 estimates the covariance matrix of the observed items and the covariates.

PRELIS/LISREL input 9: Example 2:

```
Computing Covariance Matrix
Data Ninputvars = 13
Labels
Appointment AmountTime ChooseGP Quality WaitingArea Labour Liberal
Other Female FirstAgeGroup SecondAgeGroup ThirdAgeGroup
FourthAgeGroup
Rawdata = GP.dat
Clabels Appointment-WaitingArea 1=AS 2=A 3=D 4=DS
Continuous Labour Liberal Other Female FirstAgeGroup SecondAgeGroup
ThirdAgeGroup FourthAgeGroup
Output MA=CM CM=GP.CM AC=GP.ACC WP
```

LISREL input 10 fits the model to the covariance matrix estimated from PRELIS input 9.

PRELIS/LISREL input 10: Example 2:

```
MIMIC Model
Observed Variables: Appointment AmountTime ChooseGP Quality
WaitingArea Labour Liberal Other Female FirstAgeGroup SecondAgeGroup
ThirdAgeGroup FourthAgeGroup
Covariance Matrix from File GP.CM
!Asymptotic Covariance Matrix from File GP.ACC
Sample Size: 841
Latent Variable: Eta APPOINTMENT AMOUNTTIME CHOOSEGP
QUALITY WAITINGAREA
 LABOUR LIBERAL OTHER FEMALE FIRSTAGEGROUP
SECONDAGEGROUP THIRDAGEGROUP FOURTHAGEGROUP
Relationships:
```

```
Appointment=1*APPOINTMENT
AmountTime=1*AMOUNTTIME
ChooseGP=1*CHOOSEGP
Quality=1*QUALITY
WaitingArea=1*WAITINGAREA
Labour=1* LABOUR
Liberal=1*LIBERAL
Other=1*OTHER
Female=1*FEMALE
FirstAgeGroup=1*FIRSTAGEGROUP
SecondAgeGroup=1*SECONDAGEGROUP
ThirdAgeGroup=1*THIRDAGEGROUP
FourthAgeGroup=1*FOURTHAGEGROUP
Let the error variances of Appointment-FourthAgeGroup to 0
APPOINTMENT=LABOUR LIBERAL OTHER 0.7*Eta
AMOUNTTIME CHOOSEGP QUALITY WAITINGAREA=LABOUR
LIBERAL OTHER Eta
Eta=FEMALE FIRSTAGEGROUP SECONDAGEGROUP
THIRDAGEGROUP FOURTHAGEGROUP
Path Diagram
Set the correlations of LABOUR-Eta to 0
Set the correlations of LIBERAL-Eta to 0
Set the correlations of OTHER-Eta to 0
LISREL Output: AD=OFF
End of Problem
```

The conceptual path diagram is shown in Figure 6.

Figure 6: Example 2: Conceptual path diagram for the One-factor LISREL model with covariates

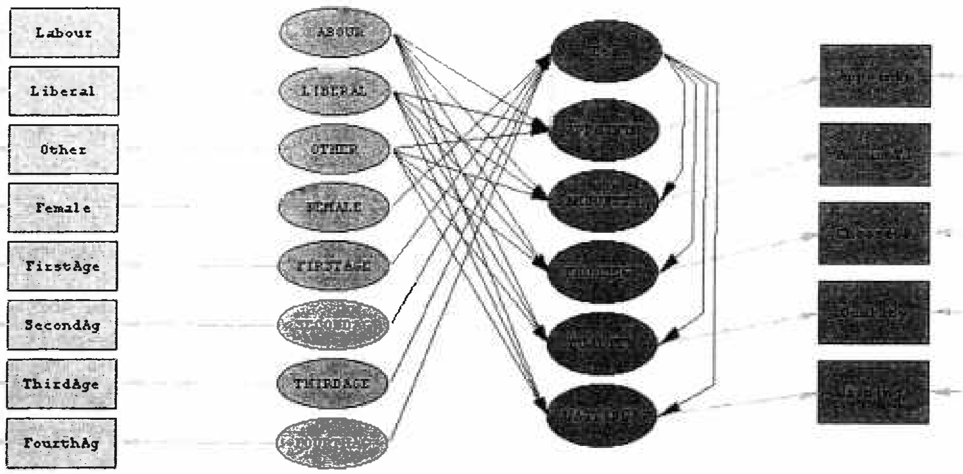


Table 65 gives the standardized (using equations 59 and 60) loadings and regression coefficients for the IRT model and the loadings and regression coefficients for the LISREL model.

