ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

SCHOOL OF INFORMATION SCIENCES AND TECHNOLOGY

DEPARTMENT OF STATISTICS

POSTGRADUATE PROGRAM

Bayesian Latent Variable Models for Ordinal Data

by

Marrieta P. Prifti

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

July 2015

“There is only one way to avoid criticism:
Do nothing, say nothing and be a nothing”

Aristotle

ACKNOWLEDGEMNTS

At this point I would like to thank some people for their encouragement and their assistance. I would like to thank all people who helped me to reach the point I am as a person and as a statistician. A huge thank belongs to my supervisor, Mr. Ioannis Ntzoufras because of his trust and his unlimited encouragement he offered to me. He helped me to improve not only my statistician skills but also my attitude as a scientist.

Furthermore, I wish to express my gratitude to Mr Dimitris Fouskakis. Mr. Fouskakis is the first who has shown me the beauty of the Statistics and especially the Bayesian. If we imagine statistics as a path, Mr. Fouskakis has shown me how to walk on it and Mr. Ntzoufras how to run on it. Both professors were more than willing to advise me to any of my concerns.

Furthermore, I would like to express my gratitude to Mrs. Katerina Tsaktani for the creation of the amazing cover of my dissertation. Once more, she did her miracles.

Last but not least, I own a great thank to my people. I am more than grateful to my wonderful parents, Paul and Maria to encourage and show me how to overcome probable difficulties. My thanks cannot end here without me referring to Chris who believed in me at the first glance.

To all of you, thank you from the bottom of my heart!

In modern science, it is common to quantify features that are not directly available. Such features are called latent variables and they need special handling in order to be estimated. An example of such variables is the political attitude or the satisfaction of a client for a new product. On the other hand, the observable features called manifest variables are be used as different indicators influenced by the unobserved latent ones. This has been generated the general class of the latent variable models. Depending on the type of the latent and the manifest variables (continuous, categorical etc.) different methods exists for each analysis. For example, when both set of features are continuous variables *factor analysis* is applied.

When the latent variables are continuous and the manifest ones are categorical variables, then the *Latent Trait Theory* is applied. Occasionally ordinal manifest variables are transformed to binary variables. Even if this practice is ususal and easy to be implemented, it is not preferable due to the loss of information. There are two main approaches for analyzing binary data with latent variable models, the *Item Response Theory (IRT)* and the *Underlying Variable Approach (UVA)*. Both approaches can be generalized in the case of the polytomous data (i.e. categorical items with more than two categories).

Latent variable models under the IRT approach are very interesting and demanding, with its theory growing more and more every year. From my point of view, it can be even more interesting if we apply it under Bayesian Approach. Through that concept it is possible to introduce our beliefs to the analysis concerning the characteristics of the under study object. As Good said: "*The subjectivist (i.e. Bayesian) states his judgments, whereas the objectivist sweeps them under the carpet by calling assumptions knowledge, and he backs in the glorious objectivity of science*" (Good, 1973).

ABSTRACT

Generally, Latent models are used in cases that the under study variables cannot be measured directly. Such variables are for example, consumer satisfaction of a new product. Here we analyze a market research dataset with discrete ordinal manifest variables and continuous latent features. Thus the methodology of latent trait models has been applied. In order to express our personal beliefs about the under study problem, the Bayesian approach has been implemented. Through the Bayesian paradigm we assume a prior distribution to express our information about the under estimation parameters. The combination of the prior distribution with the likelihood result in the posterior distribution. Estimation of the posterior distribution can be achieved through Markov Chain Monte Carlo (MCMC) algorithms. For our dataset the Gibbs sampler was implemented via WinBungs software.

Three different link functions were used, the logit, the probit and the c-loglog. Furthermore, one and two factor latent trait models were fitted. The final choice of the model was achieved through Deviance Information Criterion (DIC).

The aforementioned methodology was applied to detect a possible link between excessive consumption behaviors with schizotypy. The impulsive and the compulsive buying behaviors are considered excessive by the experts. On the other hand, schizotypy is related to a specific gene which increases the probability of schizophrenia when combined with specific environmental conditions. Such conditions are stress, anguish or even sadness (generally negative feelings). Schizotypy can be detected by its nine traits through the SPQ questionnaire. Obviously, interest lies in their association with serious psychiatric deceases. The data was collected for the purposes of a student survey (Iliopoulou, 2004) in the School of Management Sciences of the University of Aegean and the Technological Education Institutes of Crete and Piraeus.

Useful and interesting outcomes have been raised as far as the potential influence of schizotypy on impulsive and compulsive buying behaviors. Furthermore, many proposals for deeper exploration have been occurred but unfortunately are beyond the purposes of this thesis.

ΠΕΡΙΛΗΨΗ

Στην παρούσα εργασία, η μεθοδολογία των Μπεϋζιανών Μοντέλων Λανθανουσών Μεταβλητών εφαρμόζεται, τα οποία χρησιμοποιούνται σε περιπτώσεις όπου η υπό μελέτη μεταβλητές δεν μπορούν να μετρηθούν άμεσα. Παραδείγματος χάριν τέτοιες μεταβλητές είναι η ικανοποίηση του καταναλωτή για ένα νέο προϊόν. Το σετ δεδομένων που χρησιμοποιήθηκε αποτελείται από διατάξιμες διακριτές παρατηρήσιμες μεταβλητές συνεχείς λανθάνουσες μεταβλητές. Για το λόγο αυτό εφαρμόστηκε η μεθοδολογία των Λανθανουσών Χαρακτηριστικών (*Latent Trait Models*). Για να μπορέσουμε να εκφράσουμε τις προσωπικές μας πεποιθήσεις ως προς τις υπό μελέτη μεταβλητές, προτιμήθηκε η Μπεϋζιανή προσέγγιση και εφαρμόστηκε. Μέσω αυτής δύναται να υποθέσουμε την κατανομή που ακολουθούν οι εκτιμώμενες μεταβλητές, την *εκ των προτέρων κατανομή*. Ο συνδυασμός της εκ των προτέρων κατανομής με την κλασσική πιθανοφάνεια μας οδηγεί στην *εκ των υστέρων κατανομή*. Η εκτίμηση της εκ των υστέρων κατανομής γίνεται μέσω των Markov Chain Monte Carlo (*MCMC*) αλγόριθμων. Σε αυτή την εργασία ο δειγματολήπτης Gibbs χρησιμοποιήθηκε μέσω του Μπεϋζιανού λογισμικού WinBungs.

Τρεις συνδετικές συναρτήσεις χρησιμοποιήθηκαν: η logit, η probit και η c-loglog. Επιπλέον μοντέλα ενός και δύο παραγόντων προσαρμόστηκαν. Η τελική επιλογή του καταλληλότερου μοντέλου έδινε μέσω του πληροφοριακού κριτηρίου διακύμανσης (*Deviance Information Criterion -DIC*).

Όλο το παραπάνω θεωρητικό πλαίσιο εφαρμόστηκε με σκοπό να ανιχνευθεί, αν υπάρχει, πιθανή σχέση ανάμεσα σε υπερβολικές μορφές καταναλωτικής συμπεριφοράς και της σχιζοτυπίας. Ως τέτοιες μορφές καταναλωτικής συμπεριφοράς θεωρούνται η αυθόρμητη και η καταναγκαστική αγοραστική συμπεριφορά. Η σχιζοτυπία, από την άλλη, συνδέεται με την ύπαρξη ενός γονιδίου. Οι άνθρωποι που το έχουν εμφανίζουν σχιζοτυπικά συμπτώματα όταν επηρεάζονται από εσωτερικούς παράγοντες όπως είναι το στρες, το άγχος και σχεδόν κάθε αρνητικό συναίσθημα. Παράλληλα, είναι δυνατόν να ανιχνευθεί μέσω του ερωτηματολογίου SPQ που ανιχνεύει τα εννιά χαρακτηριστικά της. Είναι λοιπόν προφανές ότι υπάρχει μεγάλο ερευνητικό ενδιαφέρον ως προς την πιθανή τους σχέση διότι οι υπερβολικές αγοραστικές συμπεριφορές φαίνεται να επηρεάζονται από σοβαρές ψυχιατρικές-ψυχολογικές διαταραχές. Τα δεδομένα της εργασίας συλλέχθηκαν για τους σκοπούς

μεταπτυχιακής εργασίας στο Πανεπιστήμιο Αιγαίου στο Τμήμα Διοίκησης Επιχειρήσεων και στο ΤΕΙ Πειραιά (Ηλιοπούλου, 2004).

Ενδιαφέροντα και χρήσιμα συμπεράσματα προέκυψαν όσον αφορά την πιθανή επιρροή της σχιζοτυπίας στην αυθόρμητη και καταναλωτική συμπεριφορά. Επιπλέον από τα συμπεράσματα προέκυψε σημαντικό έναυσμα για περαιτέρω μελέτη που όμως είναι πέρα του σκοπού της συγκεκριμένης εργασίας.

CONTENTS

ACKNOWLEDGEMENTS………………………………………………………….. 2

INTRODUCTION…………………………... 3

ABSTRACT……………………………………………………………………….. 4

ΠΕΡΙΛΗΨΗ……………………………………………………………………….. 5

CONTENTS………………………………………………………………………. 7

CHAPTER 1: "LATENT VARIABLE MODELS FOR ORDINAL DATA"……… 10

1.1 Introduction………….…………………………………………….…10

1.1.1 A General Idea…………………………………………………… 10

1.1.2 The General Latent Variable Model………………………….. 11

1.1.3 The Factor Analysis Model………………………………….. 12

1.1.4 Estimation of the Parameters and Goodness of Fit………… 13

1.2 Latent Trait Models.…………………………………………………. 14

1.2.1 The Logit Latent Variable Model……………………………. 16

1.2.2 Estimation of the Parameters……………………………….... 18

1.2.3 Goodness of Fit……………………………………………….. 19

1.2.4 Factor Scores..………………….…………………………......21

1.2.5 Rotation……………………………………………………….23

1.3 Latent Variable Models for Ordinal Data…………………………....22

1.3.1 The Two Main Approaches………………………………….. 23

1.3.2 Item Response Theory (IRT)……………………………….....23

1.3.2.1 Fitting of the Model and Goodness of Fit………..26

1.3.2.2 Factor Scores……………………………………27

1.4 Example in the Latent Variable Model for Ordinal Data……………28

1.4.1 Descriptive Analysis…………………………………….…….28

1.4.2 Application of IRT……………………………………………29

1.4.2.1 The constrained model……………………………..30

1.4.2.2 The unconstrained model…………………………..33

1.5 Discussion………………………………………………………….38

CHAPTER 2: “BAYESIAN MODELS FOR LATENT VARIABLES”…………… 39
2.1 Bayesian Inference………………………………………………… 39
2.2. Sampling from the posterior…………………………………….. 42
2.3. Bayesian Model Assessment…………………………………..... 43
2.3.1 Bayes Factor (BF)………………………………………..... 43
2.3.2 Deviance Information Criterion (DIC)………………………. 45
2.4 Bayesian Models for Latent Variables……………………………. 46
2.4.1 Specification of Prior Distribution…………………………… 47
2.4.2 Sampling from the Posterior Distribution…………………….49
2.5 Example in the Bayesian Latent Variable Models for Ordinal Data.. 50
2.5.1 One Factor Latent Variable Models………………………....51
2.5.1.1 The Logit Model……………………………………...51
2.5.1.2 The Complementery Loglog Model………………….55
2.6 Discussion……………………………………………………………60

CHAPTER 3: “SCHIZOTYPY AND CONSUMER BEHAVIOR ”………………. 61
3.1 Consumer Behavior……………………………………………... 61
3.1.1 Definition of Consumer Behavior…………………………….. 61
3.1.2 Impulsive and Compulsive Behaviors………………………… 62
3.2. Schizotypy…………………………………………………….... 63
3.2.1 Characteristics of Schizotypy…………………………………. 63
3.2.2 Schizotypy and Consumer Behavior………………………….. 65
3.3 Latent Structure of Consumer Behavior……………………………. 66
3.3.1 Analysis of Two Factor Logit Model for
the Consumer Behavior...……………………………………...67
3.4 Modeling of Consumer Behavior and Total SPQ Score…………… 74
3.4.1 Influence of Total SPQ scale on Consumer Behavior……….. 74
3.4.2 Analysis of Two Factors Logit Model of Total SQP Score
on Consumer Behavior……………………………………………..77
3.5 Modeling of Consumer Behavior and Nine Traits
of Schizotypy…………………………………………………….81
3.5.1 Influence of Nine Traits of Schizotypy on Consumer
Behavior……………………………………………………..81
3.5.2 Analysis of Two Factors Logit Model of Nine Traits

of Schizotypy on Consumer Behavior…………………………..90
3.6 Discussion……………………………………………………….95

CHAPTER 4: "CONCLUSIONS-DISCUSSION"………………………………... 96
4.1 Conclusions...…………………………………………………. 96
4.2 Association of Consuming Behavior and Schizotypy…………….96
4.3 Further Research- Proposals…..…………………………………...97

APPENDIX A……………………………………………………………100

APPENDIX B…………………………………………………………….133

BIBLIOGRAPHY………………………………………………………. 138

List of Tables

1.1: Classification of latent variable models……………………………………………11

1.2: Response Patterns of binary data and its frequencies……………………….........15

1.3: Response Probabilities for each category………………………………………….24

1.4: Descriptive Information for all variables in percentages of Example 1.4………28

1.5: Cronbach's alpha for the manifest variables of Example 1.4…………………......29

1.6: Pair-wise associations for the manifest variables of Example 1.4………………30

1.7: Main Characteristics of the constrained model for Example 1.4…………….......30

1.8: Coefficients of the constrained model of Example 1.4………………………….31

1.9: Pearson Chi-squared test for two-way Margins for the
constrained model of Example 1.4…………………………………………….......32

1.10: Pearson Chi-squared test for three-way Margins for the
constrained model of Example 1.4………………………………………………...32

1.11: Main Characteristics of the unconstrained model of Example 1.4…………......33

1.12a: Coefficients of the unconstrained model of Example 1.4………….................33

1.12b: Coefficients of the unconstrained model of Example 1.4……………………......33

1.13: Pearson Chi-squared test for two-way Margins for the
unconstrained model of Example 1.4…………………………………………......34

1.14: Pearson Chi-squared test for three-way Margins for the
unconstrained model of Example1.4……………………………………………...35

1.15: Anova for Constrained and Unconstrained models of Example 1.4…………...35

1.16: The amount of test information for the fitted model based on
all variables of Example 1.4…………………………………………………………36

1.17: The amount of test information for the fitted model based on PrinCon
variable of Example 1.4……………………………………………………………...37

2.1: Bayes Factor interpretation according to Kass & Raftery (1995)……………….45

2.2: The assessment of the models via DIC of Example 2.5……………………….....51

2.3: Discrimination Parameters of each variable (10000 iterations-1000 burn-in)
for the logit model of Example2.5………………………..………………………53

2.4: Standardized Discrimination Parameters of each variable
for the logti of Example 2.5……..…………………………………………………...53

2.5: Difficulty parameters for each variable and category

for the logti of Example 2.5……………………………………………………………54
2.6a: Cumulative and level probability for each variable and category
for the logti of Example 2.5…………………………………………………………….54
2.6b: Cumulative and level probability for each variable and category
for the logti of Example ……………………………………………………………….55
2.7: Discrimination Parameters of each variable for the c-loglog model
of Example 2.5……………………………………………………………………………56
2.8: Standardized Discrimination Parameters of each variable
for the c-loglog model of Example 2.5…………………………………………….57
2.9: Difficulty parameters for each variable and category
for the c-loglog model of Example 2.5…………………………………………….58
3.1: DICs for each models (3.1) and (3.2) for schizotypic data……………………...66
3.2: DICs for models (M1) and (M2) on schizotypic data…………………………...74
3.3: DIC for third model on schizotypic data………………………………………...75
3.4: Estimations for coefficients of Total SPQ for the model (M3)
for schizotypic data……………………………………………………………..81
3.5: DICs for each model for scizotypic data………………………………………...82
3.6: Coding for the characteristics of schizotypy……………………………………83
3.7: The influence of nine schizotypic traits on each item
according to model (M5)…………………………………………………………90
3.8: DICs model (M6)………………………………………………………………..91
3.9: Estimated parameters g_{jz}^{+} and g_{jz}^{-} for model (M6)...…………………………...91

List of Figures

Figures

1.1: Several logistic response functions for various parameters…………………...17

1.2: Graphical representation of the distributions of ordinal variables
of Example 1.4……………………………………………………………..28

1.3: Item Characteristic Curves (ICC) for each variable of Example 1.4……………36

1.4: Item Characteristic Curve (ICC), Item Information Curve(IIC)
for each variable and Test Information Function of Example 1.4……………………36

2.1: Diagnostic Tests for the simulated chain of Logit Model for the first
discrimination parameter of Example 2.5……………………………………………52

2.2: Diagnostic Tests for the simulated chain of C-loglog Model
for the first discrimination parameter……………………………………………..56

3.1: Indicative Diagnostic Tests for the 7th discrimination parameter and first latent
factor with 10000 iterations (1000burn-in) for model (3.2) on schzotypal data……..67

3.2: Indicative Diagnostic Tests for the 7th discrimination parameter and
second latent factor with 10000 iterations (1000burn-in) for model (3.2)
on schzotypal data…………………………………………………………………68

3.3: Credible Intervals for discrimination parameters for each factor for
model (3.2) on schzotypal data……………………………………………………69

3.4: Standardized Discrimination Parameters of the 1st factor for
model (3.2) on schzotypal data……………………………………………………69

3.5: Standardized Discrimination Parameters of the 2nd factor for
model (3.2) on schzotypal data…………………………………………………....70

3.6: Response probabilities of a typical individual in each category for
every question for model (3.2) on schzotypal data…………………………………..71

3.7: Scatter Plot of latent scores for model (3.2) on schzotypal data………………...73

3.8: Box-plot of gj where *j*=1,…,37 number of questions based
on model (M2) for the schizotypic data……………………………………………..74

3.9: Indicative Diagnostic Tests for the third discrimination parameter
and first latent factor for model (M3) on schizotypic data…………………………...76

3.10: Indicative Diagnostic Tests for the third discrimination parameter
and second latent factor for model (M3) on schizotypic data………………………..77

3.11 : Credible Intervals for discrimination parameters for each factor
for model (M3) on schizotypic data…………………………………………………78
3.12: Standardized Discriminations for the first factor
for model (M3) on schizotypic data……………………………………………….....78
3.13: Standardized Discriminations for the second factor
for model (M3) on schizotypic data…………………………………………………79
3.14: Response probabilities of a typical individual in each category for
every question for model (M3) on schizotypic data…………………………………79
3.15: Scatter Plot of latent scores for model (M3) on schizotypic data……………...80
3.16: Box-plots for the coefficients of Ideas of Reference g[,1]
for model (M5)…………………………………………………………………….82
3.17: Box-plots for the coefficients of Excessive Social Anxiety g[,2]
for model (M5)……………………………………………………………………83
3.18: Box-plots for the coefficients of Odd Beliefs g[,3]
for model (M5)……………...84
3.19: Box-plots for the coefficients of Unusual Perceptual Experiences g[,4]
for model (M5)…………………………………………………………………….84
3.20: Box-plots for the coefficients of Odd Behavior g[,5]
for model (M5)…………………………………………………………………….85
3.21: Box-plots for the coefficients of No Close Friends g[,6]
for model (M5)………..86
3.22: Box-plots for the coefficients of Odd Speech g[,7]
for model (M5)……………...86
3.23: Box-plots for the coefficients of Constricted Affect g[,8]
for model (M5)……………………………………………………………………87
3.24: Box-plots for the coefficients of Suspiciousness g[,9]
for model (M5)…………………………………………………………………….88
3.25: Indicative Diagnostic Tests for the 12th discrimination parameter and first
latent factor for the model (M6) for schizotypic data…..……………………………91
3.26: Indicative Diagnostic Tests for the 12th discrimination parameter and second
latent factor for the model (M6) for schizotypic data………………………………..92
3.27 : Credible Intervals for discrimination parameters for each factor
for model (M6)……………………………………………………………………93
3.28: Standardized Discriminations for the first factor for model (M6)……………..93

3.29: Standardized Discriminations for the second factor for model (M6)…………..94
3.30: Response probabilities of a typical individual in each category
for every question for model (M6)…………………………………………………...94

CHAPTER 1: LATENT VARIABLE MODELS FOR ORDINAL DATA

1.1 Introduction

1.1.1 A General Idea

Frequently in social surveys or market research studies we wish to examine variables which we cannot be measured them directly, such as intelligence, political attitude, verbal ability, ambition and racial prejudice. Such variables are called *latent* and they are analyzed, summarized and studied by a family of methods and models as for example factor analysis for continuous numeric variables. The structure of the latent variables can be either univariate or multivariate. In the first case, a latent variable can summarize the whole set of observables while in the latter a more complicated structure underlies the observed variables. In the following we will assume q latent variables denoted as $y_1, y_2, \ldots, y_q$.

A widely used method for the study of latent variables is factor analysis which is a model based technique. It includes assumptions about the joint distributions over a relevant population of involved variables and allows us to extract conclusions for the population through goodness of fit, statistical significance and adequacy. We associate observations with latent scores through a probability model.

Often, we collect observable variables which we believe to be indicators of latent variables in order to indirectly measure them. These observable are called *manifest variables* (or *indicators*). Here we assume p manifest variables: $x_1, x_2, \ldots, x_p$ with $q<p$.

An example of a latent measures is the intelligence of a population. Unfortunately, intelligence cannot be measured directly like weight or volume. In this case, intelligence is the latent variable of interest which we can be introduced to a statistical model as a usual variable. The manifest variables can be the results of tests such as verbal, numerical, IQ performance test, since intelligent people can solve problems described by such tests. In any case, we use our intuition so as to measure the latent variables through suitable manifest variables. Thus, latent variables are hypothetical constructions invented by scientists in order to interpret the problem in hand and for which no direct method of measurement exists.

The relationship between a dependent observable variable and the independent latent variables is expressed via regression type model. The main issue in factor analysis or in latent variable models is the reversion of the regression equation in order to estimate the latent scores, for given values of the manifest variables.

A group of manifest variables often depends on the same latent variables. As a result, correlations structures between these manifest variables is introduced. Truly, this existence is an indicator of common source of influence. The goal of latent variable models is to specify the dependences between manifest variables and whether these dependences can be explained by a small number of latent variables.

Latent variable models have double use. We can use them either to find the hidden latent variables under a set of data, or to detect whether a set of variables is designed so as to measure specific notions. Of course, factor analysis is not the only method which can be applied in such cases. There is a variety of methods and we select the most appropriate ones according to the nature of the data; see Table 1.1. In this thesis, latent trait model will be used.

Manifest Variables	Latent Variables	Method of Analysis
Continuous	Continuous	Factor Analysis
Discrete	Continuous	Latent Trait Model
Continuous	Discrete	Latent Profile Model
Discrete	Discrete	Latent Class Model

Table 1.1 : Classification of latent variable Models

1.1.2 The General Latent Variable Model

Let **x** and **y** represent the manifest and latent variables respectively with **x'**$=[x_1, x_2, \ldots, x_p]$, **y'**$=[y_1, y_2, \ldots, y_q]$ and $q<p$. The latent variables should be less than manifest in order to produce an identifiable model. Moreover, by this way a latent variable model can be thought as a data reduction method which reduces the set of the manifest values to the set of the latent ones.

All latent variable models assume that the manifest values have a joint probability distribution conditional on latent observations: $\varphi(\mathbf{x} \mid \mathbf{y})$. When the density function of **y** is h(**y**) then the unconditional density of **x** is:

$$f(\mathbf{x}) = \int \varphi(\mathbf{x} \mid \mathbf{y})\, \mathrm{h}(\mathbf{y})\, \mathrm{d}\,\mathbf{y} \qquad (1.1)$$

From (1.1) we wish to learn how the manifest variables depend on latent ones through φ and h. Of course, it is not possible to infer about φ and h uniquely from f without some assumptions about their form. The most important assumption is the conditional independence which states that the manifest variables are independent of each other given the values of the latent variables:

$$\varphi(\mathbf{x} \mid \mathbf{y}) = \varphi_1(\mathrm{x}_1 \mid \mathbf{y})\varphi_2(\mathrm{x}_2 \mid \mathbf{y})...\varphi_p(\mathrm{x}_\mathrm{p} \mid \mathbf{y}) = \prod_{i=1}^{p} \varphi_i(\mathrm{x}_i \mid \mathbf{y}) \qquad (1.2)$$

This is the so-called "*conditional independence*" assumption which play a crucial role in latent variable models. An interpretation of (1.2) is that the latent variables create a independence between the manifest variables and when latent variables have been determined the manifest variables are basically random. Furthermore, it is assumed that h and φ_i are of known form but depend on a set of unknown parameters. So, in order to infer using h and φ_i from f we have to estimate the unknown parameters.

1.1.3 The Factor Analysis Model

The *Factor Analysis Model* for p manifest and q latent variables has the form:

$$X_i = \alpha_{i0} + a_{i1}\mathrm{y}_1 + ... + a_{iq}y_q + e_i, \quad i = 1,...p \qquad (1.3)$$

where $y_1,..., y_q$ are the latent variables, e_i are the residuals and $a_{i1},...,a_{iq}$ are the factor loadings. The constant term a_{i0} does not play any role in the fitting or interpretation of the model. If the manifest variables are measured in terms of their mean, then the constant term can be eliminated from the model. The factor loadings are the covariances between manifest and latent variables (or correlations if the manifest variables are standardized).

In factor analysis the latent variables are independent and follow the standardized normal distribution. Moreover, the residuals e_i are also independent following the $N(0,\sigma_i^2),\ \ i=1,...p$.

1.1.4 Estimation of the Parameters and Goodness of Fit

Let us assume that there is a population and a sample of size n from this population and each of them has its covariance matrix. For the population covariance matrix $\mathbf{\Sigma(\theta)}$, we know that its elements are given by specific functions of the parameters of the model: $\mathbf{\theta}'=[\theta_1, \theta_2, \ldots, \theta_k]$ where k is the number of parameters of the model. With $\mathbf{S}$ we denote the unbiased sample covariance matrix which is calculated by the above sample on the p manifest variables. The estimation of these parameters is obtained by minimizing a discrepancy function between $\mathbf{\Sigma(\theta)}$ and $\mathbf{S}$. The most common estimation method minimizing a discrepancy function between $\mathbf{\Sigma(\theta)}$ and $\mathbf{S}$ are the Ordinary Least Squares (OLS), the Generalized Least Squares (GLS) and the Maximum Likelihood (ML) which are described in the following paragraphs.

The function for the *Ordinary Least Squares (OLS)* is given by the equation (1.4) :

$$F[\mathbf{S},\mathbf{\Sigma(\theta)}]=\sum\sum_{i<j}\left[s_{ij}-\sigma_{ij}(\mathbf{\theta})\right]^2 \qquad (1.4)$$

where s_{ij} and $\sigma_{ij}(\mathbf{\theta})$ are the elements of the unbiased sample covariance matrix $\mathbf{S}$ and the population covariance matrix $\mathbf{\Sigma(\theta)}$ respectively. Although, this function is useful and extremely straightforward to apply, its implementation has major drawbacks. First of all, the function of OLS dependents on the scale of the manifest variables. This means that from using the sample covariance or correlation matrix may be produce different estimation of $\mathbf{\theta}$. Furthermore, the elements of $\mathbf{S}$ are usually correlated and have not equal variances. As a result, a simple measure of deviation between the elements of $\mathbf{\Sigma(\theta)}$ and $\mathbf{S}$ does not seem sufficient.

The *Generalized Least Squares (GLS)* minimizes the function (1.5) :

$$F[\mathbf{S},\boldsymbol{\Sigma}(\boldsymbol{\theta})] = \frac{1}{2} trace\left[(\mathbf{S}-\boldsymbol{\Sigma}(\boldsymbol{\theta}))\mathbf{S}^{-1}\right]^2 \qquad (1.5)$$

The drawbacks of OLS has lead to the definition of GLS. Here, the deviations between the elements of $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ and $\mathbf{S}$ are measured in the metric $\mathbf{S}^{-1}$; see Everitt (1984)

Last but not least, is the function of *Maximum Likelihood (ML),* which is given by the equation (1.6) :

$$F[\mathbf{S},\boldsymbol{\Sigma}(\boldsymbol{\theta})] = \ln|\boldsymbol{\Sigma}(\boldsymbol{\theta})| - \ln|\mathbf{S}| + \mathrm{trace}[\mathbf{S}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}] - p \quad (1.6)$$

This function is obtained from a transformation of the log-likelihood of the observations under the hypothesis that these have multivariate normal distribution and $\mathbf{S}$ has a Wishart distribution. All of these functions have the following properties:

a. $F(\mathbf{S},\boldsymbol{\Sigma}(\boldsymbol{\theta})) \geq 0$
b. $F(\mathbf{S},\boldsymbol{\Sigma}(\boldsymbol{\theta})) = 0$ if and only if $\mathbf{S} = \boldsymbol{\Sigma}(\boldsymbol{\theta})$
c. $F(\mathbf{S},\boldsymbol{\Sigma}(\boldsymbol{\theta}))$ is continuous in $\mathbf{S}$ and $\boldsymbol{\Sigma}(\boldsymbol{\theta})$

As far as the goodness of fit, when we can assume a multivariate normal distribution for the observed data then for (1.4) and (1.5), then

$$(n-1)\min F[\mathbf{S},\boldsymbol{\Sigma}(\boldsymbol{\theta})] \sim \mathrm{X}_v^2 \text{ where } v = \frac{1}{2}p(\mathrm{p}+1) - \mathrm{k} \qquad (1.6)$$

1.2 Latent Trait Models

In the latent trait models the manifest variables are discrete and the latent ones are continuous. These models were originally developed to solve problems in educational testing. They are based based on the perception that human abilities vary and the research subjects can be located on an ability scale under based on the answers they give to a set of questions. The essential difference between factor analysis and latent trait models is that special problems arise when the response data are binary. The goals of the analysis are the same:

a. Investigate the interdependences between manifest variables.
b. Examine if these interdependences can be explained by a small number of latent variables.
c. The assignment of scores in each object for every latent variable based on its answers.

When we working with binary data, we use one (1) to denote "success" or a positive response and zero (0) otherwise. (Moustaki et. al, 2002, p.: 326-327). This way of coding has the advantage that if we sum the answers in every row of the data matrix we can find the total number of positive responses. Many times, responses denoting by one (1) are the key-answers; a row of data is simply a string of zeros and ones e.g. 00111001001

Each row of the data matrix is called response pattern. For p manifest 2^p possible response patterns will exist. For instance, if we have p=2 then there exist 4 response patterns: 00, 01, 10, 11

If the sample size n, is large then inevitably many patterns will appear repeatedly. So, it is more convenient to present a list of possible pattern response with its respective frequencies as for example in Table 1.2 for the case p=2 manifest variables.

Patterns	Frequencies
00	35
01	42
10	2
11	18

Table 1.2: Response Patterns of binary data and its frequencies

When the number of variables p is large, some patterns may not be observed. Such simple factor analysis is not appropriate since the manifest variables cannot be considered as normal (or more generally continuous random variables) which is one of the basic assumptions in this model.

In order to surpass this difficulty, we need a modified model which will be able to correlate the latent variables with manifest binaries. It is possible achieve this using two different approaches. In the first approach a lot of characteristics from factor analysis are conserved. This can be succeeded by the use of an *underlying* variable for each *i* which is revealed partially to the binary X_i (Moustaki et. al, 2002, p.: 331). Then the factor model is maintained for this normal underlying variable. A more straightforward approach, is to adopt a logistic regression of the factor analysis model.

1.2.1 The Logit Latent Variable Model

In order to choose the regression function, we must take into consideration the regression of X_i to latent variables is $E(X_i \mid \mathbf{y})$ i.e. the expected value of X_i given **y.** In binary responses the expectation $E(X_i \mid \mathbf{y})$ is equal to the success probability (Bartholomew et. al, 2011, p.: 78). Thus, we need to specify the form of the $\pi_i(\mathbf{y})$ as a function of $y_1, y_2, \ldots, y_q$. The chosen function is known as *link function or response function* and one would expect to be monotonic. Since it is a probability we know that $0 \leq \pi_i(\mathbf{y}) \leq 1$.

The conditional distribution of x_i given $\mathbf{y}$ is:

$$g_i(X_i \mid \mathbf{y}) = \{\pi_i(\mathbf{y})\}^{Xi}\{1-\pi_i(\mathbf{y})\}^{1-Xi} \qquad (1.8)$$

The $g_i(x_i \mid \mathbf{y})$ belongs to the exponential family and the general linear latent variable model transforms to:

$$\text{logit}\,\pi_i(\mathbf{y}) = a_{i0} + \sum_{j=1}^{q} a_{ij} y_j \;\text{ where }\; \text{logit}\,\pi_i(\mathbf{y}) = \log\left(\frac{\pi_i(\mathbf{y})}{1-\pi_i(\mathbf{y})}\right) \qquad (1.9)$$

With the transformation of $\pi_i(\mathbf{y})$ using the logistic transformation we are able to write the model as linear function of the latent variables. The ratio $\dfrac{\pi_i(\mathbf{y})}{1-\pi_i(\mathbf{y})}$ is called the *odds* of the success. In psychometry, $\pi_i(\mathbf{y})$ is known as *item characteristic curve* or *item response function (IRF)*. The parameter a_{i1} defines the change of the slope of IRF along the average values. This means that, a given change in y_1, will cause a bigger change in the positive response probability when a_{i1} is big rather than be small. For this reason, in the item response theory it is known as *discrimination* parameter. The increase of the parameter a_{i0} increases the probability for all values of y_1 and is called *difficulty* parameter.

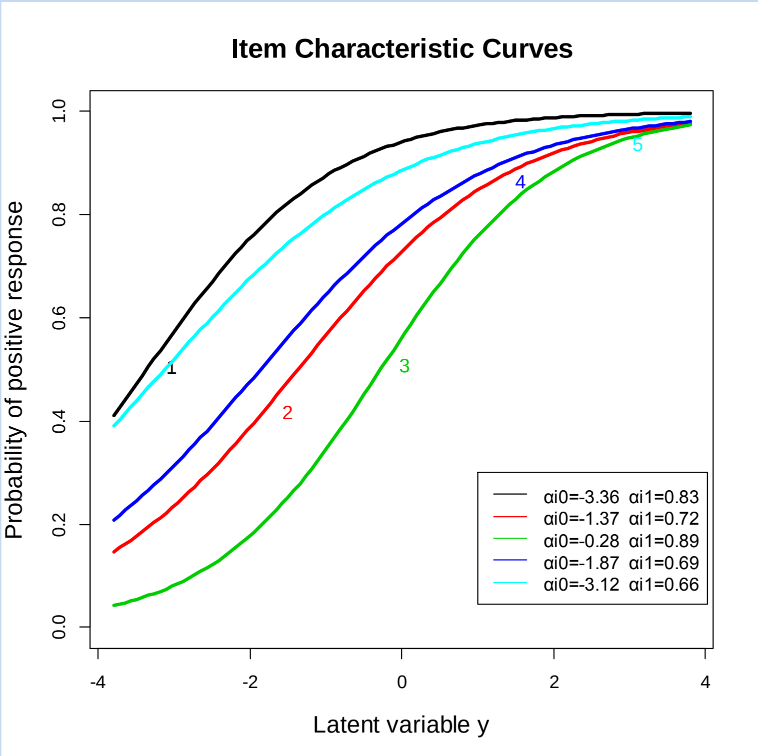


Figure 1.1: Several logistic response functions for various parameters (Source Rizopoulos,2006)

The Logit Model depends on three main assumptions:

1. Conditional independence (or Dependent independence)
2. The form of the link function
3. Independence and normality of the latent

Concerning the conditional independence we assume that the latent variables explain all correlations between the manifest variables. This assumption can be tested only by checking if the model is well-fitted to the data. A model of latent variables is well fitted when the latent variables can explain the greater part of the manifest correlation.

The selection of the link function is arbitrary and usually no or minor differences are observed for different choices. Nevertheless, the Logit and the Probit (or Normit) are the two most popular choices. Since $\text{logit}(\pi_i(\mathbf{y})) \equiv \frac{\pi}{\sqrt{3}}\Phi^{-1}(\pi_i(\mathbf{y}))$, it is clear that the Logit and Probit model are very close. Obviously, the factor loadings α s, in the Probit model will be less by $\sqrt{3}/\pi$ than the factor loadings of Logit model. These models are almost equivalent, theoretically but the Probit model has not the sufficiency property of the linear component X and is more preferable in Economics and related sciences (Bartholomew et. al, 2011, p.: 86).

Finally, the latent variables are usually assumed to be *independent* and that they follow the standard normal distribution N(0,1). We adopt this distribution due to its advantages in the rotation. Nevertheless, other distributions can be used without major differences. The selection of the distribution of the latent variables does not seem to affect the interpretation of the analysis (Bartholomew et. al, p.: 336).

1.2.2 Estimation of the Parameters

The parameters can be estimated using various approaches; see Section 1.1.3. The most widely used method to obtain the MLEs is the E-M algorithm (Expectation-Maximization algorithm). It is an iterative procedure of optimization. This method was first used for latent models by Block and Aitkin (1981). With this algorithm we can estimate parameters in models which depend on latent variables. The concept of the E-M algorithm is the following.

In the expectation step (E-step), a function is created for the expectation of the log-likelihood using the current estimate for the parameters, then in maximization step (M-step), the parameters are calculated by maximizing the expected log-likelihood from the E-step. These estimated parameters are used to define the distribution of the latent variables in the next E-step. This procedure is iterated until convergence is attained (i.e. the estimations do not change between consecutive iterations).

In some cases EM fails to obtain the MLEs. Such cases are called *Heywood case* and appear when:

1. When the sample size is small (less from a few hundred).
2. When the number of variables is small.
3. Extract more latent factors than are present (see Bartholomew et al, p.:67, 2011)

This situation is not as serious as it seems. First of all, when the estimated parameters become very large, the likelihood does not change so much and the fit of the model is hardly affected. Knott and Albanese (1992) pointed out that when large estimated values appear then, as a matter of fact, some loadings are infinite. In such cases we can adopt a cut-value in the iterative algorithm (cut-value=10) and when any parameter reaches it, the algorithm will be terminated. Today we can use many packages (in almost every statistical program) to apply the E-M algorithm. For instance, in R the appropriate package is "EMCluster". Also, other packages for Latent Trait Models have been created where the E-M algorithm is included, such as "ltm" package. More details for this algorithm can be found in Bartholomew, Knott & Moustaki (2011).

1.2.3 Goodness of Fit

The goodness of fit of a model can be checked by various ways. These ways can be applied separately or all together, complementary and are presented in the following paragraphs; see Bartholomew and Tzamoyrani (1999).

