Università degli Studi di Padova
Dipartimento di Scienze Statistiche
Corso di Laurea Magistrale in
Scienze Statistiche

Investigating uncertainty through CUB models

Relatore: Prof. Omar Paccagnella
Dipartimento di Scienze Statistiche

Correlatore: Prof. Maria Iannario
Dipartimento di Scienze Politiche
Università Federico II, Napoli

Laureanda: Serena Pavan
Matricola N. 1082512

Anno Accademico 2015/2016

Contents

List of Tables

List of Figures

Introduction

The main aim of this thesis is to introduce a new extension of CUB models, called TCUB (**CUB** models with a **T**riangular uncertainty component). CUB models are a class of mixture models for the analysis of ordinal data, collected on surveys concerning the expression of opinions on specific items or evaluation or the ranking of them. The fundamental feature of this class of models is the direct focus on the decision making process that leads respondents to choose a particular response category among the ones available. It is assumed the presence of two latent variables underling the psychological choice's mechanism: feeling and uncertainty. The first is the feeling, the liking, the attractiveness toward the item. The second dimension is defined as the natural fuzziness characterising each human choice process and therefore it is the result of several related factors. As a consequence, such uncertainty could be due, for instance, to partial understanding of the item, willingness to joke or fake, tendency to choose some response categories with respect to others (for instance, only central ones or only extreme ones), time pressure, nature of the chosen scale, question's wording and so on. The crucial idea of the solution proposed in this thesis arises from the study of the uncertainty component in these models. Looking at these examples, of uncertainty, we can detect two groups of them: one related to individual characteristics or background of the respondent (say, subjective uncertainty) and the other due to more contextual factors (say, contextual uncertainty). Therefore, it seems worthy trying to extend CUB models in order to disentangle these two kinds of uncertainty. For a researcher, in fact, it would be very useful understanding what causes uncertainty in response process of respondents. On the one hand, if it is due to subjective characteristics he/she can take them into account in any data analysis by controlling for these specific features. On the other hand, if it is due to the context and in particular to the questionnaire administration, the researcher may work on the questionnaire, (for instance in a pilot survey), or find out that the fuzziness in the choice process was not due to peculiarities of the respondents.

The thesis is organised as follows.

Chapter 1 is devoted to the introduction of CUB models, beginning with a general review of ordinal data modelling (Section 1.1), with the specification of standard CUB models, its extensions, interpretation and logical and psychological justification (Section 1.2).

Chapter 2 deals with the new approach, called TCUB, explicating its origins and characterising its formulation, inferential issues, algorithm and interpretation.

Chapter 3 is dedicated to the application of the novel TCUB to a real case study on self-evaluated work disability level, in order to show its potential in enhancing both estimate and interpretation of the uncertainty component. In particular, in Section 3.1 we describe SHARE project, that collected data used in the analysis. In Section 3.2 we provide some descriptive analysis of the sample data. Section 3.3 contains the estimated TCUB models, compared to other models of the class, on the whole dataset available (Section 3.3.1), per country (Section 3.3.2) and comparing each country with respect to the others (Section 3.3.3).

The last section reports the conclusions of this work, summarizing the main issues highlighted by the implementation of this new extended CUB model.

Chapter 1

Modelling ordinal data: CUB models

1.1 Introduction to ordinal data modelling

Qualitative measurements are very common in surveys concerning such diverse fields as sociology, public health, ecology, marketing and so on. Happiness, job or customer satisfaction, quality of life are often considered as the main responses in many survey and are characterised by phenomena where several factors affect human behaviour, in connection or apart from the usual economic variables. In these contexts, people are asked to evaluate different items (objects, services and so on), express their thoughts on a specific topic or choose between a list of them. Therefore, two schemes providing ordered responses may be identified: ranking or rating approach. In the first one, each person assigns a well defined position to items in a list, giving an indirect and compared evaluation of them. Thus, the position allocated to one of them is strongly conditioned by the characteristics of all the others. It suggests that the joint distribution of the responses should be taken into account to analyse them correctly. Then, the marginal distribution of the single item can provide a lot of information, for instance if it is a ranking of brand products for a research on effects of an advertising campaign. Classical statistical analysis contemplates the use of appropriate models of permutations or latent variables that motivate the expressed ordering (Fligner and Verducci, 1993; Marden, 1996; Jöreskog and Moustaki, 2001; Moustaki, 2003). In the rating, respondents are usually asked to select the option among a limited set of categories, which best characterise their thoughts, preferences, perceived values and so on. This set of categories may be on an ordinal or a purely nominal scale. Intermediate types of scales are also possible (Stevens, 1951, 1958, 1968), but the subject of our work concerns statistical models for categorical variables on an ordinal scale, so thereafter this one will always be considered as the reference scale. In particular, when questions are on the agreement to a statement, the usual reference is the Likert scale (Likert, 1932) that provides a verbal description of the ordered response levels. For instance, an example of a Likert scale with five response alternatives is: “strongly approve”, “approve”, “undecided”, “disapprove” and “strongly disapprove”. The number of the categories is open to manipulation like the descriptors, which do not need necessary to have negative and positive responses. To be specific, the answer that the subject assigns to an item is an integer in the support $\{1, ..., m\}$ for a given m number of categories, which are in one-to-one correspondence with the appropriate descriptor. The techniques frequently used to study this kind of data deal with Generalized Linear Models (GLM) (McKelvey and Zavoina, 1975; McCullagh, 1980; McCullagh and Nelder, 1989; Agresti, 2010). According to this line of reasoning, the most natural way to view the measurement process is to assume the existence of an underlying unobservable latent variable associated with the observed responses. This latent trait is generally considered to be drawn from a continuous distribution that changes from individual to individual and the probability of a response not superior to a given category is usually modelled as a (linear) function of selected subject's covariates. Then, these models assume that, by means of cutpoints, the latent variable can be divided in classes of values in order to obtain the discrete response.

Another approach, recently introduced in literature, tries to analyse directly the psychological process that leads the individual to the choice of a certain alternative from the m available. It is still supposed the existence of a latent variable, but it is no more requested the computation of cutpoints and this benefits model parsimony and estimation process. Moreover, individual's covariates are included in the model by means of a direct link with the parameters (generally through a *logit* transformation), simplifying the interpretation. This approach is therefore logically related to GLM, but the link function is introduced directly among parameters and covariates, instead of the expected value and the covariates (as for GLM). Furthermore, the probability of such recent class of models does not belong to the exponential family. This framework is denoted as CUB models and has been introduced by Piccolo (2003) and D'Elia and Piccolo (2005), further discussed by Iannario (2008), Piccolo and D'Elia (2008), Iannario and Piccolo (2010, 2011) among others. In the next Sections we will deepen the basic idea behind this class of models (1.2.1), explain their features (1.2.2) and main inferential issues (1.2.3), ending with an overview of the main extensions already implemented.

1.2 CUB models

1.2.1 Logical and psychological considerations

The mental process that leads the respondent of a survey to the discrete choice of an ordered modality among the m ones available is very complex. It is, indeed, the result of the interaction between several factors, each of which affects the final choice. Subject's perception of the topic is itself influenced by different aspects concerning: individual knowledge and comprehension of the argument, background, family, environment, cognitive and sensory domains in general, and so on. In order to choose a single option in a quantitative scale, like the Likert scale, respondents have to convert their perception, intrinsically continuous, into an integer in a discrete space. This whole process unavoidably creates a decisional uncertainty, that characterises each human decision. The class of CUB models was introduced exactly to explain survey responses in such contexts and the logical and psychological motivations are synthesized below, as in Iannario and Piccolo (2012a):

- Statistical models should consider that respondents, in the aforementioned selection process, are used to take the choice by pairwise comparison of the items or by sequential removals.

- The uncertainty that arises during this procedure is conceptually quite different from the randomness of the experiment and therefore they should be investigated separately and not together in a unique term that causes indecision.

- The mean value, but also higher-order moments, of a distribution do not uniquely identify its shape. Then, it is more appropriate to consider models that take into account the whole distribution of values and not average, median, mode or other similar indexes.

- In some occasions, *shelter* effects should be taken into account for effective models, since empirical evidence points out the presence of preference's distributions that vary their shapes from symmetric to highly skewed ones, with modes ranging everywhere in the support (Iannario, 2012b).

- Since the individual choice depends on several aspects, from the more subjective (like knowledge and background) to the more objective (regarding the context for example) ones, it should be considered a joint statistical model that may look at all of these factors.

- Once settled the variety of human choices, it is necessary to realise that models with the aim at measuring their latent trait have a limited predictability. Although, in order to investigate the homogeneity in people's behaviour, profile them and predict their preferences we can look at clusters, subgroups and selected categories of respondents (Corduas et al., 2009).

Such list inspired and legitimated the implementation of this kind of models as follows.

1.2.2 Specifications of CUB models

CUB models aim at analysing ordinal data using a method that refers directly to the individual's mental procedure during the choice of a single finite item between m alternatives, when a *ranking* is requested, or during the assignment of a value from 1 to m in the case of a *rating approach*. Hence, referring to the description of this psychological procedure (1.2.1), the discrete response results from a mixture of two continuous latent variables that should be modelled by discrete random variables. One variable is related to the strictly personal opinion that the respondent has to the item, the awareness and the full understanding of the problem, the liking, the attractiveness towards it, and so on. The other variable derives from the indecision, also connected with external factors, that characterises the final choice. Thus, CUB models are built on the basis of these two fundamental components, called respectively *feeling* and *uncertainty*, both modelled with discrete variables.

The *feeling*, is the sum of many unobservable subjective variables, intrinsically continuous, that become discrete at the moment of decision making between m prefixed bins. For this reason it could be thought as following a Normal distribution. According to a *latent variable approach* this is the latent trait, generally discretised by means of ordered threshold parameters (to be estimated). In this respect, CUB models do not need the computation of these cutpoints. Indeed, D'Elia (2000) proved that, selecting proper thresholds, a Shifted Binomial random variable can help to take into account the different possibilities arising by transforming a unimodal continuous random variable (i.e. a Gaussian distribution variable) into a discrete one whose support is a set of m integers. Moreover, when the respondent selects one of the m alternatives, we can think that he/she took this choice after a pairwise comparison of all the items. Hence, a Shifted Binomial distribution may be the most appropriate one, as formally justified in Iannario and Piccolo (2015). In fact, let R be the random variable generated by the selection of an ordinal category r that belongs to the interval $\{1, ..., m\}$, such that r increases with the feeling towards the item: if a subject chooses the category r, it means that he/she considers the other $r-1$ previous categories too weak, while the other $m-r$ evaluations too strong to evaluate his/her degree of feeling with respect to the object. Then, let us assume that $\xi \in [0,1]$ is a real number such that $(1-\xi)$ is the probability that an evaluation may be considered inferior to the one chosen by the respondent ("success") and vice versa for ξ. So, considering the $m-1$ possible comparisons, we can deduce that the probability to select a given category r, between the m available, has the following Shifted Binomial distribution:

$$b_r(\xi) = \binom{m-1}{r-1} \xi^{m-r} (1-\xi)^{r-1}, \quad r = 1, \ldots, m. \tag{1.1}$$

Between the two components, the *uncertainty* is the fuzzier one. It is not meant as randomness, dealing with the collection of the data, related to sampling selection, measurement errors and limited knowledge. As uncertainty we intend the indecision of the respondent that exists in any human choice and derives from different factors: degree of knowledge of the problem, partial understanding of the item, time spent for giving the response, nature of the available scale, tiredness, apathy, laziness, willingness to joke and fake and more else (Iannario and Piccolo, 2012a). In order to model this component it is considered the most random choice between the m chances, i.e. the one that assigns the same probability to each category, as it appears by a discrete Uniform random variable, whose probability mass function is:

$$U_r(m) = \frac{1}{m}, \quad r = 1, \ldots, m. \tag{1.2}$$

From a probabilistic point of view, this distribution is the one that allows to maximize entropy between all discrete distributions with a finite range $\{1, ..., m\}$ for a fixed m. Referring to the behaviour of the individual, this is the extreme solution for a totally indifferent choice, where no category is preferred to another.