Table 65 Example 2: Estimated factor loadings and regression parameters for the measurement LISREL model with direct effects and standardized loadings and regression coefficients for the IRT model.

Item	$\hat{a}_{i1}$		$\hat{\beta}_{i1}$ *(labour)*		$\hat{\beta}_{i2}$ *(liberal democrat)*		$\hat{\beta}_{i3}$ *(other)*	
	IRT	*LISREL*	*IRT*	*LISREL*	*IRT*	*LISREL*	*IRT*	*LISREL*
Appointment	0.86	0.7	0.39	0.33	0.31	0.27	0.2	0.18
AmountTime	0.93	0.84	0.38	0.35	0.17	0.14	0.19	0.14
ChooseGP	0.91	0.83	0.2	0.18	-0.11	-0.1	0.07	0.07
Quality	0.89	0.72	0.34	0.27	0.22	0.18	0.39	0.34
WaitingArea	0.8	0.61	0.28	0.23	0.01	0.03	0.35	0.26

Tables 66 and 67 give the estimated structural parameters for the IRT model and the LISREL model respectively.

Table 66: Example 2:Estimated structural parameters for the IRT model

	IRT	
	$\hat{\lambda}_l$	s.e.
constant	1.06	0.14
Female	-0.06	0.04
26-44	0.19	0.1
45-64	0.49	0.11
65+	0.7	0.11

Table 67: Example 2:Estimated structural parameters for the LISREL model

	LISREL	
	$\hat{\lambda}_l$	s.e.
Female	-0.06	0.08
18-25	1.89	1.4
26-44	2.08	1.4
45-64	2.39	1.4
65+	2.6	1.4

We cannot compare the estimates of the structural parameters of the IRT model (Table 66) with those of the LISREL model (Table 67) because the latent variable has a different distribution in each case. Though the AIC criterion shows that the model with the covariates in the IRT model is better than the model that does not contain any covariates we see that most of the coefficients are insignificant. In the LISREL model we observe the same phenomenon for the coefficients. The estimated coefficient for the covariate Female is the same in both cases. In the IRT case we used a constant and three variables for the covariate age, whereas in LISREL, instead of a constant we used four variables for the covariate age. Also we should note that we get a warning in LISREL that the covariance matrix used is not positive definite and this might have some implications in the results.

7.7 CONCLUSION

As we see from the following table (Table 68) we can get an extremely good fit for the IRT model if we consider one latent variable and two response categories. Whereas, in the LISREL model we can get an extremely satisfactory model if we consider two latent variables and two response categories. The fact that the one factor IRT model with two response categories and the two-factor LISREL model with two response categories give an excellent fit does not necessarily make them desirable because they are very complex and we also lose much information by combining categories. Table 54 shows that none of the pair of items has a larger than 64 GF contribution Also we see from Table 55 there are only two combinations of pair of items that have a GF contribution larger than 64. So we should probably prefer the one factor model both for IRT and for LISREL.

Table 68: Example 2: Total GF for various models.

MODELS	TOTAL GF CONTRIBUTION	
	IRT	LISREL
1 FACTOR-4 RESPONSE CATEGORIES	295.18	504.24
2 FACTORS-4 RESPONSE CATEGORIES	367.74	442.73
1 FACTOR-3 RESPONSE CATEGORIES	196.19	3228.49
2 FACTORS-3 RESPONSE CATEGORIES	196.19	320.64
1 FACTOR-2 RESPONSE CATEGORIES	10.22	334.52
2 FACTORS-2 RESPONSE CATEGORIES	10.22	6.97

It is very difficult to test the fit of the model when there are many covariates because we have to test the fit of the model for combinations of the values of the covariates. Also we need a large sample to do that. The LISREL results may not be correct because the covariance matrix is not positive definite.