The Global Test: In order to check the goodness of fit of a model, we compare the observed and expected frequencies of the pattern responses. In fact, the fit of the models is achieved by minimizing the observed and the expected frequenies. So, the minimum proximity is an obvious measure of goodness of fit. A test based on this, is the G^2 statistic, the logarithm of the likelihood ratios:

$$G^2 = 2\sum_{r=1}^{2^p} O(\mathrm{r}) \ln\left[\frac{O(\mathrm{r})}{E(\mathrm{r})}\right] \qquad (1.10)$$

where r is a pattern response, $O(\mathrm{r})$ is the observed frequencies of r and $E(\mathrm{r})$ is the expected frequencies of the same r.

An alternative way is to use a simple Pearson Chi-square test given by

$$\mathrm{X}^2 = \sum_{r=1}^{2^p} \frac{[O(\mathrm{r}) - \mathrm{E(r)}]^2}{E(\mathrm{r})} \qquad (1.11)$$

Both of them, follow a X^2 distribution with $\{2^p$-p(q+1)-1\} degrees of freedom. If the sample size, is much larger than 2^p, then the observed and expected frequencies will be large enough and the approximation will be valid.

However, if the number of the binary variables is large, in many response patterns we will observe low expected frequencies. These tests require observed frequencies of size five.

Suppose that p=25, then there are 2^p>5 millions of possible response pattern and even with a sample of few thousands, many expected frequencies will be small. Then, both the X^2 and G^2 statistics cannot be assumed that are chi-squared distributed.

Margins Test: Such is based on the residuals calculated from the marginal frequencies of various combinations of the oservables. It is known that a set of response pattern is equal to a set of marginal probabilities. The first order margins $P(X_i=1)$ do not contain any information about the dependencies among the manifest variables.

On the other hand, the higher order margins (such as the second order marginal probability $P(X_i=1, X_j=1)$, and all other pairwise probabilities) do contain such information about pairwise association. Thus, we can use two way margins to check the fit of the corresponding marginal frequencies. This can be achieved through the construction of 2x2 marginal tables.

The comparison is completed through Chi-square residuals $e_i = \frac{(\mathrm{O}-\mathrm{E})^2}{E}$ where O is the observed frequencies and E is the expected. The squared residuals e_i^2 are the contributions to Chi-square statistic for 2x2 margin table.

The Chi-square residuals are not independent for each cell of the 2x2 marginal tables. As a result, they cannot be summed in order to give a global Chi-square test. A rule of thumb is to test for each cell since it is Chi-square distributed with one degree of freedom (Bartholomew & Leung 2002). If the residual value is greater than 3.5 then there is an indication of poor fitting in 5% significance level. It is important to notice that through the use of a marginal test information is some truly available, we reject the null hypothesis and the fit of the model concerns some marginal tables are rejected.

The interpretable G^2 percentage: An incomplete description of a model can be useful. Even if this model can left some points without interpretation, it can capture some interesting data traits. The idea is to use the interpretable percentage of statistic of the logarithm of the likelihood ratio for independence model, which is interpreted by a model with q latent variables:

$$\%G^2 = \frac{G_0^2 - G_q^2}{G_0^2} \times 100 \qquad (1.12)$$

where G_0^2 is a measure of association among manifest variables and G_q^2 is a measure of association among of the residuals between manifest variables which has not interpreted, for a q latent variables model. Therefore equation (1.12) measures the improvement in the likelihood of the model with q latent variables and loosely speaking the percentage of the correlation structure of the data.

1.2.4 Factor Scores

The estimation of the factor scores in latent trait models is more complicated than in factor analysis. Here, we are trying to identify an appropriate prediction element for every latent variable, having observed the manifest ones. If we use regression's terms we can say that we use the conditional expectation:

$$E(\mathrm{y_i} \mid X_1, \ldots, X_p) \quad \mathrm{j} = 1, \ldots \mathrm{q} \qquad (1.13)$$

Unfortunately, these conditional expectations are no longer linear combinations of the manifest variables $X_1, \ldots, X_{\mathrm{p}}$. Nonetheless, for the logit link function are monotonic function of the component scores:

$$\mathrm{x_j} = \sum_{i=1}^{p} a_{ij} X_i\,, \quad j = 1, \ldots, q \qquad (1.14)$$

For the logit link function, the result of the component $\mathrm{x_j}$ includes all the information in the data of latent variables, regardless the hypothesis about the distribution of $\mathrm{y_j}$. On the contrary, the $E(\mathrm{y_i} \mid X_1, \ldots, X_p)$ itself will vary depending on the distribution of the latent variables. This property of *non-volatility* is a good reason to prepare the factor scores of components (Moustaki et. al, 2002, p.:350).

For a given distribution of the latents, it is also possible to calculate the conditional standard deviations: $\sigma(\mathrm{y_j} \mid X_1,\dots,X_p), \quad \mathrm{j}=1,\dots\mathrm{q}$. In the assessment of the classification of the response patterns, the estimated standard deviations should be taken into consideration in order to detect whether the factor scores are characterized by high or low uncertainty.

1.2.5 Rotation

When we are fitting a model with multiple latent variables the MLE solution is not unique. A vertical rotation of the latent variables which is connected with the corresponding rotation of the estimated loadings maintains the likelihood unchangeable. So, we are able to search for a rotation with more convenient interpretation. The rotation does not produce a new solution of the model but describe the initial solution in a different perspective. Thus, in all different rotated solutions the fit of the models remains unchangeable. Non-orthogonal rotations also exist but orthogonal rotations are preferable due to their property of independence.

However, the uncertainty of the estimations is increased rapidly with the number of the latent variables. It should be seriously considered, whether it is worthwhile to fitting models with more than two latent variables for medium sized datasets.

1.3 Latent Variable Models For Ordinal Data

Here we consider models with ordinal categorical responses. For example, when we ask about the decisions of the government about the financial crisis the suggested answers could be: “strongly agree”, “agree”, “disagree” and “strongly disagree”. The response can belong only to one category and the categories are ordered according to their *power of approval*. Any ordinal categorical variable can be transformed to a binary variable with the union of categories resulting in a loss of information. The ordinal variables with more than two categories are referred to as *polytomous ordinal variables*.

In social sciences, market surveys and psychometric tools response are usually recorded by ordinal variables. Ordinal scales with five levels are known as *Likert* scales (Moustaki, 2002). Frequently an additional category which is out of the ordinal classification exists to express denial to respond such as “I do not know” or “I do not wish to answer”.

1.3.1 The Two Main Approaches

The two most important approaches to model ordinal responses are the *Underlying Variable Approach (UVA)* and the *Item Response Theory (IRT).* The first approach is used in structural equation modeling. It is based on limited information estimation methods which is its main drawback compared to IRT. Moreover it is an extension of the normal linear factor model which uses further the polychoric correlation matrix. Polychoric correlation is a technique for estimating the correlation between two normally distributed continuous latent variables from two observed ordinal variables.

On the other hand, in IRT, the observed variables are treated as they are. In this approach, we do not have any loss of information because the item of the analysis is the whole response pattern. Within the IRT framework Verhelest, Glas and Verstralen (1994), Zwinderman (1997) and Glas (2001) discussed the one-parameter logistic model with covariate effects (*Rasch Model*). Moustaki (2003) developed a general IRT framework similar to SEM (*Structural Equation Modeling*). The SEM approach provides a general framework that also allows for covariates which are allowed to affect the manifest variables either indirectly (through the latent variables) or directly (using regression type either associations). The foremost applications of IRT can be found in educational testing in which analysts are interested in measuring examinees’ ability using a test that consists of several questions.

1.3.2 Item Response Theory (IRT)

The Item Response theory was based on the general latent variable model formulation for binary data; see Section 1.2.

The assumptions of an IRT model are:

1. The latent variables are independent standardized normal distributed.

2. The ordered responses are independent given the latent variables.

Let us assume p manifest ordinal variables $x_1, x_2, ..., x_p$ with m_i categories and q latent $y_1, y_2, ..., y_q$. Each category has its own response probability, $\pi_{i(s)}(\mathbf{y})$ which is interpreted as the probability of that a response lies in the category *s* for the x_i variable, given that we have observed **y**.

Response Probabilities						
Categories	1	2	…	*s*	…	m_i
Res. Prob.	$\pi_{i(1)}(\mathbf{y})$	$\pi_{i(2)}(\mathbf{y})$	….	$\pi_{i(s)}(\mathbf{y})$	…	$\pi_{i(mi)}(\mathbf{y})$

Table 1.3: Response Probabilities for each category

We choose randomly a category *s* such that $1 < s < m_i$ and we divide the categories into two groups. The first group has the categories from one to *s* and the second one has the rest of them. In such way, it is possible to specify the *cumulative response probabilities* for each group as:

$$\begin{aligned}\gamma_{i(s)}(\mathbf{y}) &= P(x_i \le s) \\ &= \pi_{i(1)}(\mathbf{y}) + \pi_{i(2)}(\mathbf{y}) + ... + \pi_{i(s)}(\mathbf{y})\end{aligned} \tag{1.15a}$$

$$1 - \gamma_{i(s)}(\mathbf{y}) = P(x_i > s) = \pi_{i(s+1)}(\mathbf{y}) + \pi_{i(s+2)}(\mathbf{y}) + ... + \pi_{i(mi)}(\mathbf{y}) \tag{1.15b}$$

Thus, the logistic probability of a response in *s* category can be represented by considering as success probability the $\gamma_{i(s)}(\mathbf{y})$ or the $1-\gamma_{i(s)}(\mathbf{y})$. We prefer the second one, because it is directly connected to the binary case and it is easier to handle the indicators. So, under this formulation the ordinal logit model is given by

$$\log\left[\frac{1-\gamma_{i(s)}(\mathbf{y})}{\gamma_{i(s)}(\mathbf{y})}\right] = a_{i(s)} + \sum_{j=1}^{q} a_{ij} y_j \quad \text{where } s=1,\dots,m_i-1 \text{ and } i=1,\dots,p \tag{1.16}$$

This model is called *proportional odds model*. The name comes from the fact that in the case of one factor, the difference between two cumulative logits for two persons with factor scores y_1 and y_2 i.e. $\log\left[\frac{1-\gamma_{i(s)}(y_1)}{\gamma_{i(s)}(y_1)}\right]-\log\left[\frac{1-\gamma_{i(s)}(y_2)}{\gamma_{i(s)}(y_2)}\right]$ is proportional to the latent y_1-y_2. The constants $a_{i(s)}$ for each category denote the fact that as the limit of $a_{i(s)}$ increases for a response, then the difficulty of each variable increases too. Moreover these constants are providing the log odds of being in category *s* or higher when the latent scores are zero These constants are ordered as: $a_{i(1)} < a_{i(2)} < ... < a_{i(mi)}$ or with the opposite direction of inequalities depending on the nature of each variable. Nevertheless, factor loadings a_{ij} are common across all the categories of the observable variables. In other words, the discrimination capability of each variable does not depend on the point of the division of the categories into two groups. The parameter a_{ij} is called *discrimination parameter* for y_j and has the same meaning as in the binary case.

The factor loadings a_{ij} cannot be interpreted as correlation coefficients as in usual factor analysis. This can be achieved by transforming the factor loadings. Then, they renamed to *standardized factor loadings* or *standardized discrimination parameters* and are given from:

$$st.\alpha_{ij} = \frac{a_{ij}}{\sqrt{\sum_{j=1}^{q} a_{ij}^2 + 1}} \qquad (1.17)$$

It is desirable to have all standardized discrimination parameters close to one since such values indicates strong association between the latent and the corresponding manifest variable.

Since the latent variables are assumed to be standard normal random variables, an individual with latent scores equal to zero (at the point y=0), may be described as a "median" or typical individual. Through this way, the effect of the difficulty parameter on the probability of a positive response can be understood in more straightforward manner since:

$$\begin{aligned} \gamma_{i(s)}(\mathbf{y}=0) &= \frac{1}{1+\exp(\alpha_{i(s)})} \Rightarrow \\ e^{\alpha i(s)} &= \frac{1-\gamma_{i(s)}(\mathbf{y}=0)}{\gamma_{i(s)}(\mathbf{y}=0)} \end{aligned} \qquad (1.18)$$

The probabilities $\pi_{i(s)}(\mathbf{y})$ are calculated from:

$$\pi_{i(s)}(\mathbf{y}) = \gamma_{i(s)}(\mathbf{y}) - \gamma_{\mathrm{i(s-1)}}(\mathbf{y}) \text{ with } \mathrm{s} = 2, \ldots, \mathrm{m_i} \qquad (1.19)$$

Moreover it holds that $\gamma_{i(1)}(\mathbf{y}) = \pi_{\mathrm{i(1)}}(\mathbf{y})$ and $\gamma_{i(\mathrm{mi})}(\mathbf{y}) = 1$. We refer to $\pi_{i(s)}(\mathbf{y})$ as the *category response function*.

Such models are known as the *cumulative logit model for ordinal variables* when all variables (responses and covariances) are observed; see Agresti Section 7.2, 2002.

1.3.2.1 Fitting of the Model and Goodness of Fit

The model can be fitted using the same procedure with that of latent trait model which is based on maximum likelihood method. Furthermore, the goodness of fit of the model is conducted by the same criteria as above; see equations (1.10)-(1.12). The problem of sparseness in the case of polytomous data is more evident than for models with other type of data. If there exist $\mathrm{m_i}$ categories for the *i* variable, then the total number of response patterns is $(m_1 \times m_2 \times \ldots \times m_p)$. We can overcome this situation by grouping the response patterns or by merging the categories of some variables. Usually, specific categories of responses are rarely observed in practice. As a result, merging them with neighbor categories does not cause severe loss of information. This reduces the number of categories, effectively, without deteriorating the ordinality of the variable under study.

The degrees of freedom for G^2 and X^2 (see equations 1.10 and 1.11) is equal to the number of response patterns, after the grouping, minus the number of the independent parameters decreased by one. If there is no grouping the degrees of freedom are equal:

$$d.f. = (m_1 \times m_2 \times \ldots \times m_p) - \sum_{i=1}^{p} \mathrm{m_i} - p - pq - 1 \qquad (1.20)$$

The goodness of fit can be tested by the examination of the margins of second order (or even higher). The distribution of any two variables can be presented in a two way contingency tables; see Section 1.2.3.

1.3.2.2 Factor Scores

Factor scores can be estimated for the ordinal latent logit model in a similar manner. Although the simplicity of the binary case does not exist in such cases, two different methods can be used. The first method refers to the calculation of the expected values of the latent variables, given the observed; see equation (1.13). In the second method the components scores can be used; see equation (1.14). Component scores are the linear combination of observable variables and factor loadings. Both methods give the same results.

In the general case, component scores do not include the whole information about the latent variables. So, the first method is more reliable but has implementation difficulties.

More details about the IRT will be present in Section 1.4 where a dataset is analyzed by using the R package “ltm” (Rizopoulos, 2006).

1.4 Example

The data for this example consist of five ordinal manifest variables which are measuring the attitudes to the role of government. The data set is from the 1996 British Social Attitudes Survey (BSA); see Moustaki (2003). Responders were asked if the government should or not:

1. provide a job for everyone who wants one (variable JobEvery)
2. keep prices under control (variable PrinCon)
3. provide a decent standard of living for the unemployment (variable LivUnem)
4. reduce income differences between the rich and the poor (variable IncDiff)
5. provide decent housing for those who cannot afford it (variable Housing)

The response alternatives given to the responders were: definitely should be (1), probably should be (2), probably should not be (3), definitely should not be (4). After excluding the missing values, the sample size is equal to 822 responders.

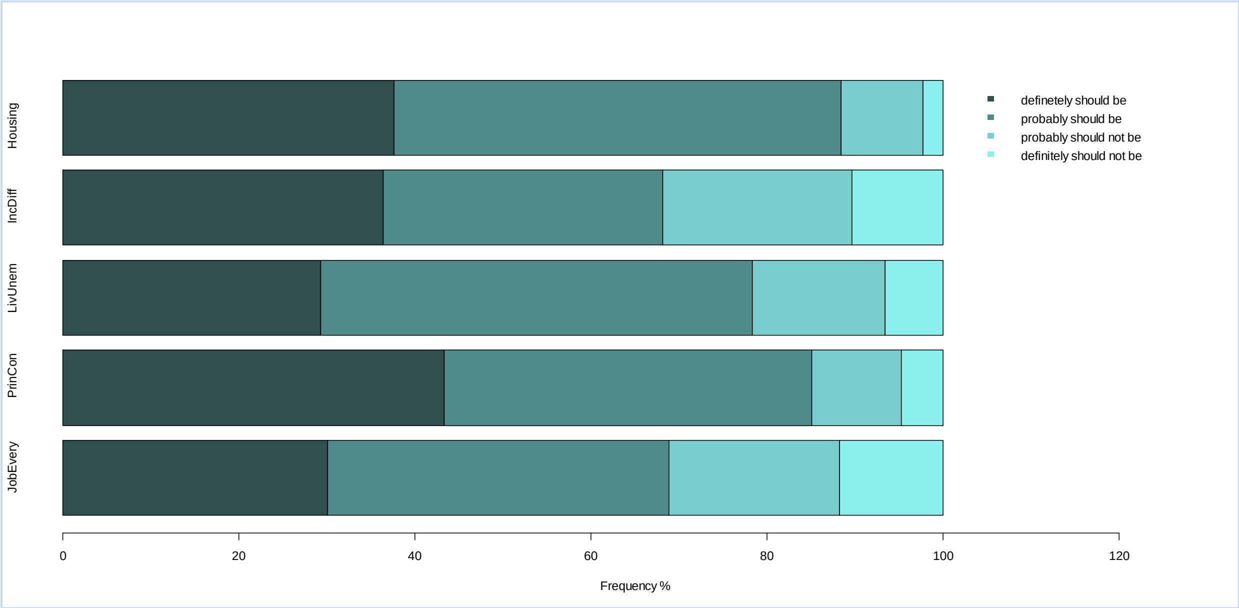
First, we provide some descriptive information about the five ordinal variables and then we will make IRT analysis in R with the help of the package "ltm". We fit both, the constrained and the unconstrained model. In the constrained model we consider all discrimination parameters equal with a constant (which is estimated). On the other hand, in the second model all discrimination parameters are unequal to each other (and also are estimated).

1.4.1 Descriptive Analysis

The percentages for each category of each variable are given from Table 1.4 and Figure 1.2. It is clear that the minority of the responses fall into "definitely should not be" for every variable. On the other hand, the majority of the sample has given positive response in every question.

	definitely should be	probably should be	probably should not be	definitely should not be
JobEvery	30.05	38.81	19.34	11.80
PrinCon	43.31	41.73	10.22	4.74
LivUnem	29.32	49.03	15.09	6.57
IncDiff	36.37	31.75	21.53	10.34
Housing	37.59	50.85	9.25	2.31

Table 1.4: Descriptive Information for all variables in percentages of Example 1.4



JobEvery: Job for everyone, PrinCon: Prices under control, LivUnem: Standard of living for the unemployed, IncDiff: Reduce income differences, Housing: Decent housing

Figure 1.2: Graphical representation of the distributions of ordinal variables of Example 1.4

1.4.2 Application of IRT

Before proceeding to the fit of the model we will provide and comment the descriptive statistics of the data set. The Cronbach's alpha values are presented in Table 1.5. This measure takes values in the interval [0,1] (mathematically can take values also out of this interval). It is a coefficient measuring the internal consistency of the questionnaire. Internal consistency refers to the intercorrelations among test items. Cronbach's alpha is calculated from the following equation:

$$a = \frac{K}{K-1}\left[\frac{\sum \sigma_k^2}{\sigma_{total}^2}\right] \qquad (1.21)$$

where K is the number of the observable variables, $\sum \sigma_k^2$ is the sum of the k item score variances and σ_{total}^2 is the variance of scores on the total measurement. From the above equation, it is clear that alpha may also take negative values. In such cases, the integrity of the scores is being disputed. For example, negative alpha can arise when the item score variance is grater that total score. Then the items are measuring different concepts and internal consistency does not exist between item scores; see Ritter (2010). High values of alpha close to one are more desirable. A rule of thumb requires a reliability of 0.7 or higher before any analysis in order to ensure that the internal consistency is high.

Cronbach's alpha	value
All Items	0.7776
Excluding JobEvery	0.7194
Excluding PrinCon	0.7726
Excluding LivUnem	0.7317
Excluding IncDiff	0.7160
Excluding Housing	0.7366

Table 1.5: Cronbach's alpha for the manifest variables of Example 1.4

Thus, from Table 1.5 we can see that internal consistency is in an acceptable level for this dataset.

Item i	Item j	p.value
1	2	0.001
1	5	0.001
2	3	0.001
2	4	0.001
2	5	0.001
3	5	0.001
4	5	0.001
1	3	$<2\times10^{-16}$
3	4	$<2\times10^{-16}$
1	4	$<2\times10^{-16}$

Table 1.6: Pair-wise associations for the manifest variables of Example 1.4

Table 1.6 depicts all pair-wise associations. It is useful to inspect the data for the evidence of positive associations. This check is performed by the constructions of all 2x2 contingency tables for all possible pairs of items and by evaluation the corresponding Chi-squared p-values. It is clear that in all pairs, the null hypothesis of independence is rejected. This means that all pairs are associated and this association structure has been modeled by one or more latent variables.

1.4.2.1 The constrained model

We fit the model which is defined by equations (1.18) and (1.19), using the "grm" command from "ltm" R package. For this model the discrimination parameters are considered constrained common for all variables.

log.Lik	AIC	BIC
-4318.79	8669.58	8744.97

Table 1.7: Main Characteristics of the constrained model for Example 1.4

Table 1.7 presents the maximized log-likelihood value, the AIC and the BIC criteria for the constrained model. These measures are used to compare constrained and the unconstrained models.

Table 1.8 provides the estimated coefficients of each variable and their corresponding standard errors. The estimated discrimination parameter is 1.858 for all variables. Therefore, for any given change in the latent variables, all manifest variables change I in a similar way in terms of probability. In the follow, the goodness of fit of the model is checked through two and three way margins.

JobEvery	**value**	**std.err**
Extrmt1	-0.696	0.066
Extrmt2	0.685	0.080
Extrmt3	1.629	0.782
Dscrmn	**1.858**	**0.073**
PrinCon		
Extrmt1	-0.188	0.060
Extrmt2	1.405	0.074
Extrmt3	2.265	0.769
Dscrmn	**1.858**	**0.073**
LivUnem		
Extrmt1	-0.719	0.066
Extrmt2	1.091	0.126
Extrmt3	2.080	0.142
Dscrmn	**1.858**	**0.073**
IncDiff		
Extrmt1	-0.463	0.062
Extrmt2	0.652	1.336
Extrmt3	1.739	1.322
Dscrmn	1.858	0.073
Housing		
Extrmt1	-0.420	0.062
Extrmt2	1.620	0.093
Extrmt3	2.781	1.537
Dscrmn	**1.858**	**0.073**

JobEvery: Job for everyone, PrinCon: Prices under control, LivUnem: Standard of living for the unemployed, IncDiff: Reduce income differences, Housing: Decent housing

Table 1.8: Coefficients of the constrained model of Example 1.4

Tables 1.9 and 1.10 assess the lack of fit for two and three way marginal tables. In Table 1.9 the upper diagonal part contains the chi-squared statistic. Obviously, the lowest value is the better. The lower diagonal part indicates the pairs for which the statistic exceed the threshold value. In both tables the triple asterisks denotes significant differences between observed and fitted values. Here since three out of six two way tables and two out of ten three way table indicate problems in the fit of the model a more elaborate model might be needed.

	JobEvery	PrinCon	LivUnem	IncDiff	Housing
JobEvery	-	30.18	48.13	19.63	21.58
PrinCon		-	90.63	31.18	61.73
LivUnem		***	-	16.74	83.90
IncDiff				-	23.58
Housing		***	***		-

*JobEvery: Job for everyone, PrinCon: Prices under control, LivUnem: Standard of living for the unemployed, IncDiff: Reduce income differences, Housing: Decent housing. The "***" indicates pairs with lack of fit*

Table 1.9: Pearson Chi-squared test for two-way Margins for the constrained model of Example 1.4

Item i	Item j	Item k	(O-E)^2/E
1	2	3	296.43 ***
1	2	4	132.02
1	2	5	146.77
1	3	4	121.74
1	3	5	220.85
1	4	5	112.16
2	3	4	217.13
2	3	5	362.54 ***
2	4	5	172.54
3	4	5	158.70

*The "***" indicates triplets with lack of fit*

Table 1.10: Pearson Chi-squared test for three-way Margins for the constrained model of Example 1.4

1.4.2.2 The unconstrained method

In the unconstrained model the discrimination parameters are different for each variable. Comparing AIC and BIC values from the Tables 1.7 and 1.11, it is obvious that the unconstrained model is clearly better.

log.Lik	AIC	BIC
-4298.05	8636.10	8730.33

Table 1.11: Main Characteristics of the unconstrained model of Example 1.4

In the Tables 1.12a and 1.12b we depict the estimated coefficients of each variable accompanied with their standard errors. The discrimination parameter for each variable is available. Using this model, we assume different effect on the response for the same change of the latent variable. The latent variable has the highest effect on the variable which records whether the government should provide or not a decent income for the unemployment (LivUnem). On the contrary, PrinCon, is influenced less by the latent variable.

JobEvery	**value**	**std.err**
Extrmt1	-0.706	0.072
Extrmt2	0.694	0.095
Extrmt3	1.655	0.848
Dscrmn	**1.779**	**0.143**
PrinCon		
Extrmt1	-0.270	0.079
Extrmt2	1.838	0.394
Extrmt3	3.040	3.286
Dscrmn	**1.175**	**0.109**
LivUnem		
Extrmt1	-0.664	0.064
Extrmt2	1.011	0.235
Extrmt3	1.911	2.995
Dscrmn	**2.242**	**0.187**

JobEvery: Job for everyone, PrinCon: Prices under control, LivUnem: Standard of living for the unemployed,

Table 1.12a: Coefficients of the unconstrained model of Example 1.4

IncDiff	value	std.err
Extrmt1	-0.430	0.061
Extrmt2	0.618	0.102
Extrmt3	1.637	1.156
Dscrmn	**2.101**	**0.169**
Housing		
Extrmt1	-0.380	0.058
Extrmt2	1.477	0.826
Extrmt3	2.517	17.385
Dscrmn	**2.311**	**0.207**

IncDiff: Reduce income differences, Housing: Decent housing

Table 1.12b: Coefficients of the unconstrained model of Example 1.4

Similar to the previous approach, Tables 1.13 and 1.14 present the lack of fit in two and three way marginal tables. Although, the unconstrained model seems to be better than the constrained one it still fails in three cases in total (two in bi-dimensional marginal table and one for three dimensional marginal table).

	JobEvery	PrinCon	LivUnem	IncDiff	Housing
JobEvery	-	52.14	56.35	18.85	29.60
PrinCon		-	50.76	24.40	31.87
LivUnem	***		-	21.29	58.72
IncDiff				-	24.47
Housing			***		-

*JobEvery: Job for everyone, PrinCon: Prices under control, LivUnem: Standard of living for the unemployed, IncDiff: Reduce income differences, Housing: Decent housing. The "***" indicates pairs with lack of fit*

Table 1.13: Pearson Chi-squared test for two-way Margins for the unconstrained model of Example 1.4

Item i	Item j	Item k	(O-E)^2/E
1	2	3	325.88 ***
1	2	4	134.44
1	2	5	188.33
1	3	4	148.75
1	3	5	211.01
1	4	5	139.17
2	3	4	139.01
2	3	5	160.98
2	4	5	119.15
3	4	5	168.96

Table 1.14: Pearson Chi-squared test for three-way Margins for the unconstrained model of Example1.4

Table 1.15 clearly suggest that the unconstrained model is better according to AIC and BIC. Moreover the significance test rejects the null hypothesis that all discriminations parameters are equal with p-value<0.001.

	AIC	BIC	log.Lik	LRT	df	p.value
Const. model (same discr.)	8669.58	8744.97	-4318.79			
Unconst. model (non same discr.)	8636.10	8730.33	-4298.05	41.48	4	<0.001

Table 1.15: Anova for Constrained and Unconstrained models of Example 1.4

The fitted unconstrained model is illustrated in the following figures. From the item characteristic curves for each variable (in Figures 1.3 and 1.4) it is obvious that there is low probability of endorsing the first category ("definitely should be") for high value of latent of scores. Therefor the questions of the survey are not considered as the main criteria about the role of the government. This conclusion is also reached by the test information curves from which we can observe that the set of five (5) questions provides 65.6% of the high latent traits. Furthermore, in the item information curve it is clear that the variable which represents whether the government should or not keep prices under control (PrinCon) provides little information in the whole latent trait field. It is possible to check this numerically by using the results in Table 1.16.

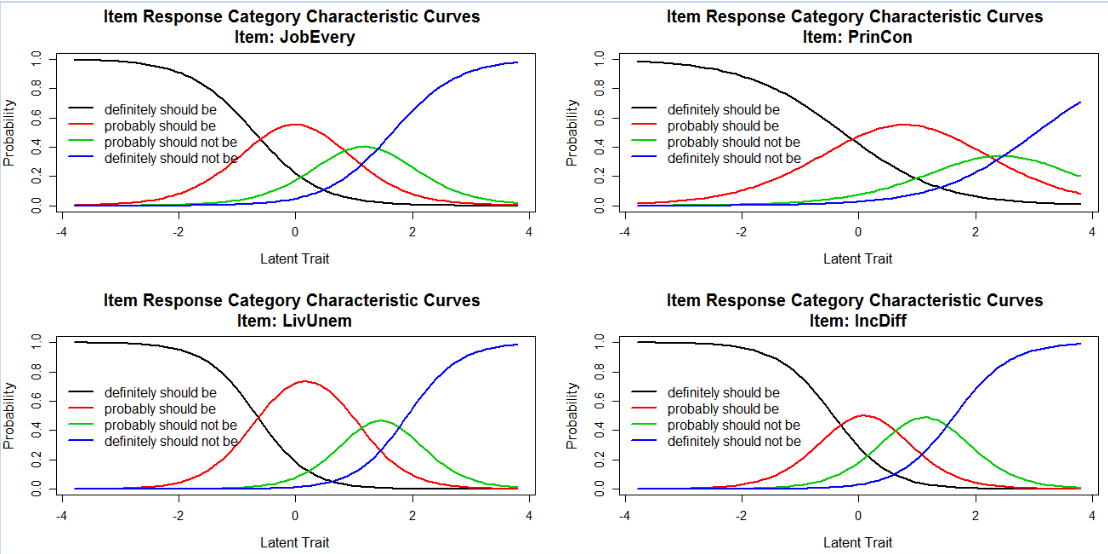


Figure 1.3: Item Characteristic Curves (ICC) for each variable of Example 1.4

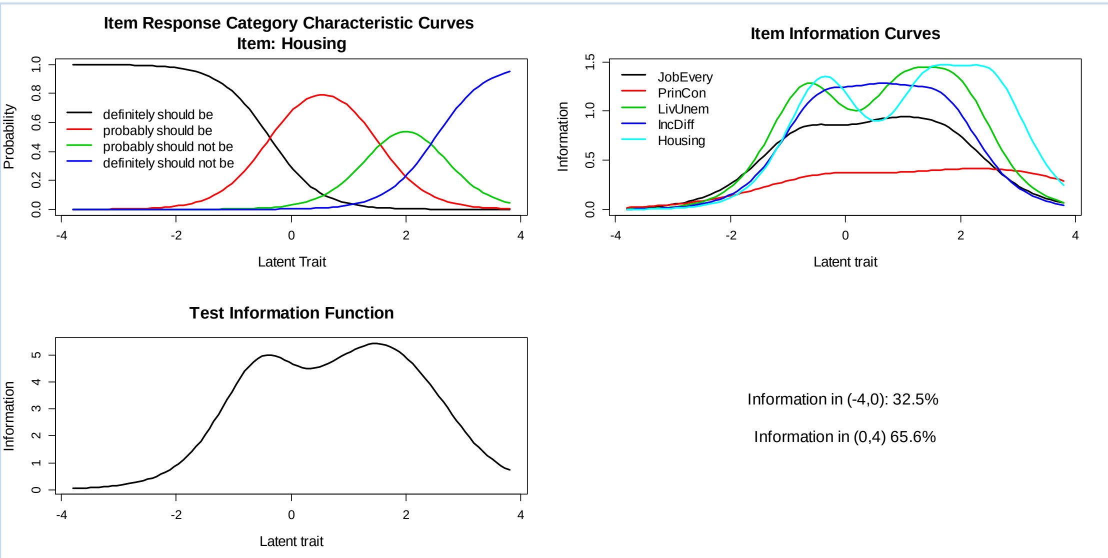


Figure 1.4: Item Characteristic Curve (ICC), Item Information Curve(IIC) for each variable and Test Information Function of Example 1.4

Test Information
Total Information = 22.64
Information in (-4, 4) = 22.19 (98.03%)

Table 1.16: The amount of test information for the fitted model based on all variables of Example 1.4

Item Information (PrinCon)
Total Information = 2.5
Information in (-4, 4) = 2.2 (87.89%)

PrinCon: Prices under control

Table 1.17: The amount of test information for the fitted model based on PrinCon variable of Example 1.4

The variable which records whether the government should or not keep prices under control (PrinCon) provides only $\frac{100 \times 2.5}{22.64} = 11.04\%$ in the total information. This variable (and the corresponding question) can be excluded from a similar future study since its contributions in the total information is minor.

Last but not least, a very useful comparison between items can be achieved by plotting the Item Response Category Characteristic Curves.

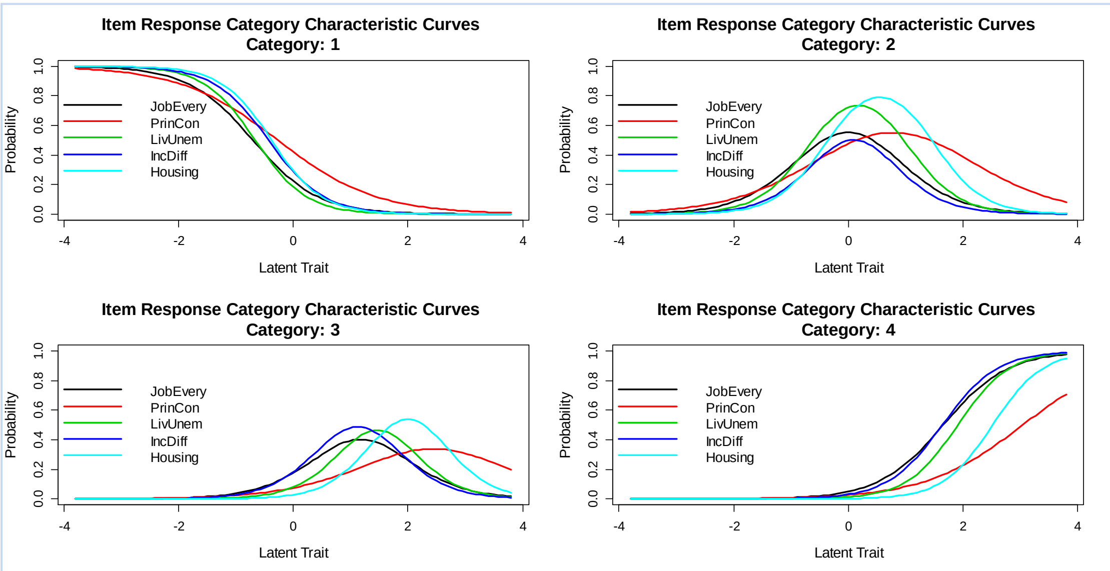


Figure 1.9: Item Response Category Characteristic Curves

Nonw of the items have similar or identical curves for all categories and this indicates that the categories clearly have not the same effect on the configuration of the attitude to the role of the government.

1.5 Discussion

The goal of the first chapter was to present the theoretical basis of this thesis through classical statistical approach. The general theoretical frame of the latent variable models was presented. The next step was to limit the general frame into special case of the latent trait models. Also, the binary and the ordinal case for the latent trait models were presented. The application was based on the field of latent trait models for ordinal data under the IRT approach, which is the main issue of this dissertation. The R package “ltm” was used for the example of this chapter.

In Chapter 2, the Bayesian approach will be presented. Under this approach latent models will also be explained. In the same illustrative example as in Section 1.4 we fit the corresponding Bayesian models using WinBugs and we compare results.

CHAPTER 2: BAYESIAN MODELS FOR LATENT VARIABLES

2.1 Bayesian Statistics

The blossom of Bayesian Statistics was at the late years of the 20th century. Until then, Bayesian Statistics were only an interesting alternative mathematical approach to the classical mainstream statistics. In classical statistics, the under estimation are considered fixed unknown quantities. On the other hand, in Bayesian statistics these parameters are considered as random variables and are characterized by a prior distribution. The combination of this prior with the classical likelihood leads to the posterior distribution of the parameters of interest on which the statistical inference is based in the Bayesian paradigm.

Similarly to any scientific approach the Bayesian approach, has both advantages and disadvantages. Its main advantages are that it is based on pure probability theory and can incorporate information from previous studies or experts via the prior distribution. However, the Bayesian approach was also criticized for the subjectivity which may be introduced via the prior distribution. Moreover, difficulties arise in the computation and the interpretation of the posterior distribution. Nevertheless, the Bayesian approach mimics the human logic. As a human can change his mind when the circumstances are changed, the same procedure is applied in the Bayesian approach.

In order to compute the posterior distribution we use the Bayes' Theorem:

$$f(\boldsymbol{\alpha} \mid \mathbf{y}) = \frac{f(\mathbf{y} \mid \boldsymbol{\alpha}) f(\boldsymbol{\alpha})}{f(\mathbf{y})}$$

$$f(\boldsymbol{\alpha} \mid \mathrm{y}_1, \ldots, \mathrm{y}_n) \propto \prod_{i=1}^{n} f(\mathrm{y}_i \mid \boldsymbol{\alpha}) f(\boldsymbol{\alpha}) \qquad (2.1\text{a})$$

where $f(\mathbf{y} \mid \boldsymbol{\alpha})$ is the likelihood of the model and contains the available information provided by the observed sample, $f(\boldsymbol{\alpha})$ is the prior distribution of parameters and $f(\mathbf{y})$ is the normalizing constant with respect to parameters and has computing difficulties. In other words, the posterior is proportional to the likelihood multiplied by the prior:

$$\text{Posterior} \propto \text{Likelihood} \times \text{Prior} \qquad (2.1b)$$

The prior distribution expresses the information which is available to the researcher before any data set is involved in the statistical analysis and its specification is important because it influences the posterior inference. Often, it is enough to specify the prior mean and the prior variance. Through the prior mean we obtain a point estimate or guess for the parameter of interest. The uncertainty of the estimate is expressed via prior variance. When a priori it is believed that this estimate is accurate then the prior variance must be set low, whereas ignorance or uncertainty about the prior mean can be expressed by a large prior variance. When the prior information is available then it should be represented via the prior distribution. The procedure of extracting information is called *elicitation of prior knowledge* from experts and transform it in prior distribution (Ntzoufras, 2009 p.5).