Identifying these two components (feeling and uncertainty) is not intended to assume there are two kinds of respondents: a more reasonable, sensible one and a more fuzzy, lazy, distracted one. On the other hand, both aspects characterise the moment of the human choice but, obviously not necessarily with

the same weight (50 % feeling and 50 % uncertainty). As a consequence of this assumption, it was introduced a mixture random variable, made of the two properly weighted components.
Continuing with the notation above, let with r_i be the observed response of the $i-th$ subject, chosen between a given number of alternatives m; in a sample of size n, r_i is the realization of a random variable R distributed as the aforementioned mixture, called ***C**ombination of a discrete **U**niform and a shifted **B**inomial distribution* (**CUB**).[1] Therefore, sample data consist of a collection of $(r_1, r_2, \cdots, r_n)$ ordered scores where each answer category $r_i \in \{1, ..., m\}$, hence the **CUB model** is defined as:

$$P_r(R_i = r_i) = \pi \binom{m-1}{r_i-1} \xi^{m-r_i} (1-\xi)^{r_i-1} + (1-\pi)\frac{1}{m}, \quad r_i = 1, \ldots, m \tag{1.3}$$

This CUB model is fully identifiable for any $m > 3$ [2] as proved by Iannario (2010) and is well defined for parameters $\theta = (\pi, \xi)'$ belonging to the parameter space $\Omega(\theta) = \{(\pi, \xi) : 0 < \pi \leq 1, 0 \leq \xi \leq 1\}$.
Expectation and variance of R are given by:

$$E(R) = \pi(m-1)\left(\frac{1}{2} - \xi\right) + \frac{(m+1)}{2}; \tag{1.4}$$

$$Var(R) = (m-1)\left\{\pi\xi(1-\xi) + (1-\pi)\left[\frac{m+1}{12} + \pi(m-1)\left(\frac{1}{2} - \xi\right)^2\right]\right\}.$$

The expected value in (1.4) moves towards the central value of the support depending on the sign of $\left(\frac{1}{2} - \xi\right)$. It means there are lower expected mean values when $\xi \to 1$ as confirmed with the skewness of the distribution which is regulated by $\left(\frac{1}{2} - \xi\right)$: the variable is symmetric if and only if $\xi = \frac{1}{2}$.
For interpreting the parameters involved in a CUB model, as already mentioned before, it is sufficient to think that each respondent decides with a *propensity* π to adhere to a reasonable alternative and $1 - \pi$ to a totally indifferent fuzzy one. When $\pi \to 0$ the inclination to a completely random choice

[1] Originally this variable was called MUB (Mixture Uniform Binomial) as in Piccolo (2003), D'Elia and Piccolo (2005) and then as CUB in further works.
[2] This constraint avoids the case of a degenerate random variable, if $m = 1$, of an indeterminate, if $m = 2$, or of a saturated model if $m = 3$.

increases; in fact it measures the uncertainty through the quantity $1-\pi$ and who it is distributed to overall the support trough:$(1-\pi)/m$. The propensity to an indifferent choice increases with $1-\pi$: uncertainty adds dispersion to the $b_r(\xi)$ and thus should be related to entropy concepts. In particular, the frequency in each class increases with $1-\pi$ modifying the heterogeneity of the distribution. Using the *normalized Gini heterogeneity index*:

$$G = \frac{m}{m-1}\left(1-\sum_{i=1}^{m} p_i^2\right)$$

where p_i is a discrete probability distribution, and defining as G_{CUB} and G_{SB} the two Gini indices respectively for a CUB and a shifted binomial, it has been proved (Iannario and Piccolo, 2010) the relation:

$$G_{CUB} = 1-\pi^2\left(1-G_{SB}\right).$$

This results points out how heterogeneity is inversely related to π and increases with the uncertainty component $1-\pi$.
Concerning the other parameter, it is worth saying that if $\xi \to 0$ the individual has a very much positive opinion on the topic because the distribution mass moves towards high ratings. The vice versa happens when $\xi \to 1$. In order to correctly interpret this parameter, it is indeed necessary to consider the direction of the rating scale since it depends on how responses were coded. If we are working on *ranking data* (where $r=1$ denotes the maximum of liking toward the items, while $r=m$ denotes the minimum), then the parameter ξ is a direct measure of feeling. On the contrary, if we are dealing with *rating data* (where $r=1$ indicates the minimum and $r=m$ the maximum satisfaction), them the parameter measuring the feeling is $(1-\xi)$. Moreover, thanks to the one-to-one correspondence among CUB models and points in the parametric space, $\Omega(\theta)$, a large dataset of observed evaluations can be synthesized by a set of points arranged in the parametric space where coordinates are expressed by (π,ξ). This chance to visualize helps the interpretation of respondent's choices in terms of variability and closeness over time, space and circumstances (Iannario, 2008; Iannario and Piccolo, 2009; Corduas et al., 2009).

Since both parameters are involved in the computation of (1.4), it is worth noting that several parameters generate the same expectation, which is not, therefore, a useful measure of feeling and uncertainty because models with the same expectation may have different probability structure with respect to selection process. Furthermore, a classical link function between covariates and expected value, as in GLM, cannot be implemented. As mentioned in Section 1.1, in CUB models with covariates there is a direct relation among feeling and/or uncertainty parameters and the feature's of the respondents. This direct relation is obtained by means of the logistic function which ensures that the real line is mapped into the unit interval[3]. Covariates may be quantitative, as age or family income, or qualitative, coded as dichotomous or polytomous variables such as gender or marital status for example. We can also consider objective covariates in order to capture potential different reactions of people in decision making process depending on the characteristics of the item that they are evaluating. Regardless of their nature, we will denote covariates by $\mathbf{y}$ if regarding uncertainty and $\mathbf{w}$ for feeling. They may partially or completely overlap. In matrix terms, they can be written as $\mathbf{Y} = ||1, y_{i1}, y_{i2}, \cdots, y_{ip}||$ and $\mathbf{W} = ||1, w_{i1}, w_{i2}, \cdots, w_{iq}||$, given a sample of n ordinal data such as $r = (r_1, r_2, \cdots, r_n)'$. Hence, for a given $m > 3$, the general formulation of a **CUB standard with $\boldsymbol{p}$ and $\boldsymbol{q}$ covariates** to explain uncertainty and feeling respectively is expressed by (Iannario, 2008):

- *a stochastic component*:

$$P_r\left(R_i = r_i|\mathbf{y}_i;\mathbf{w}_i\right) = \pi_i \binom{m-1}{r-1} \xi_i^{m-r}\left(1-\xi_i\right)^{r-1} + \left(1-\pi_i\right)\frac{1}{m},$$

- *two systematic components*:

$$\pi_i = \pi_i(\boldsymbol{\beta}) = \frac{1}{1+e^{-\boldsymbol{y}_i\boldsymbol{\beta}}}, \quad \xi_i = \xi_i(\boldsymbol{\gamma}) = \frac{1}{1+e^{-\boldsymbol{w}_i\boldsymbol{\gamma}}}, \tag{1.5}$$

where $\boldsymbol{\beta} = (\beta_0, \beta_1, \cdots, \beta_p)'$ and $\boldsymbol{\gamma} = (\gamma_0, \gamma_1, \cdots, \gamma_q)'$ are the parameters vectors, $\boldsymbol{y}_i$ and $\boldsymbol{w}_i$ are the $i-th$ rows of the matrices $\mathbf{Y}$ and $\mathbf{W}$. For instance, for

[3] It is known from extensive experimentation that the logit link is adequate, any other one-to-one function that ensure the mapping of real numbers into the unit range may be also used.

$i = 1, \cdots, n$ $\boldsymbol{y}_i = (y_{i0}, y_{i1}, \cdots, y_{ip})$ and $\boldsymbol{w}_i = (w_{i0}, w_{i1}, \cdots, w_{iq})$ with $y_{i0} = w_{i0} = 1$. According to the logistic function:

$$logit(p) = log\left(\frac{p}{1-p}\right)$$

the relationships in (1.5) are equal to:

$$logit(\pi_i) = \boldsymbol{y_i \beta}; \quad logit(\xi_i) = \boldsymbol{w_i \gamma}; \quad i = 1, \cdots, n.$$

Consequently, the probability distribution of a CUB model with covariates may be written as:

$$P_r\left(R_i = r_i | \mathbf{y}_i; \mathbf{w}_i\right) = \frac{1}{1+e^{-\boldsymbol{y}_i\boldsymbol{\beta}}}\left[\binom{m-1}{r_i-1}\frac{\left(e^{-\boldsymbol{w}_i\boldsymbol{\gamma}}\right)^{r_i-1}}{\left(1+e^{-\boldsymbol{w}_i\boldsymbol{\gamma}}\right)^{m-1}} - \frac{1}{m}\right] + \frac{1}{m} \quad (1.6)$$

where in addiction to the notation defined above $r_i \in \{1, 2, \cdots, m\}$. Generally, covariates are selected trough a stepwise procedure: they are introduced in the model one by one, for feeling and/or uncertainty component and then are chosen according to their significance levels or by means of back-forward approaches and penalized likelihood methods. Another procedure has been proposed by Iannario (2009b) for the selection of feeling covariates by means of ordinary least squares estimators of ξ, but it is a relevant issue which is currently under investigation. With respect to these variables' effects, they can be interpreted referring to the unobserved components (i.e. feeling and uncertainty), that now are individual specific, or investigating the probabilities of ordered choices. Starting from this premise, we may use for instance (but more else can be done as we will see in 3.3) plots of $1-\xi$ $(1-\pi)$ against the range of the respectively covariate w_k (y_k), (where k varies from 1 to the number of the categories) in order to visualize its effects on feeling or uncertainty. In particular, as we can notice in (1.5), if the w_k (y_k) increases positively, there is an increase in feeling (uncertainty) if $\gamma_k < 0$ $(\beta_k < 0)$ and vice versa a decrease if $\gamma_k > 0$ $(\beta_k > 0)$ (Iannario and Piccolo, 2012a).

The standard notation for CUB models, that will be followed in the next pages, is described in as in Table 1.1 of Piccolo (2006):

Table 1.1: Standard notation of CUB models, Piccolo (2006).

Models	Covariates	Parameters vectors	Parameter spaces	Number of parameters
$CUB(0,0)$	no covariates	$\boldsymbol{\theta} = (\pi, \xi)'$	$(0,1] \times [0,1]$	2
$CUB(p,0)$	covariate for π	$\boldsymbol{\theta} = (\boldsymbol{\beta}', \xi)'$	$\mathbb{R}^{p+1} \times [0,1]$	$p+2$
$CUB(0,q)$	covariates for ξ	$\boldsymbol{\theta} = (\pi, \boldsymbol{\gamma}')'$	$(0,1] \times \mathbb{R}^{q+1}$	$q+2$
$CUB(p,q)$	covariates for π and ξ	$\boldsymbol{\theta} = (\boldsymbol{\beta}', \boldsymbol{\gamma}')'$	$\mathbb{R}^{p+q+2}$	$p+q+2$

1.2.3 Main inferential issues

Considering the case of a $CUB(p,q)$ model, given a sample of size n of observed ordinal data and covariates $(\boldsymbol{r}, \boldsymbol{y}, \boldsymbol{w})'$, the log-likelihood function for the parameter vector $\theta = (\beta', \gamma')'$ derives from (1.6) and is specified by:

$$\ell(\boldsymbol{\theta}) = \sum_{i=1}^{n} log \left[\frac{1}{1+e^{-\boldsymbol{y}_i\boldsymbol{\beta}}} \left[\binom{m-1}{r_i-1} \frac{(e^{-\boldsymbol{w}_i\boldsymbol{\gamma}})^{r_i-1}}{(1+e^{-\boldsymbol{w}_i\boldsymbol{\gamma}})^{m-1}} - \frac{1}{m} \right] + \frac{1}{m} \right].$$

The estimates of parameters involved in CUB models are obtained through the *maximum likelihood estimation* (ML), pursed via E-M algorithm. Approximate variance and covariance matrices of the ML estimators are derived from asymptotic inference. In particular, in order to obtain more accurate estimates of the information matrix and therefore of the variance and covariance matrix, it has been computed the second order derivatives of $\ell(\boldsymbol{\theta})$ by analytic methods. Inferential issues are fully specified in Piccolo (2006) and implemented in R code in a package for CUB models and their extensions available in R (Iannario et al., 2016).

The validation of the estimated models relies on several point (Iannario and Piccolo, 2015):

- ***significance of parameters***: by means of Wald tests, estimates are tested and therefore significant parameters can be detected.
- ***log-likelihood comparisons***: when models are nested (as the models described in previous sections), it is possible to test if the increase in

log-likelihood is significant through deviance differences with respect to the standard χ^2_g percentiles (g degrees of freedom which derive from the difference among the number of parameters of the two considered models) and verifies if the more complex model is a valuable choice or not, as in the current literature (Agresti 2010, pp-67-75).

- ***global indices***: One fitting index used by this family of models is the dissimilarity index (***Diss***), which compares the f_r observed frequencies and the expected probabilities $\hat{p}_r = p_r(\hat{\theta})$. Its direct fitting measure normalized in $[0, 1]$, is:

$$\boldsymbol{\mathcal{F}^2} = 1 - \frac{1}{2}\sum_{r=1}^{m} |f_r - \hat{p}_r|.$$

It indicates the proportion of respondents to move in order to achieve a perfect fitting and it is often computed as a benchmark for judging the adequacy of the model: values of $\mathcal{F}^2 \geqslant 0.90$ are considered as compatible with an acceptable fitting (Iannario, 2009a). For judging goodness-of-fit and compare non-nested estimated models model selection criteria as ***BIC*** (Schwarz et al., 1978) and ***AIC*** (Akaike, 1974) are also applied. Indeed, they permit to consider both improving on likelihood and penalty given by the number of parameters involved in the model. The formulas of these indices are:

$$BIC = -2\ell(\hat{\boldsymbol{\theta}}) + (npar)log(n), \quad AIC = -2\ell(\hat{\boldsymbol{\theta}}) + 2(npar).$$

Another index, with the same purpose, is a $pseudo - R^2$ and is called ***ICON*** (=Information CONtent) (Bozdogan, 1990). It compares the log-likelihood of the estimated model with a completely uninformative distribution, the discrete Uniform:

$$ICON = 1 + \frac{\ell(\hat{\theta})/n}{log(m)}.$$

For all these indices, except of ICON, comparing several model the rule is "the smaller the better". In this context, BIC index is usually preferred with the aim of comparing different models (Iannario and Piccolo, 2009).