CHAPTER 8

GENERAL CONCLUSIONS

The results from the two examples are similar and they also agree with the results from other examples in the literature (Jöreskog and Moustaki 2000) Political Efficacy example). In both examples the interrelationships among the items can be explained by a LISREL model with two factors and two response categories. An IRT model with one factor and two response categories gives a very good fit and it has also the advantage that is less complex. It is interesting to see in both examples how better the model with three response categories and one factor fits in the IRT case. For the first example, the IRT model with the covariate improves the fit and the best model is the one with all response categories, one latent variable and the covariate. From the second example it is difficult to check the chi-squared residuals but the AIC of the IRT model suggests the introduction of the explanatory variables. In the second example in section 7.5 we show that LISREL had problems in handling so many covariates.

Both approaches have their advantages and their disadvantages. It was expected that the IRT method would give a better fit since it uses the whole response pattern and no loss of information occurs, whereas LISREL uses only the univariate and the bivariate margins. LISREL requires a large sample for the estimation of the asymptotic covariance matrix and also we do not know the effects of the violation of the bivariate normality on the estimation of the polychoric correlations. On the other hand, IRT models have been developed recently and they are not so easy to use. Especially the model with the covariates is a very recent development. If one wants to fit more than two latent variables he will probably have to use a LISREL model or Mplus or EQS. LISREL is a very easy to use program that gives much potential to the user. LISREL also allows the user to make the latent variables dependent, or even to fix the dependence among them or to fix any other parameter in the model. Also the computational burden in the IRT models is huge as the number of factors increases. If one wants to save time he may decrease the number of quadrature points but the estimates may not be precise. Also in LISREL the same covariate might have direct effects both on the manifest and the latent variables, whereas, in the IRT approach, we saw in section 2.2 that in order for the model to be identified any covariate should affect either the manifest or the latent variables but not both of them. We should note that there are not many goodness-of-fit measures or model selection criteria in the literature for latent variable model with ordinal data and especially for the case where we have covariates affecting either the manifest or the latent variables. Among the available goodness-of-fit measures and model selection criteria we might get contradictory results.

APPENDIX

As we have seen $\theta_{i,s}(z,x) = \log \dfrac{\gamma_{i,s}}{\gamma_{i,s+1} - \gamma_{i,s}}$ and

$b(\theta_{i,s}(z)) = \log \dfrac{\gamma_{i,s+1}}{\gamma_{i,s+1} - \gamma_{i,s}} = \log(1 + \exp(\theta_{i,s}(z,x)))$. We should take the partial derivatives of $\theta_{i,s}(z,x)$ and $b(\theta_{i,s}(z,x))$ with respect to τ_{is} and a_{ij} for the proportional odds model.

We have that $\gamma_{i,s} = \dfrac{e^{\tau_{is} - \sum_{j=1}^{q} a_{ij} z_j + \sum_{l=1}^{r} \beta_{il} x_l}}{1 + e^{\tau_{is} - \sum_{j=1}^{q} a_{ij} z_j + \sum_{l=1}^{r} \beta_{il} x_l}}$ and if we set

$g = e^{\tau_{is} - \sum_{j=1}^{q} a_{ij} z_j + \sum_{l=1}^{r} \beta_{il} x_l}$ we have that $\gamma_{i,s,n} = \dfrac{g}{1+g}$ and $\dfrac{dg}{\partial \tau_{i,s}} = g$. Hence,

$$\frac{d\gamma_{i,s,n}}{\partial \tau_{i,s}} = \frac{g(1+g) - g^2}{(1+g)^2} = \frac{g}{(1+g)^2} = \frac{g}{1+g}\frac{1}{1+g} = \gamma_{i,s,n}\left(1 - \frac{g}{1+g}\right) = \gamma_{i,s,n}(1 - \gamma_{i,s})$$

We should also note that $\gamma_{i,s+1,m}$ does not depend on $\tau_{i,s}$ and as a result of that

$$\frac{d\gamma_{i,s+1,m}}{\partial \tau_{i,s}} = 0$$

So we take the first derivative of $\theta_{i,s,n}(z,x) = \log \dfrac{\gamma_{i,s,n}}{\gamma_{i,s+1,n} - \gamma_{i,s,n}}$ with respect to