Unfortunately, usually no prior information is available. In such cases, we have to identify a prior distribution which will not influence the posterior inference and "let the data speak for themselves"; (Ntzoufras, 2009, p.:5). These distributions are called *noninformative, low-information* or *vague prior distributions*. A usual noninformative improper prior distribution is the uniform prior distribution over the parameter space i.e. $f(\boldsymbol{\alpha}) \propto 1$. Thos distribution is called "improper" since it does not integrate to one and can be used without any problem given that the resulting posterior will be proper.

Summary measures such as the moments of the posterior distribution can be used in order to infer, taking into consideration the uncertainty of the parameter $\boldsymbol{\alpha}' = [\alpha_1, \alpha_2, ..., \alpha_n]$. Measures of central location such as the posterior mean, median or less frequently the mode can be used as point estimations, whereas the $q/2$ and $1-q/2$ posterior quantiles can be used as $(1-q)100\%$ posterior credible intervals providing a Bayesian alternative to classical confidence intervals. The main difference between the credible and the confidence interval is its interpretation. For instance, a 95% confidence interval for a parameter α means that: if we construct 100 confidence intervals, we expect 95 out of 100 intervals to contain the true value of α. However, in a 95% credible interval, the probability that α will lie within the credible interval is equal to 95%.

Due to the Bayes rule, it is possible to infer for any parameter α of interest, even when the observed data are collected sequentially at different time-points. So, the Bayesian theory, provides an easy instrument to update the knowledge as far as the parameter α of interest.

To sum up, in a Bayesian model, we have to fully specify both the prior distribution and the likelihood. Then we focus on the description of the posterior distribution using descriptive measures and density plots. Summarizing the whole procedure, it can be divided into four main steps:

1. ***Model building***: take into consideration a model (prior, likelihood, parameters) with reasonable assumptions, appropriate to the conditions of the survey.
2. ***Calculation of the posterior distribution*** which can be achieved with suitable computational methods.
3. ***Analysis of the posterior distribution****:* the analysis is made via descriptive measures, plots and credible intervals.
4. ***Inference***: conduction of conclusions concerning the problem in hand.

Of course, after these steps, diagnostic tests must be applied, as far as the appropriateness of the adopted model. Furthermore, we have to keep in mind the robustness of the posterior distribution which can be monitored via the *sensitivity analysis,* where we evaluate differences of the posterior distribution over different prior choices.

The Bayesian approach also provides a realistic theoretical frame for the prediction of the future observations via the *predictive distribution*. This distribution is equivalent to the fitted values in classical approach but in this case we directly deal with an associated distribution. The predictive distribution is also a useful tool for checking the model rather than a single predicted value and its goodness of fit. For more details see Section 2.2.

The target posterior distribution is not always tractable. In the past (at 1970s) this intractability was surpassed through *conjugate prior distributions*. These priors are characterized by the following property: both prior and the posterior belong to the same distributional family. Later, (at 1980s) the difficulty was overcome via asymptotic approximations of the posterior. In 1990 Markov Chain Monte Carlo methods were introduced to the literature. Using these methods we can obtain samples from the posterior without its direct calculation (Gelfand and Smith, 1990; Gelfand et al., 1990).

The idea of the MCMC methods is to generate a random sample from this distribution and estimate the posterior. The estimation can be achieved using posterior summaries (posterior mean or variance), plotting marginal posteriors even estimating posterior dependencies through sample correlations.

The methodology behind MCMC methods is relatively straightforward. We construct a Markov chain which has as a stationary distribution, the posterior distribution of interest. Every iteration of the algorithm depends only on the previous one. Finally, we can use this chain to generate a sample from the stationary target posterior distribution. The most famous MCMC methods are the *Metropolis-Hastings Algorithm* and the *Gibbs Sampling*. For more details see Ntzoufras (2009, Chapter 2).

2.2 Sampling from Posterior Distribution

Bayesian inference is based on the posterior distribution of the model parameters. Unfortunately, the form of this distribution is rarely known form. Usually, the posterior is available up to a constant [see equations 2.1a & 2.1b]. Markov Chain Monte Carlo techniques are implemented in order to obtain samples from the posterior which are used for the estimation of the posterior distribution and its summaries.

The basic idea belongs to Metropolis et al. (1953). Metropolis proposed to construct a irreducible, aperiodic Markov chain whose the stationary distribution is the posterior distribution. If the chain "runs" for sufficiently long time, the resulting simulated values are obtained from the posterior. After Metropolis algorithm, many MCMC samplers have been developed which applied in various problems (see Dellaportas et al. 2001).

Most famous algorithms are Metropolis-Hastings algorithm (a generalization of Metropolis) and the Gibbs sampler. In the later, we sample iteratively and sequentially from the conditional posterior distributions of each parameter component

α_j given the rest of the parameters. So, the candidate values are sampled directly from the full conditionals instead of using the proposal density. Gibbs algorithm can be also considered as a special case of the Metropolis-Hastings algorithm when the proposal is set equal to the conditional posterior resulting to an acceptance probability equal to one.

Furthermore, when the full conditionals are not fully available but only up to a constant then the candidate values for each parameter component can be again sampled from a proposal density. This is the Metropolis-within-Gibbs algorithm in which the Metropolis step is implemented for each conditional posterior used in the Gibbs sampler.

2.3 Bayesian Model Assessment

2.3.1 Bayes Factor (BF)

The assessment and the check of the goodness of fit in a Bayesian model can be achieved in various ways. One of them, is the implementation of *measures of surprise*. Via these measures, it is possible to quantify the degree of disagreement between the data and the under assessment model, without specifying alternative models. Measures of surprise are the traditional p-values, which via Bayesian approach can be modified to prior predictive p-values, posterior predictive p-values, conditional predictive p-values and partial posterior predictive p-values. Moreover, future observations are considered either through *prior predictive distribution*:

$$f(\mathbf{y}) = \int f(\mathbf{y} \mid \boldsymbol{\alpha}) f(\boldsymbol{\alpha}) \, \mathrm{d}\boldsymbol{\alpha} \qquad (2.3)$$

or via *posterior predictive distribution*:

$$f(\mathbf{y}' \mid \mathbf{y}) = \int f(\mathbf{y}' \mid \boldsymbol{\alpha}) f(\boldsymbol{\alpha} \mid \mathbf{y}) \, \mathrm{d}\boldsymbol{\alpha} \qquad (2.4)$$

where $\mathbf{y}'$ are unobserved observations, i.e. the future data. The posterior predictive distribution is the likelihood of the future data, averaged over the posterior distribution.

In the latent variable models, the assessment of the models is carried out via the posterior predictive distribution. Also (2.4) can be applied in latent variable models for categorical responses: item response models (see Sinharay 2005 & Sinharay et al. 2006).

The prior predictive distribution (2.4) is applied so as to calculate the *Posterior Odds* (PO). The PO is defined to be the ratio of the posterior odds of two competing models m_1 and m_2 multiplied by their corresponding prior odds:

$$PO_{12} = BF_{12}\frac{f(m_1)}{f(m_2)} = \frac{f(m_1 \mid \mathbf{y})}{f(m_2 \mid \mathbf{y})} \qquad (2.5)$$

$$\text{where } BF_{12} = \frac{f(\mathbf{y} \mid m_1)}{f(\mathbf{y} \mid m_2)} \qquad (2.6)$$

BF_{12} is the Bayes Factor (BF) of model m_1 against model m_2 and is defined as the ratio of the marginal likelihoods $f(\mathbf{y} \mid m_1)$ and $f(\mathbf{y} \mid m_2)$. The Bayes factor "plays" an important role in the Bayesian approach. Equal prior model probabilities are usually considered as a default choice when no information is available concerning the structure of the model. When a model comparison is carried out, is desirable to evaluate model m_1 against model m_2. The procedure is similar to a hypothesis testing problem, with hypothesis H_0 is corresponding to model m_1 and the alternative H_1 is corresponding to model m_2. Interest lies in evaluating the null hypothesis H_0. The posterior model odds PO_{12} and the corresponding Bayes Factor BF_{12} evaluate the evidence against H_0 (as classical significance tests). However, PO_{21} and BF_{21} evaluate the evidence in favor of H_0 (this is not attainable in classical significance tests). Moreover, using PO and BF, it is possible to conduct inferences without ignoring the uncertainty of the model and determine which set of explanatory variables gives better predictive results (Ntzoufras 2009, p.:390, Fox 2010, p.:53). Kass and Raftery (1995) suggested to interpret Bayes Factors according to the scale presented in Table 2.1.

Bayes Factor B12	Evidence against model m1
1-3	Negligible
3-20	Positive
20-150	Strong
>150	Very Strong

Table 2.1: Bayes Factor interpretation according to Kass & Raftery (1995)

2.3.2 Deviance Information Criterion (DIC)

The *Deviance Information Criterion (DIC)* has been proposed by Spiegelhalter et al. (2002) for model comparison when the number of parameters is not clearly defined (Fox, 2010 p.: 60). It is given by the equation (2.7):

$$\mathrm{DIC(m)} = 2\overline{D(\boldsymbol{\theta}_{\mathrm{m}},\mathrm{m})} - D(\bar{\boldsymbol{\theta}}_{\mathrm{m}},\mathrm{m}) = D(\bar{\boldsymbol{\theta}}_{\mathrm{m}},\mathrm{m}) + 2p_m \quad (2.7)$$

Where $D(\boldsymbol{\theta}_{\mathrm{m}},\mathrm{m})$ is the measure of model deviance, given by

$$D(\boldsymbol{\theta}_{\mathrm{m}},\mathrm{m}) = -2\log f(\mathbf{y}\,|\,\boldsymbol{\theta}_{\mathrm{m}},\mathrm{m}) \qquad (2.8)$$

Moreover, $\overline{D(\boldsymbol{\theta}_{\mathrm{m}},\mathrm{m})}$ is its posterior mean and p_m can be interpreted as the number of "*effective*" parameters for model *m* (dimensions) given by

$$p_m = \overline{D(\boldsymbol{\theta}_{\mathrm{m}},\mathrm{m})} - D(\bar{\boldsymbol{\theta}}_{\mathrm{m}},\mathrm{m}) \qquad (2.9)$$

where $\bar{\boldsymbol{\theta}}_{\mathrm{m}}$ is the posterior mean of the parameters involved in the model m.

The best model is associated with the smallest DIC value. The main hypothesis when DIC is applied is that the posterior mean can be used as an adequate summary of central location for description of the posterior distribution. DIC must be used very carefully because problems may arise when the posterior distributions are not symmetric or with it is multimodal.

A problem had been arise in the application of the classic DIC when negative dimensions ($p_d < 0$) appeared. This indicates that the posterior mean is a poor summary statistic of central location and as a result we obtain large values of deviance. To surpass this difficulty, it is possible to use the maximum likelihood estimate of the parameters of interest instead of the posterior mean in the case of low information priors; see Gelman 2003. Thus, DIC in that case is computed through:

$$\text{DIC*} = \log(p(y|\hat{\theta})) - p_{DIC} \qquad (2.10)$$

where: $$p_{DIC} = 2\left(\log(p(y|\hat{\theta})) - E_{post}(\log(p(y|\hat{\theta})))\right)$$

In equation (2.10) the posterior mean of the MLEs will coincidence with the maximum log predictive density. This is the equivalent to using the posterior mode in the case of non-informative priors. Thus, p_{DIC} can be estimated via a different approach: $p_{DIC}' = 2\,\mathrm{var}_{post}(\log(p(y\,|\,\theta)))$. Given that $\text{deviance} = -2\log(p(y|\hat{\theta}))$ we can derive alternative general equation for the deviance information criterion:

$$\text{DIC} = -2\log(p(y|\hat{\theta})) + 2p_{DIC} \text{ where } p_{DIC} = 2\,\mathrm{var}_{post}(\log(p(y\,|\,\theta))) \qquad (2.11)$$

From (2.11) the following equations are derived

$$\text{DIC}_1 = \min\{\text{deviance}\} + \mathrm{var}\{\text{deviance}\} \qquad (2.12)$$

$$\text{DIC}_2 = \text{mean}\{\text{deviance}\} + 0.5\,\mathrm{var}\{\text{deviance}\} \qquad (2.13)$$

2.4 Bayesian Models for Latent Variables

Latent variable models use information available from the manifest variables in order to extract knowledge about the unobserved (latent) part. In other words, the joint distribution of the manifest variables is applied so as to quantify and assess the distribution of the latent ones. This is fulfilled through the Bayes Theorem:

$$f(\mathbf{y}\,|\,\mathbf{x}) = \frac{f(\mathbf{x}\,|\,\mathbf{y})f(\mathbf{y})}{f(\mathbf{x})} \qquad (2.14)$$

Where $\mathbf{x}$ and $\mathbf{y}$ represent the manifest and latent variables respectively with $\mathbf{x}' = [x_1, x_2, \ldots x_p]$, $\mathbf{y}' = [y_1, y_2, \ldots y_q]$ and q<p. So, the Bayes theorem is used in order to estimate the latent variables. It is obvious that latent variables are manipulated into the

Bayesian field of statistics, starting from a prior density distribution $f(\mathbf{y})$ and closing to the posterior $f(\mathbf{y} \mid \mathbf{x})$. So, there is not a purely classical approach for the latent variable models. The approach is actually, partially Bayesian or entirely Bayesian. The difference is on the way the vector of the parameters $\boldsymbol{\alpha}$ is treated. In the purely Bayesian approach, the parameter vector $\boldsymbol{\alpha}$ is stochastic and is associated with a prior distribution.

2.4.1 Specification of Prior Distribution

The posterior distribution is proportional to the product of the likelihood times the prior. The prior can be either informative (subjective) or non- informative (vague). Thus, the posterior quantities are directly associated with the specification of the prior. When we wish to use a non-informative prior then flat improper or proper distributions are used with large variances in order to express our uncertainty about the parameters. In such cases, the influence of the data becomes dominant since the likelihood contributes more in the structure of the posterior than the prior. As a result, the posterior estimates are closer to the corresponding maximum likelihood ones. When the parameter space is discrete, a discrete uniform distribution may be used to express ignorance. This is known as the *principle of insufficient reason* (Hans-Werner Sinn, 1980, p.493). Equivalently, when the parameter space is continuous, then flat (and sometimes improper) prior was used instead.

In the context of latent variable models, the prior distribution plays an important role for an additional reason: It ensures the uniqueness of the solution. Truncated or constant priors are applied in order to choose item parameters in a similar way that constraints are imposed in order to fix rotation, in the classical approach. As far as the latent variables, independent standard normal distributions are used as a standard default choice.

In the Item Response Theory (IRT) there are two scientific schools. The first school suggests to apply the probit response (for more information, see Mislevy, 1986). In this approach, a conjugate or conditional conjugate prior exists that facilitates the Bayesian implementation. To be more specific, normal priors are used for the difficulty parameters, whereas truncated normal priors for the discrimination parameters which must be positive and beta priors for the guessing parameters. The second school prefers the logistic regression approach (see Patz and Junker, 1999a & 1999b).

Similar priors are used in models with multilevel structure either on the ability parameters (see Fox & Glas, 2001) or on the item parameters (see Janssen et al., 2000), on person fit analysis IRT models (see Glas & Meijer, 2003). In the logistic IRT models (second school), there are no priors which lead to conjugate forms therefore MCMC techniques are used instead. Generally, normal $N(0,\sigma^2_{\alpha i0})$ priors are used for the difficulty parameter and lognormal priors $LN(0,\sigma^2_{\alpha i1})$ for the discrimination parameter with close choices for the prior variances (see equation (1.16)).

Four criteria exist for the construction of the prior distribution:

a. The prior distribution should be non informative but proper in order to be able to compute the Bayes Factors.
b. Constraints should be imposed in order to achieve unique solution.
c. The prior distribution should be suitable for Bayesian model comparison.
d. The rior distribution should be potentially generalized to other members of the Generalized Linear Latent Variable Models (GLLVM).

As far as the unconstrained discrimination and difficulty parameters normal priors are considered to be appropriate for the model parameter vector $\boldsymbol{\alpha}' = [\alpha_1, \alpha_2, ..., \alpha_k]$, Fouskakis et al. (2009) suggest a normal prior of the general form:

$$f(\boldsymbol{\alpha}) = N(0, \boldsymbol{\Sigma}) \qquad (2.15)$$

where $\boldsymbol{\Sigma} = \mathrm{N}[\mathrm{I}(\boldsymbol{\alpha})]^{-1}$ is the prior covariance matrix, N the total sample size and $\mathrm{I}(\boldsymbol{\alpha})$ is the information matrix:

$$\mathrm{I}(\boldsymbol{\alpha}) = \mathbf{X'WX} \qquad (2.16)$$

The matrix $\mathbf{W}$ is diagonal and its form depends on the link function. In the same paper, it is stated that in the absence of prior information, the probability of correct response can be denoted a-priori equal ½. Then, (2.15) is transformed to:

$$f(\boldsymbol{\alpha}) = N(0, 4\,\mathrm{N}[\mathbf{X'X}]^{-1}) \qquad (2.17)$$

For the multivariate case the prior discrimination parameters is summarized in the equation (2.18):

$$a_{ij} = \begin{cases} 1 & \text{if } i<j \\ LN(0,1) & \text{if } i=j \\ N(0,4) & \text{if } i>j \end{cases} \qquad \text{where i=1,..,p and j=1,...,q} \quad (2.18)$$

As far as the prior difficulty parameters, they follow normal distribution with prior mean equal to zero in order to express our ignorance and high prior variance so as to express our uncertainty (Vitoratou, 2013).

2.4.2 Sampling from the Posterior Distribution

In order to sample from the posterior distribution and estimate model parameters (1.16), the Metropolis-within-Gibbs (MG) sampler is used (Patz and Janker, 1999b) with stationary distribution $f(\boldsymbol{\alpha}, \mathbf{y} \mid \mathbf{x})$. Before MG algorithm is applied, we must consider some main points:

1. In order to construct an efficient algorithm which will fastly converge the true posterior , the parameters should be grouped in blocks.

   This methodology is applied in high dimensional problems and minimizes the required computational time (Chib and Greenberg, 1995). The general rule in the construction of the blocks is to group together parameters that are expected to be a posteriori dependent. Thus, one block is created for each item and one for each individual (Patz and Janker, 1999b). In a q-factor model the parameter components that are updated (accepted or not) simultaneously are the p components $\alpha_i = \{\alpha_{i0}, \alpha_{i1}, \dots, a_{iq}\}$ and N components $y_i = \{\mathrm{y}_{i1}, \mathrm{y}_{i2}, \dots \mathrm{y}_{iq}\}$.

2. The choice of the proposal density.

The future candidate points are generated by distributions centered at the current state (Patz and Janker, 1999b). To be more specific, normal proposal distributions are used for the latent variables $\mathbf{y}_i$:

$$\pi(\mathbf{y}'_i \mid \mathbf{y}_i) = \prod_{l=1}^{q} \pi(y'_{il} \mid y_{il}) \text{ where } \pi(y'_{il} \mid y_{il}) = N(y_{il}, c_y^2) \text{ with } i = 1,\ldots,N \quad (2.19)$$

For the item difficulties:

$$\pi(\alpha'_{i0} \mid a_{i0}) = N(\alpha_{i0}, c_a^2) \text{ with } i = 1,\ldots,p \qquad (2.20)$$

For the item discriminations:

Log-normal proposal distributions are used. When $q > 1$, i.e. it is the multivariate case, the log-normal proposals are considered to be the diagonal elements of the loadings matrix and normal proposal distributions for the unconstrained elements:

$$\pi(\alpha'_{il} \mid \alpha_{il}) = \begin{cases} LN(\log \alpha_{il}, c_a^2) \text{ if } i = l \\ N(\log \alpha il, c_a^2) \text{ if } i > l \end{cases} \qquad (2.21)$$

The variance of the proposal density is called *tuning parameter* (Fox, 2010, p.:84) because it affects the acceptance rate of the MCMC algorithm. The recommended acceptance rate for univariate parameter is about 50% and for higher dimensional blocks 25% (Gelman et al. 1996). In conclusion, an advantage of General Linear Latent Trait Models is that the acceptance probabilities are simplified directly, due to the prior and local independence assumptions.

2.5 Example in the Bayesian Latent Variable Models for Ordinal Data

The data set which used for this application consists of five ordinal manifest variables and 822 observations after excluding all missing values (NAs). The manifest variables measure the attitudes to the role of government; see Moustaki (2003). Responders were asked if they consider government's responsibility to:

1. provide a job for everyone who wants one (variable JobEvery)
2. keep prices under control (variable PrinCon)

3. provide a decent standard of living for the unemployment (variable LivUnem)
4. reduce income differences between the rich and the poor (variable IncDiff)
5. provide decent housing for those who cannot afford it (variable Housing)

After constructing the model based on equations (1.16) and (1.19) for one factor, we have run the MG algorithm for 11,000 iterations having considered the first 1,000 as burn-in values. Three different link functions were used in equation (1.16) in each case: the logit, the c-loglog and the probit. The final model is selected using DIC. We have visually checked convercence using trace plots, ergodic mean plot and the autocorrelation plots. Moreover, we have used package "CODA" in R to formally check the convergence of the chain ; see Ntoufras, 2009. All diagnostic test, were passed in this example.

2.5.1 One Factor Latent Variable Models

From Table 2.2, we observe that the Probit model is the worst due to its high DIC value (equal to 8010). The Logit and the C-loglog were close to each other. In the following we focus on the interpretation of Logit and c-loglog model.

DIC	
Logit Model	7959
C-loglog Model	7924
Probit Model	8010

Table 2.2: The assessment of the models via DIC of Example 2.5

2.5.1.1 The Logit Model

Figure 2.1 depicts the visual diagnostic convergence tests for the first discrimination parameter. While the plot of the trace plot should be within a sensible range of value without trends and seasonalities, the ergodic mean plot must stabilized after few iterations. This is clearly happens here, with diagnostic mean to be close to 1.75. Similarly, the autocorrelation plots are fastly deteriorating and tend to zero. There for there is no indication against the convergence to the true posterior.

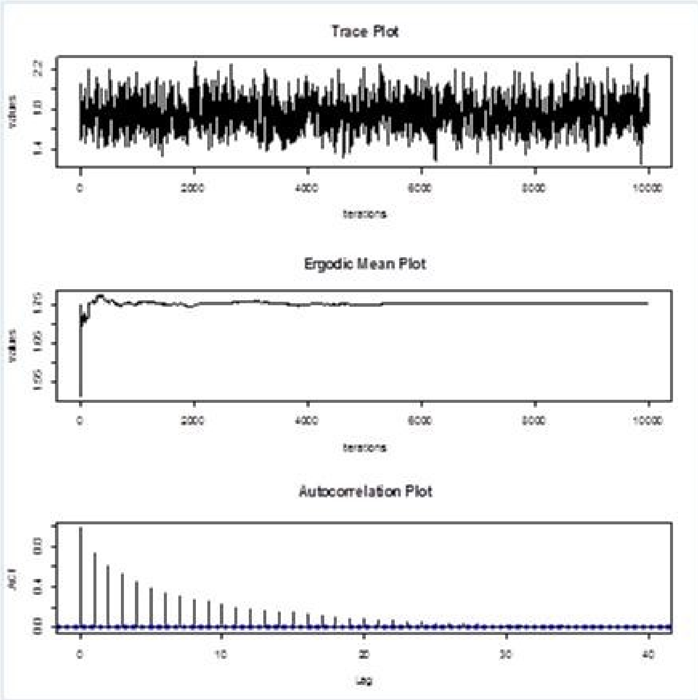


Figure 2.1: Diagnostic Tests for the simulated chain of logit Model for the first discrimination parameter of Example 2.5

The estimated discrimination and difficulty parameters with their standard errors and Monte Carlo errors are provided in Table 2.3. The discrimination parameters are all positive. So, we can assume one common factor for all variables. In that point, taking into consideration this assumtion, someone can consider this latent variable as the attitude of citizens as far as the protection of the lowest income from the government (in favor of or against). The variable which denotes whether the government should provide or not decent housing for those who cannot afford it (Housing) makes the clearest discrimination between a positive and a negative attitude. On the contrary, the variable concerning the government control of the prices (PrinCon) discriminates this attitude in the least way, by far. We have keep in mind that the standard errors of discrimination parameters are not small in relation to the respective estimations. As far as the Monte Carlo errors (MC), it is clear that they are low comparison to the corresponding estimated posterior standard deviations. Thus, the estimated posterior mean has been estimated with high precision. In order to decrease the Monte Carlo error the number of iterations must be increased (Ntzoufras, 2009).

Discrimination Param.	Mean	Sd	MC error
JobEvery	1.757	0.1419	0.005
PrinCon	1.144	0.1119	0.005
LivUnem	2.227	0.1892	0.007
IncDiff	2.082	0.1691	0.006
Housing	2.298	0.2076	0.008

JobEvery: Job for everyone, PrinCon: Prices under control, LivUnem: Standard of living for the unemployed, IncDiff: Reduce income differences, Housing: Decent housing

Table 2.3: Discrimination Parameters of each variable (10000 iterations-1000 burn-in) for the logit model of Example 2.5

Unfortunately, factor loadings as discriminations parameters cannot be considered as correlation coefficients as in factor analysis. Alternative we can use equation (1.17) to obtain the standardized discrimination parameters which have similar interpretation.

From table 2.4 we observe that that all standardized discrimination parameters are close to one. This indicates a strong link between the common latent variable and the manifest variable. Again the variable which explores the attitude towards government control prices appears the weakest link (0.75). On the other hand, the strongest link (0.92) is between the latent and the potential of the government to offer decent housing to people who cannot afford it.

Standardized disc. Parameters		Sd
JobEvery	0.87	0.017
PrinCon	0.75	0.032
LivUnem	0.91	0.013
IncDiff	0.90	0.014
Housing	0.92	0.013

JobEvery: Job for everyone, PrinCon: Prices under control,
LivUnem: Standard of living for the unemployed,
IncDiff: Reduce income differences, Housing: Decent housing

Table 2.4: Standardized Discrimination Parameters of each variable for the logit model of Example 2.5

From Table 2.5, it is clear that for every variable the ordinality is preserved. The category "definitely should be" is the less "difficult" and the most "difficult" is the third category.. In order to understand better how the difficulty parameter influences the positive response probability we will use equation (1.18), to study how a "median" individual behaves. Furthermore, response probabilities from each variable are also presented in Table 2.6.

Variables	Categories	Mean	Sd	MC error
JobEvery	1	-0.715	0.073	0.002
	2	0.705	0.071	0.003
	3	1.683	0.115	0.004
PrinCon	1	-0.278	0.081	0.002
	2	1.899	0.172	0.007
	3	3.156	0.293	0.012
LivUnem	1	-0.670	0.066	0.002
	2	1.023	0.072	0.002
	3	1.940	0.118	0.004
IncDiff	1	-0.435	0.061	0.001
	2	0.628	0.063	0.002
	3	1.662	0.105	0.003
Housing	1	-0.383	0.059	0.002
	2	1.497	0.092	0.003
	3	2.560	0.169	0.005

JobEvery: Job for everyone, PrinCon: Prices under control, LivUnem: Standard of living for the unemployed, IncDiff: Reduce income differences, Housing: Decent housing. Categories: 1:definitely should be, 2: probably should be, 3: probably should not be, 4:reference category

Table 2.5: Difficulty parameters for each variable and category for the logit model of Example 2.5

Variables	Categories	Cumulative Probability $\gamma_{i(s)}(y=0)$	Level Probability $\pi_{i(s)}(y=0)$
JobEvery	1	0.33	0.33
	2	0.67	0.34
	3	0.84	0.17
	4	1.00	0.16
PrinCon	1	0.43	0.43
	2	0.87	0.44
	3	0.96	0.09
	4	1.00	0.04
LivUnem	1	0.34	0.34
	2	0.74	0.40
	3	0.87	0.13
	4	1.00	0.13

JobEvery: Job for everyone, PrinCon: Prices under control, LivUnem: Standard of living for the unemployed, Categories: 1:definitely should be, 2: probably should be, 3: probably should not be, 4: definitely should not be (reference category)

Table 2.6a: Cumulative and level probability for each variable and category for the logit model of Example 2.5

Variables	Categories	Cumulative Probability $\gamma_{i(s)}(\mathbf{y}=\mathbf{0})$	Category Probability $\pi_{i(s)}(\mathbf{y}=\mathbf{0})$
IncDiff	1	0.39	0.39
	2	0.65	0.26
	3	0.84	0.19
	4	1.00	0.16
Housing	1	0.41	0.41
	2	0.82	0.41
	3	0.93	0.11
	4	1.00	0.07

IncDiff: Reduce income differences, Housing: Decent housing. Categories: 1:definitely should be, 2: probably should be, 3: probably should not be, 4:reference category

Table 2.6b: Cumulative and category probability for each variable and category for the logit model of Example 2.5

Due to the definition of our model (see equation 1.16) cumulative probabilities $\gamma_{i(s)}(\mathbf{y})$ are considered as "failure probability". So, the category "probably should not be" outstands from the others in every variable as the most likely non-response category from a "median" individual. For each variable, the rest of the categories appear to have lower probabilities than those of the third category. In other words, the ordinality is preserved and a "median" individual is positive inclined. Moreover, from the third column of Table 2.6 it is clear that for a "median" individual has higher probability to respond positively to such questions (the third and forth categories appear extremely low probabilities in relation to the others in every variable). Also, an individual at the median gives more easily a positive response to the question if the government should provide a job for everyone and to the question if the government should keep the prices under control.

2.5.1.2 The Complementary Loglog Model

Similarly to Figure 2.1, the trace plot is within a sensible range of the first discrimination parameter without trends and seasonalities. The ergodic mean plot has been stabilized close to 1.20. From the autocorrelation plot it is obvious that autocorrelations tend to zero. Consequently there is no indication again the convergence to the true posterior.

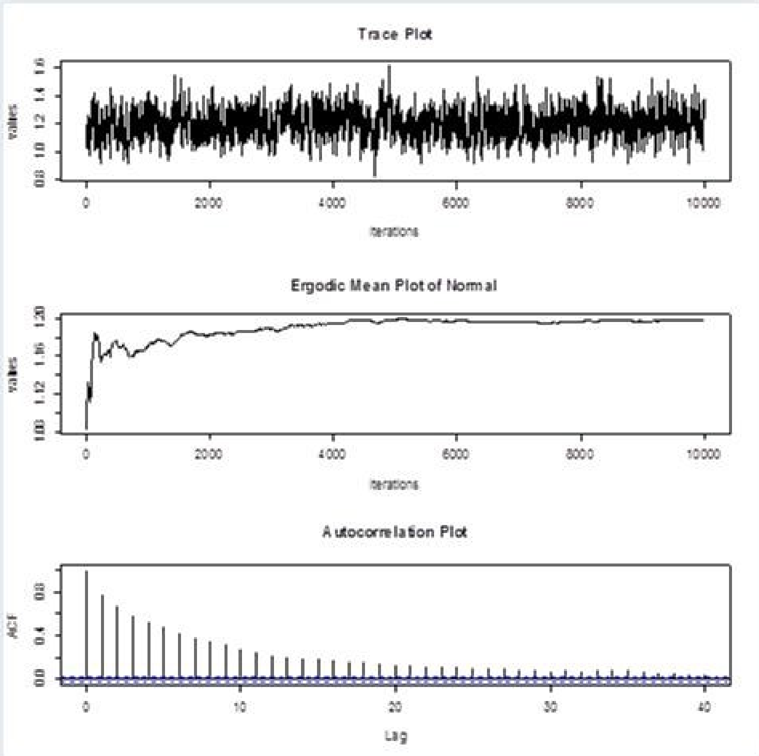


Figure 2.2: Diagnostic Tests for the simulated chain of C-loglog Model for the first discrimination parameter

The estimated discrimination and difficulty parameters with their standard errors and Monte Carlo errors are presented below. The discrimination parameters are all positives again. So, similar to the logit model we can suppose the existence of one common factor for all variables. The variable which denotes whether the government should provide or not decent housing for those who cannot afford it (Housing) makes clearer the discrimination between a positive and a negative attitude. On the other hand, the variable which denotes if the government should keep or not prices under control (PrinCon) distinguish this attitude in the least way (PrinCon variable has the lowest discrimination value). Furthemore, the standard errors of discrimination parameters are not small in relation to the respective estimations. Concerning the Monte Carlo errors (MC), they are low in comparison to the corresponding estimated posterior standard deviation. As a consequence, the estimated posterior mean has been estimated with high precision.

Discrimination Param.	Mean	Sd	MC error
JobEvery	1.119	0.097	0.004
PrinCon	0.809	0.074	0.003
LivUnem	1.491	0.123	0.005
IncDiff	1.445	0.117	0.005
Housing	1.663	0.165	0.007

JobEvery: Job for everyone, PrinCon: Prices under control, LivUnem: Standard of living for the unemployed, IncDiff: Reduce income differences, Housing: Decent housing

Table 2.7: Discrimination Parameters of each variable for the c-loglog model of Example 2.5

Once again, discrimination parameters will be transformed to standardized discrimination parameters through equation (1.17) so as to they are interpreted as correlation coefficients of the common latent variable with each manifest one.

All standardized discrimination parameters are close to one but not as near as those of the logit model. This indicates a strong connection between the common latent variable with each manifest variable. The variable which indicates whether the government should keep or not the prices under control has the weakest link (0.63) with the latent one. On the contrary, the strongest link appears between the latent and the variable which denoted whether the government should provide or not decent housing for people who cannot afford it (0.85).

Standardized disc. Parameters		Sd
JobEvery	0.77	0.026
PrinCon	0.63	0.035
LivUnem	0.83	0.022
IncDiff	0.82	0.022
Housing	0.85	0.023

JobEvery: Job for everyone, PrinCon: Prices under control, LivUnem: Standard of living for the unemployed, IncDiff: Reduce income differences, Housing: Decent housing

Table 2.8: Standardized Discrimination Parameters of each variable for the c-loglog model of Example 2.5

The estimated difficulty parameters for each variable and category, their standard errors and Monte Carlo errors are below, in Table 2.9

Variables	Categories	Mean	Sd	MC error
JobEvery	1	-0.307	0.062	0.002
	2	1.096	0.085	0.003
	3	2.214	0.152	0.006
PrinCon	1	0.255	0.074	0.002
	2	2.582	0.221	0.009
	3	4.174	0.383	0.015
LivUnem	1	-0.345	0.058	0.002
	2	1.381	0.091	0.004
	3	2.446	0.162	0.006
IncDiff	1	-0.091	0.056	0.002
	2	0.949	0.074	0.003
	3	2.087	0.133	0.005
Housing	1	-0.065	0.054	0.005
	2	1.841	0.122	0.009
	3	3.049	0.232	0.001

JobEvery: Job for everyone. PrinCon: Prices under control. LivUnem: Standard of living for the unemployed. IncDiff: Reduce income differences. Housing: Decent housing. Categories: 1:definitely should be. 2: probably should be. 3: probably should not be. 4:reference category

Table 2.9: Difficulty parameters for each variable and category for the c-loglog model of Example 2.5

For the complementary loglog (cloglog) link function the model (1.16) is replaced by:

$$\log(-\log(1-\gamma_{i(s)}(\mathbf{y}))) = a_{i(s)} + \sum_{j=1}^{q} a_{ij} y_j \quad (2.22) \quad \text{where s=1….. m}_i\text{ -1 and i=1…..p}$$

For **y**=0, the probability of positive response from a median person $\gamma_{i(s)}(\mathbf{y}=0)$ is given by the following equation:

$$\gamma_{i(s)}(\mathbf{y}=0) = \mathrm{e}^{-e^{a i(s)}} \qquad (2.23)$$

Variables	Categories	$\gamma_{i(s)}(\mathbf{y}=0)$	$\pi_{i(s)}(\mathbf{y}=0)$
JobEvery	1	0.42	0.42
	2	0.75	0.33
	3	0.90	0.15
	4	1.00	0.10
PrinCon	1	0.56	0.56
	2	0.93	0.37
	3	0.98	0.05
	4	1.00	0.02
LivUnem	1	0.41	0.41
	2	0.80	0.39
	3	0.92	0.12
	4	1.00	0.08
IncDiff	1	0.48	0.48
	2	0.72	0.24
	3	0.89	0.17
	4	1.00	0.11
Housing	1	0.48	0.48
	2	0.86	0.38
	3	0.95	0.09
	4	1.00	0.05

JobEvery: Job for everyone. PrinCon: Prices under control. LivUnem: Standard of living for the unemployed. IncDiff: Reduce income differences. Housing: Decent housing. Categories: 1:definitely should be. 2: probably should be. 3: probably should not be. 4: definitely should not be (reference category)

Table 2.10: Cumulative and level probability for each variable and category for the c-loglog model of Example 2.5

Similar to the logit model. due to its definition; see equation (2.23), $\gamma_{i(s)}(\mathbf{y})$ considered as “failure probability”. So. the category “probably should not be” sticks out from the others in every variable as the most likely non-response category from a “median” individual. In each variable. the other categories appear lower probabilities than those of the third category. In other words. the ordinality is preserved and a “median” individual is positive inclined.

Both models, the logit and the c-loglog give results with the same interpretation but the numeric results in the logit model are higher compared to those of the c-loglog model. Even if the “oldest” link function is the complementary c-loglog (Fisher, 1922) and works best with extremely skewed distribution, the logit function (Berkson, 1944) is preferred because it leads to simpler mathematics due to complexity of the standard normal cumulative distribution function and it is easier to be interpreted.

2.6 Discussion

The goal of this chapter was to present the classic statistical approach and the Bayesian paradigm for the latent variable models with ordinal data, under the IRT approach. The essentials of Bayesian theory was presented such as inference and the estimation of the posterior distribution based on MCMC. Model via Bayes Factor and Deviance Information Criterion is of great interest.

The implementation of Bayesian theory on latent variable models is presented in detail. In that step. Both the prior and the posterior estimation for latent variable were presented. Finally, the Bayesian approach was implemented (via WinBungs). The same data set as in chapter 1 was used in order to make results comparable.

In the real data application three different link functions were used: the logit. the probit and the complementary log-log. The results-estimations of probit model were excluded from further analysis due to lower DIC values.

CHAPTER 3: Schizotypy and Consumer Behavior

The aim of this chapter is to detect whether relationships among consumer behaviors can be characterized as impulsive or compulsive and and whether these can be further associated with the nine subscales of schizotypy.