- ***residuals diagnostic***: a classical analysis based on the residuals of the model may be computed considering for example Pearson and relative residuals or further analyses based on generalized residuals as in Di Iorio and Iannario (2012).

1.2.4 Extended CUB models

Moving from standard CUB models, some extensions and variations concerning the kind of the involved distributions were implemented.
A first relevant extension is the **CUB model with a shelter effect**, that allows to take into account the presence of a sort of "refuge" category (Corduas et al., 2009; Iannario, 2012b). When people are asked to choose a category of response between m alternatives, they may select frequently a category, called c, in order to avoid more elaborate decisions. It could happen, for instance, because of time pressure, privacy issues or laziness, that induct respondents to choose a simplified option or a central category of "no-choice". Yet, there are also psychological mechanisms that lead a subject to prefer for example rounded number in a set of real ones or be attracted by a specific word. A category c is therefore defined as a *shelter choice* if it receives an observed preference higher than the one expected by the standard model. This *shelter effect* is added into the class of CUB models by means of a dummy variable $D_r^{(c)}$ that for $c \in [1, m]$ is:

$$D_r^{(c)} = \begin{cases} 1, & \text{if } r = c; \\ 0, & \text{otherwise.} \end{cases}$$

Thus, this extended CUB model is defined by:

$$P_r\left(R = r\right) = \pi_1 b_r(\xi) + \pi_2 U_r(m) + \left(1 - \pi_1 - \pi_2\right) D_r^{(c)},$$

where $b_r(\xi)$ and $U_r(m)$ are the classical constituents of this family of models. For a given order of components and a number of categories $m > 4$, such model is identifiable and its parameter space is:

$$\Omega(\boldsymbol{\theta}) = \{(\pi_1, \pi_2, \xi) : \pi_1 > 0, \pi_2 \geq 0, \pi_1 + \pi_2 \leq 1, 0 \leq \xi \leq 1\},$$

where $\boldsymbol{\theta}$ indicates the parameter vector $(\pi_1, \pi_2, \xi)'$. The weight of the *shelter choice* at $R = c$ with respect to a CUB(0,0) is provided by $\delta = (1 - \pi_1 - \pi_2)$. So, when no *shelter effect* is present, δ would be equal to 0 $(\pi_1 + \pi_2 = 1)$ and the model will collapse to the standard one. Parameters' estimates are computed through maximum likelihood, via E-M algorithm and the routine is implemented in the software R (R Core Team, 2016), as for the standard formulation (Iannario, 2012b; Iannario and Piccolo, 2014). Moreover, covariates were added into a *CUB model with shelter effect*, developing **GeCUB** models (=**G**eneralized **CUB** model) (Iannario and Piccolo, 2012b, 2015). Similarly to a CUB(p,q), this framework allows to insert covariates for each component (feeling, uncertainty and shelter) by means of the logistic link. The code useful for the estimate has been written in *GAUSS*© language and it is available from the authors upon request. If π_1 is null, a *CUB model with shelter effect* becomes a **CUSH** model, i.e. a **C**ombination of only a discrete **U**niform random variable with a **SH**elter effect, as designed by Capecchi and Piccolo (2015). It models a context of extreme heterogeneity, when respondents either choose a shelter category (for the aforementioned reasons) or are totally fuzzy and give the same weight to each alternative.

Another proposed extension are *Hierarchical CUB models*, implemented to take into account of the behaviour of subgroups, including random effects as intercept and/or slope, as it is shown in Iannario (2012a).

Regarding other formulations of CUB models, which are made up of probability distributions which are different from the standard one, it is worth mentioning CUBE (and their specific case IHG) and VCUB models. **CUBE** models (Iannario, 2014) are a **C**ombination of a **U**niform and a **BE**ta-binomial distribution and they allow to capture a possible overdispersion. Their particular case is the model which supposes that the data generating process follows an **I**nverse **H**yper**G**eometric distribution, that is adequate if the mode is a extreme value of the support. It has obviously no more the structure of a standard CUB model because, to ensure identifiability, it is not added the discrete Uniform variable for the uncertainty component (D'Elia, 2003).

VCUB (= **V**arying Uncertainty in **CUB** models) are the most recent generalization of CUB model introduced by Gottard et al. (2016). The aim of this kind of models is to provide the possibility to consider a specification of the uncertainty component different from the discrete Uniform distribution included in the standard formulation. Therefore, the structure of the model remains the same:

$$P(R = r) = \pi b_r(\xi) + (1 - \pi)p_r^V \quad r = 1, \ldots, m,$$

except for the distribution of the second component p_r^V, that may be specified "ad hoc" choosing among different distributions, as we will better explain hereafter. Hence, with a VCUB it is possible to take into account, with greater effectiveness, with respect to standard CUB model the indecision of subjects, but also the *response styles*. With *response styles* we mean the phenomenon whereby respondents may have the tendency, consciously or unconsciously, to answer choosing a specific category irrespective of the content of the question (Baumgartner and Steenkamp, 2001). Each subject may have his/her own response style, depending on his/her culture, country, propensity to adhere or to dissent, but also on situational factors such as questionnaire's wording, response scale, time pressure and similar aspects (Baumgartner and Steenkamp, 2001). Then, this behaviour can lead to misleading interpretations of the topic, because it may hide actual responses. Referring to psychological literature and substantive arguments, Gottard et al. (2016) identified the most common *response styles* in the following:

- ***Resoluteness in the extremes***: it characterises a subject who chooses the extreme values of the scale and usually retains the choice also if intermediate alternatives are modified.

- ***Acquiescence response style***: it occurs when one tends to adhere to the questions asked to him/her (specifically called *yeasaying*) or, on the other hand, to disagree with them (*naysaying*). It is strongly affected by the context, in particular when socially desirable answers are given.

- ***Response contraction bias***: it indicates a behaviour that is the opposite of *resoluteness in the extremes*, because in this case the respondent tends to avoid extreme categories when several modalities are available. People may refrain from selecting only one of the sides of the scale, say *one side contraction*, so that the effective scale is left or right truncated. In this cases dubious people may be probably affected by optimistic or pessimistic moods.

- ***Spike responses***: it happens when some categories of response are chosen for some special reasons: the preference for odd numbers, central categories (*midpoint response style*), round numbers and for other similar tendencies, as described above for the model with *shelter effect* (Iannario, 2012b).

Referring to these behaviours, Gottard et al. (2016) proposed four alternative and more selective specifications of the uncertainty component. These are:

1) ***Trimmed uniform distribution***: for a known integer $0 \leq k < m/2$, its probability mass is defined by:

$$p_r^V = \begin{cases} \dfrac{1}{m-2k}, & \text{if } r = k+1, k+2, \ldots, m-k; \\ 0, & \text{if } r = 1, 2, \ldots, k, m-k+1, \ldots, m. \end{cases}$$

   It is useful, for instance, to take into account of *response contraction bias* and *resoluteness in the extremes.*

2) ***Left/right bounded Uniform distribution*** is instead indicated for the treatment of *one side contraction* behaviour and is specified for a given integer $0 \leq k < m$ as follows:

$$p_r^V = \begin{cases} \dfrac{1}{m-k}, & \text{if } r = k+1, k+2, \ldots, m; \\ 0, & \text{if } r = 1, 2, \ldots, k; \end{cases}$$

or

$$p_r^V = \begin{cases} \dfrac{1}{m-k}, & \text{if } r = 1, 2, \ldots, m-k; \\ \\ 0, & \text{if } r = m-k+1, m-k+2, \ldots, m; \end{cases}$$

3) ***Triangular distribution*** according to Kokonendji and Zocchi (2010):

$$p_r^V = \begin{cases} \dfrac{2(r-1)}{(m-1)(k-1)}, & \text{if } r = 1, 2, \ldots, k; \\ \\ \dfrac{2(r-m)}{(m-1)(k-m)}, & \text{if } r = k+1, k+2, \ldots, m; \end{cases}$$

where k is the mode and belongs to the open interval $(1, m)$. This distribution may be chosen to specify both *acquiescence response style* and *response contraction bias.*

4) ***Symmetric parabolic distribution*** can be preferred to the previous one if it may be more convenient to use a smoother distribution such as:

$$p_r^V = \frac{6(r-1)(m-r)}{m(m-1)(m-2)}, \quad r = 1, 2, \ldots, m.$$

It is also possible to assign no zero probability to the extreme modalities by adding a constant $c > 0$ in the following way:

$$p_r^{V*} = \frac{p_r^V + c}{1 + mc}, \quad r = 1, 2, \ldots, m.$$

For the sake of parsimony and simplicity the varying uncertainty component has to be chosen a priori among these distributions, according to information available to the researcher. If in doubt, various models may be computed in order to select the one which seems to better explain the variability of the data. Definition and formulation of VCUB are derived in a totally similar way of CUB models, while estimation is obtained by means of maximum likelihood. The code, useful to compute this kind of models has been written in R language by the authors and it is available upon request.

After this review of CUB models, in the next chapter we will suggest an extended model that allows to further investigate the uncertainty concept. The aim of this proposal is to better understand its nature and its relationships with the characteristics of respondents.

Chapter 2

An extension of CUB models in order to disentangle uncertainty

As aforementioned in section 1.2.2, CUB models are built on a basic formulation that refers directly to the psychological mechanism of a subject answering to a question with m possible categories of response. Indeed, the response is modelled as the mixture of two latent components: one related to individual feeling towards the item and another to the uncertainty in the decision making process. The latter may be defined as the fuzziness surrounding the final choice, result of possible convergent and related factors. To recap, these were identified trough the different generalizations and extensions in: willingness to joke and fake, respondent knowledge/ignorance, partial understanding of the item, response styles, as well as questionnaire administration, number of questions, nature of the response scale, amount of time devoted to the response, tiredness or fatigue, laziness and so on. Looking at this list of factors we could detect two kinds of uncertainty: one related to external factors, another one related to individual characteristics and background. In fact, features such as time pressure, number of questions, questionnaire administration depend on the context in which respondents is required to answer. On the other hand, response styles, willingness to joke and fake, knowledge, ignorance, misunderstanding and so on are subjective sources of uncertainty. With respect to this possible split of uncertainty, we are going to suggest, an extension of CUB models that may allow to disentangle it in these two parts: a more contextual and a more subjective uncertainty.

2.1 Specification of a TCUB model

Let us assume that the mental process, that leads the individual to the response, is now a mixture of three latent variables that can be modelled by discrete random variables. One latent trait is the one related to the perception, the feeling toward the object. The other two may be denoted globally as uncertainty but one is subject-related, say S, the other one is context- and content-related, say U. Feeling is therefore defined as standard CUB models and can be modelled in the same way by a Shifted Binomial distribution. With respect to the *Subjective Uncertainty*, it is related to the characteristics and generally to the background of respondents. In particular, being the result of the aforementioned factors, it may be interpreted as if each subject has his/her own way to use the measurement scale available, depending on his/her characteristics. According to the theory of response styles, explained at the end of the previous chapter, each respondent has "a systematic tendency to respond to a range of questionnaire items on some other basis than the specific item content" (Paulhus, 1991). However, such subjective uncertainty deals also with the concept that there is an individual heterogeneity that leads respondents to interpret, understand or use the response categories for the same questions differently (Holland and Wainer, 2012). For instance, belonging to different socio-economic groups can lead individuals to interpret or understand an identical question in different ways, because people apply different scales to evaluate themselves or, more simply, because they differ in world view, mood, propensity to use extreme categories and more else, even if they are similar regarding to economic and non-economic conditions. This evidence of response scale inter-personal and inter-cultural differences is known as Differential Item Functioning (DIF) (Holland and Wainer, 2012). In order to take into account these different ways to use the response scale, we suggest the use of a discrete random variable that is not Uniform across all the answer categories, but a more flexible one, like a Triangular distribution. To enhance interpretation of this component, since it is strongly related with individual characteristics, it

seems suitable to further extend this model specifying that the parameter of this component is associated with some subjective covariates. To reach this goal we can introduce a logistic link between the parameter in object and covariates (as generally implemented in the class of CUB models).
The remaining part of uncertainty is the one related to external factors or general characteristics of the questionnaire. For this reason, it seems suitable keeping an Uniform random variable so that we can assign the same probability to each response category available. This is an extreme solution for a totally indifferent choice, as it happens for uncertainty component in the standard CUB modelling.
Thus, such extended CUB model is defined for $r = 1, \ldots, m$ by:

$$P_r\,(R_i = r_i) = \pi b_{r_i}(\xi) + \delta_i S_{r_i}(m) + (1 - \pi - \delta_i)\,U_{r_i}(m), \tag{2.1}$$

where $\delta_i S_{r_i}(m)$ is the new introduced third component. $S_{r_i}(m)$ is the aforementioned flexible discrete random variable that depends on r_i, as in the case of a triangular distribution. The parameter δ_i is the i-th component, for $i = 1, \ldots, n$ number of respondents, of the parameter vector $\boldsymbol{\delta} = (\delta_0, \delta_1, \cdots, \delta_n)'$. According to the logistic function ($logit\,(p) = log\,(p/(1-p))$) each parameter is computed as follows:

$$\delta_i = \delta_i(\boldsymbol{\omega}) = \frac{1}{1 + e^{-\boldsymbol{x_i}\boldsymbol{\omega}}}, \text{ so that } logit\,(\delta_i) = \boldsymbol{x_i}\boldsymbol{\omega},$$

where $\boldsymbol{\omega} = (\omega_0, \omega_1, \ldots, \omega_s)'$ and $\boldsymbol{x_i} = (x_{i0}, x_{i1}, \ldots, x_{is})'$ is the i-th rows of the matrix $\boldsymbol{X}$ containing the observed subject's covariates for explaining $\boldsymbol{\delta}$. Then, if we denote with $\boldsymbol{\theta} = (\pi, \boldsymbol{\delta'}, \xi)'$ the parameter vector with length $(s + 3)$ (where s is the number of introduced covariates) characterizing this new mixture random variable, for a given order of components, such models are identifiable for $m > 4$, as it is for CUB models with shelter effect, according to Iannario (2010, 2012b). The parametric space is defined by:

$$\Omega(\boldsymbol{\theta}) = \{(\pi, \boldsymbol{\delta'}, \xi) : \pi > 0, \boldsymbol{\delta'} \in \mathbb{R}^{s+1}, \pi + \delta_i \leq 1, 0 \leq \xi \leq 1\}.$$

With respect to interpretation, it is worth noting that the quantity δ_i is the weight of the different way to use of the response scale, a way specified trough

$S_{r_i}(m)$, for the i-th individual, given his/her characteristics provided by the covariates. It indicates, therefore, the proportion of the so called subjective uncertainty.