τ_{is}

$$\frac{d\theta_{i,s,n}}{\partial \tau_{i,s}} = \frac{\gamma_{i,s+1,n} - \gamma_{i,s,n}}{\gamma_{i,s,n}} \frac{\gamma_{i,s,n}(1 - \gamma_{i,s,n})(\gamma_{i,s+1,n} - \gamma_{i,s,n}) + \gamma_{i,s,n}(1 - \gamma_{i,s,n})\gamma_{i,s,n}}{(\gamma_{i,s+1,n} - \gamma_{i,s,n})^2}$$

$$= \frac{(1 - \gamma_{i,s,n})\gamma_{i,s+1,n}}{\gamma_{i,s+1,n} - \gamma_{i,s,n}}$$

$$b\left(\theta_{i,s,n}(z,x)\right)=\log\frac{\gamma_{i,s+1,n}}{\gamma_{i,s+1,n}-\gamma_{i,s,n}}=\log\left(1+\exp\left(\theta_{i,s,n}(z,x)\right)\right)$$

$$\frac{db\left(\theta_{i,s,n}(z,x)\right)}{\partial\tau_{is}}=\frac{1}{1+\exp\left(\theta_{i,s,n}(z,x)\right)}\frac{d\theta_{i,s,n}(z,x)}{\partial\tau_{is}}e^{\theta_{i,s,n}(z,x)}$$
$$=\frac{\gamma_{i,s+1,n}-\gamma_{i,s,n}}{\gamma_{i,s+1,n}}\frac{\left(1-\gamma_{i,s,n}\right)\gamma_{i,s+1,n}}{\gamma_{i,s+1,n}-\gamma_{i,s,n}}\frac{\gamma_{i,s,n}}{\gamma_{i,s+1,n}-\gamma_{i,s,n}}=\frac{\gamma_{i,s,n}\left(1-\gamma_{i,s,n}\right)}{\gamma_{i,s+1,n}-\gamma_{i,s,n}}$$

Now we find the derivatives of $\theta_{i,s}(z)$ and $b\left(\theta_{i,s}(z)\right)$ with respect to a_{ij}

We have that $\frac{dg}{\partial a_{ij}}=-z_j g$ and

$$\frac{d\gamma_{i,s,m}}{\partial a_{ij}}=\frac{-z_j g(1+g)+z_j g^2}{(1+g)^2}=\frac{-z_j g}{(1+g)^2}=-z_j\frac{g}{1+g}\frac{1}{1+g}$$
$$=-z_j\gamma_{i,s,n}\left(1-\frac{g}{1+g}\right)=-z_j\gamma_{i,s,n}\left(1-\gamma_{i,s,n}\right)$$

If we set

$$y=e^{\tau_{i,s+1}-\sum_{j=1}^{q}a_{ij}z_j+\sum_{l=1}^{r}b_{ij}x_l}$$

and using the same logic we have that

$\frac{dy}{\partial b_{ij}}=-z_j\gamma_{i,s+1,n}\left(1-\gamma_{i,s+1,n}\right)$. So

$$\frac{d\theta_{i,s,n}(z)}{\partial a_{ij}}=\frac{\gamma_{i,s+1,n}-\gamma_{i,s,n}}{\gamma_{i,s,n}}$$
$$\frac{-z_j\gamma_{i,s,n}\left(1-\gamma_{i,s,n}\right)\left(\gamma_{i,s+1,n}-\gamma_{i,s,n}\right)+\left(z_j\gamma_{i,s+1,n}\left(1-\gamma_{i,s+1,n}\right)-z_j\gamma_{i,s,n}\left(1-\gamma_{i,s,n}\right)\right)\gamma_{i,s,n}}{\left(\gamma_{i,s+1,n}-\gamma_{i,s,n}\right)^2}$$
$$=\frac{-z_j\left(1-\gamma_{i,s,n}\right)\gamma_{i,s+1,n}+z_j\gamma_{i,s+1,n}\left(1-\gamma_{i,s+1,n}\right)}{\gamma_{i,s+1,n}-\gamma_{i,s,n}}$$
$$=\frac{-z_j\gamma_{i,s+1,n}\left(\gamma_{i,s+1,n}-\gamma_{i,s,n}\right)}{\gamma_{i,s+1,n}-\gamma_{i,s,n}}=-z_j\gamma_{i,s+1,n}$$