3.1 Consumer Behavior

3.1.1 Definition of Consumer Behavior

The term *consumer behavior*, refers to a person's behavior with reference to his purchase habits and use of products and services. William Wilkie (1994) defined consumer behavior as "*the mental, emotional and physical activities that people engage in when selecting purchasing, using and disposing of products and services so as to satisfy needs and desires*". Similar definitions have been denoted by other researchers such as Siomkos (1994).

Nowadays, consumer behavior does not include only the process of decision making about buying a product but also all further consumer's activities which take place after the purchase of a product or a service. Such activities are the use, the assessment and the rejection of a product or a service.

In order to understand the consumer behavior, we have to take into consideration the factors which influence the decision making process. These factors are the seven main characteristics of consumer behavior (see for details Wilkie, 1994):

1. Motivations
2. Activities
3. Process of Consumer Behavior
4. Diversification of Consumer Behavior in time and complexity
5. Different roles of Consumer Behavior
6. Exogenous factors which influence Consumer Behavior
7. Diversification of consumer's personality and how react on products consumption

Most of the consumers desire to fulfill more than one needs or goals. So, we cannot refer only to a simple motivation, but to a group of them which impels the consumer behavior. Furthermore, some motivations are clear to consumers in converse to others that may not be completely obvious decisions they are based on emotions of the consumer.

A part of a consumer behavior comes from *functional motivations*. For example, when someone buys clothes. Another case arise from *self-expressive motivations*; when someone buys a present for a friend (Wilkie, 1994, p.10).

On the other hand, Blackwell, Miniard and Engel (2001, p. 233-245) claim that the needs of consumers cannot be divided into two groups, but only to subgroups. These subgroups should include and declare different cases of needs and desires of a consumer. Some of them are the need of safety and health, love and comradeship, wealthy and pleasure, the need of creation a social image, the need of information and the need of possessing.

In the near future, consumer behavior will be described by the same way as today. On the contrary, a few changes are expected to happen in the distant future. Furthermore, these changes are expected to be intensified with the passing of time. The behaviors that will concern us are two. Those which are described as impulsive behaviors and those which are characterized as compulsive ones.

3.1.2 Impulsive and Compulsive Consumer Behaviors

The *impulsive consumer behavior* is expressed through a spontaneous buying. This buying is unplanned by default and has the element of impulsion as its main ingredient. It is strong, sudden and almost always irresistible (Beatty and Ferrell, 1998).

According to Blackwell, Miniard and Engel (2001), impulsive buying has the following features:

1. An emerge, sudden and impulsive desire of action.
2. A situation of psychological unbalance, where the consumer may feel temporary out of control.
3. An inner fighting which can calmed down by immediate action.
4. Domination of the feelings and not of the objective logic.

5. The consumer does not take into consideration the consequences of his action.

Alternatively, *compulsive consumer behavior* is expressed through a uncontrollable buying. According to related studies, compulsive buying is directly related to emotions such as anger, sadness and stress which take place in the inner psychological worlds of consumers. The compulsive consumption may last years with repeated, sometimes excessive, episodes. It is developed when the consumer has undergone negative feelings and events. Thus, a situation like compulsive buying may resulting unfortunate psychological and financial effects (O' Guinn and Faber, 1992 and Shoham and Brenic, 2003).

These two types of consumer behaviors have obvious differences but also share some common features. First of all, the psychological mood plays an important role in both behaviors. Mood is the impeller of impulsive and compulsive buying (Dittmar, Beattie and Friese, 1996). Moreover, women are more vulnerable to such behaviors than men. This sounds reasonable due to they are more emotional personalities of woman compared to men.

3.2 Schizotypy

3.2.1 Characteristics of Schizotypy

Rado (1960) introduced the term "schizotype" as the shortening of words "phenotype" and "schizophrenic". It is used to describe the observable propensity of a person to schizophrenia before the outbreak of psychosis.

Two years later, Meehl (1962) connected schizotypy to the presence of a gene which called schizogene and leads to schizotypic personality. Although, schizotypy linked to schizophrenia, only 10% of schizptypic people finally develop the symptoms of the disease. Thus, schizotypy is a necessary but not certain settlement for the development of schizophrenia. In order to show symptoms of the decease, inner and environmental factors take place such as stress and anguish.

Environmental factors have an effect on the disturbance of balance of a schizotype person. These factors are divided into two groups: stressed environmental factors occurred during childhood and during adultness.

The features (or dimensions) of schizotypy are the following:

1. ***Ideas of reference:*** it is related to misinterpretation of certain events which have a special importance for each person.

2. ***Magical thinking-Odd beliefs:*** beliefs which are incompatible to social status, such as superstition, soothsaying ability, telepathy etc.

3. ***Unusual perceptual experiences:*** the feeling of some abstract presence, voice or shadow is close to you.

4. ***Odd speech:*** it is expressed via idiosyncratic phrases or construction of worlds, vagueness speech, compacted or abstract thinking.

5. ***Suspiciousness:*** a constant fear and belief that dangerous thoughts, conspiracies and plans from other people exist and are related to you.

6. ***Constricted affect:*** the inability to adopt and be member of a social group because you feel different.

7. ***Odd behavior:*** the feeling that their behavior is odd or different from others.

8. ***No close friends:*** to the degree of social behavior and the ability to have friends outside the family environment.

9. ***Excessive social anxiety:*** paranoid fears and negative feelings when socializing with a group of people.

The above criteria are called DSM-IV. To sum up, characteristics of schizotypy are an intense failure in interpersonal relations, eccentricities and quirks of thought, perception, behavior, speech and appearance, which are not severe enough to meet criteria for schizophrenia.

In 1991, Raine constructed the SPQ questionnaire of schizotypic personality. It includes nine subscales which represent the nine aspects of a schizotypic personality. The SPQ was translated and used in Greece by the ASPIS team (Stefanis et al, 2002).

3.2.2 Schizotypy and Consumer Behavior

The potential relationship between schizotypy and consumer behavior presents an increased interest which lies in the detection of the effect of psychiatric diseases.

There are two extreme consumer behaviors. The impulsive and the compulsive buying. Compulsive consumption have been found to be closely related to schizotypy. This consumer behavior can be considered parallel with other spontaneous, dependent or extremely compulsive disturbances of human behavior, such as stress, phobia, mental ribs, bulimia nervosa. Schlosser et al (1994) concluded that compulsive buying is a clinically identifiable syndrome, which cumbers patients both psychologically socially.

So, behaviors which are characterized as compulsive (such as compulsive consumer behavior) are faced as clinically syndromes from some researchers (Roth and Baribeau, 2000) and are examined in comparison to schizotypy and its characteristics.

Here we will investigate the connection between the impulsive and compulsive consumer behaviors and the nine features of schizotypy. Additionally, we will consider the four main groups of subscales which are the following:

1. ***Negative Characteristics***: suspiciousness, extremely social stress, lack of close friends.

2. ***Positive Characteristics***: odd beliefs, unusual perceptual experiences

3. ***Characteristics of Insanity***: correlation ideas, suspiciousness

4. ***Characteristics of Disorganization***: strange behavior, odd speech

3.3 Latent Structure of Consumer Behavior

In this section we use the student survey data of Iliopoulou (2004). A total of 108 complete cases were collected. The data were collected in the School of Management Sciences of the University of Aegean and Technological Education Institutes of Crete and Piraeus. The questionnaire was divided in five parts including three different scales for measuring the variables. In our investigation of consumer behavior we focus on items 2-11 and 14a-16i. As a result a total of 37 responses obtained from questions 2-11 and 14a-16i responses. All responses measure the consumer behavior using a Linkert ordinal scale (1-5); see at the Appendix B for the questionnaire.

A logit type of Item Response Model (see 1.16) is used to analyze the data assuming one and two factors. The linear predictor for the one factor model is given by

$$\text{logit}(p_{i,j,k}) = a_{j} * (\text{theta}_i - b_{j,k}) \quad (3.1)$$

while for the two factor model is given by:

$$\text{logit}(p_{i,j,k}) = -b_{j,k} + a_{j,1} * \text{theta1}_i + a_{j,2} * \text{theta2}_i \quad (3.2)$$

In both cases, i denotes the observation (i=1,…,108), j is the number of items (j=1,…,37) and k is the number of categories minus 1 (k=1,…,4).

The values of the deviance information criterion (DIC) for the two models are given in Table 3.1. The two factor model is indicated by the DIC expression (equation 2.7) negative dimensions produced (further details in Section 2.3.2). That's why DIC_1 was calculated. The second one is preferable than the classic DIC because it is more stabilized from the MCMC output (through Winbugs).

Models	DIC	DIC_1
1 Factor	10420	11965
2 Factors	10520	**10937**

DICs have been derived from equations (2.7) and (2.12)

Table 3.1: DICs for each models (3.1) and (3.2) for schizotypic data

3.3.1 Inference for the Two Factor Logit Model for the Consumer Behavior

In this model, there are 294 parameters (which become equal to 510 when we include 294 discrimination parameters and 108 parameters for each factor score $theta_{i1}$ and $theta_{i2}$) for estimation. Because of the huge number of estimated parameters an indicative diagnostic test will be present having rejected the burn-in iterations. Thus the diagnostic test for the seventh discrimination parameters is shown below. All other diagnostic tests are in similar level. Obviously, the algorithm converges to a stationary chain. The estimated discrimination parameters with their standard errors and Monte Carlo Errors are presented in Appendix A.

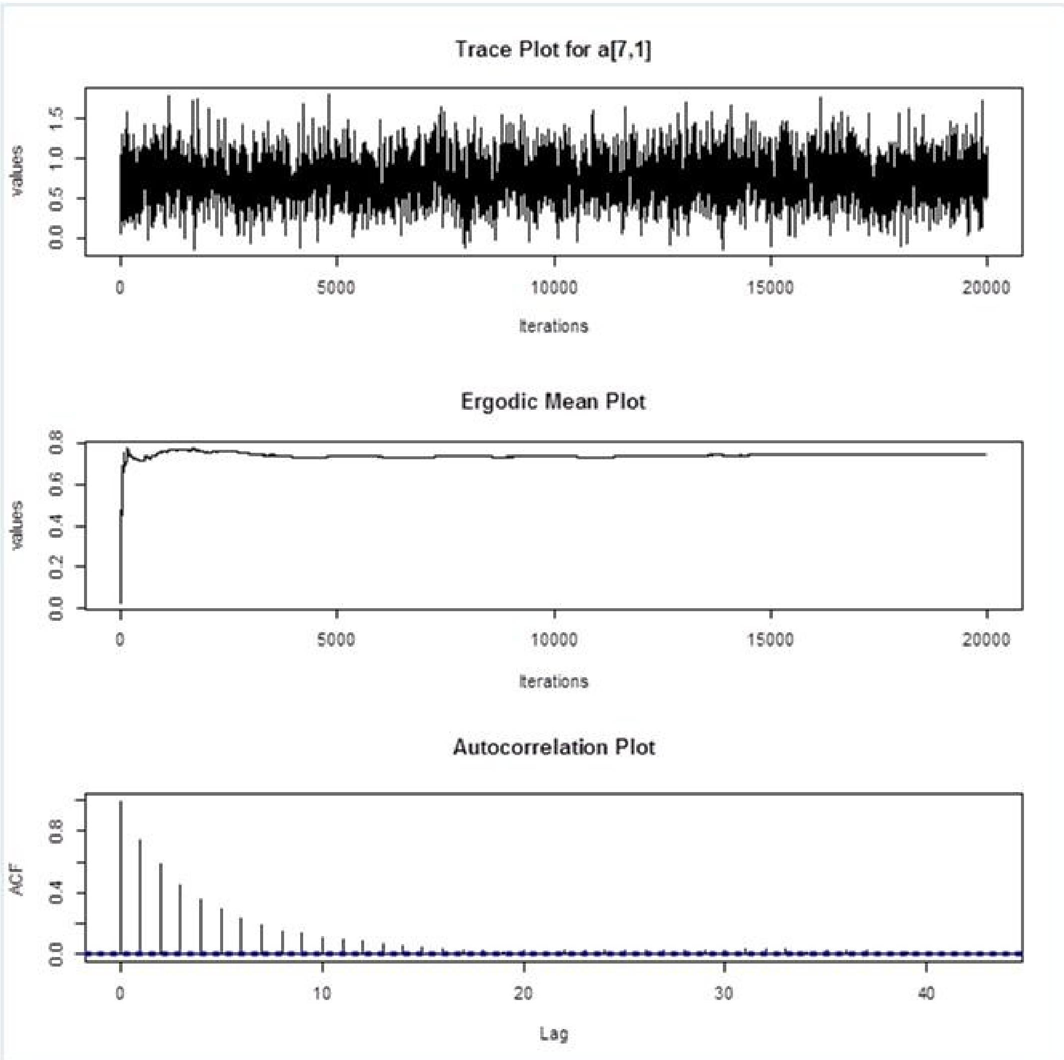


Figure 3.1: Indicative Diagnostic Tests for the 7th discrimination parameter and first latent factor with 10000 iterations (1000burn-in) for model (3.2) on schzotypal data

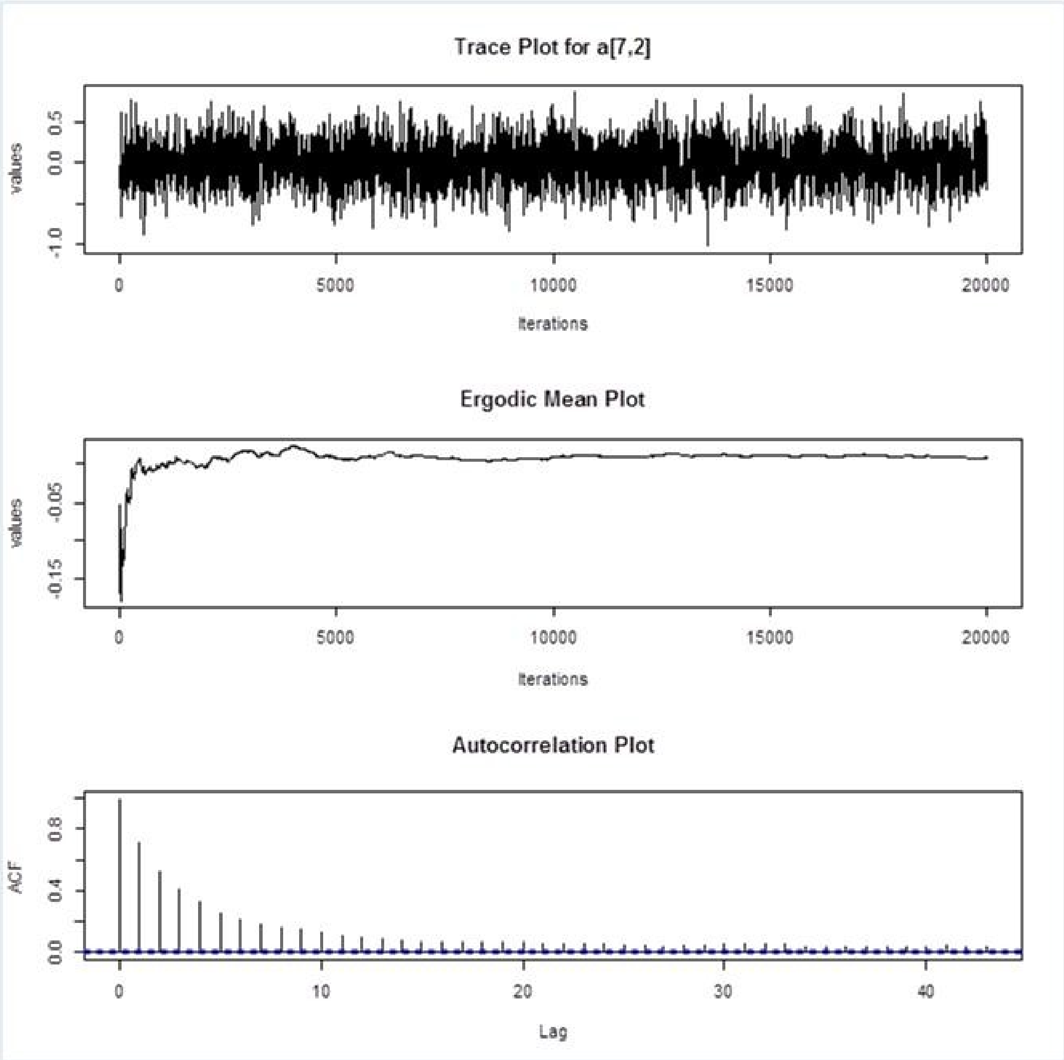


Figure 3.2: Indicative Diagnostic Tests for the 7th discrimination parameter and second latent factor with 10000 iterations (1000burn-in) for model (3.2) on schzotypal data

From Figure 3.3, we observe that only 13 out of 37 discrimination parameters of the first factor are statistical important (i.e. their 95% credible intervals do not include the zero). More specifically, the items one to nine, 32 and 33 examine impulsive and compulsive buying behaviors. The items 12 and 19 examine buying habits. As far as the second factor, 24 out of 37 discrimination parameters are statistical important. Items one, two, four, five, seven to ten and 31 to 36 examine impulsive and compulsive behaviors. All other statistical important discrimination parameters respond to items which check buying habits. The main difference between two factors is that the second one includes all items which investigate the impulsive buying.

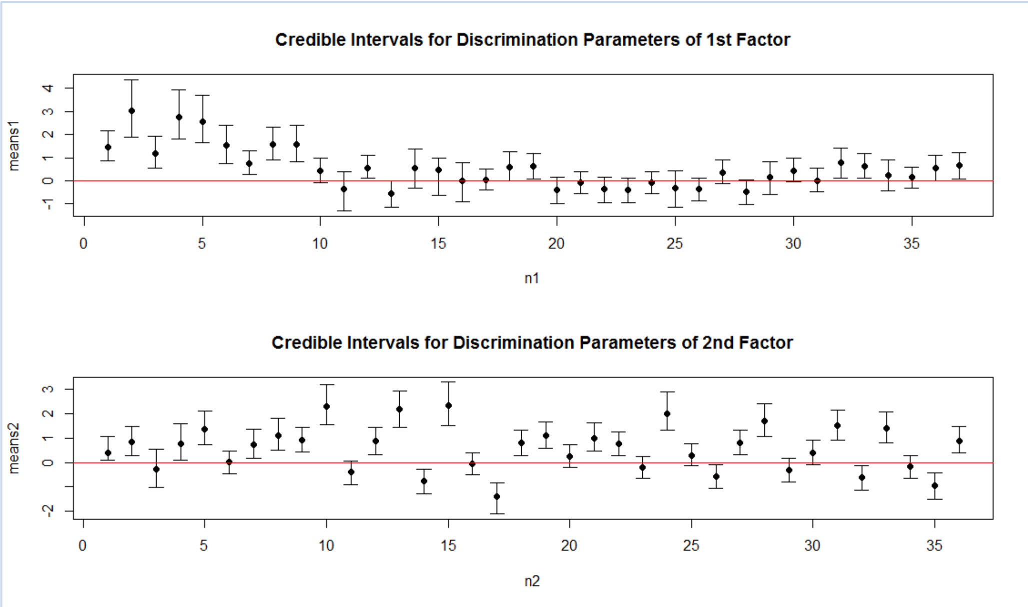


Figure 3.3: Credible Intervals for discrimination parameters for each factor for model (3.2) on schzotypal data

Standardized discrimination parameters present particular interest because of their interpretation; see Figure 3.4 and 3.5 for a graphical representation (credible intervals are included) and Table A.2 for detailed estimations.

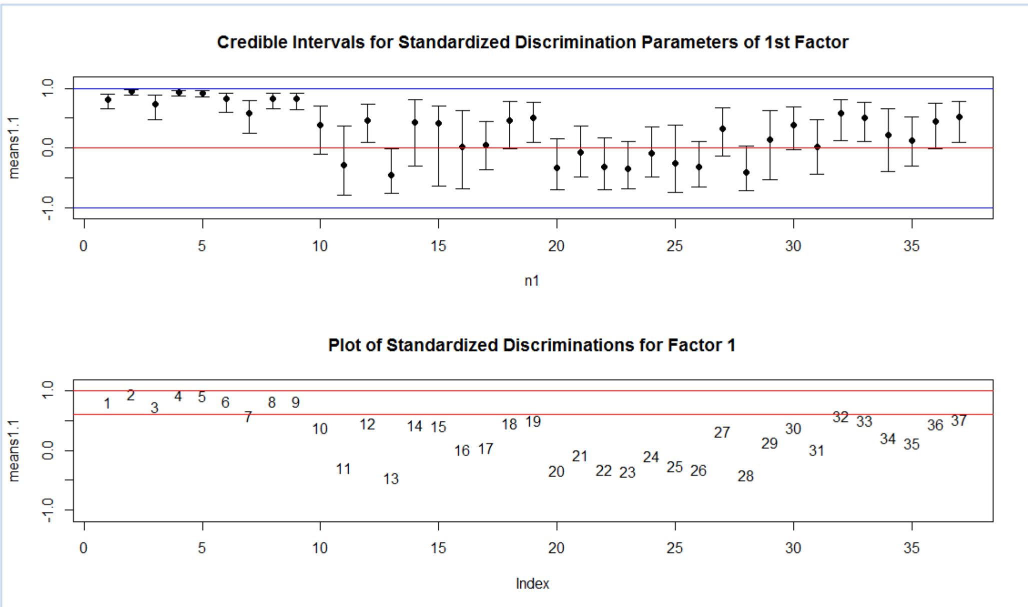


Figure 3.4: Standardized Discrimination Parameters of the 1^{st} factor for model (3.2) on schzotypal data

Figure 3.4 depicts which items have strong association with the first latent variable. Clearly the first nine items has medium to strong association (the red zone 0.6-1). These queries are linked with both the impulsive and compulsive buying. Principally, standardized discrimination parameters with the highest values are those of items two, four and five. Items two and five investigate whether the consumer behavior of a responder is related to impulsive buying. Their object to assess is whether statements such as "just do it" and "buy now and think later" express the consumer behavior of the responder. As far as, item four, it is related to compulsive buying and its goal is to check if the responder feels anxious when he is not going shopping.

On the other hand, items that investigate the general consumption of products (such as queries 13, 20, 28) do not seem to be associated with this latent factor since this parameters are very low.

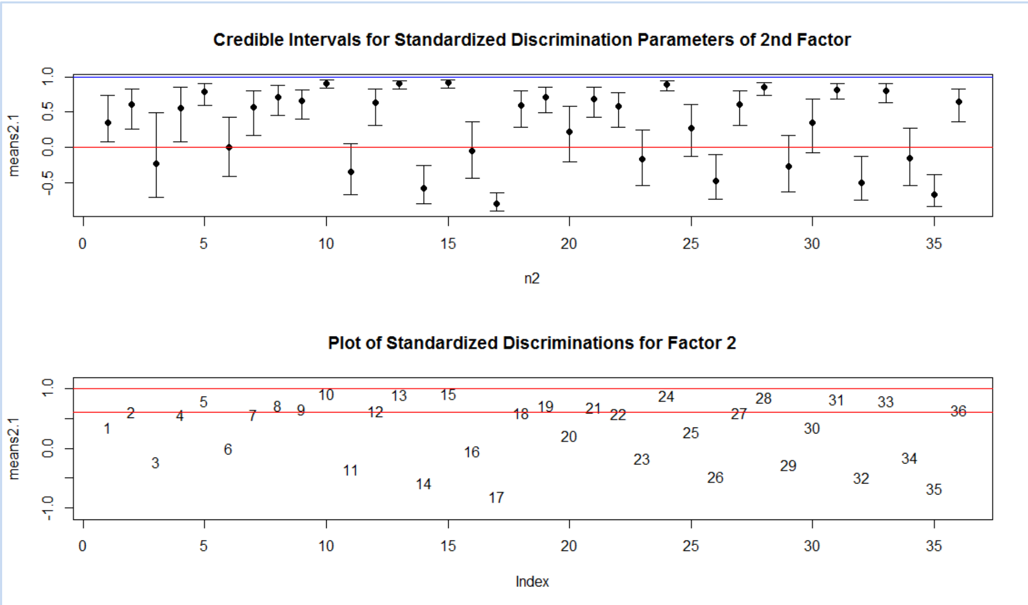


Figure 3.5: Standardized Discrimination Parameters of the 2nd factor
for model (3.2) on schzotypal data

From Figure 3.5 we may identify which manifest variables have strong association with the second latent factor. All items which lie in the red zone (such as items 5, 8, 9, 10, 15 e.t.c) or at in borderlines (such as items 2, 18, 27) have standardized discrimination values from 0.6 to one. Comparing Figures 3.4 and 3.5 a larger number of manifest variables demonstrate strong association with the second factor rather than the first one. Queries 2, 5, 7, 8 and 9 show almost the same level of association with both, the first and the second latent factors. Let us examine these items in more detail. Items 2 and 5 have been previously described. Item 7 investigates spontaneous buys after a visual contact with a product or service. This obviously lies in the field of the impulsive buying. On the contrary, item eight assess buying products which it cannot be afford it while item nine quantifies the uncertainty about a purchase. Both of these queries lie in the area of compulsive behavior.

Main interest lies on the study of a typical person, whose latent scores is equal to zero. Its numerical estimations for each category and also its estimated response probabilities arise from equations (1.18) and (1.19) and are listed in Appendix A (Tables A.4 and A5).

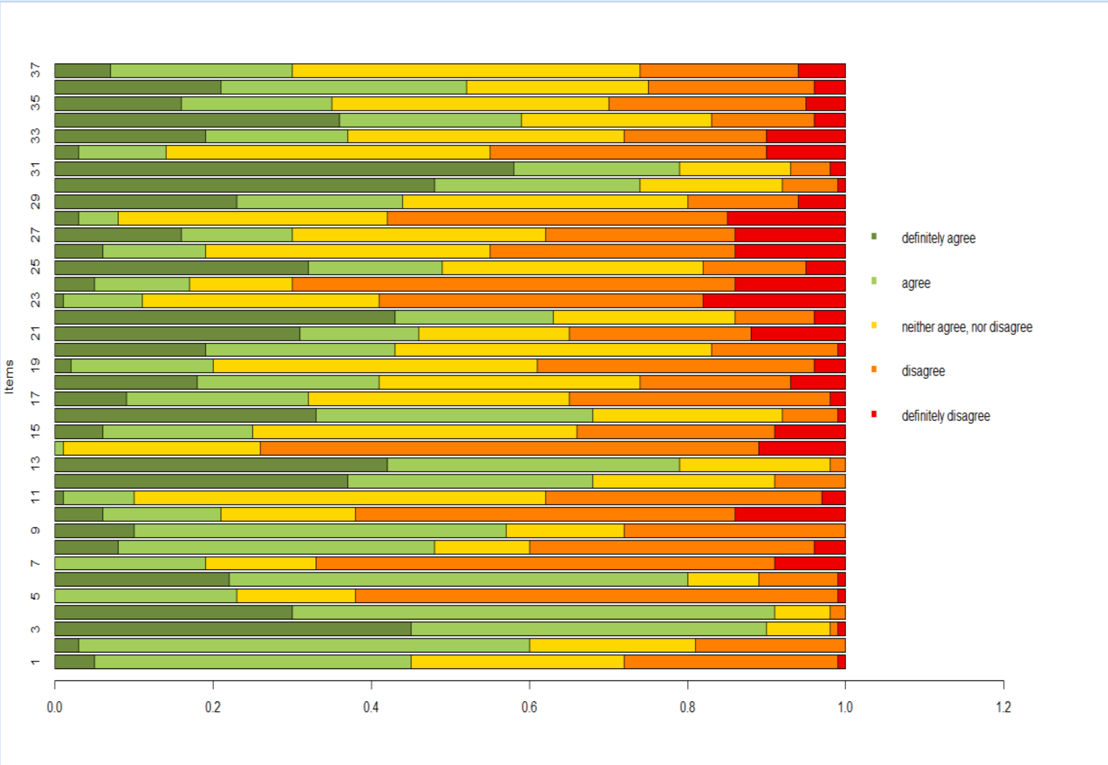


Figure 3.6: Response probabilities of a typical individual in each category for every question for model (3.2) on schzotypal data

From Figure 3.6, response probabilities in each category are presented for a "median" individual (i.e. a person with zero latent score). For items which imply strong impulsive buying such as question 1 (which tests if the statement "Just do it" expresses the participant) a typical individual with zero latent score has the same probability to give either a positive or a negative response. For the rest of the impulsive buying queries such as:

- buying products without thinking (item 2)
- buying now and thinks later (item 4)
- is carelessness concerning shopping (item 5)
- spontaneous buys after visual contact (item 7)

a ‘median” person has higher probability for a positive response rather than for a negative one for items 2 and 4 and the opposite items 5 and 7.

On the other hand, in the compulsive buying queries which assess:

- anxiety when he does not go shopping (item 3)
- whether they buy something, whatever it is (item 6)
- buying even if the financial conditions does not allow it (item 8)
- uncertainty after buying a product (item 9)
- buying something to cheer up (item 10)

a typical person has high probability for a strong positive response for queries 3 and 6, positive for 9^{th} and a negative response for 10^{th} . For query 8, the odds of positive response is approximately equal to one.

Finally, a scatter plot the latent factors (first versus second) is provided in. Figure 3.8. It is obvious that latent scores are independently distributed (cloud shaped scatter plot). Outlier points indicate individuals with problems. So, if the first and the second latent factor represent compulsive and impulsive buying behaviors respectively, individuals who are upper from the red horizontal line considered as individuals with intense impulsive buying behavior (individuals 32 and 37). Similarly, individuals who are rightmost of the vertical blue line considered as individuals with strong compulsive buying behavior (individuals 40 and 7). Individuals 32 and 7 are of special interest since they demonstrate high compulsive and impulsive buying behavior.

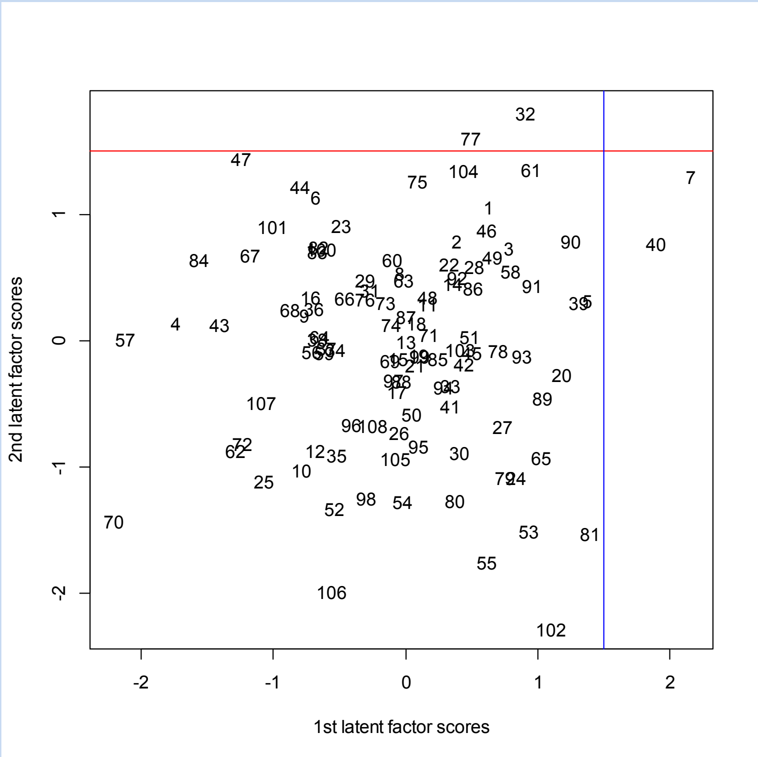


Figure 3.7: Scatter Plot of latent scores for model (3.2) on schzotypal data

3.4 Modeling of Consumer Behavior and Total SPQ Score

In order to study the effect of schizotypy (via the total SPQ score) on consumer behavior, we construct the following models for one and two latent factors respectively:

$$\mathrm{logit}(p_{i,j,k}) = a_{j*}(\theta_i - b_{j,k}) + g_j * SPQ_i \quad \text{(M1)}$$

$$\mathrm{logit}(p_{i,j,k}) = -b_{j,k} + a_{j,1} * \theta_{1i} + a_{j,2} * \theta_{2i} + g_j * SPQ_i \quad \text{(M2)}$$

where *i, j, k* have the same meaning as before: i=1,..,108, j=1,…,37, and k=1,…,4 and SPQ_i is the total SPQ score for every participant. For g_j, which is the coefficient vector of each item, normal prior distribution with mean equal to zero and variance equal to 1000 was considered. Through Deviance Information Criterion (DIC) we receive:

Models	DIC	DIC_1
1 Factor (M1)	8660	12004.69
2 Factors (M2)	10520	**11025.00**

DICs have been derived from equations (2.7), (2.12)

Table 3.2: DICs for models (M1) and (M2) on schizotypic data

As in case of consumer behavior, negative dimensions for classic DIC have been arise in one factor model. Based on Table 3.2 the most appropriate model is the second one.

3.4.1 The Influence of Total SPQ Scale on Consumer Behavior

From the boxplot of coefficient of variable total, g, we can see which g_j are a-posteriori away from zero. Clearly, from Figure 3.16 g_9, g_{22}, g_{29} and g_{33} are a-posteriori distributed away from zero.

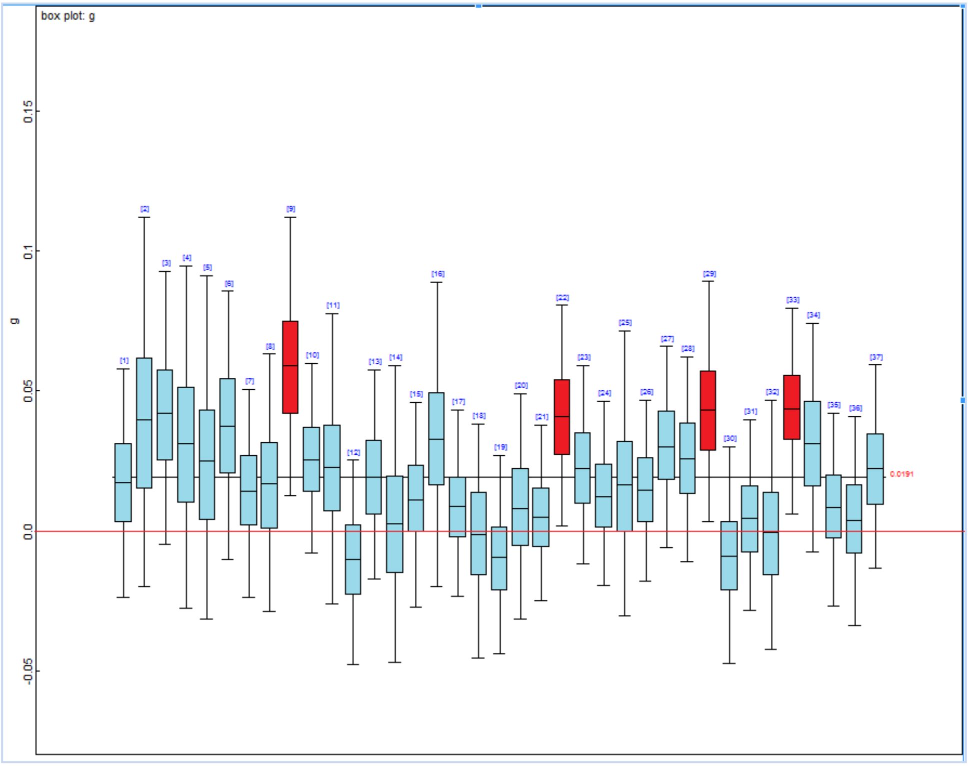


Figure 3.8: Box-plot of gj where *j*=1,…,37 number of questions based on model (M2) for the schizotypic data

As a result we will construct a third model:

$$\text{logit}(p_{i,j,k}) = -b_{j,k} + a_{j,1}*\theta_{1i} + a_{j,2}*\theta_{2i} + g_i*\text{total}_i \quad \text{(M3)}$$

where *i, j, k* have the same meaning as before: i=1,..,108, j=1,…,37, k=1,…,4 and $g_l \neq 0$ for l=9, 22, 29, 33 and zero otherwise. All g were considered equal to a constant as if m, where this constant m has prior distribution N(0,1000).

Model	DIC	DIC1
2 Factors (M3)	10520	**10929.08**

DICs have been derived from equations (2.7) and (2.12)

Table 3.3: DIC for third model on schizotypic data

From Table 3.3 and 3.2 it is clear that the last form of two factors model (M3) is preferable.

3.4.2 Analysis of Two Factors Logit Model of Total SQP Score on Consumer Behavior

Figures 3.9 and 3.10 present visual diagnostic tests for the third discrimination parameter, for the first and second factor of the fitted model (M3). The diagnostic tests for the rest of the parameters are similar and no convergence problems are evident.