Instead, $\nu_i = 1 - \pi - \delta_i$ measures a sort of residual component, the so called contextual uncertainty, indicating the weight of a totally indifferent choice due to the context in which the decision making mental process of the subject has to take place. It is then easy to understand that this model can be very useful to enhance the estimate of the uncertainty components and their nature, but not exclusively. By means of this division we can capture the weight of the subjective and contextual part so that, if the latter predominates, we may suppose that there are problems concerning questionnaire administration. Such model may therefore be estimated also to investigate if there were critical issues in the instruments and the modalities used to collect individual's evaluations. In pilot surveys this feature might allow to be aware of the fact that the questionnaire may need to be adjusted. As a consequence the submitted one would be more robust.

Of course, if $\pi + \delta_i = 1$ the extended CUB models collapses to a VCUB model, but with the addition of covariates to explain the uncertainty component, if the latter is specified in the same way as the $S_{r_i}(m)$ component. Otherwise, if this extension is not significant, i.e. the parameter δ_i is null, the model collapses to the standard CUB specification. With respect to the choice of the distribution of $S_{r_i}(m)$, as we already mentioned, it could be any suitable flexible discrete random distribution, which assigns a different probability to each modality of the response scale. We suggest the use of a Triangular distribution as the one defined by Kokonendji and Zocchi (2010) because it permits to consider no symmetric distributions. Moreover, this is also the same distribution proposed by Gottard et al. (2016) aiming at taking into account for response style such as *acquiescence response style* and *response contraction bias* as defined at the end of the previous chapter. Then specifically, with the term $S_{r_i}(m)$ we will hereafter mean the following distribution:

$$S_{r_i}(m) = \begin{cases} \dfrac{2(r-1)}{(m-1)(k-1)}, & \text{if } r = 1, 2, \cdots, k; \\ \\ \dfrac{2(r-m)}{(m-1)(k-m)}, & \text{if } r = k+1, k+2, \cdots, m; \end{cases}$$

where k is the mode and belongs to the open interval $(1, m)$. Several reasons led to this choice, some for the sake of interpretation, others for the sake of identifiability. Indeed, such Triangular distribution satisfies the request of flexibility and allows to assign a different probability to each category, with the possibility to choose the mode of the distribution. These features are in line with the definition of subjective uncertainty. One could also choose another discrete distribution with similar features, for example in a way closer to other response styles such as *resoluteness in the extremes*. However, every choice has to ensure model identifiability. In fact, other response styles, according to Gottard et al. (2016), may be included by means of modified Uniform distributions, such as the Trimmed Uniform distribution for instance. But, as a consequence, it is easy to understand that such distributions are extremely similar to the Uniform one that is considered for the contextual uncertainty component. Therefore, it would become meaningless to insert one of them and moreover it would be impossible to identify the model.

Thus, since we chose the Triangular distribution for the modelling of $S_{r_i}(m)$, this extended model is called **TCUB** (=**CUB** models with a **T**riangular uncertainty component). If we keep the notation of CUB models, this extension may be denoted more specifically as TCUB(0, 0, s) to express the number of covariates added to explain each component.

2.2 Inferential issues

2.2.1 ML inference and fitting measures

Let $\boldsymbol{r} = (r_1, r_2, \ldots, r_n)'$ be the observed sample of ordinal values, considered as realization of the random sample $\boldsymbol{R} = (R_1, R_2, \ldots, R_n)'$ where each R_i is independently distributed as a discrete random variable over the support

$\{1, 2, \ldots, m\}$. The integers $1 \leqslant r_i \leqslant m$, for $i = 1, 2, \ldots, n$ are the expressed individual's ratings towards the specific item. Given the sample $\boldsymbol{r}$ and the matrix of subject's covariates $\boldsymbol{X}$, where each i-th row $\boldsymbol{x_i} = (x_{i0}, x_{i1}, \cdots, x_{is})'$ contains all other available sample information on the i-th subject, the log-likelihood function for inferring on $\boldsymbol{\theta}$ is defined by:

$$
\begin{aligned}
\ell(\boldsymbol{\theta}) &= \sum_{i=1}^{n} log\left[P_r\left(R_i = r_i | \boldsymbol{x_i}, \boldsymbol{\theta}\right)\right] \\
&= \sum_{i=1}^{n} log\left[\pi b_{r_i}(\xi) + \delta_i S_{r_i}(m) + (1 - \pi - \delta_i)\, U_{r_i}(m)\right] \\
&= \sum_{i=1}^{n} log\left[\pi b_{r_i}(\xi) + \frac{1}{1 + e^{-\boldsymbol{x_i}\boldsymbol{\omega}}} S_{r_i}(m) + \left(1 - \pi - \frac{1}{1 + e^{-\boldsymbol{x_i}\boldsymbol{\omega}}}\right) U_{r_i}(m)\right].
\end{aligned}
$$

As for the whole family of CUB models and for mixture models in general, parameter estimate is achieved by means of maximum likelihood (ML) estimation, pursued by E-M algorithm (McLachlan and Krishnan, 2007; McLachlan and Peel, 2004). Such a procedure will be detailed for TCUB models in the next section. Asymptotic inference requires the knowledge of the observed information matrix, in order to derive approximate variance-covariance matrix of ML estimators, that in this case is obtained by numerical computations.
The validation of the estimated models follows the same guidelines described in 1.2.3. The usefulness of this extended model may be pointed out by means of some graphical tools which may also simplify the interpretation. Indeed, for a standard CUB model the visualization of its marginal distributions in the parametric space is useful for interpreting data in terms of closeness, similarity and peculiarity. Aiming at reproducing this graphic for a TCUB, we may exploit the relation among parameters of the mixture: $\pi + \delta_i + \nu_i = 1$ to drawing a simplex plot in a way similar to the one used to visualise CUB models with shelter choice (Iannario, 2012b). In this plot, for each three-dimensional parameter vector (π, δ_i, ξ), coordinates of each point indicate the weight of the uncertainty components, while the size of the point shows the value of the parameter ξ. As a consequence, it become easier to compare different models or subgroups of subjects in terms of feeling level but also stability, relative

weights or changing in the components given specific covariates' values. Other graphical tools may be plotted as for classical CUB models, among others: observed frequencies against predicted ones or the representation of how the estimate of a parameter varies with the value of a given covariate.

2.2.2 E-M algorithm for a TCUB model

In this section we discuss the computational steps necessary to implement E-M algorithm for the ML estimation of the parameters in the TCUB model, in a way similar to (among others) Piccolo (2006) and Iannario and Piccolo (2015). As already mentioned in the previous section, the collected ratings $\boldsymbol{r} = (r_1, r_2, \ldots, r_n)'$ are the realization of the random sample $\boldsymbol{R} = (R_1, R_2, \ldots, R_n)'$ where each random variable is identically and independently distributed, for a given $m > 4$, as a discrete random variable R on the support $\{1, 2, \ldots, m\}$. For a given i-th subject ($i = 1, 2, \ldots, n$), we define the three distributions included in the mixture as:

$$b_{r_i}(\xi) = \binom{m-1}{r_i - 1} \xi^{m-r_i} (1-\xi)^{r_i - 1}; \qquad U_{r_i}(m) = \frac{1}{m}$$

$$S_{r_i}(m) = \begin{cases} \dfrac{2(r_i - 1)}{(m-1)(k-1)}, & \text{if } r_i = 1, 2, \cdots, k; \\[2ex] \dfrac{2(r_i - m)}{(m-1)(k-m)}, & \text{if } r_i = k+1, k+2, \cdots, m. \end{cases}$$

Then, adopting the logistic function as link between parameters and covariates, a TCUB model is fully specified by:

$$\begin{cases} P_r\left(R_i = r_i | \boldsymbol{x_i}, \boldsymbol{\theta}\right) = \pi b_{r_i}(\xi) + \delta_i S_{r_i}(m) + (1 - \pi - \delta_i)\, U_{r_i}(m); \\[2ex] \delta_i = \dfrac{1}{1 + e^{-\boldsymbol{x_i}\boldsymbol{\omega}}}; \end{cases}$$

where $\boldsymbol{\theta} = (\pi, \boldsymbol{\omega'}, \xi)'$ is the parameter vector, $\boldsymbol{\omega} = (\omega_0, \omega_1, \ldots, \omega_s)'$ and $\boldsymbol{x_i} = (x_{i0}, x_{i1}, \ldots, x_{is})'$ is the i-th row of the matrix:

$$
\boldsymbol{X} = \begin{pmatrix} 1 & x_{11} & \cdots & x_{1s} \\ 1 & x_{21} & \cdots & x_{2s} \\ \cdot & \cdots & \cdots & \cdots \\ 1 & x_{i1} & \cdots & x_{is} \\ \cdot & \cdots & \cdots & \cdots \\ 1 & x_{n1} & \cdots & x_{ns} \end{pmatrix},
$$

for $i = 1, 2, \dots, n$ and $x_{i0} = 1$ for all the sample units, specifying the constant (baseline) of the model.

To simplify notation and to denote the parameters in line with their roles, as in McLachlan and Peel (2004), we indicate by $\boldsymbol{\theta} = (\boldsymbol{\psi}', \boldsymbol{\eta}')$ the full parameter vector of a TCUB model. With $\boldsymbol{\psi}$ we characterise the parameter vector of weights α_g and with $\boldsymbol{\eta}$ the one of the probability distributions $\mathcal{P}_g$. Indeed, hereafter for $g = 1, 2, 3$ number of the three components of the mixture, we will use the notation synthesised in Table 2.1.

Table 2.1: Component's notation of TCUB models.

g	$\alpha_{gi} = \alpha_{gi}(\boldsymbol{\psi}_g)$	$p_{gi} = p_g(r_i; \boldsymbol{\eta}_g)$	$\boldsymbol{\psi}_g$	$\boldsymbol{\eta}_g$
1	π	$b_{r_i}(\xi)$		$\eta_1 = \xi$
2	δ_i	$S_{r_i}(m)$	$\psi_2 = \boldsymbol{\omega}$	
3	$1 - \pi - \delta_i$	$U_r(m)$	$\psi_3 = \boldsymbol{\omega}$	

We introduce then the unobservable vector $\boldsymbol{z} = (z_1, z_2, \dots, z_n)'$ where each $\boldsymbol{z_i} = (z_{1i}, z_{2i}, z_{3i})'$ is a three-dimensional vector such that, for $g = 1, 2, 3$:

$$
z_{gi} = \begin{cases} 1, & \text{if the i-th subject belongs to the g component } \mathcal{P}_g; \\ \\ 0, & \text{otherwise.} \end{cases}
$$

As a consequence, the likelihood function of the complete-data vector $(\boldsymbol{r}', \boldsymbol{z}')'$ is obtained as:

$$
L_c(\boldsymbol{\theta}) = \prod_{g=1}^{3} \prod_{i=1}^{n} [\alpha_{gi}(\boldsymbol{\psi}_g) p_g(r_i; \boldsymbol{\eta_g})]^{z_{gi}},
$$

and the complete-data log-likelihood function is:

$$\ell_c(\boldsymbol{\theta}) = \sum_{g=1}^{3}\sum_{i=1}^{n}\left[z_{gi}\log\left(\alpha_{gi}(\boldsymbol{\psi}_g)\right) + z_{gi}\log\left(p_g(r_i;\boldsymbol{\eta_g})\right)\right].$$

If we specify starting values $\boldsymbol{\theta}^{(0)}$, the $(k+1)$-th iteration of the E-M algorithm consists of the following steps:

- E-step:

  The conditional expectation of the indicator random variable Z_{gi}, given the observed sample $\boldsymbol{r}$, is:

  $$\mathbb{E}\left(Z_{gi}|\boldsymbol{r},\boldsymbol{\theta}^{(k)}\right) = P_r\left(Z_{gi}=1|\boldsymbol{r},\boldsymbol{\theta}^{(k)}\right) = \frac{\alpha_{gi}(\boldsymbol{\psi}_g^{(k)})p_g\left(r_i;\boldsymbol{\eta_g}^{(k)}\right)}{\sum_{j=1}^{3}\alpha_{ji}(\boldsymbol{\psi}_j^{(k)})p_j\left(r_i;\boldsymbol{\eta_j}^{(k)}\right)} = \tau_{gi}^{(k)},$$

  for $g = 1,2,3$ and $i = 1,\dots,n$. This quantity, according to the Bayes' theorem, is the posterior probability that the i-th subject of the sample with the observed r_i belongs to the g-th component $\mathcal{P}_g$ of the mixture. Since we can notice that $\alpha_{3i} = 1-\alpha_{1i}-\alpha_{2i}$, then $\tau_{3i} = 1-\tau_{1i}-\tau_{2i}$. The expected log-likelihood of complete-data vector is then given by:

  $$\begin{aligned}
  \mathbb{E}\left(\ell_c(\boldsymbol{\theta}^{(k)})\right) &= \sum_{g=1}^{3}\sum_{i=1}^{n}\tau_{gi}\left[\log\left(\alpha_{gi}\left(\boldsymbol{\psi}_g^{(k)}\right)\right) + \log\left(p_g\left(r_i;\boldsymbol{\eta}_g^{(k)}\right)\right)\right] \\
  &= \sum_{i=1}^{n}\tau_{1i}\log\left(\alpha_{1i}\left(\boldsymbol{\psi}_1^{(k)}\right)\right) + \sum_{i=1}^{n}\tau_{2i}\log\left(\alpha_{2i}\left(\boldsymbol{\psi}_2^{(k)}\right)\right) \\
  &+ \sum_{i=1}^{n}\tau_{3i}\log\left(\alpha_{3i}\left(\boldsymbol{\psi}_3^{(k)}\right)\right) + \sum_{i=1}^{n}\tau_{1i}\log\left(p_1\left(r_i;\boldsymbol{\eta}_1^{(k)}\right)\right) \\
  &+ \sum_{i=1}^{n}\tau_{2i}\log\left(p_2\left(r_i;\boldsymbol{\eta}_2^{(k)}\right)\right) + \sum_{i=1}^{n}\tau_{3i}\log\left(p_3\left(r_i;\boldsymbol{\eta}_3^{(k)}\right)\right) \\
  &= \sum_{i=1}^{n}\tau_{1i}\log\left(\pi^{(k)}\right) + \sum_{i=1}^{n}\tau_{2i}\log\left(\delta_i\left(\boldsymbol{\omega}^{(k)}\right)\right) \\
  &+ \sum_{i=1}^{n}(1-\tau_{1i}-\tau_{2i})\log\left[\left(1-\pi^{(k)}-\delta_i\left(\boldsymbol{\omega}^{(k)}\right)\right)\right] + Q^*
  \end{aligned} \tag{2.2}$$

  where Q^* is independent from $\alpha_{gi}^{(k)}$ parameters and therefore we can rewrite the expected value in 2.2 as:

  $$\mathbb{E}\left(\ell_c(\boldsymbol{\theta}^{(k)})\right) = Q_1\left(\pi^{(k)},\boldsymbol{\omega}^{(k)}\right) + Q^*$$

- M-step:

  At the $(k+1)$-th iteration of the E-M algorithm, the function $Q_1\left(\pi^{(k)}, \boldsymbol{\omega}^{(k)}\right)$ has to be maximized with respect to the parameter vector $(\pi^{(k)}, \boldsymbol{\omega}'^{(k)})'$. For obtaining a solution for this expression, numerical methods are generally required. Instead, the estimate of ξ, for a given k, can be obtained from the maximization of the following function:

  $$\sum_{i=1}^{n} \tau_{1i} \log\left(p_1\left(r_i; \eta_1^{(k)}\right)\right)$$

  by solving the system:

  $$\sum_{i=1}^{n} \tau_{1i} \frac{\partial \log\left(p_1\left(r_i; \xi^{(k)}\right)\right)}{\partial \xi} = 0.$$

  The solution produced is:

  $$\xi^{(k+1)} = \frac{m - \bar{R}_n(\boldsymbol{\theta}^{(\boldsymbol{k})})}{m-1}$$

  where $\bar{R}_n(\boldsymbol{\theta}^{(k)}) = \frac{\sum_{i=1}^{n} r_i \tau_1^{(k)}}{\sum_{i=1}^{n} \tau_1^{(k)}}$ is the average of the posterior probability that r_i is a realization of the first component of the mixture, i.e. the Shifted Binomial distribution.

These two steps (E-step and M-step) have to be repeated with the new estimated parameter vector $\boldsymbol{\theta}^{(k+1)}$ until a convergence criterion is satisfied. For instance:

$$\left|\ell\left(\boldsymbol{\theta}^{(k+1)}\right) - \ell\left(\boldsymbol{\theta}^{(k)}\right)\right| < \varepsilon, \qquad \text{for a small } \varepsilon > 0.$$

For the initial values $\boldsymbol{\theta}^{(\mathbf{0})}$ we could choose arbitrary values, but an accurate choice of these numbers permit to accelerate the convergence of the procedure towards the maximum likelihood estimation as stated by McLachlan and Peel (2004) and by Iannario (2012c) in the specific case of CUB models. In the previous algorithm we set the starting values for π and ξ equal to the one used for a standard CUB model. They are obtained, for a given relative frequency distribution of ordinal responses (f_r), as follows:

- $\pi^{(0)}$: the minimum between $\sqrt{\frac{\sum_{r=1}^{m} (f_r)^2 - \frac{1}{m}}{\sum_{r=1}^{m} b_r(\xi)^2 - \frac{1}{m}}}$ and 0.99, in order to avoid initial values on the border of the parameter space. Another choice can be the midrange of the parameter space: $\pi^{(0)} = 0.5$;

- $\xi^{(0)}$: $1 + \frac{(0.5 - maximum(f_r))}{m}$. Another initial value can be $\xi^{(0)} = \dfrac{m - \bar{R}_n}{m-1}$ that is the moment estimator, given $\pi = 1$.

Instead, for $\boldsymbol{\omega}^{(0)}$ we select arbitrary small values e.g. $\boldsymbol{\omega}^{(0)} = (0.1, \dots, 0.1)$. The asymptotic variance-covariance matrix $V\left(\hat{\boldsymbol{\theta}}\right)$ of the ML estimates of the parameter vector is derived by means of the *observed information matrix* $\mathcal{I}(\hat{\boldsymbol{\theta}})$, that is computed as the negative of the Hessian matrix at the estimated parameters.

This step-by-step formulation of the E-M algorithm was programmed using the software R, (R Core Team, 2016).

Chapter 3

A real case study

We apply the model introduced in the previous chapter to a real case study related to self-reported measures of work disability collected within the SHARE survey. We introduce briefly the SHARE project and the data involved in the analysis in Sections 3.1 and 3.2 respectively. Then, in Section 3.3 we report some results concerning the application of TCUB models on this dataset. Through this analysis we would like to highlight the potential of this approach in investigating uncertainty, compared to CUB models with and without covariates.

3.1 SHARE project

Ageing population is one of the greatest social transformations occurred in the 21st century for wealthy countries. Among them, Europe has the older population structure: over 65s are indeed 18.9% of the population and their share will further increase reaching 28.7% by 2080[1]. This phenomenon, caused by low fertility rate and longer life expectancy, is bringing significant changes to the structure of European society, which impacts on economy, health care systems and many other aspects of life. The challenge of these years is then understanding how this process will affect all of us, trying to deal with the novelty of these changes constructively. The main task of SHARE (Survey of Health, Ageing and Retirement in Europe) is exactly understanding demo-

[1]Source: Eurostat, Statistics Explained (http://ec.europa.eu/eurostat/statistics-explained/) - 19/08/2016

graphic ageing. This longitudinal survey is in fact a unique panel database of multidisciplinary micro data covering about 28,000 individuals aged 50 and over in several European countries. SHARE has been designed after the role models of the US Health and Retirement Study (HRS) and the English Longitudinal Study of Ageing (ELSA) and has become itself a role model for several ageing surveys worldwide. Differing from the other ones, it has the advantage of encompassing cross-national variation of public policies, cultures and histories. Data are collected by means of a CAPI (Computer Assisted Personal Interviewing) program, written in Blaise language and supplemented by a self-completion paper and pencil questionnaire. The interviewers conduct face-to-face interviews using a laptop computer on which the CAPI instrument is installed. Personal interviews are necessary for SHARE because they also make the execution of physical tests (Börsch-Supan et al., 2013). To date, SHARE has collected four panel waves (2004, 2006, 2010, 2013) of current living conditions and retrospective life histories (2008, SHARELIFE).

In this work we will focus on the first wave (2004), attended by eleven countries representing the various regions of Europe: Scandinavia (Denmark and Sweden), Central Europe (Austria, France, Germany, Switzerland, Belgium and the Netherlands) and Mediterranean (Spain, Italy and Greece). All household members over 50s plus their spouse/partner, independently of age, were eligible to be interviewed. Data collected measures on physical and mental health (e.g. self-reported health, physical and cognitive functioning, health behaviour, use of health care facilities), economic (e.g. income, current work activity, wealth and consumption, housing) as well as non-economic activities, life satisfaction and well-being (Börsch-Supan et al., 2005; Börsch-Supan, 2005). In particular, we will consider information collected in vignette sub-sample of SHARE: it involved individuals of eight different European countries (Germany, Sweden, the Netherlands, Spain, Italy, France, Greece and Belgium) evaluating their work disability[2]. In the next Section (3.2) we will introduce this data

[2]Vignette data information is not used in this work, but the self-reported measure of interest was collected only in this sub-sample of SHARE.

providing some descriptive statistics.

3.2 Data

The suggested TCUB models will be applied in the next Section (3.3) to the subset of SHARE-wave1 regarding work disability measures. This issue affects economic and social policies in several important ways. Indeed, people with disabilities usually have difficulties in entry into labour force and if in, they may encounter architectural obstacles which do not let them have an adequate accommodation during work. Some difficulties may also be encountered to obtain professional training and education in general. As a consequence, they have low wage levels, low work force participation rates and hence high enrollment rates in public benefits. Trying to measure work disability may help policy makers in their job, however in practice this is quite complicated. Indeed, measuring their physical diseases can be expensive and difficult, because each one has different features related specifically to each health domain. Moreover, the knowledge of the health level of a subject may not directly provide information to his/her work capacity. Hence, the most common solution is to ask people with disabilities to declare if they are limited in their abilities at work, as it happens in this wave: "Do you have any impairment or health problem that limits the kind or amount of work you can do?". Respondents have then to choose an alternative between the five available on a Likert scale: "None", "Mild", "Moderate", "Severe" and "Exteme". However, this approach, involving subjective judgements, may leads to response bias. On the one hand, because of a kind of *Justification Bias*: it may happen when an individual evaluates his/her diseases in a more severe way in order to obtain early retirement or disability benefits or justify his/her employment status (Bound, 1989). On the other hand, because of the tendency of the subject to use the response scale in a way irrespective of the content but according to his/her characteristics or background that lead him/her to develop a response style, as previously described. Therefore, comparing self-assessed evaluations could be problematic since we may notice differences across European countries that may not be in the effective work incapability of the individuals, but in the differences in reporting styles across individuals and countries. With respect to this, Börsch-Supan et al. (2005) pointed out that, comparing the enrolment rates in disability insurance among SHARE countries, it could be found large cross-national differences in disability insurance rates that are not justified by equally large differences in health. He found that this phenomenon can be explained by differences in the institutional rules among countries. In some countries, in fact, there are specifically enrolment and eligibility rules that make disability insurance benefits easier to receive and more generous than in other countries. Furthermore, Banks et al. (2009) find that the self-reported work disability rate in the Netherlands is more than 50% higher than in the United States, even though some analyses suggest that Dutch people are generally healthier than U.S. people.

Thus, to summarize, self-reported work disability, as well as any self-assessed evaluation, may be affected by response styles bias among individuals and countries (DIF). Therefore, it seems useful to implement a TCUB model for this data, in order to show its potential in disentangling the uncertainty component, enhancing cross-national and cross-individual interpretation of differences in work disability evaluation. As already said, we deal with eight European countries, representing the various region in Europe, i.e. Sweden (North), Germany, France, Belgium, the Netherlands (Central) and Spain, Italy, Greece (Mediterranean). The whole sample of respondents involves 4511 individuals. As it is shown in Table 3.1, the majority of them comes from France (877) and Greece (718) while the least are Swedish people (411). In the analyses we will include two demographic characteristics of each respondent: gender (female as reference) and age. We can observe that for each country there is a slight predominance of women, but the difference never exceeds 8%. The age of interviewees varies from 36 to 102 years old, because, as aforementioned, not only over 50s individuals, but also their spouse/partners (regardless of age) were eligible to be interviewed. Nevertheless, only 2.86% of the respondents is younger than 50 years old. By country, the age is quite homogeneous on average: it ranges from 62 (the Netherlands and Greece) to 65 years old (France and Spain).