and

$$\frac{d(\theta_{i,s,n}(z))}{\partial a_{ij}}=\frac{d(1+\exp(\theta_{i,s,n}(z)))}{\partial a_{ij}}=\frac{1}{1+\exp(\theta_{i,s,n}(z))}\frac{d\theta_{i,s,n}(z)}{\partial a_{ij}}\exp(\theta_{i,s,n}(z))$$

$$=\frac{1}{1+\frac{\gamma_{i,s,n}}{\gamma_{i,s+1,n}-\gamma_{i,s,n}}}(-z_j\gamma_{i,s+1,n})\frac{\gamma_{i,s,n}}{\gamma_{i,s+1,n}-\gamma_{i,s,n}}$$

$$=\frac{\gamma_{i,s+1,n}-\gamma_{i,s,n}}{\gamma_{i,s+1,n}}(-z_j\gamma_{i,s+1,n})\frac{\gamma_{i,s,n}}{\gamma_{i,s+1,n}-\gamma_{i,s,n}}=-z_j\gamma_{i,s,n}$$

We have that $\frac{dg}{\partial\beta_{il}}=x_l g$ and

$$\frac{d\gamma_{i,s,n}}{\partial\beta_{il}}=\frac{x_l g(1+g)-x_l g^2}{(1+g)^2}=\frac{x_l g}{(1+g)^2}=x_l\frac{g}{1+g}\frac{1}{1+g}$$

$$=x_l\gamma_{i,s,n}\left(1-\frac{g}{1+g}\right)=x_l\gamma_{i,s,n}(1-\gamma_{i,s,n})$$

$\frac{dy}{\partial\beta_{il}}=x_l\gamma_{i,s+1,n}(1-\gamma_{i,s+1,n})$. So

$$\frac{d\theta_{i,s,n}(z)}{\partial\beta_{il}}=\frac{\gamma_{i,s+1,n}-\gamma_{i,s,n}}{\gamma_{i,s,n}}$$

$$\frac{x_l\gamma_{i,s,n}(1-\gamma_{i,s,n})(\gamma_{i,s+1,n}-\gamma_{i,s,n})+(-x_l\gamma_{i,s+1,n}(1-\gamma_{i,s+1,n})+x_l\gamma_{i,s,n}(1-\gamma_{i,s,n}))\gamma_{i,s,n}}{(\gamma_{i,s+1,n}-\gamma_{i,s,n})^2}$$

$$=\frac{x_l(1-\gamma_{i,s,n})\gamma_{i,s+1,n}-x_l\gamma_{i,s+1,n}(1-\gamma_{i,s+1,n})}{\gamma_{i,s+1,n}-\gamma_{i,s,n}}$$

$$=\frac{x_l\gamma_{i,s+1,n}(\gamma_{i,s+1,n}-\gamma_{i,s,n})}{\gamma_{i,s+1,n}-\gamma_{i,s,n}}=x_l\gamma_{i,s+1,n}$$

and

$$\frac{db(\theta_{i,s,n}(z))}{\partial\beta_{il}}=\frac{d(1+\exp(\theta_{i,s,n}(z)))}{\partial\beta_{il}}=\frac{1}{1+\exp(\theta_{i,s,n}(z))}\frac{d\theta_{i,s,n}(z)}{\partial\beta_{il}}\exp(\theta_{i,s,n}(z))$$

$$=\frac{1}{1+\frac{\gamma_{i,s,n}}{\gamma_{i,s+1,n}-\gamma_{i,s,n}}}(x_l\gamma_{i,s+1,n})\frac{\gamma_{i,s,n}}{\gamma_{i,s+1,n}-\gamma_{i,s,n}}$$

$$=\frac{\gamma_{i,s+1,n}-\gamma_{i,s,n}}{\gamma_{i,s+1,n}}(x_l\gamma_{i,s+1,n})\frac{\gamma_{i,s,n}}{\gamma_{i,s+1,n}-\gamma_{i,s,n}}=x_l\gamma_{i,s,n}$$