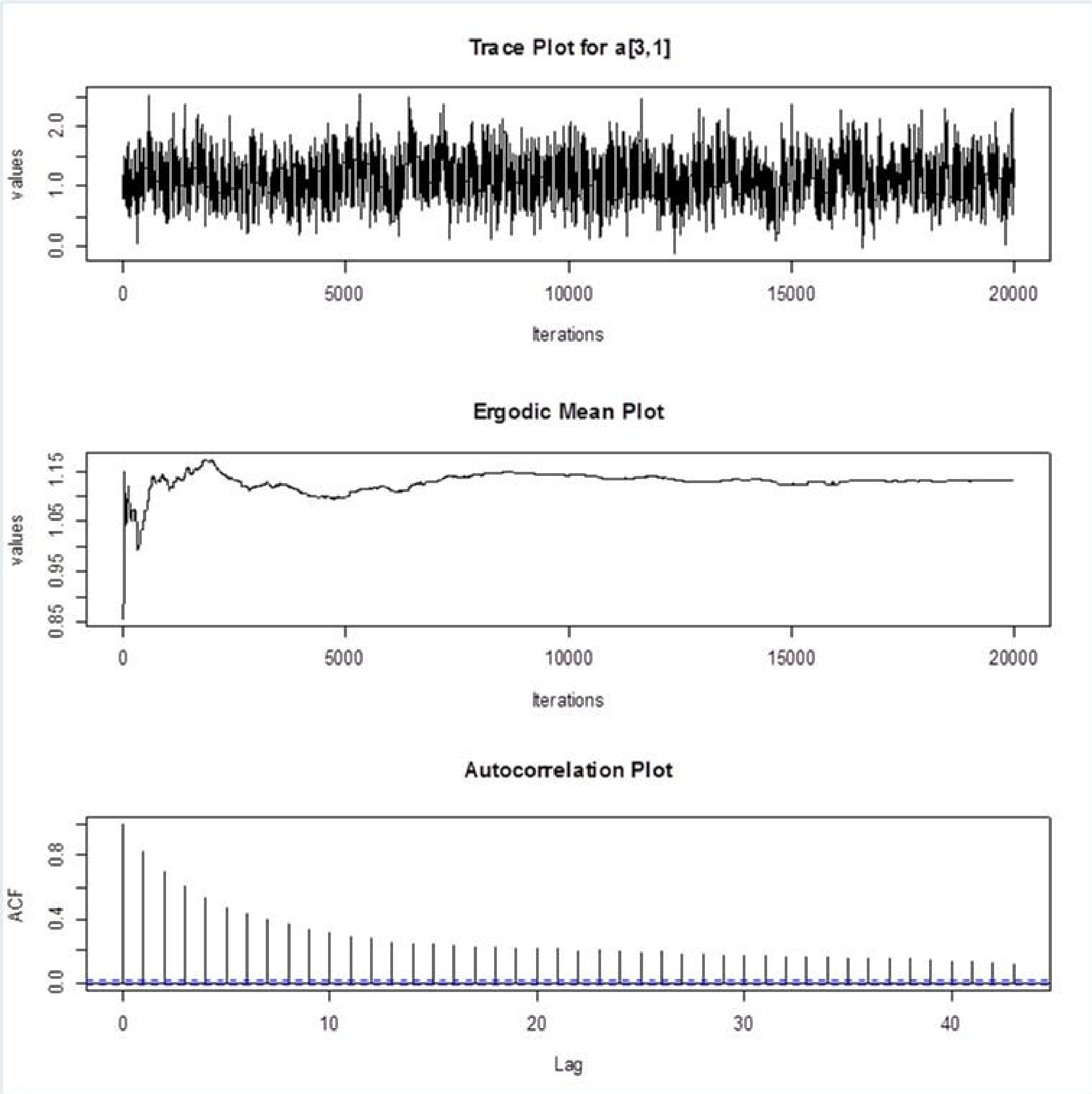


Figure 3.9: Indicative Diagnostic Tests for the third discrimination parameter and first latent factor for model (M3) on schizotypic data

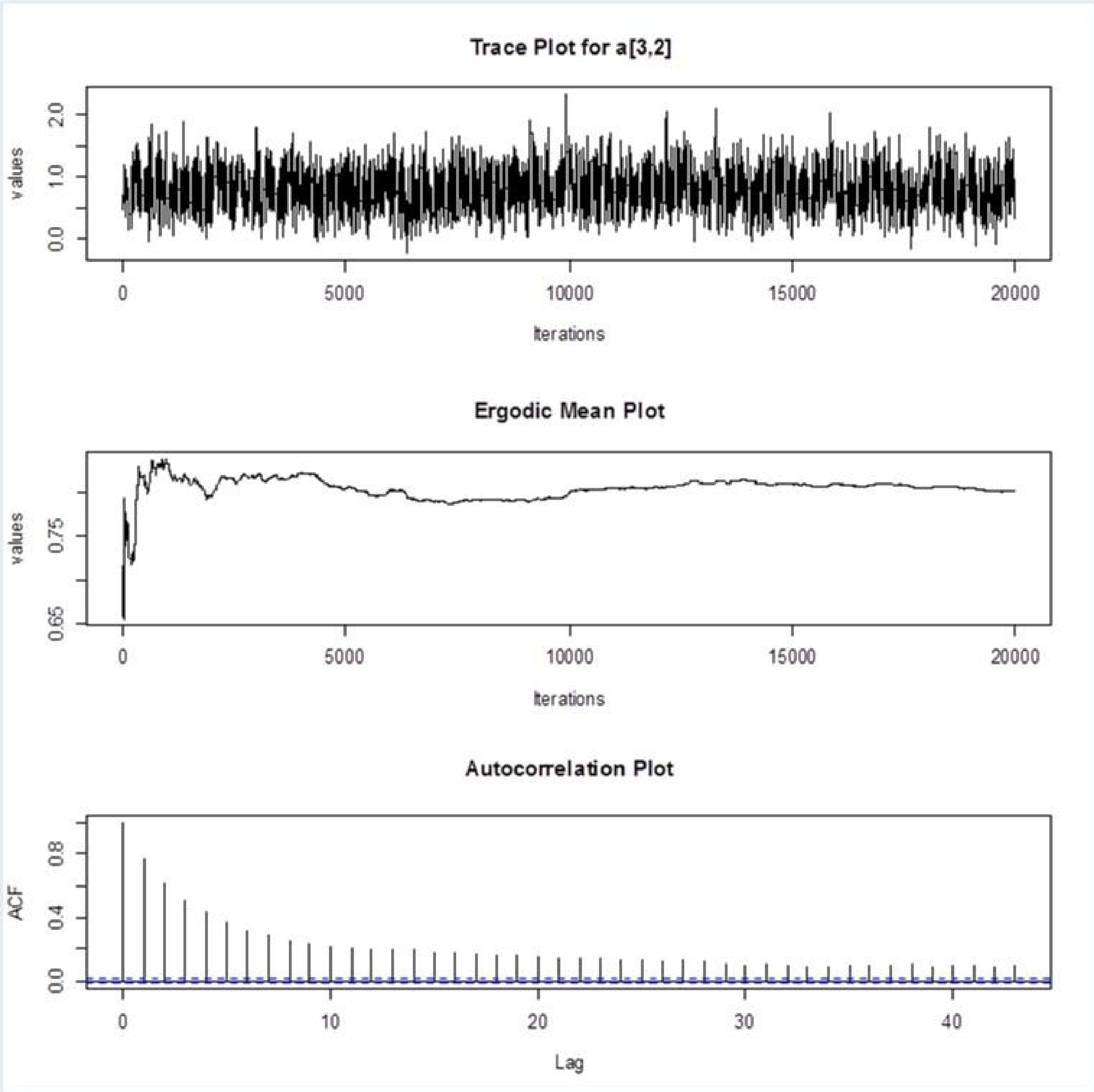


Figure 3.10: Indicative Diagnostic Tests for the third discrimination parameter and second latent factor for model (M3) on schizotypic data

The estimated discrimination parameters with their standard errors and Monte Carlo Errors are presented in Appendix A. Although, the estimated parameters are slight different to that of consumer behavior model, their credible intervals are almost same as their interpretation too.

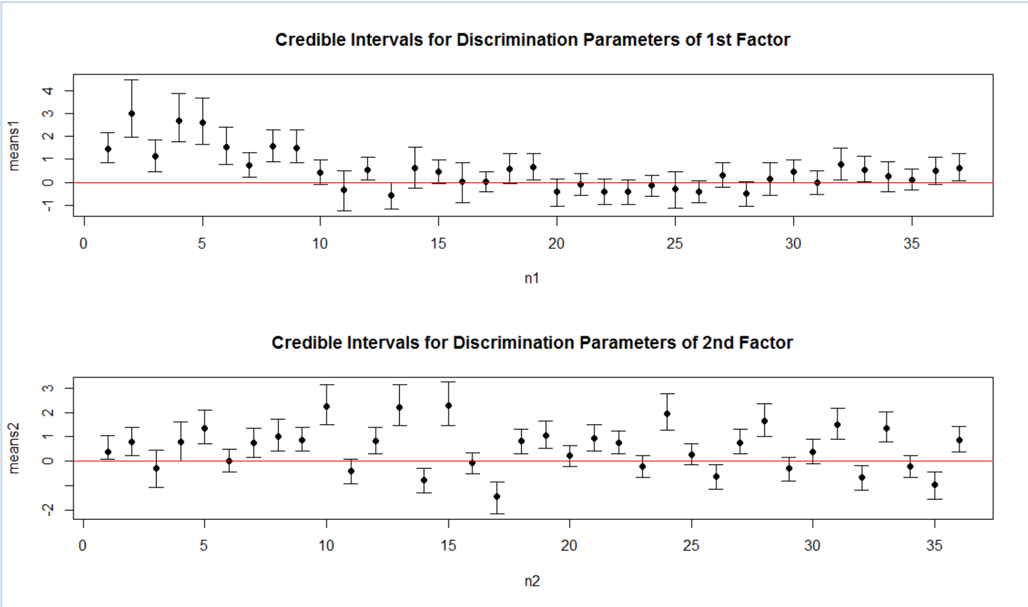


Figure 3.11 : Credible Intervals for discrimination parameters for each factor for model (M3) on schizotypic data

Standardized discrimination parameters are provided in Figure 3.12 and 3.13 while posterior summaries can be found at Table A6 in Appendix A. The pictures are similar to those of consumer behavior model for one and two factors.

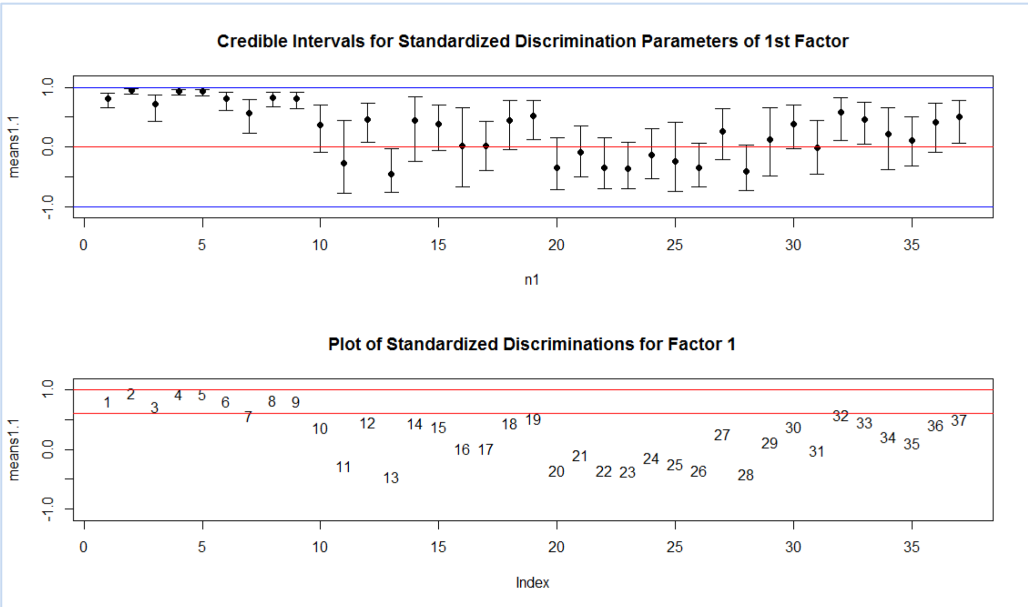


Figure 3.12: Standardized Discriminations for the first factor for model (M3) on schizotypic data

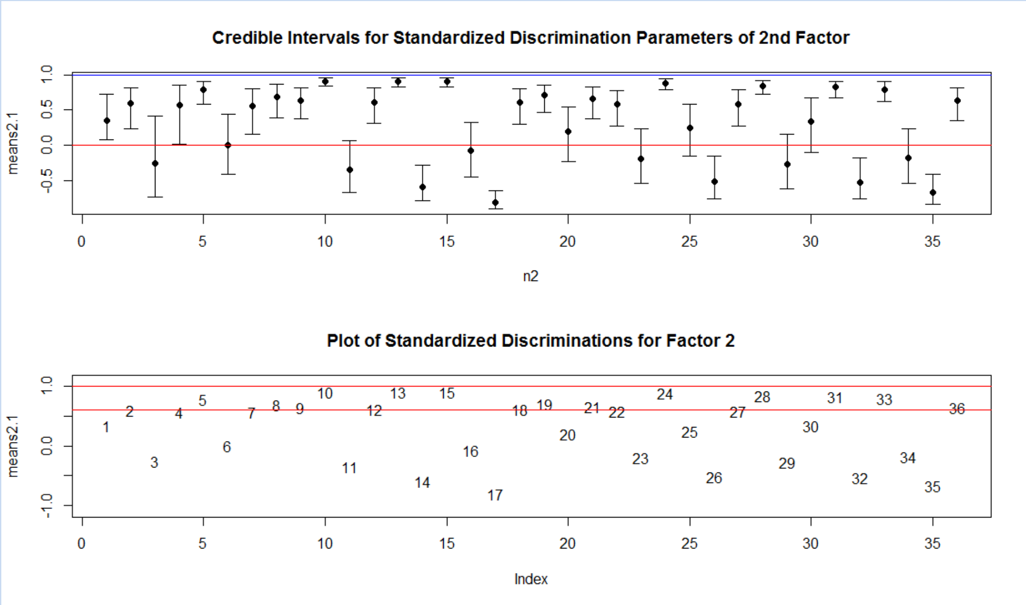


Figure 3.13: Standardized Discriminations for the second factor
for model (M3) on schizotypic data

The probability of the responses of a typical person (whith latent score is equal to zero) expected to present low difference to the previous model. Its numerical estimations for each category and also its estimated response probabilities come from equations (1.18) and (1.19) and are listed in Appendix A (Tables A.9 and A.10). So, in both figures do not observed notable differences between model of consumer behavior and model of total SQP scale on consumer behavior.

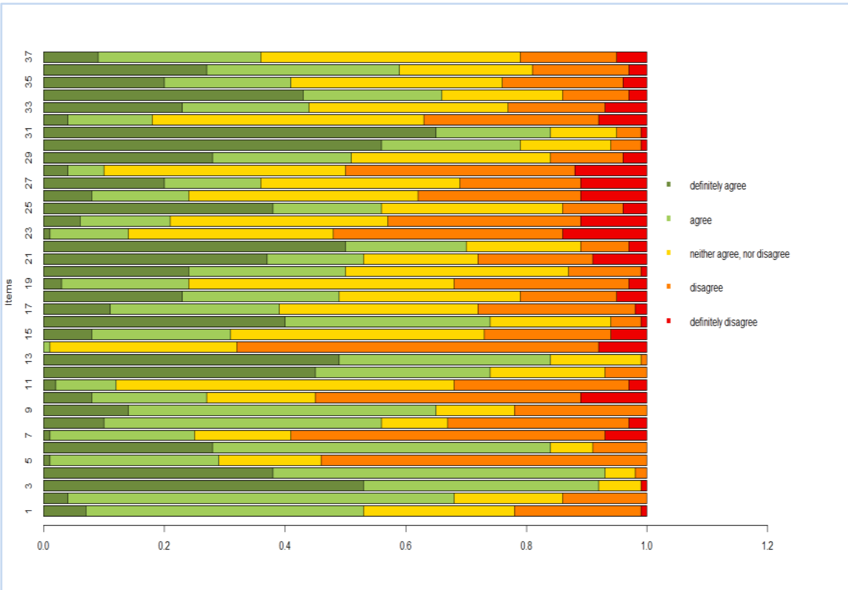


Figure 3.14: Response probabilities of a typical individual in each category for every question
for model (M3) on schizotypic data

Similar is the picture of the scatter plot of latent scores to the responding of consumer behavior model. Cloud shape is clear and the same persons appear as outliers.

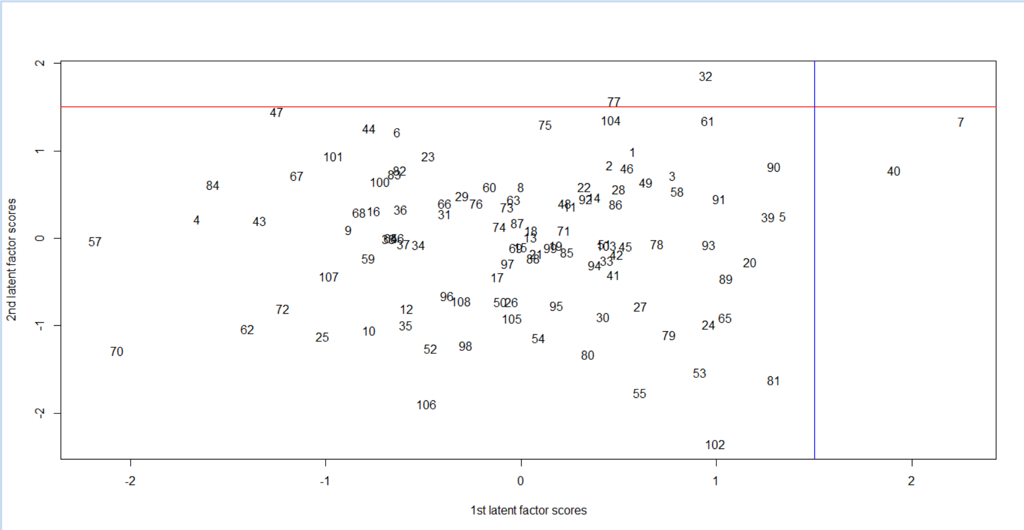


Figure 3.15: Scatter Plot of latent scores for model (M3) on schizotypic data

The credible interval for g_l was: (0.006,0.020). The effect of total SPQ (g_l) on the consumer behavior can be considered as important since zero in not contained to the 95% credible interval under model (M3).

Coefficients	Mean	Sd	MC_error
g_l	0.01332	0.00356	2,31E-01

Table 3.4: Estimations for coefficients of Total SPQ for the model (M3) for schizotypic data

As far as the interpretation of g_l , the response probability $p_{i,j,k}$ will increased by 1% ($e^{g_l} = 1.01$) if the total SPQ score is increased by one unit. From Figure 3.9 only four items seem to be influenced by total SPQ score. More specifically, item 9 which expresses the uncertainty after a buying products is associated with compulsive buying. Items 29 and 33 which are investigate the spontaneous and the frequent buying behavior are associated with impulsive buying and item 22 regards to product preferences.

3.5 Modeling of Consumer Behavior and Nine Traits of Schizotypy

So as to detect the influence of the nine traits of schizotypy on consumer behavior, we construct the following models for one and two latent factors respectively:

$$logit(p_{i,j,k})=a_j*(\theta_i-b_{j,k})+ g_{j,z}*SPQscales_{i,z} \quad (M4)$$

$$logit(p_{i,j,k})=-b_{j,k}+a_{j,1}*\theta_{1i}+a_{j,2}*\theta_{2i}+g_{j,z}*SPQscales_{i,z} \quad (M5)$$

where *i, j, k* have the same meaning as before: i=1,..,108, j=1,…,37, k=1,…,4, z=1,…,9. SPQscales is a matrix with 9 columns, one for each schizotypic trait and number of rows equal to sample size. For $g_{j,z}$ normal prior distribution with mean equal to zero and variance equal to 1000 was considered. Via Deviance Information Criterion (DIC) we received:

Models	DIC	DIC_1
1 Factor (M4)	10400.00	11554.77
2 Factors (M5)	10760.00	**11482.21**

DICs have been derived from equations (2.7) and (2.12)

Table 3.5: DICs for each model for scizotypic data

As in case of consumer behavior, negative dimensions for classic DIC arise in one factor model. Based on Table 3.5 the most appropriate model is the second one.

3.5.1 The Influence of Nine Traits of Schizotypy on Consumer Behavior

Having taken schizotypic traits as covariates and after choosing M5 as the optimal model using DIC we checked which coefficients of schizoypy have important effect on the consumer behavior. To facilitate the process for nine traits of schizotypy we use the coding of Table 3.6.

Nine Traits of Schizotypy	Coding for the corresponding coefficient
Ideas of Reference	g[,1]
Excessive Social Anxiety	g[,2]
Odd Beliefs	g[,3]
Unusual perceptual experiences	g[,4]
Odd behavior	g[,5]
No close friends	g[,6]
Odd speech	g[,7]
Constricted affect	g[,8]
Suspiciousness	g[,9]

Table 3.6: Coding for the characteristics of schizotypy.

The red box-plots correspond to coefficients which are a-posteriori distributed away from zero. In other words their effect is important. The rest of them can be excluded from the model. Consequently, the posterior distribution of each SPQ subscale on the items is given in Figures 3.16-3.24:

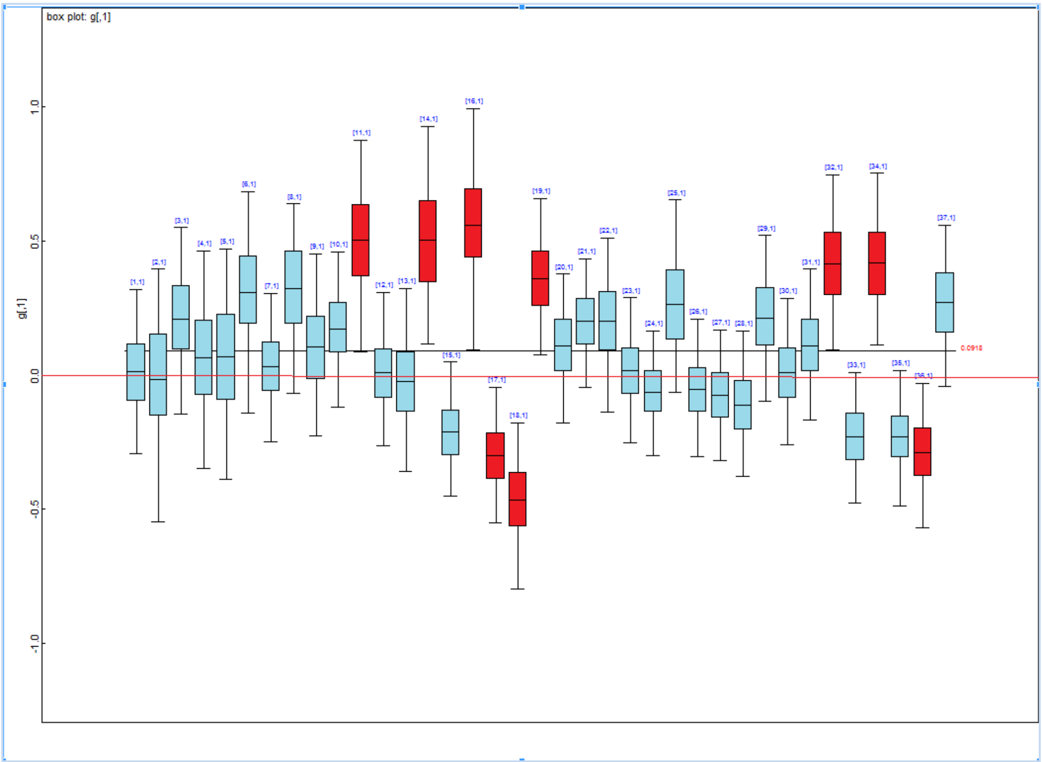


Figure 3.16: Box-plots for the coefficients of Ideas of Reference g[,1] for model (M5)

So, for the Ideas of Refernece statistical important seem to be g[11,1], g[14,1], g[16,1], g[17,1], g[18,1], g[19,1], g[32,1], g[34,1] and g[36,1]. This trait influences 9 querries out of 37. From these, only queries 32, 34 and 36 are manifestation of impulsive buying.

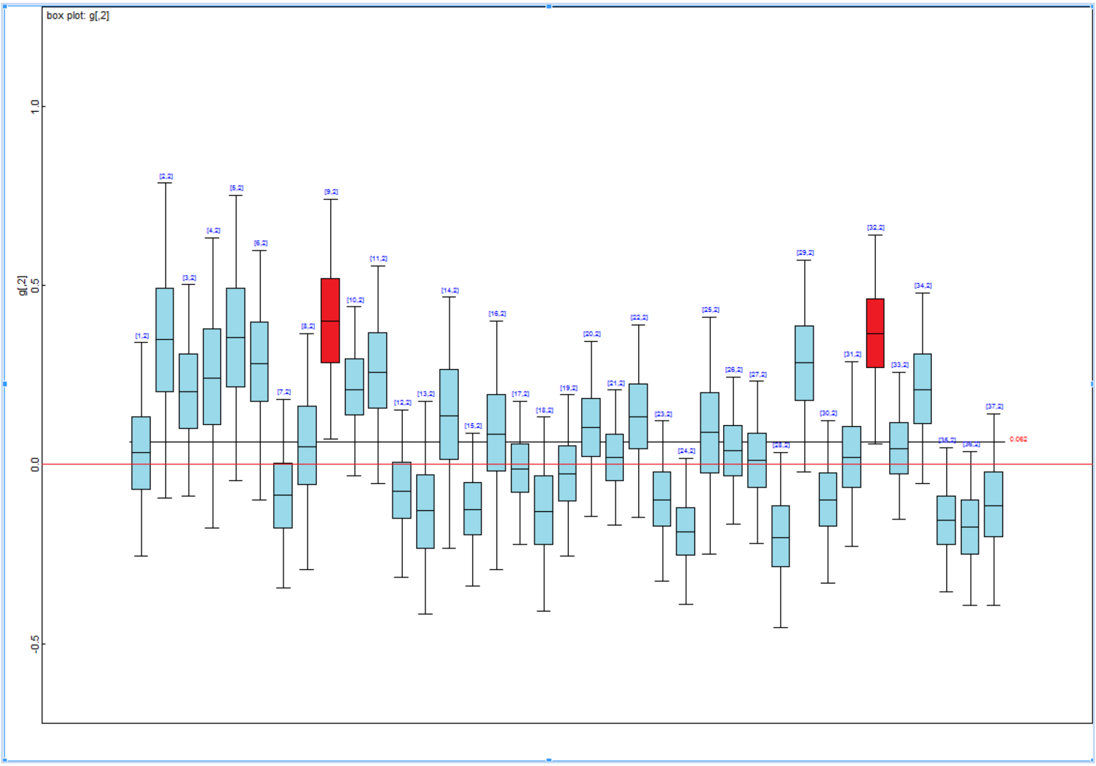


Figure 3.17: Box-plots for the coefficients of Excessive Social Anxiety g[,2] for model (M5)

For the second trait, according to Figure 3.17, statistical important are g[9,2] and g[32,2]. The consumer behavior which are under detection from each item is for the first one the compulsive buying and for the second one the impulsive buying. Only two items out of 37 are influenced by this trait.

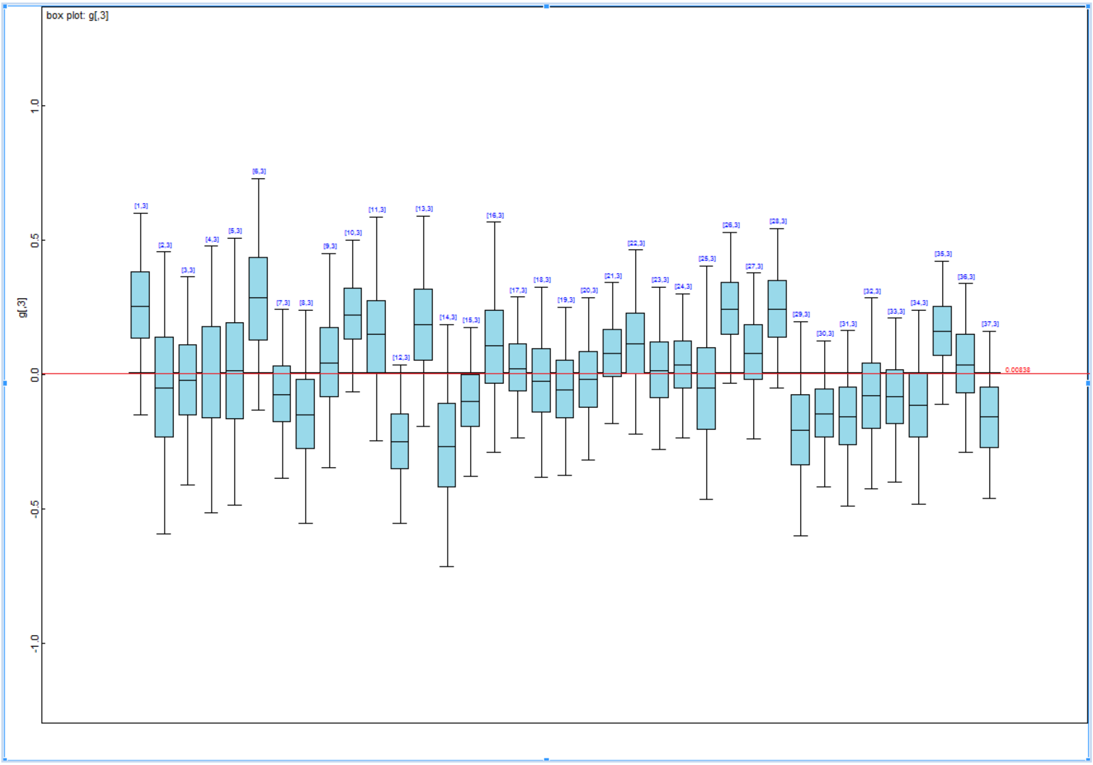


Figure 3.18: Box-plots for the coefficients of Odd Beliefs g[,3] for model (M5)

None of the coefficients of third trait is statistical important. This trait could be excluded from a future survey.

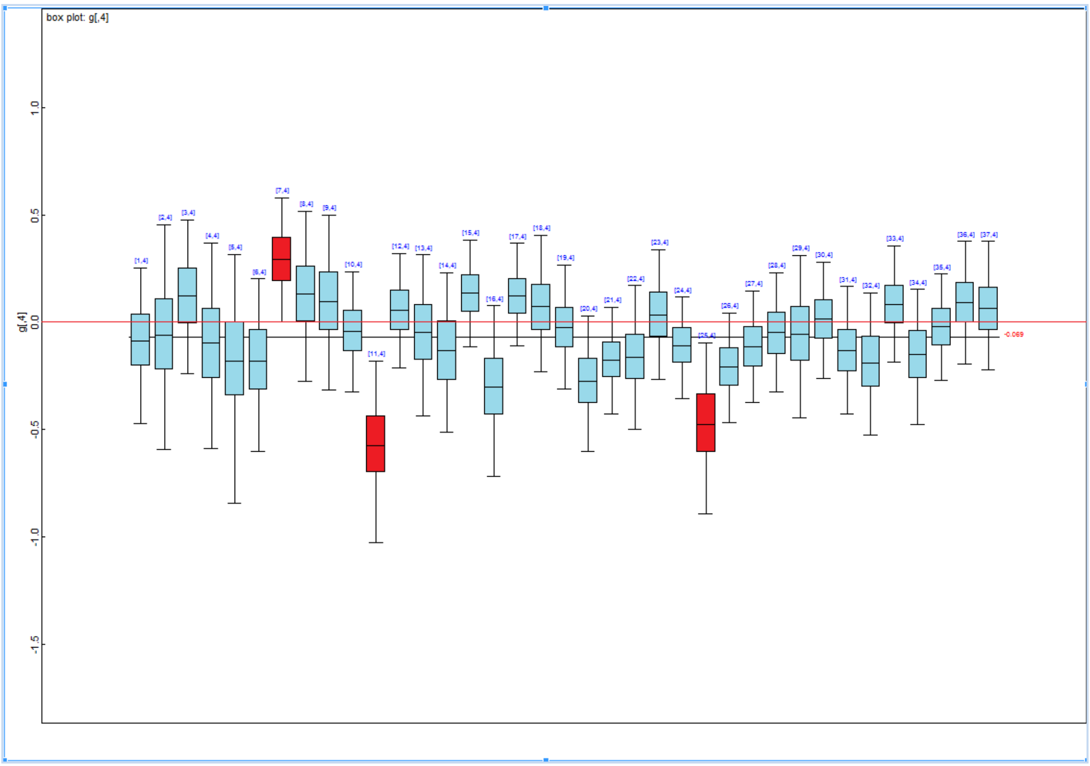


Figure 3.19: Box-plots for the coefficients of Unusual Perceptual Experiences g[,4] for model (M5)

From Figure 3.19, only three coefficients of unusual perceptual experiences are a-posteriori distributed away from zero. Those are statistical important. They are g[7,4], g[11,4] and g[25,4] and only the items 7 and 11 have to do with impulsive and compulsive consumer behavior.

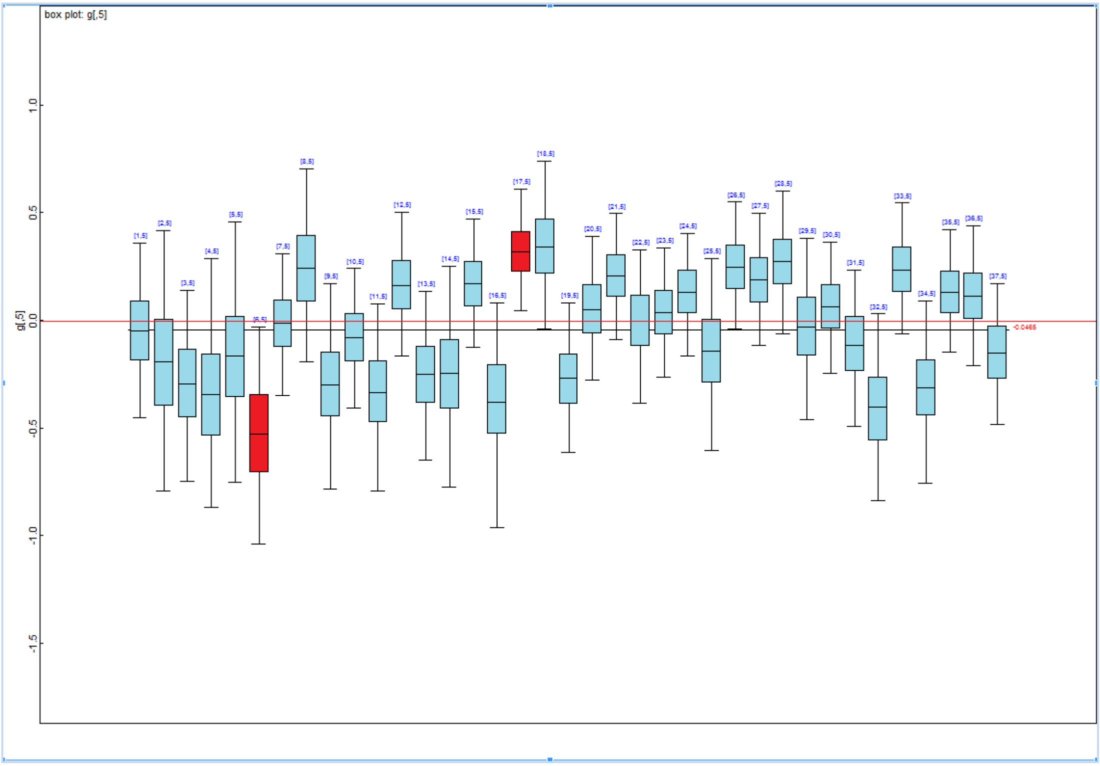


Figure 3.20: Box-plots for the coefficients of Odd Behavior g[,5] for model (M5)

For the odd behavior trait only two coefficients are statistical important. These are g[6,5] and g[17,5]. From these only the item 6 has as object a strange consumer behavior which is the compulsive buying. This characteristic appear to influence only two items out of 37.

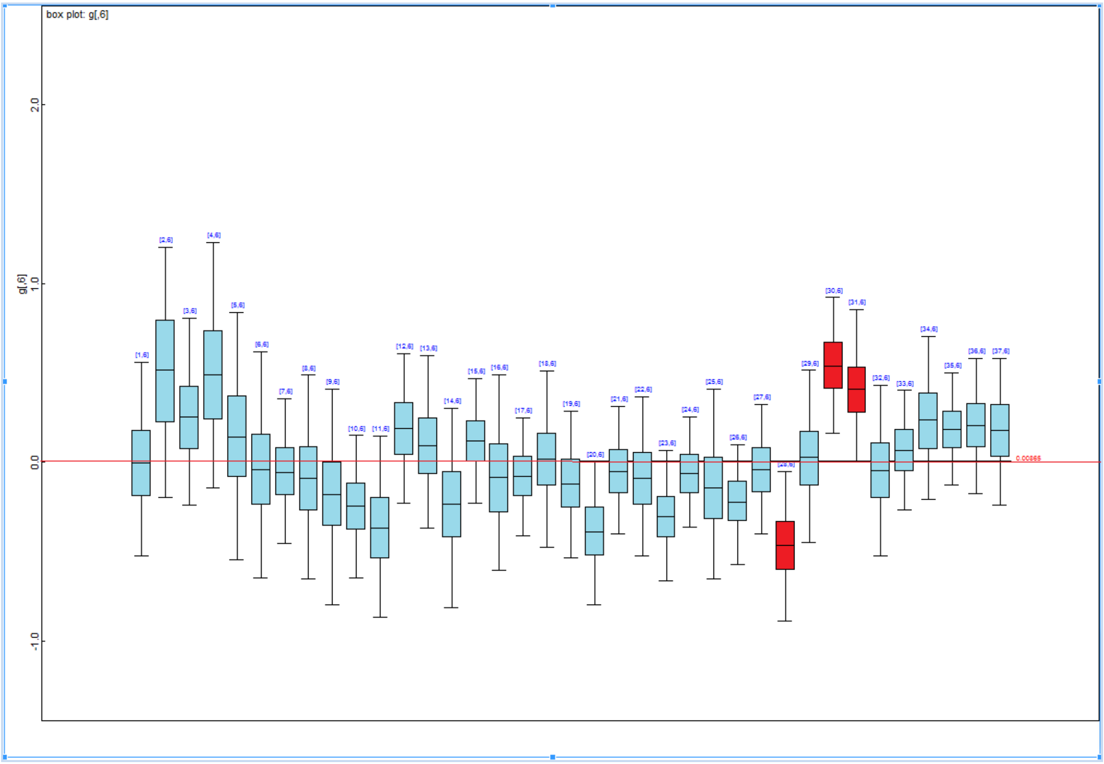


Figure 3.21: Box-plots for the coefficients of No Close Friends g[,6] for model (M5)

Figure 3.21, depicts only three red box-plots. That means, only 3 queries out of 37 are influenced by this schizotypic trait. These queries are 28, 30 and 31 and only the last two have to do with impulsive buying behavior.