Table 3.1: Main characteristics of the sample, by country.

Country	Respondents #	Respondents (%)	Women (%)	Mean Age	Work disability (%)
Germany	506	11.22	56.52	63	29.05
Sweden	411	9.11	52.80	64	22.10
The Netherlands	534	11.84	52.25	62	15.92
Spain	463	10.25	58.10	65	33.69
Italy	441	9.78	56.01	64	24.04
France	877	19.44	57.24	65	26.68
Greece	718	15.92	54.18	62	13.79
Belgium	561	12.44	56.15	64	26.02
Total	4511	100.00	55.51	64	24.19

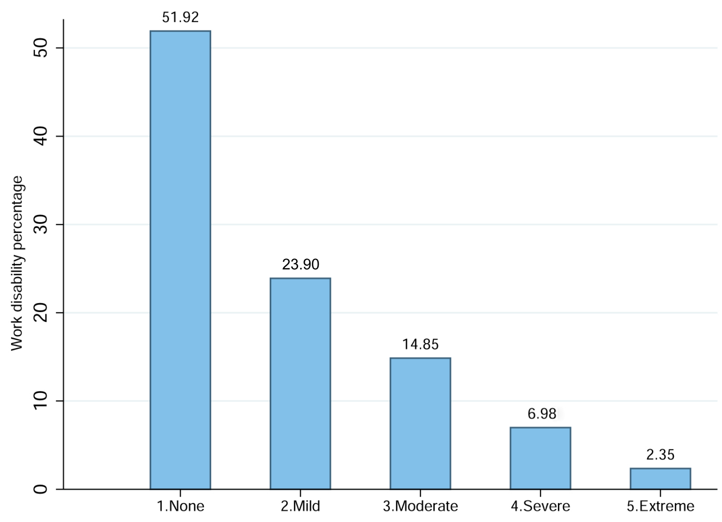

Figure 3.1: Percentage of self-reported work disability, per response category.

Regarding the self-reported work disability of the whole sample, we can notice in Figure 3.1 that most of individuals declares of not beeing disabled (52%). Adding them the almost 24% claiming to have mild health problems that limit the kind or the amount of work that they can do, we can summarize that 76% of the sample evaluates itself as not work incapable. The rest of the individuals classifies themselves as limited in their labour activities, but among them only 2% in an extreme level. In order to provide a more direct measure of the amount of people with work disabilities, according to their self-evaluations, we report in Table 3.1 the work disability percentage per the whole sample and per each country. We assume that a person can be defined work disabled if he/she claims to be "Mild", "Moderate" or "Extreme" limited in the quantity and quality of work that he/she can do because of his/her diseases. These percentages are generally high (more than 24% overall) and different between countries. The greatest percent of self-assessed work disability can be found in Spain, where about a third of interviewees claims to be work disabled, whereas in Greece and in the Netherlands this percentage is less than one half of the previous one (about 14% in Greece and about 16% in the Netherlands).

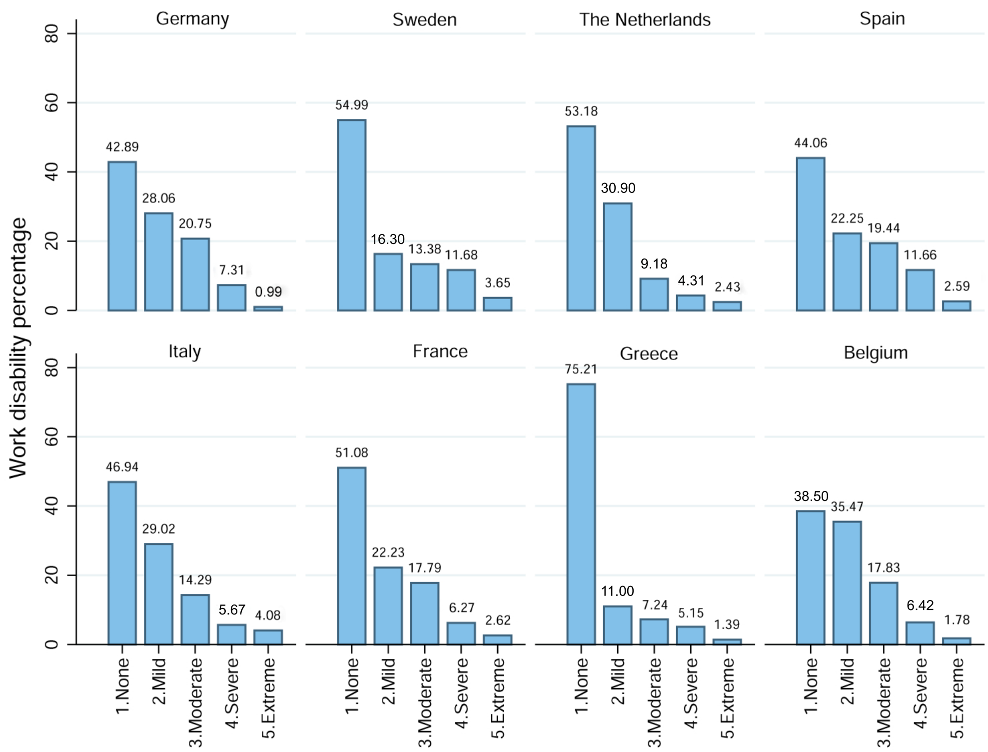

Figure 3.2: Percentage of self-reported work disability per response category, by country.

Looking at the original five-point Likert scale for each country, as reported in Figure 3.2, the differences among countries are even more evident. The shape of the observed distributions is substantially heterogeneous: some are quite flat, as in the case of Belgium, others are sharper as in the case of Greece, where more than 75% of the individuals are concentrated in the first modality. Looking at the work disability distribution by gender, we can notice from Figure 3.3 that males have a more asymmetric distribution than females. Indeed, they choose the first category more than women, who instead locate themselves with non trivial percentages also in central categories.

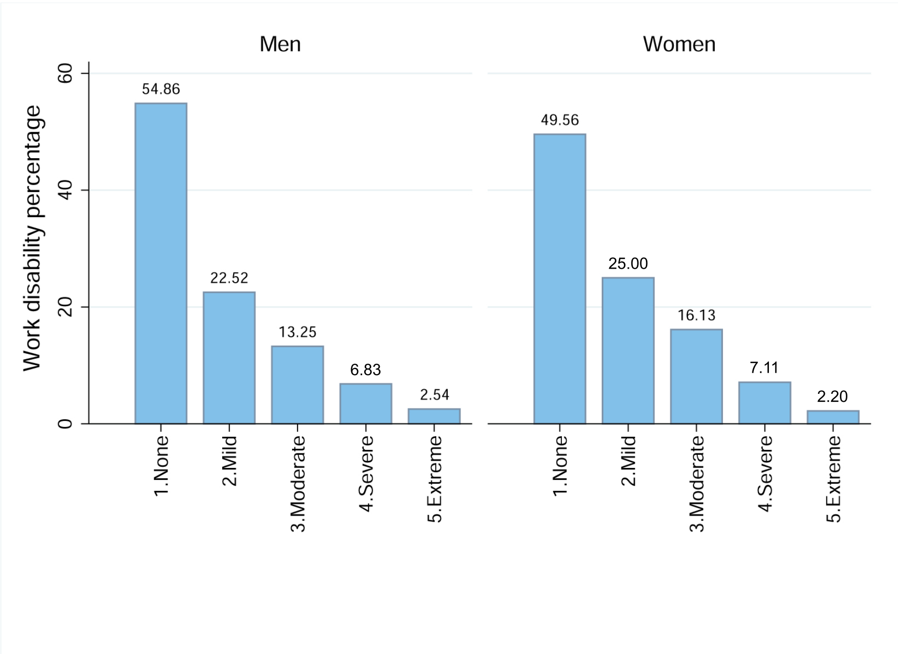

Figure 3.3: Percentage of self-reported work disability per response category, by gender.

Dividing the observed frequencies for each category by age classes (respondents younger than 65 years old, between 65 and 79 years and older than 79 years), we can also capture some interesting findings. According to Figure 3.4, individuals of the latter two classes evaluate themselves as more work disabled with respect to the first one. On the one hand, this might depend on the fact that older people might have more severe health problems than younger one. On the other hand, the majority of people older than 65 years is composed by retired persons. Therefore, we cannot know whether they may answer according to their current level of disability or to the one that they had during their working period.

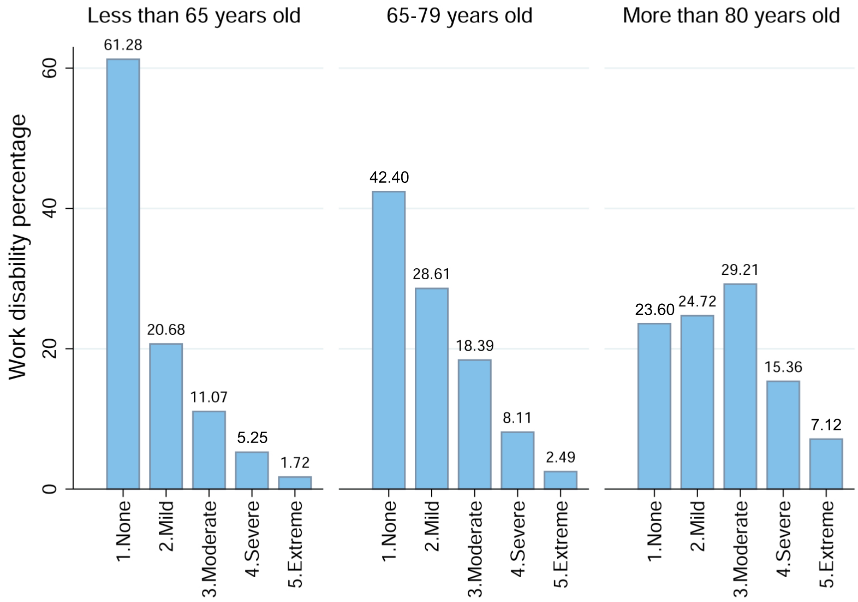

Figure 3.4: Percentage of self-reported work disability per response category, by three classes of respondent's age (less than 65, 65-79, more than 80).

At this point a question raises: does this heterogeneity in self-evaluations point out real dissimilarities between Europeans due to health diseases or does it hide the fact that people with different characteristics (gender, age and so on) and backgrounds, living in different countries, have different response styles and norms for what should be called extreme, severe, etc.? In this latter case, the extension of CUB models suggested in Chapter 2 can be a solution to answer to this question.

3.3 Some empirical evidence

Understanding if the observable heterogeneity among countries and individuals is real or depends on a kind of bias is a very important task. As already mentioned, this is a relevant issue not only for the topic of our analysis, but for each one involving self-reported evaluations. Therefore, providing a tool which can help discriminating the feeling part from the uncertainty one can be very useful. With respect to the specific terminology of CUB models, the aim of a researcher in such cases is trying to separate real feeling toward the item from

the uncertainty that characterises any human choice. This separation can be obtained by means of standard CUB models. But in addition, the uncertainty component can be due to individual traits, responding styles or the context in which respondents have to answer. To discriminate between these different kinds of uncertainty we can apply a TCUB model.
The features of TCUB models that we will estimate in this Section are the ones described in Chapter 2. In particular, for the specification of the $S_{r_i}(m)$ component, we will consider a Triangular distribution with mode beeing in the third response category ($k = 3$). The reasons that lead us to select a Triangular distribution are already been mentioned describing TCUB. To summarize: flexibility of the distribution, possibility to assign different probabilities to each modality and taking into account for response styles such as *acquiescence response style* and *contraction bias*. What's new is the choice of the mode $k = 3$, that does not deal with the observed response distribution, as verifiable in Figure 3.2. In fact, this choice is taken a priori because, in addition to the aforementioned advantages, such Triangular distribution would allow to take into account potential "no choice" responses, which are usually observable in the middle class. Furthermore, as tests for robustness, we estimated TCUB models also with $k = 2$, that is the response category nearest to the observed mode.[3] Such estimates were very close to the ones obtained considering $k = 3$, but the latter provides better fitting. Hence, hereafter as $S_{r_i}(m)$ component in our TCUB model (2.1) it is meant the following probability mass function:

$$S_{r_i}(m = 5) = \begin{cases} 0 & \text{if } r_i = 1, 5; \\ \frac{1}{4} & \text{if } r_i = 2, 4; \\ \frac{1}{2} & \text{if } r_i = 3. \end{cases} \tag{3.1}$$

In the first instance we are going to analyse the whole dataset (Section 3.3.1), then we focus on each country (Section 3.3.2) and finally we compare each

[3]The modes of the observed distributions are $k = 1$, but this modality cannot be chosen as mode in a Triangular specification as the one we consider, according to Kokonendji and Zocchi (2010).

country with respect to the others (Section 3.3.3). At first, we apply a standard CUB model in order to estimate the proportion of feeling and uncertainty components in the decision making process of respondents. Then, observed a relevant presence of uncertainty, it would be desirable trying to enhance its evaluation. In this case, the implementation of a TCUB model would allow to improve the estimate and the interpretation of this component, disentangling it in two parts: one related to the characteristics/background of respondents (*subjective uncertainty*), the other one related to external factors or general characteristics, i.e. questionnaire administration (*content/context uncertainty*). A researcher then will be able to drawn his/her own conclusions concerning the nature and not only the amount of the uncertainty detected. In fact, if uncertainty is mainly subjective, he/she can take it into account in the subsequent analysis by including some relevant subjective characteristics. On the contrary, if it is dominated by the contextual component, it means that there was something fuzzy regarding questionnaire administration. It is evident how much this information can help in understanding if answers are biased and how, but also in enhancing questionnaire realisation. So, let us assume that there is a researcher who has to organise a pilot survey in order to spent resources, time and effort in the most efficient way possible to achieve success in conducting the main survey. He/she has to select a smaller sample size compared to the one planned for the real survey. Usually, information available to select the sample pilot are demographic, i.e. gender and age. If the study is cross-national, he/she might also know the respondent's home country. For these reasons, in the next Sections we will analyse our data controlling for these variables only: the idea is to hypothesise of beeing a researcher who is analysing the results of a pilot survey to verify if the questionnaire may need to be adjusted.