REFERENCES

Akaike, H. (1987). Factor analysis and AIC, *Psychometrica, 52,* 317-332

Bartholomew, D. J. (1980). Factor analysis for categorical data, *Journal of the Royal Statistical Society, Series B 42*, 293-321

Bartholomew, D. J. (1988). The sensitivity of latent trait analysis to choice of prior distribution, *British Journal of Mathematical and Statistical Psychology* 41, 101-107

Bartholomew, D. J. and Knott, M (1999). *Latent Variable Models and Factor Analysis.* (Second Edition ed.), Volume 7 of Kendall Library of Statistics, London: Arnold

Bartholomew, D. J. and Tzamourani, P. (1999). The goodness-of-fit of latent trait models in attitude measurement, *Sociological Methods and Research 27*, 525-546

Bentler, P. M.(1992). EQS: Structural Equation Program Manual. BMDP Statistical Software

Grier, J. B. (1975). The number of alternatives for optimum test reliability, *Journal of Educational Measurement,* 12, 109-112

Heath, A., Jowell, R., Curtice, J. and Witherspoon, S. (1986). *End of Award to the ESRC: Methodological Aspects of Attitude Research,* London: SCPR

Jöreskog, K. G. (1990). New developments in LISREL: analysis of ordinal variables using polychoric correlations and weighted least squares, *Quality and Quantity 24*, 387-404

Jöreskog, K. G. (1994). On the estimation of polychoric correlations and their asymptotic covariance matrix, *Psychometrica 59*, 381-389

Jöreskog, K. G. (2001). Analysis of Ordinal Variables 2: Cross-Sectional Data, (Found in the Internet address www.ssicentral.com/LISREL/corner.htm)

Jöreskog, K. G. (2002). Analysis of Ordinal Variables 5: Covariates, (Found in the Internet in the address www.ssicentral.com/LISREL/corner.htm)

Jöreskog, K. G and Goldberger, A. (1975). Multiple causes of a single latent variable, *Journal of the American Statistical Association* 70, 631-638

Jöreskog, K. G. and Rayment, R. (1996). *Applied factor analysis in the natural sciences.* Cambridge university press

Jöreskog, K. G. and Moustaki, I. (2001). Factor analysis of ordinal variables: a comparison of three approaches, *Multivariate Behavioral Research 36*, 347-387

Jöreskog, K. G. and Sörbom, D. (1993). LISREL 8 User's Reference Guide. Chicago: Scientific Software International

Lord, F. (1980). *Applications of item response theory to practical testing problems.* Lawrence Erlbaum Associates, New Jersey

McCullagh, P. (1980). Regression models for ordinal data, *Journal of the Royal Statistical Society, Series B, 42*, 109-142

Moustaki, I. (2000). A latent variable model for ordinal variables, *Applied Psychological Measurement, 24*, 211-224

Moustaki, I. (2001). A review of Exploratory Factor Analysis for Ordinal Categorical Data, *Structural equation Modeling: Present and Future.* 461-476, (eds. R. Cudeck, S. du Toit, D. Sörbom), Scientific Science International (SSI)

Moustaki, I. (2002). GENLAT 1.1: A computer program for fitting one- or two-factor latent variable model to categorical, metric and mixed observed items with missing values. Technical report, Statistics Department, London School of Economics and Political Science (To receive a copy of the software send an email to i.moustaki@lse.ac.uk)

Moustaki, I (2003). A general class of latent variable models for ordinal manifest variables with covariate effects on the manifest and latent variables (To appear in the British Journal of Mathematical and Statistical Psychology).

Muthén, B. O. (2002). Beyond SEM : General latent variable modeling. Behaviormetrica 29, 81-117

Muthén, B. O. and Muthén, L. (2000). Mplus: The comprehensive modeling to applied researchers . 11965 Venice Boulevard, Suite 407, Los Angeles, CA 90066

Quiroga, A. (1992). Studies of the polychoric correlation and other correlation measures for ordinal variables. Doctoral dissertation at Uppsala University

Sammel, R.D., Ryan, L. M. and Legler, J. M.(1997). Latent Variable models for mixed discrete and continuous outcomes, *Journal of the Royal Statistical Society, Series B, 59*, 667-678

Straud, A. H. and Sechrest, D. (1966). *Gaussian Quadrature Formulas.* Prentice-Hall