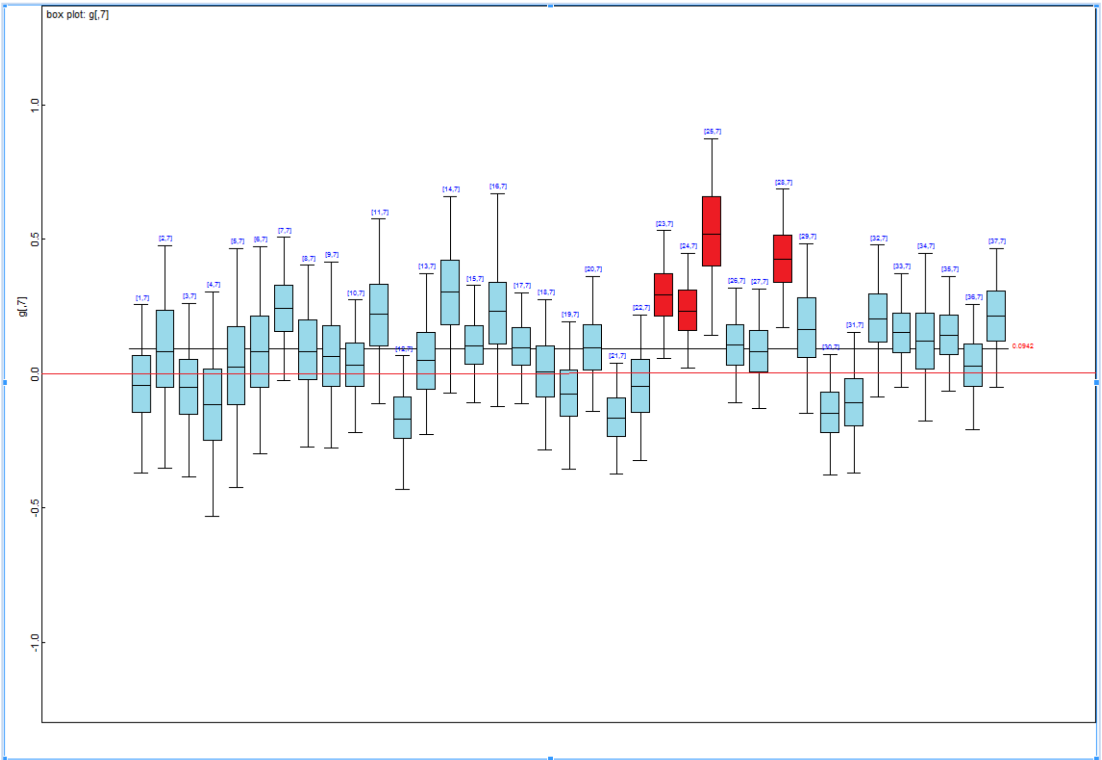


Figure 3.22: Box-plots for the coefficients of Odd Speech g[,7] for model (M5)

In Figure 3.22 there are four red box-plots. So, the corresponding queries are influenced by the odd speech trait. These queries are 23-25 and 28. All of them examine a general consumer behavior as far as the customer's preferences.

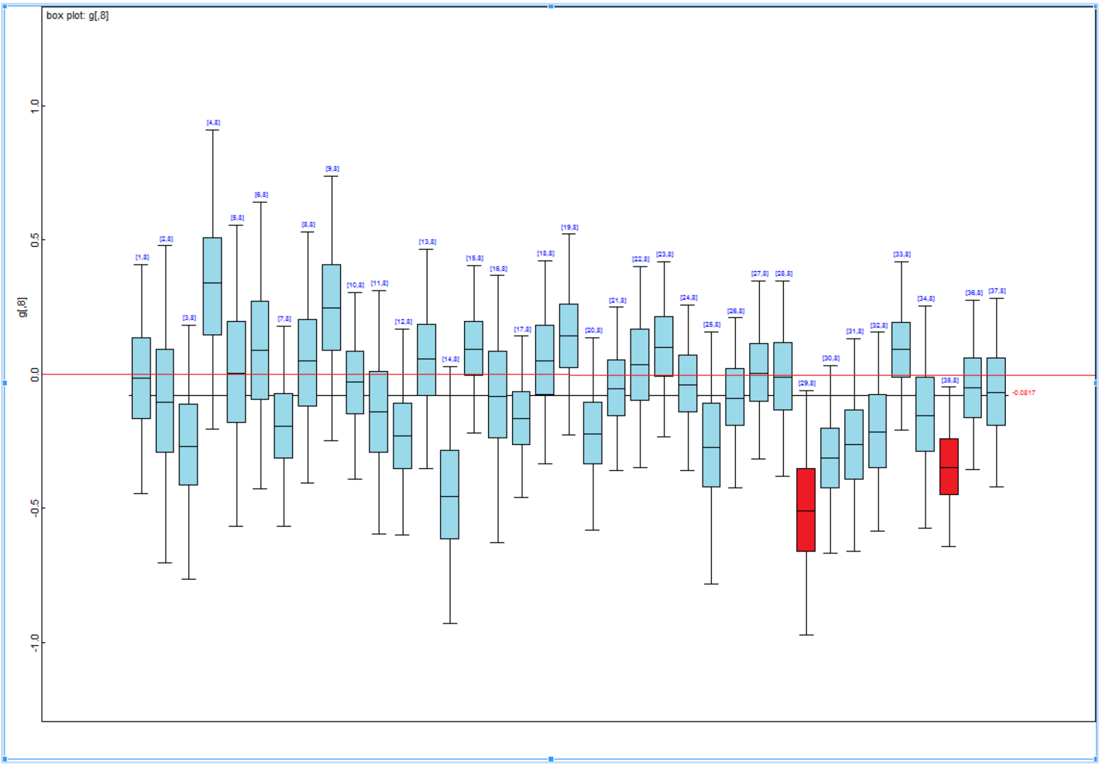


Figure 3.23: Box-plots for the coefficients of Constricted Affect g[,8] for model (M5)

Constricted affect trait clearly influences only two items, which have the corresponding red box-plots. These items are 29 and 35. Both of them are examine the impulsive consumer behavior.

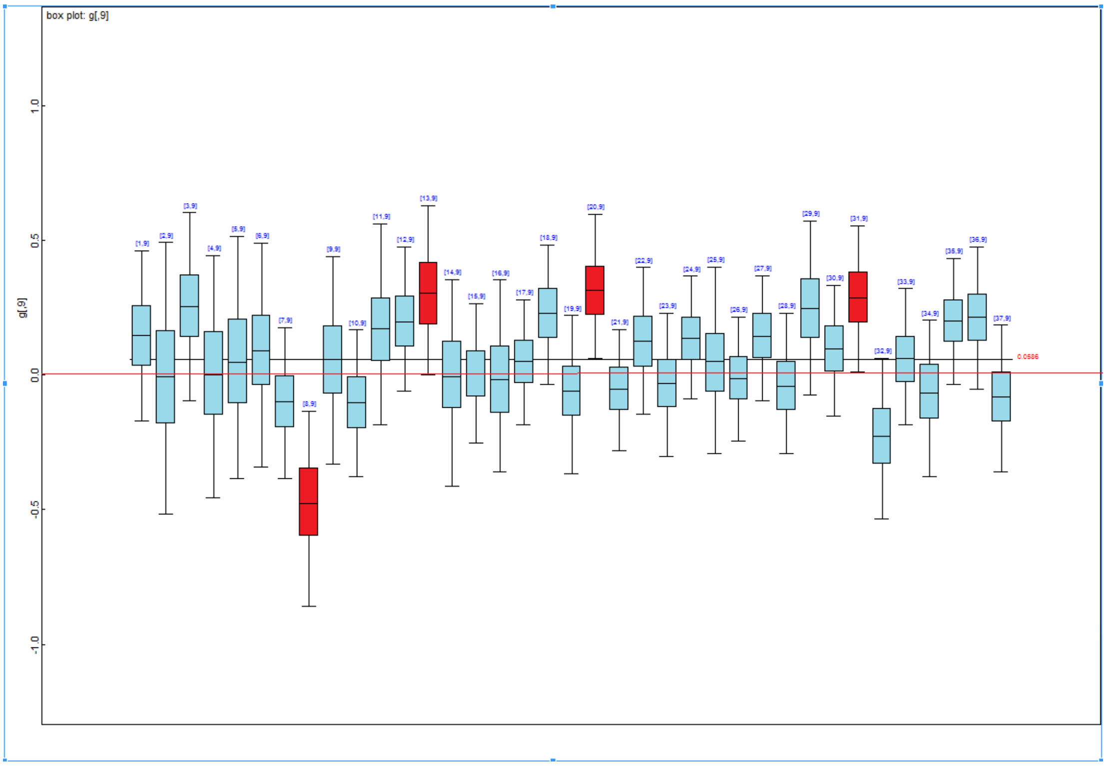


Figure 3.24: Box-plots for the coefficients of Suspiciousness g[,9] for model (M5)

The schizotypic trait of suspiciousness influences four items out of 37. These items are 8, 13, 20 and 31. From all of them, only items 8 and 31 have to do with excessive consumer behavior, i.e. compulsive and impulsive consumer behavior respectively.

Through the Table 3.7 we can see the possible influence of each schizotypic trait on impulsive and compulsive consumer behaviors. So, the excessive social anxiety (g[,2]), the odd behavior (g[,5]) and the suspiciousness (g[,9]) influence the compulsive consumers. On the contrary, the unusual perceptual experiences (g[,4]) influence the impulsive consumer behaviors. Moreover for the impulsive consumers we can conclude if the nine schizotypic traits influence their buying preferences (items 29 to 37). Thus, the impulsive consumer's preferences are influenced by the ideas of reference (g[,1]), the excessive social anxiety (g[,2]), the trait of no close friends (g[,6]), the constricted affect trait (g[,8]) and the suspiciousness (g[,9]).

It is obvious that, even if the nine schizotypic traits seem to influence both impulsive and compulsive consumer behaviors, the last consumer behavior is influenced to a higher grade. This conclusion is expected in same way because experts consider compulsive behavior as a clinic disorder.

As far as items 11 to 28 clearly are influenced by the schizotypic traits but these items do no indicate an excessive consumer behavior.

Traits Items	g[,1]	g[,2]	g[,3]	g[,4]	g[,5]	g[,6]	g[,7]	g[,8]	g[,9]
1									
2									
3									
4									
5									
6					✓				
7				✓					
8									✓
9		✓							
10									
11	✓			✓					
12									
13									✓
14	✓								
15									
16	✓								
17	✓				✓				
18	✓								
19	✓								
20									✓
21									
22									
23							✓		
24							✓		
25				✓			✓		
26									
27									
28						✓	✓		
29								✓	
30						✓			
31						✓			✓
32	✓	✓							
33									
34	✓								
35								✓	
36	✓								
37									

Red and blue colored items denote impulsive and compulsive consumer behaviors respectively
and black items the consumer preferences

Table 3.7: The influence of nine schizotypic traits on each item according to model (M5)

3.5.2 Analysis of Two Factors Logit Model of Nine Traits of Schizotypy on Consumer Behavior

After the study of each trait separately, it is worth to exclude all non-important coefficients of schizotypic traits. In order to succeeded it we construct the following model with the structure of (M5):

$$logit(\gamma_{i,j,k})=-b_{j,k}+a_{j,1}*theta1_i+a_{j,2}*theta2_i+g_{jz}*schizotypy_{i,z} \quad (M6)$$

$$\text{where } g_{jz}=\begin{cases} g_{jz}^{+}, \text{ when the 95\% credible inerval is upper than 0} \\ 0, \text{ when the 95\% credible interval concludes the 0} \\ g_{jz}^{-}, \text{ when the 95\% credible interval is lower than 0} \end{cases}$$

where *i, j, k* have the same meaning as before: i=1,..,108, j=1,…,37, k=1,…,4, z=1,…,9. Schizotypy matrix has the same structure as before. For both g_{jz}^{+} and g_{jz}^{-} normal prior distribution with mean equal to zero and variance equal to 1000 was considered. Via Deviance Information Criterion (DIC) we received:

Models	DIC	DIC1
2 Factor (M6)	10470.00	10881.67

DICs have been derived from equations (2.7) and (2.12)

Table 3.8: DICs model (M6)

Taking into consideration Tables 3.5 and 3.8 the current model (M6) seems to be sufficient. The estimated parameters g_{jz}^{+} and g_{jz}^{-} are away from zero as expected by construction.

Parameters	Mean	SDs	MC_error	Low CI limit	Upper CI limit
g_{jz}^{+}	0.184	0.025	0.001	0.135	0.233
g_{jz}^{-}	-0.136	0.038	0.001	-0.214	-0.063

Table 3.9: Estimated parameters g_{jz}^{+} and g_{jz}^{-} for model (M6)

Diagnostic visual convergence tests are provided in Figures 3.27-3.28. No evidence of no convergence is presented.

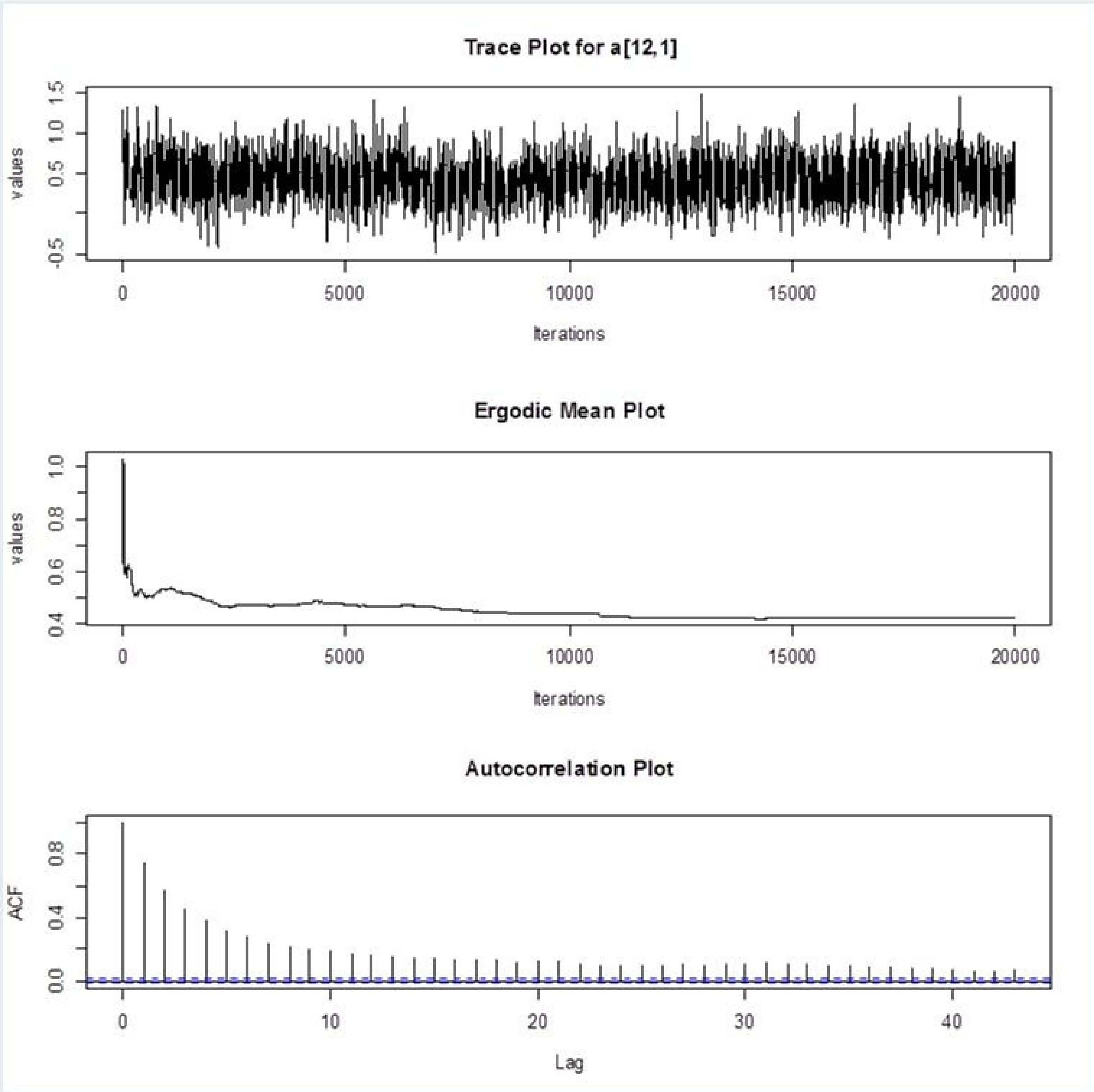


Figure 3.25: Indicative Diagnostic Tests for the 12th discrimination parameter and first latent factor for the model (M6) for schizotypic data

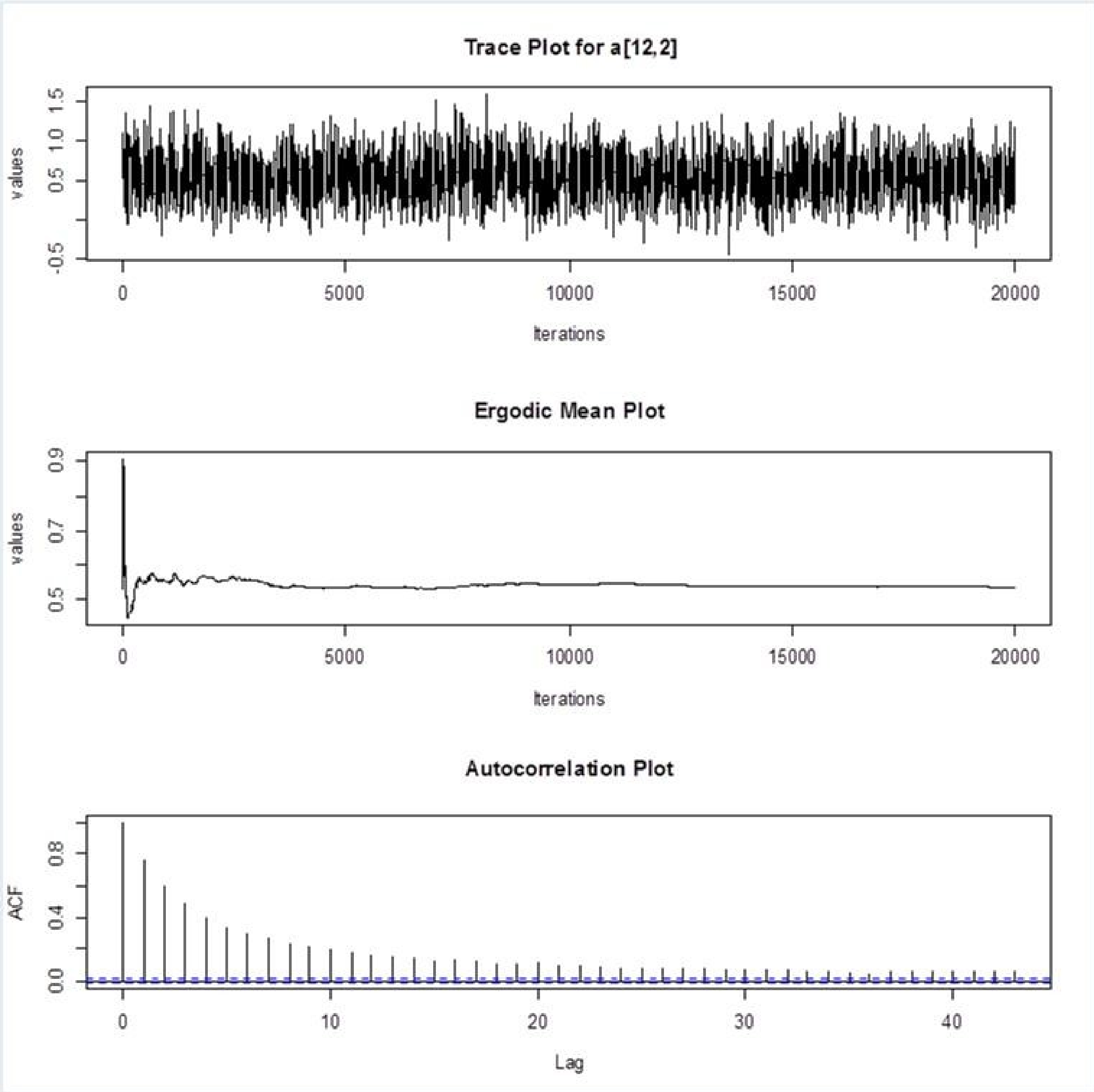


Figure 3.26: Indicative Diagnostic Tests for the 12th discrimination parameter and second latent factor for the model (M6) for schizotypic data

The diagnostic test for remaining parameter are equivalent. It is evident, the algorithm converges to its stationary chain. The estimated discrimination parameters with their standard errors and Monte Carlo Errors are presented in Appendix A. Although, the estimated parameters are slightly different to the ones of consumer behavior model, their credible intervals are almost same resulting to similar interpretation.

The credible intervals for the discrimination parameters for each factor are presented visually in Figure 2.27 and the numeric estimation in Table A.11 at

Appendix A. The outcomes are very close to the corresponding model of consumer behavior (Model 3.2) and the same happens to their interpretation.

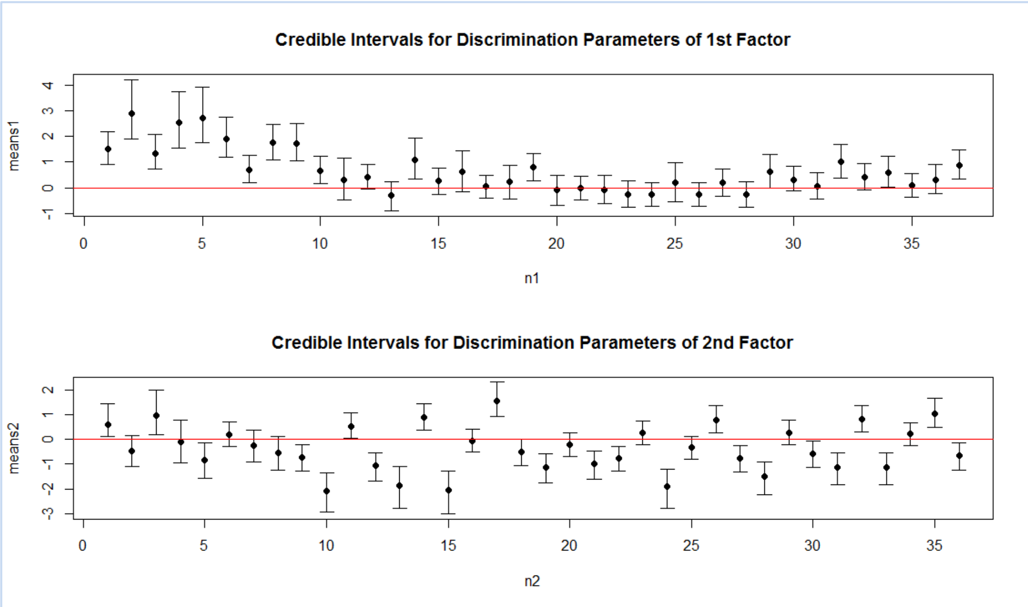


Figure 3.27 : Credible Intervals for discrimination parameters for each factor for model (M6)

In Figures 3.28 and 3.29 standardized discrimination parameters are depicted for each factor. Once again, the results are in similar level with those of consumer behavior model.

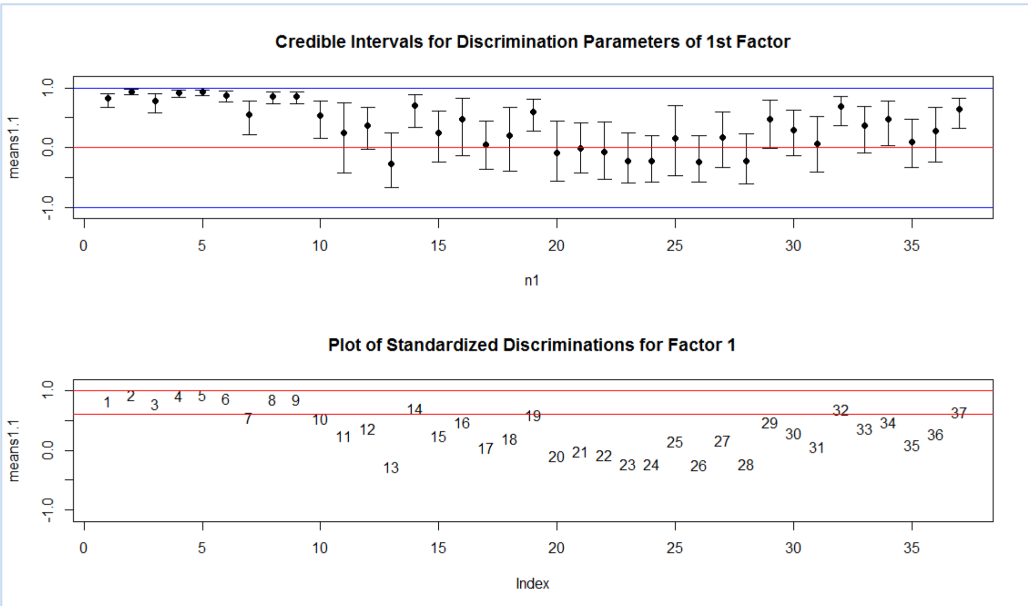


Figure 3.28: Standardized Discriminations for the first factor for model (M6)

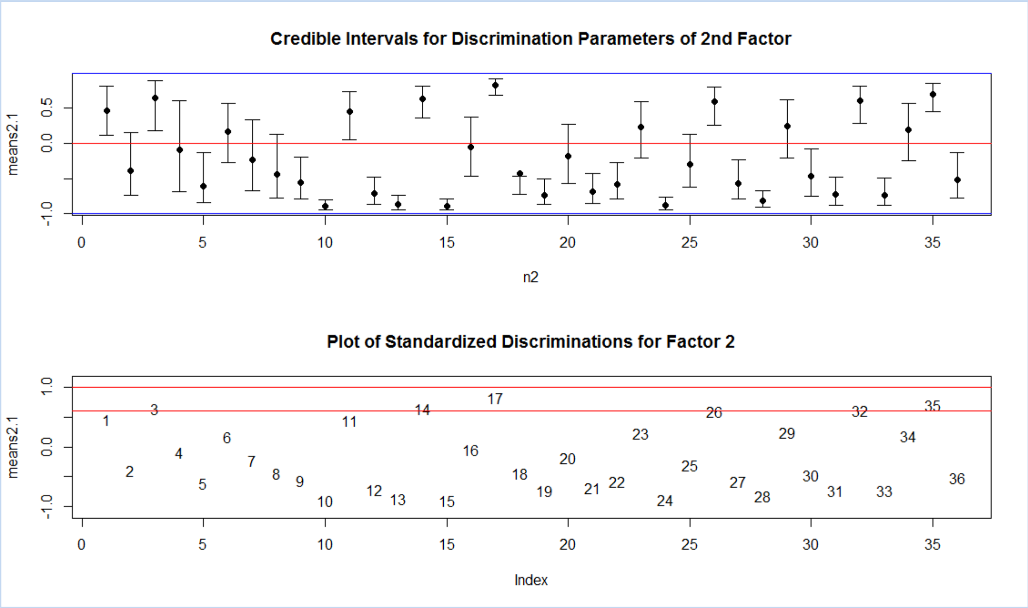


Figure 3.29: Standardized Discriminations for the second factor for model (M6)

A typical person shows high probability to express disagreement to the queries. That conclusion in evident from Figure 3.30.

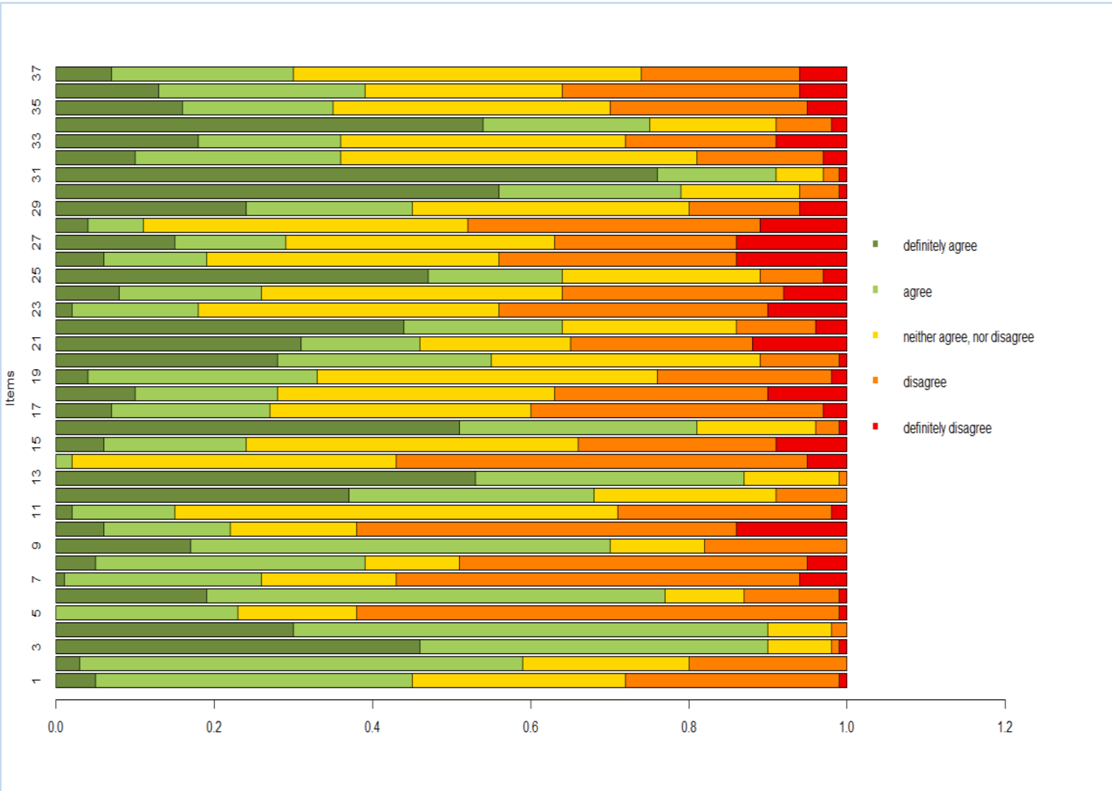
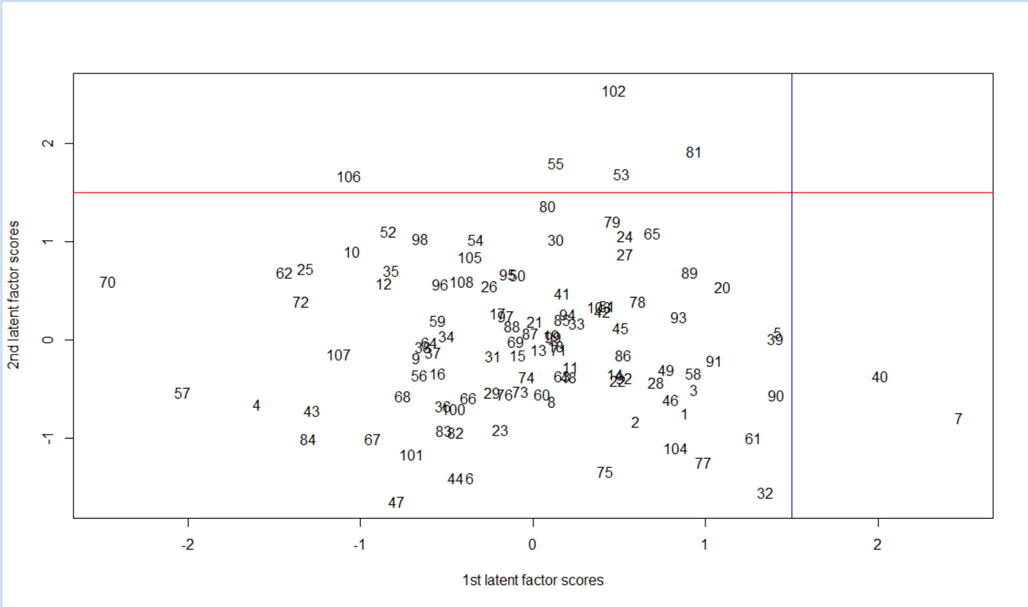


Figure 3.30: Response probabilities of a typical individual in each category for every question for model (M6)



Each number represents the latent score of each sample - member

Figure 3.31: Scatter Plot of latent scores

In Figure 3.31 the scatter plot of latent factors (first versus second) is presented. From this figure it is obvious that latent scores are independently distributed (shape "cloud"). The outlier points indicate individuals with problems. So, if the first and second latent factor represent compulsive and impulsive buying behaviors respectively, individuals who are upper from the red horizontal line considered as individuals with intense impulsive buying behavior (individuals 53, 55, 81, 102, 106). Similarly, individuals who are rightmost of the vertical blue line considered as individuals with strong compulsive buying behavior (individuals 40 and 7). Comparatively, between these cases and previous ones, impulsive buying behavior is affected to a greater extent than compulsive buying behavior.

3.6 Discussion

The main theme of this chapter was the connection between excessive forms of consumer behavior and schizotypy. Firstly, the theoretical frame of consumer behavior was presented. In that point the definition and the excessive forms of consumer behavior i.e. impulsive and compulsive buying were described. As far as schizotypy, its main characteristics were listed with a short description. In conclusion of the theoretical part of this chapter the relationship between schizotypy and consumer behavior was presented.

In the next step of this chapter, 108 fully completed questionnaires by university students were used to study the latent structure of consumer behavior under Item Response Theory approach. From the whole questionnaire, only questions with ordered answers from one two to five, were used. So, two models were constructed (with one and two factors) and the better was chosen via DIC.

Finally, in last model total SPQ scale and nine traits of schizotypy were added, separately of each other. By these models, it was attempted to look for probable relationship between schizotypy and consuming behavior. From all above process seems to exist a strong association between excessive consumer behaviors and especially with compulsive consumption and schizotypy.

As far as the fitness of the models for the scope of this thesis, from my point of view, the model which combines consumer behavior with the nine traits of schizotypy (model M6) is the most appropriate. Through that model it is possible to detect a probable association between consumer behavior and the nine traits schizotypy in detail.

CHAPTER 4: CONCLUSIONS - DISCUSSION

4.1 Conclusions

In this thesis we have studied the latent structure of the consumer behavior and whether is influenced by schizotypy through total SPQ score and via the nine schizotypal traits. These schizotypal traits combined all the available information coming from the 74 items of the schizotypal personality questionnaire (SPQ). Since the SPQ expresses the schizotypal personality disorder, our target was to examine whether or not this psychological disorder influences the consumer behavior.

We constructed latent variable models for ordinal data under the item response theory (IRT) for one and two latent factors. In all models, the logit link function was considered as the most appropriate one due to its straightforward interpretation. In the whole analysis the categorical distribution has been taken as response distribution due to the nature of our data. We have constructed one and two factor model in order to study the consumer behavior and consequently we have also created the further models by adding separately as covariates the total SPQ score and the nine schizotypal traits in the corresponding one and two factor model structure.

The model assessment was carried out via the deviance information criterion. A usual problem which arise in the application of this assessment tool is to result negative number of "effective" parameters for the model. In such cases classic estimate of DIC does not work and alternative but equivalent forms of DIC were used. Hence, in every case of models (consumer behavior, consumer behavior and total SPQ score, consumer behavior and nine traits of schizotypy) a two factor model was considered as the most appropriate.

4.2 Association of Consuming Behavior and Schizotypy

In the questions of the consuming behavior were included questions which detect excessive behaviors such as impulsive and compulsive consumptions. In the construction of the models there is no separation of these two types of consumer behaviors.

A general conclusion is that there is an association between extreme consuming behaviors and schizotypy. A more strong association was observed between compulsive consumption and schizotypy. It must be referred that a person may responds positively to one or more of nine schizotypal characteristics this does not implies that this person has a strange or excessive consumer behavior. In other words, a schizotypal person it is possible to show normal consumer behavior. On the contrary, a schizotypal person with excessive consumer behaviors may show these symptoms more intense than a non-schizotypal person.

4.3 Further Research - Proposals

Further research concerning the association between consumer behavior and schizotypy can be drawn to several directions. First of all, it would be useful to look for a possible interaction between impulsive and compulsive buying behaviors. Especially, major lies in the influence between compulsive behavior with schizotypy and vice versa since such a behavior is considered as a clinical condition by the experts.

As far as the model assessment, Bayes Factor could be used instead of DIC. The disadvantage of Bayes Factor is its completition in high dimensional models. Furthermore, Watanabe - Akaike information criterion (WAIC) can be also considered as an alternative. WAIC is a fully Beyesian approach for estimating the out of sample expectations. It starts with the computed log pointwise posterior predictive density and consequently adds a corrections for the effective number of parameters to adjust for overfitting (more information can be found in Gelman et. al 2013).

In this thesis, the sample is not representative of the general Greek population. The subjects of the sample consist of students in the School of Management Sciences of the University of Aegean and Technological Education Institutes of Crete and Piraeus. Concerning the age of the sample 54% was of the age 18-21, 38% between 22-25 years old, 7% between 26-29 and only one percent older than 30 years old. Even if the financial dependence of the sample does not seem to be a problem a survey with a much higher age interval group would provide a more suitable approach to infer about the association between consumer behavior and SPQ. More specifically concerning the financial dependence, 80% of the subjects answered that are independent financially and only 12% are fully dependent by their parents. Furthermore, all participants have high educational level with 91% enrolled in a B.Sc course and 9% in a M.Sc course (Oikonomou, 2008). Hence, this survey is focused only in a very specific subset of the Greek population.

Concerning the study of the relation of the consumer behavior and schizotypy, it would be interesting to study the inverse association. To see how and if consumer behavior influences the schizotypal traits. In that case, latent variable models under IRT approach would be applied. The difference lies in the nature of data which are binary. Then, the number of factors could be from one to five and consumer behavior could be added as covariates in the model. So, the separation of schizotypal traits from the latent factors can be held with the following way:

- One factor: All scizotypal traits
- Two factors (Kendler's et. al, 1991): The first one is the *positive factor* which includes ideas of reference, odd beliefs, unusual perceptual experiences, suspiciousness, social anxiety and odd speech. The other factor, the *negative* includes suspiciousness, social anxiety, no close friends, constricted affect and odd behavior.
- Three factors (Disorganized three factor model- Raine et al, 1994): The first factor is called *cognitive*. In that factor belong ideas of reference, odd beliefs, unusual perceptual experiences and suspiciousness. The second factor is named *interpersonal*. It contains suspiciousness, social anxiety, no close friends and constricted affect. The rest of the traits, i.e. odd behavior and odd speech, consist the third factor which is the *disorganized* factor.
- Four factors (Paranoid four factor model- Stefanis et al, 2004): The first factor is the *cognitive* which has the traits odd beliefs and unusual perceptual experiences. Next is the *negative* factor. There are suspiciousness, social anxiety, no close friends and constricted affect. Third is the *disorganized* factor which includes the odd behavior and speech. The *paranoid* factor is the fourth and consists of ideas of reference, suspiciousness and social anxiety.
- Five factors (Fogelson et. al, 1999): The *paranoid* factor is the first one and has ideas of reference and suspiciousness. The second is called *positive*

factor and includes ideas of reference, odd beliefs, unusual perceptual experiences. The *schizoid* is the next factor and consists of no close friends, constricted affects and odd speech. The fourth factor is the *avoidant*. There belong the following traits, ideas of reference and social anxiety. The last factor is the *disorganized* and take account of suspiciousness, constricted affect and odd behavior.

It is more than clear that a trait can belong to more than one factor. All of them have advantages and disadvantages too. The three and four factor models are the standard way to model such data.

Thank you for your interest and the time you devote for my thesis.