Therefore, the aim of the next Sections is to provide some examples that permit to show the potential of these new approach in such cases. There are some limits yet, because the behaviour and the application of this approach have to be further studied. However, the following real case study allows to appreciate that TCUB estimates could be worthwhile for improving both interpretation of uncertainty component and model fitting, compared with respect to other CUB model findings.

3.3.1 Investigating uncertainty in the whole data set

At first, we report the analysis arranged on the behaviour of all respondents, i.e. considering the whole working disability dataset. In Table 3.2 we provide the estimates of:

- a CUB(0,0) model;
- a VCUB with uncertainty specified as in (3.1) for the $S_{r_i}(m)$ component;
- a CUB with shelter effect at $c = 1$.

The CUB(0,0) is the basic solution, our starting point. The VCUB allows to check if a standard CUB model without covariates, considering the same Triangular distribution of our TCUB, could be preferable to the standard one and/or to our extension. In addition, the CUB with shelter effect permits to compare the aforementioned models with one that control for the presence of a "refuge" category, since in our sample the first was the more selected one (52%). We will exploit these models as benchmarks to compare each estimated TCUB(0,1) not only with the corresponding CUB(0,1), as reported in the following tables, but also with these three solutions without covariates. First of all, it is noticeable that, among the results reported in Table 3.2, all estimated parameters are significant, but the best fitting is obtained with the CUB model with shelter effect. In fact, AIC e BIC criteria suggest that this model is preferable, since they are the lowest. As far as components are concerned, it is worthy to notice that both CUB standard and VCUB detect an high estimate of ξ, i.e. a small estimated work disability level. This because, in this case, the opinion of the individual is expressed as a direct evaluation of the work disability, so the disability level increases with $1 - \xi$.

Table 3.2: CUB standard, VCUB and CUB Shelter estimated on the whole data set.

Parameter	CUB(0,0)		VCUB		CUB Shelter	
ξ	0.877	$***$	0.854	$***$	0.697	$***$
δ					0.375	$***$
π	0.743	$***$	0.841	$***$	0.517	$***$
$L(\hat{\boldsymbol{\theta}})$	-5713.38		-6051.923		-5596.988	
AIC	11430.75		12107.85		11199.98	
BIC	11443.58		12120.67		11219.22	

The estimated feeling level increases in the CUB Shelter (from 0.123/0.146 to 0.303). Uncertainty is higher when we consider the extended model (0.311 with respect to 0.239 and 0.274 of CUB and VCUB). In particular, looking at the VCUB, it seems that specifying uncertainty by means of the triangular distribution (3.1) reduces its estimate with respect to CUB(0,0). However, the fact that this model has a lower log-likelihood with respect to the one with the standard formulation should not lead to signal only a worse fitting. In fact, both estimates are significant in the two models. This leads us to suspect that both behaviours of uncertainty may exist with different weights. This will be further investigated through TCUB models. Regarding the CUB with shelter effect, instead, we can notice that the presence of a substantial effect of the shelter choice (δ), at the first category, is confirmed and has an impact greater than 37%. Moreover, the introduction of a shelter component increases the estimated uncertainty and feeling. This means that most of the respondents identified themselves as workers with "No impairment or health problem that limits their work capacity". Beeing this modality the most selected one, we may conceptually cannot define it properly as a shelter choice. In fact, according with the known health status of the Europeans, we expect that the majority of interviewees would be placed in the first category, because reasonably the majority is just not work disabled.

Therefore, in order to interpreting in a more functional way the behaviour of our respondents, we estimated standard CUB model and TCUB with covariates respectively for the uncertainty and for the subjective uncertainty component. We insert at first one by one, as previously justified, gender and age of individuals. In particular, we include:

- Gender as dummy variable with female as reference category (Table 3.3);
- Age as continuous variable (Table 3.4). More specifically, $\widetilde{age_i} = \log(age_i) - \overline{\log(age_i)}$, where $\overline{\log(age_i)}$ denotes the average of age's logarithm. We use this transformation in order to reduce the variability by means of logarithm and centre the variable thanks to the deviation from its mean. It allows also to obtain a unimodal distribution of the variable under investigation.

Table 3.3 reports the estimated parameters of CUB(0,1) and TCUB(0,1) with gender as covariate, the log-likelihood and the model selection criteria AIC and BIC. We can note that the estimate of a gender-specific uncertainty in CUB(1,0) is not significant (π_{i_F}). In fact looking at log-likelihood and selection criteria we can understand that the performances of this model are similar to the one of the standard CUB solution (log-likelihood and AIC) or even worse, as attests BIC criterion. This happens because, beeing not significant the introduction of this covariate, this model degenerates to the standard one. However, for the computation of the BIC criterium its log-likelihood is penalized with an additional parameter and therefore it results worse. Instead, all parameters of the TCUB are significant. The estimates of feeling and uncertainty dimensions are slightly superior and inferior, respectively, to the one of the CUB(0,0), but not substantially different. The estimates of δ_i, i.e. the parameter related to the triangular component, are different among men (0.170) and women (0.220). It means that for women the weight of the subjective uncertainty is greater (and significant) than for the men. Therefore, women seem more inclined to use central categories, that allow to take no extreme positions (according to *response contraction bias*), or to adhere to the question (as for *acquiescence response style*). Regarding the other uncertainty component, the contextual one, we can affirm that women give "more reasoned" answers, since the weight of the uniform component (the totally indifferent choice) is 9.1%

(14% for men).

Table 3.3: CUB(0,1) and TCUB(0,1) with respondent's gender as covariate.
$***$ p-value$<$ 0.001, $**$ p-value$<$ 0.01, $*$ p-value$<$ 0.05, $\cdot$ p-value$<$ 0.1

Parameter	**CUB(0,1)**		**TCUB(0,1)**	
ξ	0.878	$***$	0.901	$***$
π			0.685	$***$
β_0	1.160	$***$		
β_1	-0.187			
ω_0			-1.574	$***$
ω_1			0.303	$*$
$L(\hat{\boldsymbol{\theta}})$	-5712.102		-5589.845	
AIC	11430.200		11187.690	
BIC	11449.450		11213.350	
Uncertainty			**Gender**	
	M	**F**	**M**	**F**
Subjective			0.170	0.220
Contextual			0.140	0.091
Standard	0.240	0.274		

Two important findings result from this analysis. First, the contextual uncertainty has not a large proportion, so we may think that the questionnaire administration is not a serious problem. Second, even if not large, significant differences between women and men in the decision making process appear. Indeed, for women the uncertainty is clearly subjective-oriented, while for men, contextual uncertainty has a not trivial role. In the hypothesis of a pilot study, a researcher should take into account this finding, because it may underline the presence of some problems in the questionnaire administration (i.e. understanding of questions? tiredness?) for men.
Looking at the performance of this model, we notice an improving in the log-likelihood such that AIC and BIC are lower with respect to the ones of the previously analysed models.
Implementing these two kinds of models including the age (specified as aforementioned), we obtain significant estimates for both.

Table 3.4: CUB(0,1) and TCUB(0,1) with respondent's $\widetilde{age}$ as covariate.
$***$ p-value$<$ 0.001, $**$ p-value$<$ 0.01, $*$ p-value$<$ 0.05, $\cdot$ p-value$<$ 0.1

Parameter	**CUB(0,1)**		**TCUB(0,1)**	
ξ	0.891	$***$	0.901	$***$
π			0.681	$***$
β_0	1.024	$***$		
β_1	-5.734	$***$		
ω_0			-1.473	$***$
ω_1			0.942	$**$
$L(\hat{\boldsymbol{\theta}})$	-5602.901		-5587.720	
AIC	11211.800		11183.440	
BIC	11231.050		11209.100	

The results are reported with their significance level and log-likelihood, AIC and BIC in Table 3.4. The estimates are significant for both models, but log-likelihood is larger for the TCUB, allowing AIC e BIC to be lower than CUB(0,1) and to the previous models. Since this covariate is inserted as continuous, there are presented the estimates of parameters directly related to it, instead of the ones transformed with the logistic link. The relation among them is $logit\,(\pi_i) = \boldsymbol{x_i\beta}$ for CUB(0,1) and $logit\,(\delta_i) = \boldsymbol{x_i\omega}$ for TCUB(0,1). The estimates of ξ and π are coherent with the previous one. The ones of the age-specific uncertainty in CUB(0,1) point out that to an increase in the age corresponds an increase in uncertainty too. Old people are more fuzzy in the decision making process than young one. It can be justified both by health reasons, and by the fact that older respondents are retired and therefore they could be more indecisive or indifferent giving their answers. In order to see in a easier way the impact that respondent's age has on the uncertainty's parameters, we plot in Figure 3.5 these estimates as function of age (more specifically, $\widetilde{age}$). This graph permits to visualize dynamically how uncertainty in the CUB(0,1) and subjective and contextual uncertainty in TCUB(0,1) are modified by different age of respondents and simplifies comparisons. The CUB(0,1) is drawn on the top of Figure 3.5. This chart shows the aforementioned increasing trend with respect to the age of respondents. But, in addition, it points out that the uncertainty's growth is small (ranging from 0 to 0.2), if age is lower than the $\overline{\log(age_i)}$. The other way around, if it is higher than $\overline{\log(age_i)}$.

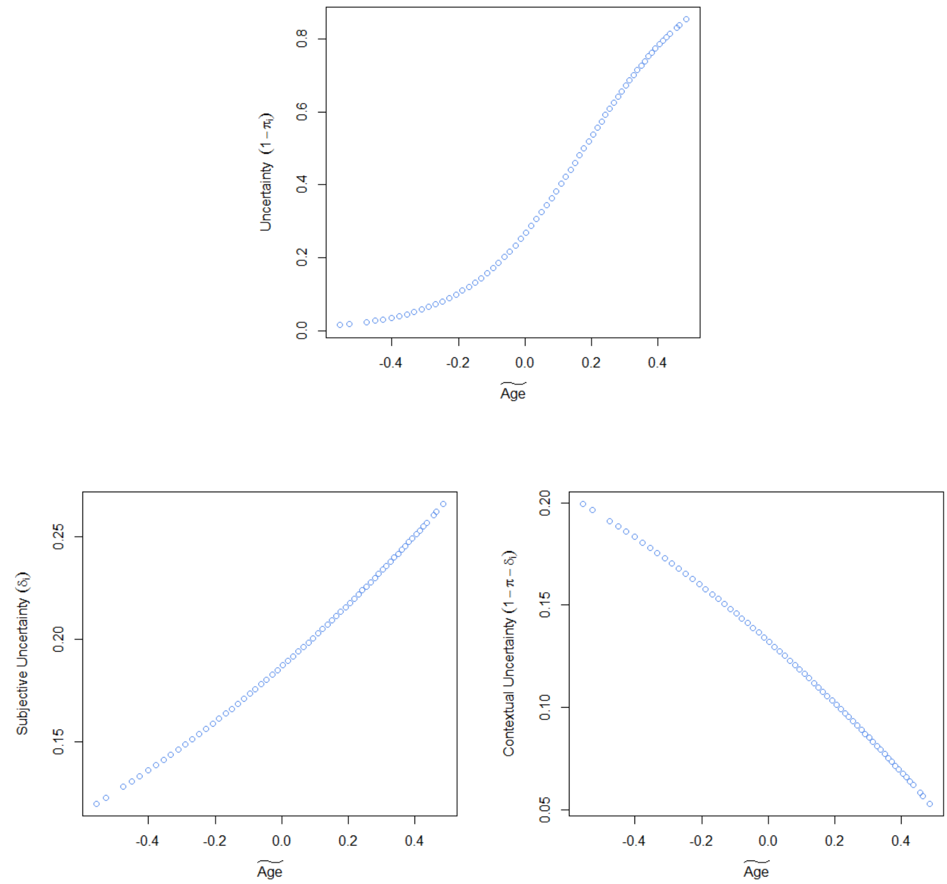

Figure 3.5: Uncertainty estimates of CUB(0,1) and TCUB(0,1) at different values of $\widetilde{age}$.