Mariatta P. Prifti

Appendix A

Discrimination Parameters of Consumer Behavior

Discrimination Parameters	Mean	Sd	MC_error
a[1,1]	1.468	0.3312	0.009164
a[2,1]	3.031	0.6318	0.03682
a[2,2]	0.4091	0.2652	0.01281
a[3,1]	1.177	0.3534	0.01508
a[3,2]	0.8413	0.3114	0.01054
a[4,1]	2.774	0.5612	0.02882
a[4,2]	-0.276	0.3978	0.02216
a[5,1]	2.545	0.5207	0.02743
a[5,2]	0.7708	0.3933	0.02009
a[6,1]	1.553	0.418	0.02191
a[6,2]	1.378	0.3463	0.01389
a[7,1]	0.7585	0.2683	0.006268
a[7,2]	0.007466	0.2368	0.005844
a[8,1]	1.574	0.3664	0.01585
a[8,2]	0.7425	0.2989	0.01145
a[9,1]	1.573	0.3866	0.01852
a[9,2]	1.087	0.3229	0.01325
a[10,1]	0.4478	0.2735	0.01132
a[10,2]	0.9151	0.2569	0.007201
a[11,1]	-0.3587	0.4356	0.02674
a[11,2]	2.311	0.4408	0.02205
a[12,1]	0.5623	0.2546	0.007563
a[12,2]	-0.3955	0.2464	0.006626
a[13,1]	-0.556	0.2874	0.01109
a[13,2]	0.8696	0.2862	0.008308
a[14,2]	2.168	0.3863	0.01393
a[15,1]	0.4882	0.2518	0.009197
a[15,2]	-0.7567	0.258	0.007802
a[16,1]	0.005971	0.4333	0.02698
a[16,2]	2.325	0.4422	0.01886

a[17,1]	0.05228	0.2254	0.004521
a[17,2]	-0.04554	0.2215	0.00404
a[18,1]	0.5813	0.3273	0.01628
a[18,2]	-1.411	0.3251	0.01247
a[19,1]	0.63	0.2759	0.011
a[19,2]	0.7949	0.2556	0.007676
a[20,1]	-0.395	0.2936	0.01285
a[20,2]	1.087	0.2841	0.008978
a[21,1]	-0.07336	0.2415	0.005301
a[21,2]	0.2383	0.2338	0.005072
a[22,1]	-0.3689	0.2945	0.01244
a[22,2]	0.9949	0.2886	0.008334
a[23,1]	-0.411	0.2663	0.009219
a[23,2]	0.7524	0.2483	0.007146
a[24,1]	-0.09461	0.2396	0.005288
a[24,2]	-0.1851	0.2289	0.004633
a[25,1]	-0.3077	0.3916	0.02256
a[25,2]	2.007	0.3973	0.01596
a[26,1]	-0.3648	0.2486	0.006104
a[26,2]	0.3014	0.2278	0.005531
a[27,1]	0.3642	0.2659	0.008174
a[27,2]	-0.5708	0.246	0.006411
a[28,1]	-0.4831	0.269	0.009952
a[28,2]	0.8039	0.2577	0.007428
a[29,1]	0.1592	0.3654	0.02082
a[29,2]	1.696	0.346	0.01245
a[30,1]	0.4441	0.2526	0.006315
a[30,2]	-0.3029	0.2426	0.006323
a[31,1]	0.0112	0.26	0.007703
a[31,2]	0.4098	0.2569	0.005308
a[32,1]	0.7739	0.3206	0.01666
a[32,2]	1.501	0.3139	0.0108
a[33,1]	0.6194	0.2672	0.008409
a[33,2]	-0.6059	0.2557	0.007464
a[34,1]	0.2505	0.3278	0.01692
a[34,2]	1.41	0.319	0.01028
a[35,1]	0.1416	0.2309	0.004871
a[35,2]	-0.1644	0.2316	0.005399

a[36,1]	0.5324	0.2842	0.01123
a[36,2]	-0.9409	0.2741	0.008564
a[37,1]	0.6575	0.2881	0.01209
a[37,2]	0.8944	0.2735	0.007855

Table A.1: a[,1] , a[,2] represent discrimination parameters for the first and second factor respectively

Standardized Discrimination Parameters of Consumer Behavior

Stand. Discrimination Parameters	Mean	Sd	MC_error
st.discr1[1,1]	0.8126	0.06555	0.001798
st.discr1[2,1]	0.9438	0.02333	0.001299
st.discr1[2,2]	0.3518	0.1768	0.007903
st.discr1[3,1]	0.7369	0.1093	0.004576
st.discr1[3,2]	0.6132	0.1488	0.004768
st.discr1[4,1]	0.9346	0.02475	0.001199
st.discr1[4,2]	-0.2304	0.3193	0.01787
st.discr1[5,1]	0.9237	0.02897	0.001536
st.discr1[6,1]	0.8201	0.08424	0.004838
st.discr1[6,2]	0.7918	0.07945	0.002936
st.discr1[7,1]	0.5793	0.1405	0.003157
st.discr1[7,2]	0.00672	0.2198	0.005387
st.discr1[8,1]	0.829	0.0689	0.00306
st.discr1[8,2]	0.5651	0.1615	0.005786
st.discr1[9,1]	0.8275	0.07016	0.003543
st.discr1[9,2]	0.7121	0.1091	0.004145
st.discr1[10,1]	0.3811	0.2073	0.008734
st.discr1[10,2]	0.6556	0.1093	0.00297
st.discr1[11,1]	-0.2801	0.3179	0.01896
st.discr1[11,2]	0.9108	0.03093	0.00152
st.discr1[12,1]	0.4649	0.1669	0.00478
st.discr1[12,2]	-0.3454	0.1911	0.005103
st.discr1[13,1]	-0.4544	0.191	0.00713
st.discr1[13,2]	0.6309	0.1308	0.003706
st.discr1[14,1]	0.4315	0.2796	0.01705
st.discr1[14,2]	0.9015	0.03168	0.001121
st.discr1[15,1]	0.4142	0.1808	0.006313

st.discr1[15,2]	-0.5802	0.1352	0.003896
st.discr1[16,1]	0.01239	0.3517	0.02157
st.discr1[16,2]	0.9118	0.03087	0.001267
st.discr1[17,1]	0.04882	0.2105	0.004238
st.discr1[17,2]	-0.04273	0.207	0.003765
st.discr1[18,1]	0.463	0.2086	0.009765
st.discr1[18,2]	-0.8014	0.06813	0.002599
st.discr1[19,1]	0.5043	0.1729	0.007101
st.discr1[19,2]	0.6002	0.1287	0.0038
st.discr1[20,1]	-0.337	0.2266	0.00974
st.discr1[21,1]	-0.06761	0.2228	0.004921
st.discr1[21,2]	0.2168	0.2044	0.004414
st.discr1[22,1]	-0.3162	0.2288	0.009485
st.discr1[22,2]	0.6835	0.11	0.003072
st.discr1[23,1]	-0.3543	0.2052	0.007077
st.discr1[23,2]	0.5795	0.1308	0.00367
st.discr1[24,1]	-0.08789	0.22	0.004856
st.discr1[24,2]	-0.1704	0.2058	0.004151
st.discr1[25,1]	-0.2506	0.3044	0.01705
st.discr1[25,2]	0.8866	0.03811	0.00149
st.discr1[26,1]	-0.3209	0.1982	0.00475
st.discr1[26,2]	0.2718	0.1921	0.004703
st.discr1[27,1]	0.3176	0.2105	0.006342
st.discr1[27,2]	-0.472	0.162	0.004077
st.discr1[28,1]	-0.4074	0.1929	0.006924
st.discr1[28,2]	0.6042	0.1282	0.003638
st.discr1[29,1]	0.1389	0.3109	0.01775
st.discr1[29,2]	0.8509	0.04994	0.001686
st.discr1[30,1]	0.3816	0.1871	0.004631
st.discr1[30,2]	-0.2709	0.2024	0.005285
st.discr1[31,1]	0.01009	0.2386	0.007076
st.discr1[31,2]	0.3548	0.197	0.003999
st.discr1[32,1]	0.5772	0.1745	0.009518
st.discr1[32,2]	0.8202	0.05964	0.00195
st.discr1[33,1]	0.4995	0.1673	0.005045
st.discr1[33,2]	-0.493	0.1617	0.004676
st.discr1[34,1]	0.2174	0.2741	0.01429
st.discr1[34,2]	0.8014	0.06902	0.002119

st.discr1[35,1]	0.1309	0.21	0.004431
st.discr1[35,2]	-0.1516	0.2087	0.004825
st.discr1[36,1]	0.4391	0.1936	0.007293
st.discr1[36,2]	-0.6638	0.1139	0.003624
st.discr1[37,1]	0.5187	0.1774	0.007717
st.discr1[37,2]	0.6442	0.12	0.003316

Table A.2: st.discr1[,1] , st.discr1[,2] represent standardized discrimination parameters for the first and second factor respectively

Difficulty Parameters of Consumer Behavior

Dif. Parameters	Mean	Sd	MC_error
b[1,1]	-2.908	0.4306	0.00937
b[1,2]	-0.1955	0.2545	0.007706
b[1,3]	0.9527	0.2725	0.007797
b[1,4]	4.444	0.7019	0.01048
b[2,1]	-3.501	0.6234	0.02633
b[2,2]	0.4043	0.4075	0.01739
b[2,3]	1.451	0.452	0.02029
b[2,4]	6.28	1.101	0.0447
b[3,1]	-0.1897	0.263	0.008421
b[3,2]	2.176	0.362	0.009679
b[3,3]	4.092	0.6386	0.01497
b[3,4]	4.752	0.7785	0.01766
b[4,1]	-0.8468	0.3819	0.01385
b[4,2]	2.257	0.4781	0.01974
b[4,3]	3.838	0.6454	0.02601
b[4,4]	7.753	1.54	0.03921
b[5,1]	-5.502	0.923	0.03134
b[5,2]	-1.191	0.3884	0.01579
b[5,3]	-0.4766	0.3666	0.01495
b[5,4]	5.018	0.842	0.02988
b[6,1]	-1.265	0.3437	0.01237
b[6,2]	1.378	0.3551	0.01378
b[6,3]	2.053	0.3894	0.0143
b[6,4]	5.068	0.8217	0.02004
b[7,1]	-5.49	1.292	0.01218
b[7,2]	-1.425	0.2614	0.005571
b[7,3]	-0.7006	0.2289	0.005096

b[7,4]	2.289	0.3413	0.005251
b[8,1]	-2.506	0.4022	0.01212
b[8,2]	-0.07285	0.2812	0.01081
b[8,3]	0.4087	0.2843	0.01083
b[8,4]	3.27	0.4943	0.01292
b[9,1]	-2.156	0.3895	0.01312
b[9,2]	0.3021	0.297	0.01163
b[9,3]	0.945	0.3118	0.01174
b[9,4]	5.598	0.9723	0.0178
b[10,1]	-2.818	0.3978	0.008476
b[10,2]	-1.316	0.2645	0.007553
b[10,3]	-0.4977	0.2362	0.007092
b[10,4]	1.811	0.3006	0.006797
b[11,1]	-4.339	0.669	0.02387
b[11,2]	-2.25	0.4335	0.01649
b[11,3]	0.4738	0.3492	0.01433
b[11,4]	3.38	0.5435	0.02112
b[12,1]	-0.5135	0.2199	0.00437
b[12,2]	0.7486	0.2261	0.004283
b[12,3]	2.359	0.3459	0.005016
b[12,4]	5.496	1.275	0.01228
b[13,1]	-0.3368	0.2347	0.005938
b[13,2]	1.354	0.2715	0.006376
b[14,3]	4.095	0.6649	0.008592
b[13,4]	28.9	18.07	0.182
b[14,1]	-6.202	1.02	0.02249
b[14,2]	-4.691	0.7093	0.01964
b[14,3]	-1.03	0.3587	0.01365
b[14,4]	2.083	0.4103	0.01562
b[15,1]	-2.751	0.3921	0.006528
b[15,2]	-1.119	0.2513	0.005513
b[15,3]	0.6749	0.2344	0.005574
b[15,4]	2.347	0.348	0.006791
b[16,1]	-0.7048	0.3524	0.01437
b[16,2]	0.7553	0.3539	0.0149
b[16,3]	2.387	0.4379	0.0168
b[16,4]	4.298	0.6342	0.02037
b[17,1]	-2.365	0.3414	0.003729

b[17,2]	-0.7564	0.2116	0.002596
b[17,3]	0.6285	0.2067	0.002401
b[17,4]	3.771	0.6436	0.004899
b[18,1]	-1.54	0.315	0.01048
b[18,2]	-0.3728	0.2742	0.009802
b[18,3]	1.027	0.2902	0.00991
b[18,4]	2.647	0.4043	0.01139
b[19,1]	-3.802	0.5694	0.008458
b[19,2]	-1.407	0.2757	0.00676
b[19,3]	0.4536	0.2387	0.006178
b[19,4]	3.282	0.4727	0.007283
b[20,1]	-1.428	0.2846	0.007477
b[20,2]	-0.2628	0.2432	0.007224
b[20,3]	1.619	0.2922	0.007685
b[20,4]	4.666	0.8416	0.01096
b[21,1]	-0.8169	0.215	0.003677
b[21,2]	-0.1509	0.199	0.003482
b[21,3]	0.6391	0.2068	0.003331
b[21,4]	1.99	0.2958	0.003434
b[22,1]	-0.2691	0.2351	0.006535
b[22,2]	0.5521	0.2402	0.006875
b[22,3]	1.822	0.3016	0.007646
b[22,4]	3.282	0.4749	0.008852
b[23,1]	-4.566	0.8263	0.00976
b[23,2]	-2.072	0.3148	0.006328
b[23,3]	-0.3766	0.225	0.005362
b[23,4]	1.527	0.2732	0.005114
b[24,1]	-3.012	0.4479	0.004374
b[24,2]	-1.605	0.2635	0.002994
b[24,3]	-8,61E-01	0.1979	0.002342
b[24,4]	1.815	0.28	0.002798
b[25,1]	-0.7737	0.3276	0.01291
b[25,2]	-0.04373	0.3174	0.01288
b[25,3]	1.509	0.3544	0.01372
b[25,4]	2.985	0.4446	0.01467
b[26,1]	-2.707	0.3913	0.005051
b[26,2]	-1.445	0.2575	0.004017
b[26,3]	0.2165	0.2057	0.003107

b[26,4]	1.834	0.285	0.003183
b[27,1]	-1.683	0.2743	0.005275
b[27,2]	-0.8524	0.2277	0.004948
b[27,3]	0.5101	0.2195	0.004732
b[27,4]	1.799	0.2854	0.005045
b[28,1]	-3.514	0.5232	0.008206
b[28,2]	-2.449	0.3624	0.006717
b[28,3]	-0.3038	0.2295	0.005278
b[28,4]	1.711	0.2905	0.00599
b[29,1]	-1.229	0.3109	0.01085
b[29,2]	-0.2382	0.2882	0.01079
b[29,3]	1.402	0.3228	0.0115
b[29,4]	2.813	0.4237	0.01295
b[30,1]	-0.08184	0.2082	0.003663
b[30,2]	1.021	0.2314	0.004007
b[30,3]	2.432	0.353	0.004816
b[30,4]	4.441	0.8166	0.008404
b[31,1]	0.3095	0.212	0.003533
b[31,2]	1.32	0.2502	0.00425
b[31,3]	2.549	0.3691	0.005622
b[31,4]	3.934	0.642	0.008228
b[32,1]	-3.404	0.4884	0.01274
b[32,2]	-1.803	0.3414	0.01152
b[32,3]	0.2118	0.2879	0.01068
b[32,4]	2.167	0.3632	0.01138
b[33,1]	-1.479	0.2694	0.005471
b[33,2]	-0.5359	0.2265	0.005189
b[33,3]	0.9348	0.2393	0.005377
b[33,4]	2.246	0.3337	0.005898
b[34,1]	-0.5923	0.2745	0.009936
b[34,2]	0.367	0.2705	0.01004
b[34,3]	1.56	0.3098	0.0104
b[34,4]	3.173	0.4584	0.01173
b[35,1]	-1.681	0.2698	0.003397
b[35,2]	-0.6274	0.2107	0.002759
b[35,3]	0.8545	0.2189	0.002711
b[35,4]	2.978	0.4431	0.004357
b[36,1]	-1.328	0.2678	0.00699

b[36,2]	0.06165	0.2353	0.006781
b[36,3]	1.106	0.2635	0.007023
b[36,4]	3.177	0.459	0.008363
b[37,1]	-2.578	0.3783	0.008432
b[37,2]	-0.8705	0.2555	0.007521
b[37,3]	1.023	0.2646	0.007351
b[37,4]	2.686	0.3982	0.007961

Table A.3: Categories b[,i] where i: 1: definitely agree, 2: agree, 3: neither agree, nor disagree, 4:disagree, 5: definitely disagree. Questions: b[j,] where j=1,…,37

Cumulative Probabilities of a "Median" Individual in Each Category of Consumer Behavior

Questions	Categories				
	1	2	3	4	5
[1]	0.05	0.45	0.72	0.99	1
[2]	0.03	0.60	0.81	1.00	1
[3]	0.45	0.90	0.98	0.99	1
[4]	0.30	0.91	0.98	1.00	1
[5]	0.00	0.23	0.38	0.99	1
[6]	0.22	0.80	0.89	0.99	1
[7]	0.00	0.19	0.33	0.91	1
[8]	0.08	0.48	0.60	0.96	1
[9]	0.10	0.57	0.72	1.00	1
[10]	0.06	0.21	0.38	0.86	1
[11]	0.01	0.10	0.62	0.97	1
[12]	0.37	0.68	0.91	1.00	1
[13]	0.42	0.79	0.98	1.00	1
[14]	0.00	0.01	0.26	0.89	1
[15]	0.06	0.25	0.66	0.91	1
[16]	0.33	0.68	0.92	0.99	1
[17]	0.09	0.32	0.65	0.98	1
[18]	0.18	0.41	0.74	0.93	1
[19]	0.02	0.20	0.61	0.96	1
[20]	0.19	0.43	0.83	0.99	1
[21]	0.31	0.46	0.65	0.88	1
[22]	0.43	0.63	0.86	0.96	1
[23]	0.01	0.11	0.41	0.82	1

[24]	0.05	0.17	0.30	0.86	1
[25]	0.32	0.49	0.82	0.95	1
[26]	0.06	0.19	0.55	0.86	1
[27]	0.16	0.30	0.62	0.86	1
[28]	0.03	0.08	0.42	0.85	1
[29]	0.23	0.44	0.80	0.94	1
[30]	0.48	0.74	0.92	0.99	1
[31]	0.58	0.79	0.93	0.98	1
[32]	0.03	0.14	0.55	0.90	1
[33]	0.19	0.37	0.72	0.90	1
[34]	0.36	0.59	0.83	0.96	1
[35]	0.16	0.35	0.70	0.95	1
[36]	0.21	0.52	0.75	0.96	1
[37]	0.07	0.30	0.74	0.94	1

Categories: 1: definitely agree, 2: agree, 3: neither agree, nor disagree, 4:disagree, 5: definitely disagree (reference category)

Table A.4: Cumulative probabilities of consumer behavior for a typical person using equation (1.18)

Response Probabilities of a "Median" Individual in Each Category of Consumer Behavior

Questions	Categories				
	1	2	3	4	5
[1]	0.05	0.40	0.27	0.27	0.01
[2]	0.03	0.57	0.21	0.19	0.00
[3]	0.45	0.45	0.08	0.01	0.01
[4]	0.30	0.61	0.07	0.02	0.00
[5]	0.00	0.23	0.15	0.61	0.01
[6]	0.22	0.58	0.09	0.10	0.01
[7]	0.00	0.19	0.14	0.58	0.09
[8]	0.08	0.40	0.12	0.36	0.04
[9]	0.10	0.47	0.15	0.28	0.00
[10]	0.06	0.15	0.17	0.48	0.14
[11]	0.01	0.09	0.52	0.35	0.03
[12]	0.37	0.31	0.23	0.09	0.00
[13]	0.42	0.37	0.19	0.02	0.00
[14]	0.00	0.01	0.25	0.63	0.11
[15]	0.06	0.19	0.41	0.25	0.09

[16]	0.33	0.35	0.24	0.07	0.01
[17]	0.09	0.23	0.33	0.33	0.02
[18]	0.18	0.23	0.33	0.19	0.07
[19]	0.02	0.18	0.41	0.35	0.04
[20]	0.19	0.24	0.40	0.16	0.01
[21]	0.31	0.15	0.19	0.23	0.12
[22]	0.43	0.20	0.23	0.10	0.04
[23]	0.01	0.10	0.30	0.41	0.18
[24]	0.05	0.12	0.13	0.56	0.14
[25]	0.32	0.17	0.33	0.13	0.05
[26]	0.06	0.13	0.36	0.31	0.14
[27]	0.16	0.14	0.32	0.24	0.14
[28]	0.03	0.05	0.34	0.43	0.15
[29]	0.23	0.21	0.36	0.14	0.06
[30]	0.48	0.26	0.18	0.07	0.01
[31]	0.58	0.21	0.14	0.05	0.02
[32]	0.03	0.11	0.41	0.35	0.10
[33]	0.19	0.18	0.35	0.18	0.10
[34]	0.36	0.23	0.24	0.13	0.04
[35]	0.16	0.19	0.35	0.25	0.05
[36]	0.21	0.31	0.23	0.21	0.04
[37]	0.07	0.23	0.44	0.20	0.06

Categories: 1: definitely agree, 2: agree, 3: neither agree, nor disagree, 4:disagree, 5: definitely disagree (reference category)

Table A.5: Response probabilities of consumer behavior for a typical person using equation (1.19)

Discrimination Parameters of Total SPQ Scale on Consumer Behavior (Model 3)

Discrimination Parameters	mean	sd	MC_error
a[1,1]	1.453	0.327	0.007348
a[2,1]	2.984	0.6326	0.03029
a[2,2]	0.3947	0.249	0.0104
a[3,1]	1.129	0.348	0.01268
a[3,2]	0.801	0.2978	0.009501
a[4,1]	2.694	0.5378	0.02189
a[4,2]	-0.3037	0.3819	0.01811
a[5,1]	2.584	0.5262	0.02223

a[5,2]	0.7853	0.41	0.01975
a[6,1]	1.529	0.4151	0.01905
a[6,2]	1.347	0.3524	0.01313
a[7,1]	0.7426	0.2707	0.006328
a[7,2]	0.00744	0.2425	0.00616
a[8,1]	1.564	0.3601	0.01394
a[8,2]	0.736	0.3082	0.01254
a[9,1]	1.502	0.3665	0.01523
a[9,2]	1.024	0.3298	0.0141
a[10,1]	0.4355	0.2763	0.01121
a[10,2]	0.8702	0.2543	0.006337
a[11,1]	-0.3402	0.4455	0.02643
a[11,2]	2.247	0.4146	0.01521
a[12,1]	0.5542	0.256	0.006745
a[12,2]	-0.4066	0.2477	0.006191
a[13,1]	-0.5673	0.2905	0.01057
a[13,2]	0.8286	0.2808	0.006898
a[14,1]	0.6048	0.4469	0.02632
a[14,2]	2.217	0.4325	0.01682
a[15,1]	0.4485	0.2568	0.009488
a[15,2]	-0.7775	0.2493	0.005752
a[16,1]	0.01363	0.4376	0.02582
a[16,2]	2.275	0.4469	0.01576
a[17,1]	0.0195	0.2296	0.004461
a[17,2]	-0.07875	0.2182	0.004192
a[18,1]	0.5645	0.3268	0.01564
a[18,2]	-1.445	0.3254	0.01174
a[19,1]	0.6677	0.2848	0.01107
a[19,2]	0.8174	0.2579	0.008259
a[20,1]	-0.4056	0.3038	0.0131
a[20,2]	1.06	0.2841	0.007618
a[21,1]	-0.09582	0.2461	0.005302
a[21,2]	0.2125	0.2299	0.004688
a[22,1]	-0.4013	0.2898	0.01196
a[22,2]	0.9282	0.2797	0.007351
a[23,1]	-0.428	0.2691	0.009831
a[23,2]	0.7463	0.243	0.005645
a[24,1]	-0.1498	0.2373	0.005404

a[24,2]	-0.2056	0.2309	0.004421
a[25,1]	-0.2935	0.4041	0.02251
a[25,2]	1.936	0.383	0.01231
a[26,1]	-0.3961	0.2442	0.00617
a[26,2]	0.2732	0.2199	0.004178
a[27,1]	0.291	0.2703	0.008334
a[27,2]	-0.6322	0.2496	0.005905
a[28,1]	-0.4963	0.2782	0.01069
a[28,2]	0.7686	0.2554	0.00649
a[29,1]	0.1392	0.3596	0.01913
a[29,2]	1.637	0.3358	0.00994
a[30,1]	0.4503	0.2543	0.005834
a[30,2]	-0.3068	0.242	0.005508
a[31,1]	-0.006771	0.2581	0.005927
a[31,2]	0.3819	0.254	0.004348
a[32,1]	0.7871	0.3536	0.01781
a[32,2]	1.508	0.3237	0.01161
a[33,1]	0.5495	0.2717	0.009465
a[33,2]	-0.6682	0.2589	0.006696
a[34,1]	0.2456	0.3243	0.01563
a[34,2]	1.365	0.3148	0.009429
a[35,1]	0.1132	0.234	0.005869
a[35,2]	-0.2017	0.2259	0.00422
a[36,1]	0.4896	0.2947	0.01187
a[36,2]	-0.9706	0.2802	0.007895
a[37,1]	0.6361	0.3009	0.01199
a[37,2]	0.8735	0.2648	0.008788

Table A.6: a[,1] , a[,2] represent discrimination parameters for the first and second factor respectively

Standardized Discrimination Parameters of Total SPQ Scale on Consumer Behavior (Model 3)

Stand. Discrimination Parameters	Mean	Sd	MC_error
st.discr1[1,1]	0.81	0.0655	0.001383
st.discr1[2,1]	0.9424	0.02211	9,24E-01
st.discr1[2,2]	0.3432	0.1725	0.007063

st.discr1[3,1]	0.7219	0.116	0.004059
st.discr1[3,2]	0.5958	0.1498	0.004659
st.discr1[4,1]	0.9313	0.02585	0.001053
st.discr1[4,2]	-0.2516	0.3012	0.01427
st.discr1[5,1]	0.9257	0.02845	0.00116
st.discr1[5,2]	0.5639	0.217	0.01026
st.discr1[6,1]	0.8165	0.08178	0.003725
st.discr1[6,2]	0.7838	0.08388	0.00316
st.discr1[7,1]	0.5704	0.1456	0.00343
st.discr1[7,2]	0.006644	0.225	0.005741
st.discr1[8,1]	0.8281	0.06451	0.002402
st.discr1[8,2]	0.5596	0.1687	0.006684
st.discr1[9,1]	0.8159	0.07193	0.00291
st.discr1[9,2]	0.6884	0.1218	0.005026
st.discr1[10,1]	0.3711	0.2079	0.008441
st.discr1[10,2]	0.6363	0.1156	0.00275
st.discr1[11,1]	-0.268	0.3358	0.01962
st.discr1[11,2]	0.907	0.03101	0.001085
st.discr1[12,1]	0.4592	0.1705	0.00439
st.discr1[12,2]	-0.3539	0.191	0.004746
st.discr1[13,1]	-0.4613	0.1918	0.006785
st.discr1[13,2]	0.6126	0.1339	0.003257
st.discr1[14,1]	0.4505	0.2817	0.0166
st.discr1[14,2]	0.9041	0.03343	0.001249
st.discr1[15,1]	0.3843	0.1926	0.00709
st.discr1[15,2]	-0.5924	0.1285	0.002889
st.discr1[16,1]	0.01245	0.359	0.02112
st.discr1[16,2]	0.9079	0.03372	0.00113
st.discr1[17,1]	0.0183	0.2142	0.004152
st.discr1[17,2]	-0.0737	0.203	0.003904
st.discr1[18,1]	0.4519	0.2189	0.01028
st.discr1[18,2]	-0.8086	0.06484	0.002181
st.discr1[19,1]	0.5253	0.1696	0.006547
st.discr1[19,2]	0.6107	0.128	0.004018
st.discr1[20,1]	-0.3436	0.2303	0.009799
st.discr1[20,2]	0.7079	0.09961	0.002605
st.discr1[21,1]	-0.08826	0.2253	0.004787
st.discr1[21,2]	0.1949	0.2044	0.00416

st.discr1[22,1]	-0.3428	0.2235	0.009121
st.discr1[22,2]	0.6578	0.1176	0.002933
st.discr1[23,1]	-0.3666	0.2034	0.007218
st.discr1[23,2]	0.5771	0.1294	0.00297
st.discr1[24,1]	-0.1381	0.2155	0.004888
st.discr1[24,2]	-0.1887	0.2055	0.003929
st.discr1[25,1]	-0.2381	0.3162	0.01757
st.discr1[25,2]	0.8797	0.0404	0.001223
st.discr1[26,1]	-0.3464	0.1899	0.004754
st.discr1[26,2]	0.2489	0.1899	0.003596
st.discr1[27,1]	0.2569	0.2244	0.006805
st.discr1[27,2]	-0.5105	0.1519	0.003535
st.discr1[28,1]	-0.4151	0.1972	0.007441
st.discr1[28,2]	0.5867	0.1336	0.00336
st.discr1[29,1]	0.1197	0.3053	0.01632
st.discr1[29,2]	0.8425	0.05253	0.001469
st.discr1[30,1]	0.3858	0.1882	0.004257
st.discr1[30,2]	-0.2745	0.2019	0.004539
st.discr1[31,1]	-0.006039	0.2367	0.005435
st.discr1[31,2]	0.3335	0.199	0.003369
st.discr1[32,1]	0.5772	0.1874	0.009318
st.discr1[32,2]	0.8208	0.06037	0.002109
st.discr1[33,1]	0.4531	0.1812	0.006133
st.discr1[33,2]	-0.5305	0.1541	0.003926
st.discr1[34,1]	0.213	0.2715	0.01312
st.discr1[34,2]	0.792	0.07048	0.002045
st.discr1[35,1]	0.1045	0.2135	0.005314
st.discr1[35,2]	-0.1855	0.2013	0.003759
st.discr1[36,1]	0.4073	0.212	0.008452
st.discr1[36,2]	-0.6751	0.1111	0.003017
st.discr1[37,1]	0.5031	0.1855	0.007396
st.discr1[37,2]	0.6362	0.1203	0.003904

Table A.7: st.discr1[,1] , st.discr1[,2] represent standardized discrimination parameters for the first and second factor respectively

Difficulty Parameters of Total SPQ Scale on Consumer Behavior (Model 3)

Dif. Parameters	Mean	Sd	MC_error
b[1,1]	-2.583	0.4314	0.01013
b[1,2]	0.1248	0.2733	0.009521
b[1,3]	1.275	0.2872	0.009755
b[1,4]	4.776	0.7186	0.01193
b[2,1]	-3.152	0.6521	0.02682
b[2,2]	0.7364	0.4274	0.01868
b[2,3]	1.776	0.4765	0.021
b[2,4]	6.593	1.157	0.04213
b[3,1]	0.1225	0.2672	0.008843
b[3,2]	2.468	0.3612	0.01065
b[3,3]	4.364	0.6184	0.01623
b[3,4]	5.018	0.7504	0.01859
b[4,1]	-0.5046	0.3957	0.01545
b[4,2]	2.563	0.4822	0.01875
b[4,3]	4.13	0.6391	0.02314
b[4,4]	8.019	1.53	0.03545
b[5,1]	-5.234	0.938	0.02865
b[5,2]	-0.8807	0.394	0.01629
b[5,3]	-0.1564	0.3747	0.01562
b[5,4]	5.426	0.865	0.02544
b[6,1]	-0.9475	0.3413	0.01162
b[6,2]	1.694	0.3508	0.013
b[6,3]	2.369	0.3854	0.01356
b[6,4]	5.371	0.8424	0.01845
b[7,1]	-5.162	1.274	0.01291
b[7,2]	-1.105	0.2727	0.007162
b[7,3]	-0.3809	0.2394	0.006916
b[7,4]	2.61	0.349	0.007374
b[8,1]	-2.194	0.4001	0.01162
b[8,2]	0.2401	0.2872	0.01135
b[8,3]	0.723	0.2911	0.01148
b[8,4]	3.588	0.4971	0.01268

b[9,1]	-1.811	0.3832	0.01211
b[9,2]	0.6223	0.2939	0.01108
b[9,3]	1.255	0.3053	0.01111
b[9,4]	5.832	0.9604	0.01726
b[10,1]	-2.505	0.3941	0.007937
b[10,2]	-1.01	0.2687	0.007482
b[10,3]	-0.1922	0.2414	0.007202
b[10,4]	2.11	0.3044	0.007364
b[11,1]	-4.001	0.6484	0.01774
b[11,2]	-1.946	0.4276	0.01392
b[11,3]	0.7432	0.3393	0.01237
b[11,4]	3.645	0.5306	0.01657
b[12,1]	-0.2134	0.2341	0.00697
b[12,2]	1.049	0.2401	0.00708
b[12,3]	2.662	0.3553	0.007811
b[12,4]	5.787	1.265	0.01403
b[13,1]	-0.05933	0.2441	0.007113
b[13,2]	1.634	0.2782	0.007297
b[13,3]	4.385	0.671	0.009568
b[13,4]	29.15	18.02	0.1591
b[14,1]	-5.992	1.049	0.02464
b[14,2]	-4.473	0.7331	0.02029
b[14,3]	-0.7579	0.3581	0.01274
b[14,4]	2.417	0.4209	0.01448
b[15,1]	-2.448	0.4041	0.008282
b[15,2]	-0.8189	0.2621	0.007481
b[15,3]	0.9806	0.2482	0.007503
b[15,4]	2.669	0.3587	0.008362
b[16,1]	-0.4151	0.3486	0.01285
b[16,2]	1.039	0.3455	0.01315
b[16,3]	2.67	0.4263	0.01465
b[16,4]	4.574	0.6224	0.01721
b[17,1]	-2.075	0.3532	0.006399
b[17,2]	-0.4622	0.2296	0.005992
b[17,3]	0.9229	0.2261	0.005911
b[17,4]	4.073	0.6519	0.007672
b[18,1]	-1.226	0.3254	0.01198
b[18,2]	-0.04227	0.2898	0.01095

b[18,3]	1.354	0.3052	0.01033
b[18,4]	2.988	0.4197	0.01173
b[19,1]	-3.529	0.5789	0.009165
b[19,2]	-1.127	0.2836	0.007576
b[19,3]	0.7538	0.2469	0.007099
b[19,4]	3.615	0.4896	0.008363
b[20,1]	-1.151	0.2924	0.008319
b[20,2]	0.0106	0.2489	0.007765
b[20,3]	1.899	0.2977	0.007927
b[20,4]	4.955	0.8345	0.01137
b[21,1]	-0.5371	0.2269	0.006101
b[21,2]	0.1268	0.2119	0.006093
b[21,3]	0.9206	0.2199	0.005928
b[21,4]	2.285	0.3053	0.00606
b[22,1]	0.01962	0.2444	0.007383
b[22,2]	0.838	0.2489	0.007626
b[22,3]	2.102	0.3066	0.008185
b[22,4]	3.569	0.4761	0.00935
b[23,1]	-4.298	0.8319	0.01027
b[23,2]	-1.797	0.3245	0.007556
b[23,3]	-0.09624	0.2382	0.006951
b[23,4]	1.822	0.2833	0.007118
b[24,1]	-2.719	0.4494	0.00717
b[24,2]	-1.314	0.2744	0.006322
b[24,3]	0.3012	0.217	0.006111
b[24,4]	2.131	0.2928	0.006161
b[25,1]	-0.488	0.3181	0.01101
b[25,2]	0.2308	0.305	0.01076
b[25,3]	1.775	0.346	0.0117
b[25,4]	3.243	0.4398	0.01288
b[26,1]	-2.433	0.4007	0.006994
b[26,2]	-1.165	0.2689	0.006322
b[26,3]	0.5037	0.2209	0.005776
b[26,4]	2.133	0.2996	0.006384
b[27,1]	-1.393	0.2911	0.007897
b[27,2]	-0.5568	0.2462	0.007602
b[27,3]	0.8187	0.2372	0.007436
b[27,4]	2.12	0.2984	0.007596

b[28,1]	-3.235	0.5278	0.009394
b[28,2]	-2.171	0.3696	0.008074
b[28,3]	-0.0195	0.2394	0.006998
b[28,4]	2.001	0.2971	0.00789
b[29,1]	-0.9298	0.3095	0.01024
b[29,2]	0.05379	0.2805	0.009723
b[29,3]	1.683	0.3123	0.009867
b[29,4]	3.088	0.4079	0.01051
b[30,1]	0.2356	0.2263	0.006681
b[30,2]	1.34	0.2495	0.007079
b[30,3]	2.752	0.3653	0.007568
b[30,4]	4.76	0.8245	0.01064
b[31,1]	0.6154	0.2187	0.005704
b[31,2]	1.624	0.2552	0.005907
b[31,3]	2.849	0.3728	0.006666
b[31,4]	4.231	0.649	0.008811
b[32,1]	-3.121	0.4897	0.01152
b[32,2]	-1.518	0.341	0.01048
b[32,3]	0.5125	0.2897	0.009611
b[32,4]	2.484	0.3711	0.01042
b[33,1]	-1.19	0.281	0.007362
b[33,2]	-0.2452	0.2429	0.007356
b[33,3]	1.235	0.2567	0.007846
b[33,4]	2.563	0.348	0.008758
b[34,1]	-0.2945	0.271	0.008572
b[34,2]	0.6638	0.2671	0.008656
b[34,3]	1.854	0.3065	0.008796
b[34,4]	3.461	0.4503	0.00992
b[35,1]	-1.395	0.277	0.00597
b[35,2]	-0.3436	0.2213	0.005799
b[35,3]	1.143	0.2296	0.005954
b[35,4]	3.283	0.4477	0.007005
b[36,1]	-1.02	0.2881	0.009217
b[36,2]	0.3752	0.2534	0.008959
b[36,3]	1.426	0.2797	0.009222
b[36,4]	3.518	0.4696	0.01033
b[37,1]	-2.268	0.3792	0.008438
b[37,2]	-0.5647	0.2562	0.007614

b[37,3]	1.33	0.2657	0.007674
b[37,4]	2.988	0.3938	0.008504

Table A.8: Categories b[,i] where i: 1: definitely agree, 2: agree, 3: neither agree, nor disagree, 4:disagree, 5: definitely disagree. Questions: b[j,] where j=1,…,37

Cumulative Probabilities of a "Median" Individual in Each Category Total SPQ Scale on Consumer Behavior (Model 3)

Questions	Categories				
	1	2	3	4	5
[1]	0.07	0.53	0.78	0.99	1
[2]	0.04	0.68	0.86	1.00	1
[3]	0.53	0.92	0.99	0.99	1
[4]	0.38	0.93	0.98	1.00	1
[5]	0.01	0.29	0.46	1.00	1
[6]	0.28	0.84	0.91	1.00	1
[7]	0.01	0.25	0.41	0.93	1
[8]	0.10	0.56	0.67	0.97	1
[9]	0.14	0.65	0.78	1.00	1
[10]	0.08	0.27	0.45	0.89	1
[11]	0.02	0.12	0.68	0.97	1
[12]	0.45	0.74	0.93	1.00	1
[13]	0.49	0.84	0.99	1.00	1
[14]	0.00	0.01	0.32	0.92	1
[15]	0.08	0.31	0.73	0.94	1
[16]	0.40	0.74	0.94	0.99	1
[17]	0.11	0.39	0.72	0.98	1
[18]	0.23	0.49	0.79	0.95	1
[19]	0.03	0.24	0.68	0.97	1
[20]	0.24	0.50	0.87	0.99	1
[21]	0.37	0.53	0.72	0.91	1
[22]	0.50	0.70	0.89	0.97	1
[23]	0.01	0.14	0.48	0.86	1
[24]	0.06	0.21	0.57	0.89	1
[25]	0.38	0.56	0.86	0.96	1
[26]	0.08	0.24	0.62	0.89	1
[27]	0.20	0.36	0.69	0.89	1

[28]	0.04	0.10	0.50	0.88	1
[29]	0.28	0.51	0.84	0.96	1
[30]	0.56	0.79	0.94	0.99	1
[31]	0.65	0.84	0.95	0.99	1
[32]	0.04	0.18	0.63	0.92	1
[33]	0.23	0.44	0.77	0.93	1
[34]	0.43	0.66	0.86	0.97	1
[35]	0.20	0.41	0.76	0.96	1
[36]	0.27	0.59	0.81	0.97	1
[37]	0.09	0.36	0.79	0.95	1

Categories: 1: definitely agree, 2: agree, 3: neither agree, nor disagree, 4:disagree, 5: definitely disagree (reference category)

Table A.9: Cumulative probabilities of consumer behavior for a typical person using equation (1.18)

Response Probabilities of a "Median" Individual in Each Category Total SPQ Scale on Consumer Behavior (Model 3)

Questions	Categories				
	1	2	3	4	5
[1]	0.07	0.46	0.25	0.21	0.01
[2]	0.04	0.64	0.18	0.14	0.00
[3]	0.53	0.39	0.07	0.00	0.01
[4]	0.38	0.55	0.05	0.02	0.00
[5]	0.01	0.28	0.17	0.54	0.00
[6]	0.28	0.56	0.07	0.09	0.00
[7]	0.01	0.24	0.16	0.52	0.07
[8]	0.10	0.46	0.11	0.30	0.03
[9]	0.14	0.51	0.13	0.22	0.00
[10]	0.08	0.19	0.18	0.44	0.11
[11]	0.02	0.10	0.56	0.29	0.03
[12]	0.45	0.29	0.19	0.07	0.00
[13]	0.49	0.35	0.15	0.01	0.00
[14]	0.00	0.01	0.31	0.60	0.08
[15]	0.08	0.23	0.42	0.21	0.06
[16]	0.40	0.34	0.20	0.05	0.01
[17]	0.11	0.28	0.33	0.26	0.02
[18]	0.23	0.26	0.30	0.16	0.05

[19]	0.03	0.21	0.44	0.29	0.03
[20]	0.24	0.26	0.37	0.12	0.01
[21]	0.37	0.16	0.19	0.19	0.09
[22]	0.50	0.20	0.19	0.08	0.03
[23]	0.01	0.13	0.34	0.38	0.14
[24]	0.06	0.15	0.36	0.32	0.11
[25]	0.38	0.18	0.30	0.10	0.04
[26]	0.08	0.16	0.38	0.27	0.11
[27]	0.20	0.16	0.33	0.20	0.11
[28]	0.04	0.06	0.40	0.38	0.12
[29]	0.28	0.23	0.33	0.12	0.04
[30]	0.56	0.23	0.15	0.05	0.01
[31]	0.65	0.19	0.11	0.04	0.01
[32]	0.04	0.14	0.45	0.29	0.08
[33]	0.23	0.21	0.33	0.16	0.07
[34]	0.43	0.23	0.20	0.11	0.03
[35]	0.20	0.21	0.35	0.20	0.04
[36]	0.27	0.32	0.22	0.16	0.03
[37]	0.09	0.27	0.43	0.16	0.05

Categories: 1: definitely agree, 2: agree, 3: neither agree, nor disagree, 4:disagree, 5: definitely disagree (reference category)

Table A.10: Response probabilities of consumer behavior for a typical person using equation (1.19)

Discrimination Parameters of Nine Traits of Schizotypy on Consumer Behavior (Model 6)

Discrimination Parameters	mean	sd	MC_error
a[1,1]	1.499	0.32	0.007161
a[2,1]	2.899	0.588	0.02631
a[2,2]	0.584	0.3495	0.01774
a[3,1]	1.33	0.3407	0.01075
a[3,2]	-0.4678	0.3166	0.01257
a[4,1]	2.538	0.5522	0.02558
a[4,2]	0.97	0.4614	0.02475
a[5,1]	2.716	0.5367	0.02114
a[5,2]	-0.1092	0.4274	0.02297
a[6,1]	1.895	0.4005	0.01417

a[6,2]	-0.8519	0.3696	0.01784
a[7,1]	0.7003	0.2631	0.005711
a[7,2]	0.1887	0.25	0.007102
a[8,1]	1.756	0.3572	0.01141
a[8,2]	-0.269	0.3225	0.01527
a[9,1]	1.725	0.3718	0.01237
a[9,2]	-0.5429	0.3405	0.01612
a[10,1]	0.6727	0.2737	0.009953
a[10,2]	-0.7192	0.2696	0.008542
a[11,1]	0.3037	0.4105	0.02439
a[11,2]	-2.076	0.4177	0.01534
a[12,1]	0.4237	0.2445	0.007381
a[12,2]	0.5351	0.2593	0.00714
a[13,1]	-0.3139	0.2969	0.01248
a[13,2]	-1.069	0.2928	0.007681
a[14,1]	1.093	0.4126	0.02218
a[14,2]	-1.866	0.4302	0.01773
a[15,1]	0.2661	0.2573	0.01056
a[15,2]	0.8789	0.2653	0.006712
a[16,1]	0.6356	0.4092	0.02314
a[16,2]	-2.058	0.4354	0.01597
a[17,1]	0.05077	0.2245	0.003903
a[17,2]	-0.05314	0.2341	0.004511
a[18,1]	0.2373	0.3403	0.01771
a[18,2]	1.54	0.3488	0.013
a[19,1]	0.7916	0.2723	0.008259
a[19,2]	-0.5192	0.2657	0.009226
a[20,1]	-0.0908	0.297	0.01387
a[20,2]	-1.14	0.2947	0.007706
a[21,1]	-0.01307	0.2375	0.005169
a[21,2]	-0.2051	0.2512	0.005335
a[22,1]	-0.08092	0.2824	0.01213
a[22,2]	-0.9978	0.2922	0.007643
a[23,1]	-0.2476	0.2573	0.009952
a[23,2]	-0.7611	0.2512	0.005727
a[24,1]	-0.2468	0.2299	0.00461
a[24,2]	0.2618	0.2408	0.005339
a[25,1]	0.19	0.3893	0.02223

a[25,2]	-1.909	0.4047	0.01306
a[26,1]	-0.2661	0.2372	0.005565
a[26,2]	-0.3308	0.235	0.004654
a[27,1]	0.1869	0.2712	0.009053
a[27,2]	0.7822	0.2755	0.00758
a[28,1]	-0.2526	0.2536	0.008907
a[28,2]	-0.749	0.2692	0.006032
a[29,1]	0.6146	0.3415	0.01749
a[29,2]	-1.5	0.336	0.01058
a[30,1]	0.3216	0.2391	0.005577
a[30,2]	0.2703	0.2519	0.005399
a[31,1]	0.06425	0.2684	0.008309
a[31,2]	-0.5748	0.2659	0.005986
a[32,1]	1.019	0.3359	0.01496
a[32,2]	-1.143	0.3217	0.01195
a[33,1]	0.4222	0.2636	0.009706
a[33,2]	0.8168	0.2767	0.006688
a[34,1]	0.6056	0.3099	0.01346
a[34,2]	-1.148	0.3181	0.009871
a[35,1]	0.09843	0.2286	0.005177
a[35,2]	0.2075	0.2354	0.005214
a[36,1]	0.3103	0.294	0.01229
a[36,2]	1.042	0.2965	0.008953
a[37,1]	0.8825	0.2906	0.009738
a[37,2]	-0.6562	0.277	0.009827

Table A.11: a[,1] , a[,2] represent discrimination parameters for the first and second factor respectively

Standardized Discrimination Parameters of Nine Traits of Schizotypy on Consumer Behavior (Model 6)

Stand. Discrimination Parameters	Mean	Sd	MC_error
st.discr1[1,1]	0.8195	0.06011	0.001318
st.discr1[2,1]	0.9396	0.02316	0.001009
st.discr1[2,2]	0.4603	0.1968	0.009423
st.discr1[3,1]	0.7811	0.08103	0.002482
st.discr1[3,2]	-0.3875	0.2301	0.009349

st.discr1[4,1]	0.9223	0.03171	0.001428
st.discr1[4,2]	0.6423	0.1867	0.009211
st.discr1[5,1]	0.9324	0.02502	9,40E-01
st.discr1[5,2]	-0.09237	0.3512	0.01893
st.discr1[6,1]	0.874	0.04541	0.001572
st.discr1[6,2]	-0.6062	0.183	0.008967
st.discr1[7,1]	0.5484	0.1475	0.003161
st.discr1[7,2]	0.1714	0.2215	0.0062
st.discr1[8,1]	0.8586	0.0487	0.001471
st.discr1[8,2]	-0.2323	0.2656	0.01272
st.discr1[9,1]	0.8536	0.05218	0.001685
st.discr1[9,2]	-0.4349	0.2331	0.0112
st.discr1[10,1]	0.5305	0.1611	0.005694
st.discr1[10,2]	-0.5577	0.1506	0.004697
st.discr1[11,1]	0.2449	0.3168	0.01837
st.discr1[11,2]	-0.8922	0.03875	0.001368
st.discr1[12,1]	0.3676	0.1866	0.005593
st.discr1[12,2]	0.4456	0.1768	0.004749
st.discr1[13,1]	-0.2724	0.2425	0.01029
st.discr1[13,2]	-0.7099	0.1016	0.002465
st.discr1[14,1]	0.6993	0.1449	0.007461
st.discr1[14,2]	-0.8686	0.05342	0.002127
st.discr1[15,1]	0.2388	0.2201	0.009103
st.discr1[15,2]	0.6385	0.1197	0.00298
st.discr1[16,1]	0.4785	0.2529	0.01386
st.discr1[16,2]	-0.8898	0.04135	0.001441
st.discr1[17,1]	0.04757	0.2095	0.003624
st.discr1[17,2]	-0.04928	0.217	0.004191
st.discr1[18,1]	0.2042	0.2834	0.01481
st.discr1[18,2]	0.8252	0.06144	0.002184
st.discr1[19,1]	0.5957	0.1374	0.004032
st.discr1[19,2]	-0.4336	0.1865	0.006526
st.discr1[20,1]	-0.0817	0.2633	0.01233
st.discr1[20,2]	-0.7331	0.09205	0.002344
st.discr1[21,1]	-0.01212	0.2208	0.004778
st.discr1[21,2]	-0.1859	0.2207	0.004637
st.discr1[22,1]	-0.0733	0.2546	0.01098
st.discr1[22,2]	-0.6841	0.1107	0.0028

st.discr1[23,1]	-0.2227	0.2227	0.008674
st.discr1[23,2]	-0.5835	0.1326	0.002896
st.discr1[24,1]	-0.2249	0.2006	0.004035
st.discr1[24,2]	0.2366	0.2072	0.004541
st.discr1[25,1]	0.1573	0.3201	0.01812
st.discr1[25,2]	-0.8755	0.04534	0.001413
st.discr1[26,1]	-0.2409	0.205	0.004753
st.discr1[26,2]	-0.2955	0.1934	0.003795
st.discr1[27,1]	0.1683	0.2376	0.007928
st.discr1[27,2]	0.5905	0.1393	0.003628
st.discr1[28,1]	-0.227	0.2178	0.007686
st.discr1[28,2]	-0.574	0.1449	0.003181
st.discr1[29,1]	0.4812	0.215	0.0107
st.discr1[29,2]	-0.8185	0.0627	0.001949
st.discr1[30,1]	0.287	0.1965	0.004547
st.discr1[30,2]	0.2422	0.2137	0.004533
st.discr1[31,1]	0.05794	0.2441	0.007545
st.discr1[31,2]	-0.4711	0.1726	0.003763
st.discr1[32,1]	0.6849	0.1289	0.005599
st.discr1[32,2]	-0.7303	0.1044	0.00379
st.discr1[33,1]	0.3634	0.202	0.007426
st.discr1[33,2]	0.6073	0.1371	0.003202
st.discr1[34,1]	0.4821	0.1958	0.008298
st.discr1[34,2]	-0.7328	0.09873	0.002966
st.discr1[35,1]	0.09158	0.2115	0.004804
st.discr1[35,2]	0.1897	0.2088	0.004587
st.discr1[36,1]	0.2696	0.2408	0.01007
st.discr1[36,2]	0.7001	0.1056	0.003113
st.discr1[37,1]	0.636	0.1305	0.004283
st.discr1[37,2]	-0.52	0.1668	0.005859

Table A.12: st.discr1[,1] , st.discr1[,2] represent standardized discrimination parameters for the first and second factor respectively

Difficulty Parameters of Nine Traits of Schizotypy on Consumer Behavior (Model 6)

Dif. Parameters	Mean	Sd	MC_error
b[1,1]	-2.919	0.4378	0.009329
b[1,2]	-0.1813	0.2609	0.007789
b[1,3]	0.9673	0.2761	0.008029
b[1,4]	4.507	0.7206	0.01045
b[2,1]	-3.429	0.6276	0.02476
b[2,2]	0.3833	0.392	0.01603
b[2,3]	1.409	0.4204	0.01659
b[2,4]	6.165	1.033	0.03025
b[3,1]	-0.1682	0.2541	0.007804
b[3,2]	2.174	0.3539	0.009073
b[3,3]	4.079	0.6167	0.01419
b[3,4]	4.734	0.7455	0.01659
b[4,1]	-0.8468	0.3916	0.01551
b[4,2]	2.226	0.4704	0.01831
b[4,3]	3.794	0.6385	0.02413
b[4,4]	7.673	1.538	0.03727
b[5,1]	-5.536	0.9483	0.02908
b[5,2]	-1.211	0.3906	0.01574
b[5,3]	-0.4883	0.3703	0.01511
b[5,4]	5.132	0.8729	0.02532
b[6,1]	-1.465	0.3439	0.01157
b[6,2]	1.198	0.3522	0.01273
b[6,3]	1.87	0.3872	0.0133
b[6,4]	4.912	0.8359	0.0177
b[7,1]	-5.114	1.286	0.01222
b[7,2]	-1.027	0.2642	0.005595
b[7,3]	-0.2901	0.2349	0.005406
b[7,4]	2.766	0.3466	0.005784
b[8,1]	-2.89	0.4202	0.01275
b[8,2]	-0.432	0.3035	0.01154
b[8,3]	0.05755	0.3066	0.01158
b[8,4]	2.976	0.5128	0.01268

b[9,1]	-1.617	0.38	0.01099
b[9,2]	0.861	0.2965	0.01078
b[9,3]	1.522	0.3121	0.01106
b[9,4]	6.238	0.9853	0.01669
b[10,1]	-2.781	0.3962	0.007199
b[10,2]	-1.287	0.263	0.006569
b[10,3]	-0.4737	0.2354	0.006131
b[10,4]	1.815	0.2965	0.005941
b[11,1]	-3.814	0.6429	0.01783
b[11,2]	-1.773	0.4275	0.01422
b[11,3]	0.911	0.3439	0.01288
b[11,4]	3.755	0.5181	0.01661
b[12,1]	-0.522	0.2224	0.004326
b[12,2]	0.7375	0.227	0.004544
b[12,3]	2.347	0.3428	0.005456
b[12,4]	5.475	1.269	0.0123
b[13,1]	0.1211	0.2524	0.007178
b[13,2]	1.89	0.2948	0.00853
b[13,3]	4.712	0.6861	0.01128
b[13,4]	29.42	17.86	0.182
b[14,1]	-5.485	1.015	0.02133
b[14,2]	-3.996	0.7176	0.0178
b[14,3]	-0.2866	0.3552	0.01278
b[14,4]	2.867	0.4325	0.01476
b[15,1]	-2.779	0.3935	0.006921
b[15,2]	-1.144	0.2549	0.006036
b[15,3]	0.6549	0.236	0.005582
b[15,4]	2.338	0.3459	0.005804
b[16,1]	0.03867	0.3446	0.01248
b[16,2]	1.48	0.3436	0.01355
b[16,3]	3.1	0.427	0.01565
b[16,4]	5.011	0.63	0.01874
b[17,1]	-2.634	0.3776	0.007034
b[17,2]	-1.0	0.2611	0.006564
b[17,3]	0.4124	0.256	0.006265
b[17,4]	3.598	0.6427	0.007919
b[18,1]	-2.145	0.3554	0.01194
b[18,2]	-0.9373	0.3173	0.01127

b[18,3]	0.5115	0.3328	0.01154
b[18,4]	2.177	0.4483	0.0137
b[19,1]	-3.11	0.5798	0.00827
b[19,2]	-0.7115	0.285	0.006898
b[19,3]	1.18	0.2533	0.006983
b[19,4]	4.052	0.4926	0.0088
b[20,1]	-0.9584	0.2863	0.007303
b[20,2]	0.2044	0.2492	0.007419
b[20,3]	2.126	0.3001	0.008173
b[20,4]	5.22	0.8294	0.01108
b[21,1]	-0.8146	0.2134	0.003079
b[21,2]	-0.1486	0.1975	0.002935
b[21,3]	0.6397	0.2058	0.003016
b[21,4]	1.989	0.2978	0.00351
b[22,1]	-0.234	0.2321	0.006225
b[22,2]	0.5837	0.2392	0.006649
b[22,3]	1.842	0.3001	0.00745
b[22,4]	3.292	0.4694	0.008775
b[23,1]	-3.993	0.8156	0.00936
b[23,2]	-1.496	0.32	0.006761
b[23,3]	0.2545	0.2396	0.006615
b[23,4]	2.206	0.292	0.007415
b[24,1]	-2.48	0.4543	0.006264
b[24,2]	-1.062	0.2739	0.004843
b[24,3]	0.5908	0.2154	0.004692
b[24,4]	2.477	0.2973	0.005482
b[25,1]	-0.1397	0.3245	0.01216
b[25,2]	0.5703	0.3147	0.01224
b[25,3]	2.13	0.357	0.0132
b[25,4]	3.628	0.4488	0.0144
b[26,1]	-2.704	0.3893	0.005162
b[26,2]	-1.437	0.2514	0.003998
b[26,3]	0.2234	0.2059	0.003602
b[26,4]	1.834	0.2861	0.004055
b[27,1]	-1.735	0.2831	0.006146
b[27,2]	-0.885	0.2358	0.005615
b[27,3]	0.5162	0.2269	0.005021
b[27,4]	1.83	0.293	0.005235

b[28,1]	-3.19	0.5246	0.009623
b[28,2]	-2.123	0.3687	0.007786
b[28,3]	0.07452	0.2445	0.006207
b[28,4]	2.139	0.306	0.006801
b[29,1]	-1.159	0.2999	0.009494
b[29,2]	-0.1934	0.276	0.009375
b[29,3]	1.406	0.3084	0.009786
b[29,4]	2.804	0.4126	0.01043
b[30,1]	0.2367	0.2113	0.004114
b[30,2]	1.355	0.2314	0.00456
b[30,3]	2.777	0.3479	0.005497
b[30,4]	4.784	0.814	0.007996
b[31,1]	1.178	0.2472	0.007097
b[31,2]	2.258	0.2873	0.008076
b[31,3]	3.514	0.3987	0.009087
b[31,4]	4.9	0.6648	0.01129
b[32,1]	-2.175	0.4939	0.01287
b[32,2]	-0.5795	0.3607	0.01196
b[32,3]	1.471	0.3228	0.0119
b[32,4]	3.514	0.4085	0.01326
b[33,1]	-1.513	0.2737	0.006114
b[33,2]	-0.564	0.2315	0.005916
b[33,3]	0.9254	0.247	0.006078
b[33,4]	2.255	0.3387	0.00657
b[34,1]	0.1407	0.2811	0.009015
b[34,2]	1.104	0.2799	0.009249
b[34,3]	2.294	0.3179	0.009872
b[34,4]	3.905	0.461	0.01089
b[35,1]	-1.693	0.2722	0.003564
b[35,2]	-0.6402	0.2118	0.003133
b[35,3]	0.8476	0.2164	0.00279
b[35,4]	2.978	0.4443	0.004027
b[36,1]	-1.889	0.3159	0.009737
b[36,2]	-0.4657	0.2819	0.009675
b[36,3]	0.5967	0.3045	0.009895
b[36,4]	2.692	0.4894	0.01102
b[37,1]	-2.55	0.3733	0.00691
b[37,2]	-0.8485	0.2492	0.005955

b[37,3]	1.04	0.2555	0.005993
b[37,4]	2.7	0.3942	0.006686

Table A.13: Categories b[,i] where i: 1: definitely agree, 2: agree, 3: neither agree, nor disagree, 4:disagree, 5: definitely disagree. Questions: b[j,] where j=1,…,37

Cumulative Probabilities of a "Median" Individual in Each Category Nine Traits of Schizotypy on Consumer Behavior (Model 6)

Questions	Categories				
	1	2	3	4	5
[1]	0.05	0.45	0.72	0.99	1
[2]	0.03	0.59	0.80	1.00	1
[3]	0.46	0.90	0.98	0.99	1
[4]	0.30	0.90	0.98	1.00	1
[5]	0.00	0.23	0.38	0.99	1
[6]	0.19	0.77	0.87	0.99	1
[7]	0.01	0.26	0.43	0.94	1
[8]	0.05	0.39	0.51	0.95	1
[9]	0.17	0.70	0.82	1.00	1
[10]	0.06	0.22	0.38	0.86	1
[11]	0.02	0.15	0.71	0.98	1
[12]	0.37	0.68	0.91	1.00	1
[13]	0.53	0.87	0.99	1.00	1
[14]	0.00	0.02	0.43	0.95	1
[15]	0.06	0.24	0.66	0.91	1
[16]	0.51	0.81	0.96	0.99	1
[17]	0.07	0.27	0.60	0.97	1
[18]	0.10	0.28	0.63	0.90	1
[19]	0.04	0.33	0.76	0.98	1
[20]	0.28	0.55	0.89	0.99	1
[21]	0.31	0.46	0.65	0.88	1
[22]	0.44	0.64	0.86	0.96	1
[23]	0.02	0.18	0.56	0.90	1
[24]	0.08	0.26	0.64	0.92	1
[25]	0.47	0.64	0.89	0.97	1
[26]	0.06	0.19	0.56	0.86	1
[27]	0.15	0.29	0.63	0.86	1

[28]	0.04	0.11	0.52	0.89	1
[29]	0.24	0.45	0.80	0.94	1
[30]	0.56	0.79	0.94	0.99	1
[31]	0.76	0.91	0.97	0.99	1
[32]	0.10	0.36	0.81	0.97	1
[33]	0.18	0.36	0.72	0.91	1
[34]	0.54	0.75	0.91	0.98	1
[35]	0.16	0.35	0.70	0.95	1
[36]	0.13	0.39	0.64	0.94	1
[37]	0.07	0.30	0.74	0.94	1

Categories: 1: definitely agree, 2: agree, 3: neither agree, nor disagree, 4:disagree, 5: definitely disagree (reference category)

Table A.14: Cumulative probabilities of consumer behavior for a typical person using equation (1.18)

Questions	Categories				
	1	2	3	4	5
[1]	0.05	0.40	0.27	0.27	0.01
[2]	0.03	0.56	0.21	0.20	0.00
[3]	0.46	0.44	0.08	0.01	0.01
[4]	0.30	0.60	0.08	0.02	0.00
[5]	0.00	0.23	0.15	0.61	0.01
[6]	0.19	0.58	0.10	0.12	0.01
[7]	0.01	0.25	0.17	0.51	0.06
[8]	0.05	0.34	0.12	0.44	0.05
[9]	0.17	0.53	0.12	0.18	0.00
[10]	0.06	0.16	0.16	0.48	0.14
[11]	0.02	0.13	0.56	0.27	0.02
[12]	0.37	0.31	0.23	0.09	0.00
[13]	0.53	0.34	0.12	0.01	0.00
[14]	0.00	0.02	0.41	0.52	0.05
[15]	0.06	0.18	0.42	0.25	0.09
[16]	0.51	0.30	0.15	0.03	0.01
[17]	0.07	0.20	0.33	0.37	0.03
[18]	0.10	0.18	0.35	0.27	0.10
[19]	0.04	0.29	0.43	0.22	0.02
[20]	0.28	0.27	0.34	0.10	0.01
[21]	0.31	0.15	0.19	0.23	0.12
[22]	0.44	0.20	0.22	0.10	0.04

[23]	0.02	0.16	0.38	0.34	0.10
[24]	0.08	0.18	0.38	0.28	0.08
[25]	0.47	0.17	0.25	0.08	0.03
[26]	0.06	0.13	0.37	0.30	0.14
[27]	0.15	0.14	0.34	0.23	0.14
[28]	0.04	0.07	0.41	0.37	0.11
[29]	0.24	0.21	0.35	0.14	0.06
[30]	0.56	0.23	0.15	0.05	0.01
[31]	0.76	0.15	0.06	0.02	0.01
[32]	0.10	0.26	0.45	0.16	0.03
[33]	0.18	0.18	0.36	0.19	0.09
[34]	0.54	0.21	0.16	0.07	0.02
[35]	0.16	0.19	0.35	0.25	0.05
[36]	0.13	0.26	0.25	0.30	0.06
[37]	0.07	0.23	0.44	0.20	0.06

Categories: 1: definitely agree, 2: agree, 3: neither agree, nor disagree, 4:disagree, 5: definitely disagree (reference category)

Table A.15: Response probabilities of consumer behavior for a typical person using equation (1.19)

APPENDIX B

The whole questionnaire is in the end of Appendix B. The under analysis questions for the consuming behavior are:

1. Η έκφραση «Just do it» μπορεί να περιγράψει την αγοραστική συμπεριφορά μου.

Συμφωνώ	Συμφωνώ	Αδιαφορώ	Διαφωνώ	Διαφωνώ
Απόλυτα(1)	(2)	(3)	(4)	απόλυτα(5)
☐	☐	☐	☐	☐

2. Συχνά αγοράζω προϊόντα χωρίς να σκεφτώ.

Συμφωνώ	Συμφωνώ	Αδιαφορώ	Διαφωνώ	Διαφωνώ
Απόλυτα(1)	(2)	(3)	(4)	απόλυτα(5)
☐	☐	☐	☐	☐

3. Νιώθω ανήσυχα τις μέρες που δεν πηγαίνω για ψώνια.

Συμφωνώ	Συμφωνώ	Αδιαφορώ	Διαφωνώ	Διαφωνώ
Απόλυτα(1)	(2)	(3)	(4)	απόλυτα(5)
☐	☐	☐	☐	☐

4.. « Αγόρασέ το τώρα και σκέψου το αργότερα» περιγράφει τον τρόπο με τον οποίο αγοράζω.

Συμφωνώ Απόλυτα(1)	Συμφωνώ (2)	Αδιαφορώ (3)	Διαφωνώ (4)	Διαφωνώ απόλυτα(5)
☐	☐	☐	☐	☐

5. Κάποιες φορές είμαι λίγο απερίσκεπτος όσον αφορά τις αγορές μου.

Συμφωνώ Απόλυτα(1)	Συμφωνώ (2)	Αδιαφορώ (3)	Διαφωνώ (4)	Διαφωνώ απόλυτα(5)
☐	☐	☐	☐	☐

6. Απλά ήθελα να αγοράσω κάτι και δεν με ενδιέφερε τι θα ήταν αυτό

Συμφωνώ Απόλυτα(1)	Συμφωνώ (2)	Αδιαφορώ (3)	Διαφωνώ (4)	Διαφωνώ απόλυτα(5)
☐	☐	☐	☐	☐

7. Όταν δω κάτι που θέλω το αγοράζω

Συμφωνώ Απόλυτα(1)	Συμφωνώ (2)	Αδιαφορώ (3)	Διαφωνώ (4)	Διαφωνώ απόλυτα(5)
☐	☐	☐	☐	☐

8. Αγόρασα πράγματα ακόμα κι όταν ήξερα ότι τα οικονομικά μου δεν επαρκούσαν.

Συμφωνώ Απόλυτα(1)	Συμφωνώ (2)	Αδιαφορώ (3)	Διαφωνώ (4)	Διαφωνώ απόλυτα(5)

9. Αγόρασα κάτι και όταν επέστρεψα σπίτι δεν ήμουν σίγουρος/ η γιατί το αγόρασα

Συμφωνώ Απόλυτα(1)	Συμφωνώ (2)	Αδιαφορώ (3)	Διαφωνώ (4)	Διαφωνώ απόλυτα(5)

10 O Elliott, R υποστηρίζει σε προηγούμενη έρευνά του, ότι όταν η διάθεσή μας δεν είναι καλή αγοράζουμε καταναλωτικά αγαθά για να νιώσουμε καλύτερα. Συμφωνείτε?

Συμφωνώ Απόλυτα(1)	Συμφωνώ (2)	Αδιαφορώ (3)	Διαφωνώ (4)	Διαφωνώ απόλυτα(5)

Ποια/ ες από τις παρακάτω κατηγορίες προϊόντων αγοράζετε και πόσο συστηματικά;

		Καθόλου	Πολύ λίγο	Λίγο	Πολύ	Πάρα πολύ
11	Προϊόντα περιποίησης σώματος (σαμπουάν, αρώματα κλπ)					
12	Αθλητικό εξοπλισμό (ρακέτες, μπάλες, αθ. μπλούζες κλπ)					
13	Είδη νοικοκυριού (μαχαίρια, ποτήρια, κατσαρόλες κλπ)					
14	Είδη ρουχισμού (μπλούζες, παντελόνια, πουκάμισα κλπ)					
15	Μουσική (κασέτες, δίσκοι, Cds κλπ)					

16	Κοσμήματα (δαχτυλίδια, σκουλαρίκια, κλπ)					
17	Βιβλία (περιοδικά, λογοτεχνία, κλπ)					
18	Ηλεκτρονικά είδη ψυχαγωγίας (βιντεοταινίες, DVDs, παιχνίδια Η/ Υ, κλπ)					
19	Παπούτσια (αθλητικά, μπότες, κλπ)					
-	Άλλο: ____________________________________					

Χρονικά, πόσο σκέφτεστε την αγορά για την καθεμία από τις παρακάτω κατηγορίες προϊόντων;

		Καθόλου	Πολύ λίγο	Λίγο	Πολύ	Πάρα πολύ
20	Προϊόντα περιποίησης σώματος (σαμπουάν, αρώματα κλπ)					
21	Αθλητικό εξοπλισμό (ρακέτες, μπάλες, αθ. μπλούζες κλπ)					
21	Είδη νοικοκυριού (μαχαίρια, ποτήρια, κατσαρόλες κλπ)					
23	Είδη ρουχισμού (μπλούζες, παντελόνια, πουκάμισα κλπ)					
24	Μουσική (κασέτες, δίσκοι, Cds κλπ)					
25	Κοσμήματα (δαχτυλίδια, σκουλαρίκια, κλπ)					
26	Βιβλία (περιοδικά, λογοτεχνία, κλπ)					
27	Ηλεκτρονικά είδη ψυχαγωγίας (βιντεοταινίες, DVDs, παιχνίδια Η/ Υ, κλπ)					
28	Παπούτσια (αθλητικά, μπότες, κλπ)					

- Άλλο: ___________________________________					

Ποια / ες από τις παρακάτω κατηγορίες προϊόντων θα αγοράζατε αυθόρμητα και πόσο συχνά; (π.χ. Αν περνούσατε έξω από ένα κατάστημα και βλέπατε μια μπλούζα που σας αρέσει πολύ θα την αγοράζατε χωρίς να το σκεφτείτε ιδιαίτερα;)

	Καθόλου	Πολύ λίγο	Λίγο	Πολύ	Πάρα πολύ
29 Προϊόντα περιποίησης σώματος (σαμπουάν, αρώματα κλπ)					
30 Αθλητικό εξοπλισμό (ρακέτες, μπάλες, αθ. μπλούζες κλπ)					
31 Είδη νοικοκυριού (μαχαίρια, ποτήρια, κατσαρόλες κλπ)					
32 Είδη ρουχισμού (μπλούζες, παντελόνια, πουκάμισα κλπ)					
33 Μουσική (κασέτες, δίσκοι, Cds κλπ)					
34 Κοσμήματα (δαχτυλίδια, σκουλαρίκια, κλπ)					
35 Βιβλία (περιοδικά, λογοτεχνία, κλπ)					
36 Ηλεκτρονικά είδη ψυχαγωγίας (βιντεοταινίες, DVDs, παιχνίδια Η/ Υ, κλπ)					
37 Παπούτσια (αθλητικά, μπότες, κλπ)					
- Άλλο: ___________________________________					

BIBLIOGRAGHY

A. Gelman, Hwand J, Vehtari A. (2013) "*Understanding Predictive Information Ctiteria for Bayesian Models*".

Agresti A. (2002). "*Categorical Data Analysis*", Wiley-Interscience.

Bartholomew D. J. and Tzamourani P. (1999). "*The Goodness of Fit of Latent Trait Models in Attitude Measurement*", *Sociological Methods and Research*, **27**(4), 525-546.

Bartholomew D. J., Knott M., Moustaki I. (2011), "*Latent Variable Models and Factor Analysis*" 3rd Edition, John Wiley & Sons.

Bartholomew D. J., Steele F., Moustaki I., Galbraith J. I. (2002). "*The analysis and interpretation of multivariate data for social scientists*", Chapman & Hall.

Beatty S.E. and Ferrell M.E. (1998) "*Impulsive buying: modeling its precursors*", *Journal of Retailing*, **74**(2), 169-191.

Blackwell R.D., Miniard P.W. and Engel J.F. (2001) "*Consumer Behavior*" 9th Ed, Harcourt College Publishers.

Dellaportas P. and Roberts G. O. (2001). "*Introduction to MCMC*".

Dittmar H., Beatty J. and Friese S., (1996). "*Objects, decisions considerations and self-image in men's and women's impulse purchase*", *Acta Psychologica* , **93**, 187-206

Draper D., Fouskakis D., Ntzoufras I., (2009). "*Bayesian Variable Selection using Cost-adjusted BIC, with Application to Cost-Effective Measurement of Quality of Health Care*", *Annals of Applied Statistics,* **3**,663–690.

Everitt B.S. (1984) "An *Introduction to Latent Variable Models*", Chapman & Hall.

Fox J. and Glas C. A. W. (2001). "*Bayesian Estimation of s Multilevel IRT model using Gibbs sampling*".

Glas C. A. W. and Meijer R.R. (2003). "A *Bayesian Approach to Person Fit Analysis in Item Response Theory models*". *Applied Psychological Measurement*, **27**(3): 217-233.

Good, I.J. in Foundations of Statistical Inference (1973). V.P. Godambe and D.A. Sprott Toronto: Holt, Rinehart & Winston. (Quoted in Berger 1985).

Jeffreys H., (1961). "*Theory of Probability*", Oxford University Press.

Marcoulides G.A., Moustaki I. (2002). "*Latent Variable and Latent Structure Models*", Lawrence Erlbaum Associates Publishers Mahwash.

Mislevy R. J. (1986). "*Bayes Modal Estimation in Item Response Models*", Psychometrica **51**, 177-195.

Moustaki I. (2003). "*A General class of latent variable models for ordinal manifest variables with covariate effects on the manifest and latent variables*", British Journal of Mathematical and Statistical Psychology **56**, 337-357

Ntzoufras I., (2009). "*Bayesian Modeling Using WinBugs* ", Wiley.

Patz R. J. and Junker B. W. (1999b). "*A Straightforward Approach to Markov Chain Monte Carlo Methoda for Item Response Models*", *Journal of Educational and Behavioral Statistics*, **24,** 146-178.

Patz R. J. and Junker B. W. (1999a). "*Applications and Extensions of MCMC in IRT: Multiple Item Types, Missing Data and Rated Responses*", *Journal of Educational and Behavioral Statistics*, **24**, 342-366.

Ritter N. L. (2010). "*Understanding aWidely Misunderstood Statistic Cronbach's a* ", Texas A & M University.

Rizopoulos D. (2006). "*An R Package for Latent Variable Modeling and Item Response Theory Analyses*", *Journal of Statistical Software*, **17**, Issue 5.

Siomkos G., (1994). "*Consumer Behavior and Marketing Strategy*", *Stamoulis*.

Janssen R., Tuerlinckx F., Meulders M., De Boeck P. (2000). "*A hierarchical IRT Model for Criterion-Referenced Measurement*", *Journal of Educationa and Behavioral Statistics*, **25**, 285-306.

Kass R. and Raftery A., (1995). “*Bayes Factors*”, *Journal of American Statistician Association*, **90**, 773-795.

Sclosser S., Black D.W., Repertinger S. and Freet D., (1994). “*Compulsive buying: demography, phenomenology and comorbidity in 46 subjects*”, *General Hospital Psychiatry*, **16**, 205-212.

Shoham A. and Brenic M., (2003). “*Compulsive buying behavior*”, *Journal Consumer Marketing*, 20, 127-138.

Cagnone S., Mignani S., Moustaki I. (2009). “*Latent Variable Models for Ordinal Data*”, *Statistical Methods of Educational Services and Quality of Products*, Contribution to Statistics, Springer-Verlag Berlin Heidelberg.

Sinharay S., “Assessing Fit of Unidimensional Item Response Theory using a Bayesian Approach ”, Journal of Educational Measurement, **42**(4):375-394, 2005

Sinharay S., “Bayesian *Item Fit Analysis for Unidimensional Item Response Theory Models*”, British *Journal of Mathematical and Statistical Psychology*, **59**(2), 429-449.

Stefanis N.C., Smyrnis N., Avramopoulos D., Endokimidis I., Ntzoufras I., Stefanis C.N. (2002), “*Factorial Composition of Self-Rated Schizotypal Traits Among Young Males Undergoing Military Training*” Schizophrenia Bulletin.

Wilkie, W.L. (1994). “*Consumer Behavior*”,:*John Wiley & Sons*, New York.