Concerning the weight of the uniform distribution estimated with the TCUB, we can notice that the range is much smaller than the one of the CUB(0,1), since it varies from 0.05 to 0.20. In addition, the relation highlighted by the plot is reversed: for respondents who are older than $\overline{\log(age_i)}$ the proportion of the uniform is lower than the younger ones. However, this evidence is not a contradiction. The kind of uncertainty captured by this distribution is different. In fact, in the TCUB there is a sort of disentangling of the standard defined uncertainty. So the uniform distribution is requested to take into account only of a part of it. In particular, the part of a residual uncertainty

that, regardless from respondent's characteristics, might be related to fuzziness due to time pressure, to questionnaire administration and generally to uncertainty arises because of the context. The other uncertainty dimension, instead, has an upward trend as it was for the uncertainty in CUB(0,1), but it is more straightforward and varies from 0.10 to about 0.30. Such evidence attests that the more the respondent is old, the more he/she tends to choice central modalities, according to the respective response styles.
Since gender and age have significant estimates of the associated parameters, we computed also the TCUB that includes both of them. The estimate of the parameter related to the gender is no longer significant. That indicates that inclusion of age allows to explicate in some way also the gender of the respondent. Therefore, the model that specifies only $\widetilde{age}$ as covariate results preferable. This could be due for instance by the fact that men and women are not equally distributed by age (the proportion of women is larger than the one for men in the oldest age classes). Moreover, in the TCUB model with only gender, the p-value of the estimated parameter was not so large.

3.3.2 Investigating uncertainty per country

The descriptive analysis in Section 3.2 pointed out different self-assessed work disability levels among countries, and we now try to understand if these differences are only in term of feeling or also in term of uncertainty. The first ones are differences in the strictly personal opinion of respondents on their work incapability, the others are differences in the fuzziness arisen in the decision making process. Then, it is worth investigating if the latter concerns the so called subjective or contextual uncertainty. Following the aforementioned procedure, at first we will provide CUB(0,0) estimates computed on the whole sample and on each European country (separately). These are drawn in Figure 3.6 as points in the parameter space (reduced to simplify interpretation), whose coordinates are the estimated feeling (on the vertical axis) and uncertainty (on the horizontal axis). By means of this representation we can capture the differences among countries in a easy way. First of all, we can note that they can be grouped in approximately three levels of feeling. One containing Belgium and Germany, which are quite overlapped, having the highest feeling estimate, that is about 0.20.

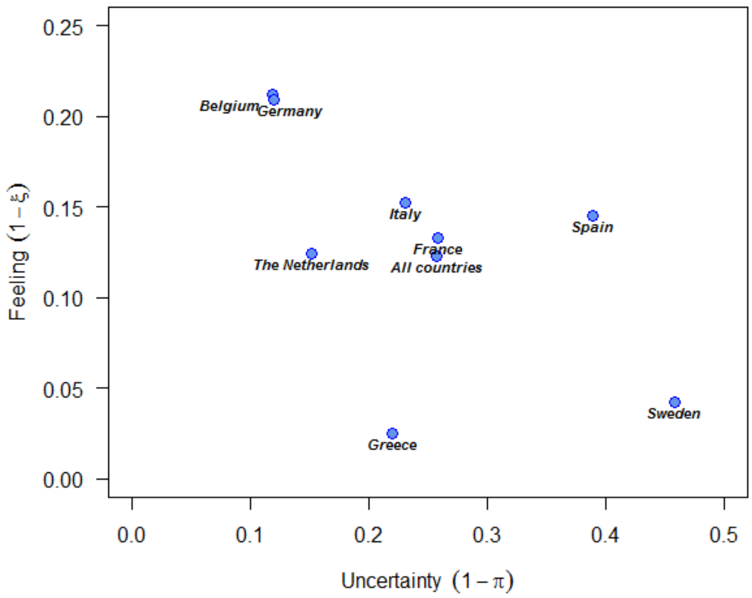

Figure 3.6: Estimated $CUB(0,0)$ for each European country and for the whole sample, represented as points in the parameter space.

In this case, the opinion of the individual is expressed as a direct evaluation of the work disability, so the disability level increases with $1-\xi$. Therefore, respondents of this first group have the greatest estimated level of work incapability. For the second set of countries, that involves Italy, France, Spain and also the $CUB(0,0)$ of the whole sample, feeling is slightly lower, ranging from 0.10 to 0.15. In the end, respondents from Greece and Sweden have the lowest estimated feeling, confirming the Greeks as the one having the lowest self-assessed work disability.

Corresponding to the highest levels of feeling dimension, for Belgian and German people there are also the lowest levels of uncertainty. This reveals that their propensity to an indifferent choice is lower than 13%. It has a similar value also for the Netherlands, and then it increases from left to right, as we can notice in Figure 3.6. The greater estimate of this component is measured

in Sweden, where the tendency to choose a response category indifferently is almost 50% (half uncertainty, half feeling).
These findings highlight how it may be meaningful the will to know what is causing uncertainty. Therefore, we implement a TCUB model for each country including gender and $\widetilde{age}$ as covariates for the subjective component, as done for the whole dataset in the previous section.
A first interesting result is that some models, not each model estimated on the single countries, do not converge (specifically, for Greece and Germany data). This could be a limit of the model, but we think it deals with the sample size and in particular, with the number of observations for each response category. First, implementing the TCUB on the whole dataset, there were no convergence problems, while they arise now, considering each country and therefore a more limited sample size. Second, only Greece and Germany have this problem, that is just the ones, with the lowest proportion of respondents in the fifth category (1.39% and 0.99% respectively- see Figure 3.2), which is the one that ensures identifiability ($m > 4$ is the identifiability condition).
This finding highlights, as expected whenever any new model is introduced in the literature, that some further studies on the TCUB are needed: on the sample size as a whole and on categories' size that allows to identify the model.
Regarding the results of the TCUB(0,1) with gender as covariate for uncertainty, we obtain that women do not use in a significant different way the response scales with respect to men. So, for the sake of brevity, we do not report such results. Moreover, in Greece and Germany the number of men and women separately in the fifth category is extremely low. Sample size involved in the country-specific analysis may not be sufficient to ensure accurate standard error estimates, particularly when the effects of such covariate do not seem so strong, compared to those of age.
Indeed, the estimation of TCUB models with $\widetilde{age}$ as covariate for subjective uncertainty, shows similar (but not identical) results. The estimated parameters associated with $\widetilde{age}$ are significant for the TCUB model computed on Swedish and French data. Such estimates are reported in Table 3.5 together

with the estimates obtained from the corresponding CUB(0,1) models. With respect ot the fitting, we can notice that the TCUB allows to enhance it, since AIC e BIC criteria are inferior to the one of CUB model with the same covariate for uncertainty component.

Table 3.5: CUB(0,1) and TCUB(0,1) estimated for Sweden and France with $\widetilde{age}$ as covariate for uncertainty and subjective uncertainty respectively.
$***$p-value$<$ 0.001, $**$p-value$<$ 0.01, $*$p-value$<$ 0.05, $\cdot$p-value$<$ 0.1

	Parameter	**Sweden**		**France**	
CUB(0,1)					
	ξ	0.965	$***$	0.892	$***$
	β_0	0.100		0.892	$***$
	β_1	-2.560	$\cdot$	-6.266	$***$
	$L(\hat{\boldsymbol{\theta}})$	-529.246		-1109.332	
	AIC	1064.492		2224.664	
	BIC	1076.548		2238.994	
TCUB(0,1)					
	ξ	0.968	$***$	0.930	$***$
	π	0.570	$***$	0.486	$***$
	ω_0	-1.437	$***$	-0.499	$**$
	ω_1	2.673	$*$	0.824	$\cdot$
	$L(\hat{\boldsymbol{\theta}})$	-519.726		-1102.817	
	AIC	1047.452		2213.634	
	BIC	1063.527		2232.740	

We report also the results obtained for the other countries, that gave non significant estimates of $\widetilde{age}$ parameters, because they can however lead to some interesting considerations (Table 3.6). Regarding CUB(0,1) estimates, they are significant with except only to Germany; however, they cannot be compared to the TCUB ones because, as aforementioned, such model does not converge for this country. The finding that parameter estimates associated with $\widetilde{age}$ are all significant for CUB(0,1) points out that the totally indifferent choice has not the same weight for respondents with different ages, as already noticed considering the whole dataset. TCUB estimates do not show significant differences concerning the use of central categories, due to response styles for example.

Table 3.6: CUB(0,1) and TCUB(0,1) estimated for each country with $\widetilde{age}$ as covariate for uncertainty and subjective uncertainty respectively. $***$p-value<0.001, $**$p-value<0.01, $*$p-value<0.05, $\cdot$p-value<0.1

	Parameter	**DE**[a]		**NL**		**BE**		**GR**[b]		**IT**		**ES**	
CUB(0,1)													
	ξ	0.798	$***$	0.882	$***$	0.797	$***$	0.975	$***$	0.862	$***$	0.855	$***$
	β_0	1.881	$***$	1.860	$***$	2.175	$***$	1.588	$***$	1.234	$***$	0.533	$*$
	β_1	-2.924		-6.659	$***$	-7.472	$**$	-6.846	$***$	-7.692	$***$	-4.533	$***$
	$L(\hat{\boldsymbol{\theta}})$	-666.944		-599.608		-719.547		-594.295		-526.238		-646.005	
	AIC	1339.888		1205.217		1445.093		1194.591		1058.475		1298.011	
	BIC	1352.568		1218.058		1458.082		1208.320		1070.567		1310.424	
TCUB(0,1)													
	ξ			0.895	$***$	0.884	$***$			0.874	$***$	0.903	$***$
	π			0.794	$***$	0.601	$***$			0.719	$***$	0.539	$***$
	ω_0			-2.382	$**$	-0.748				-2.322	$***$	-0.708	$***$
	ω_1			-0.056		-0.520				1.228		-0.228	
	$L(\hat{\boldsymbol{\theta}})$			-610.823		-724.720				-536.222		-629.440	
	AIC			1229.654		1457.439				1080.445		1266.88	
	BIC			1246.775		1474.758				1096.568		1283.431	

[a]TCUB(0,1) for Germany does not converge.
[b]TCUB(0,1) for Greece does not converge.

This could mean that on these countries either the subjective component is not sizeable or it is important (compared to the contextual component) but it is related to other individual characteristics.

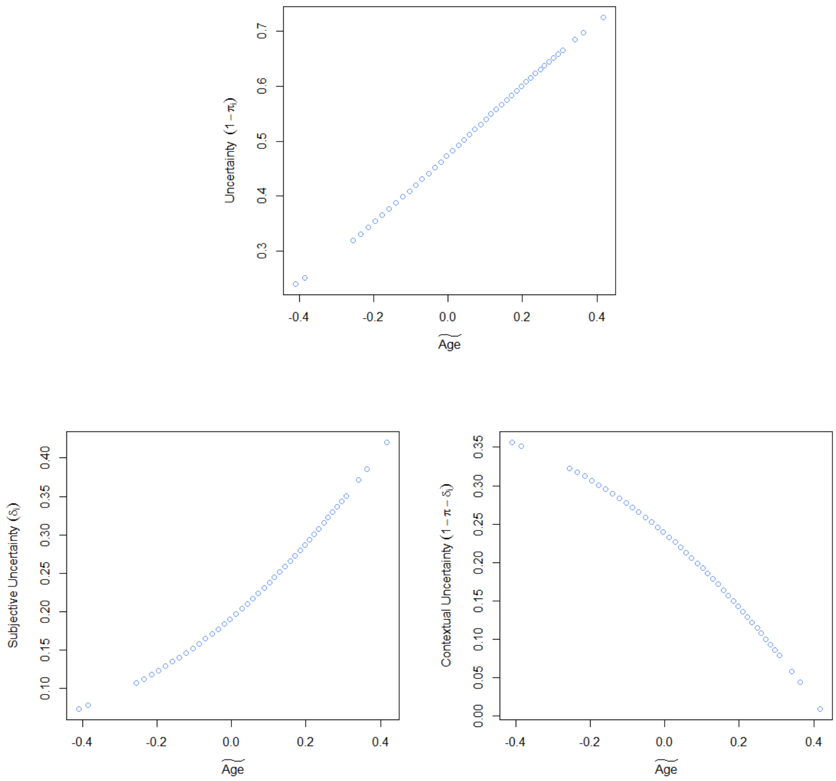

Figure 3.7: Uncertainty estimates of CUB(0,1) and TCUB(0,1) on Swedish data at different values of $\widetilde{age}$.

Regarding Sweden and France, we can see in Figures 3.7 and 3.8 the influence that $\widetilde{age}$ has on the estimated uncertainty component in CUB(0,1) and subjective and contextual ones in TCUB(0,1). Adopting the same approach used analysing the whole dataset, in the CUB(0,1) model for both countries, the weight of a totally indifferent choice increases with $\widetilde{age}$, varying between almost the whole parameter space. Differences are in term of relation's shape, that is more straightforward for France. With respect to subjective and contextual uncertainty, estimated by means of the TCUB models, we observe almost the same trends of the whole sample, with the same interpretation. More specifically, if we compare these figures to Figure 3.5, we can notice that the trends are really similar, in particular if we consider the tendencies of France with respect to the ones of the whole sample. This might happen because France is the country with the majority of respondents (almost 20%), so the overall observed behaviour seems to be driven by the answers of French respondents.

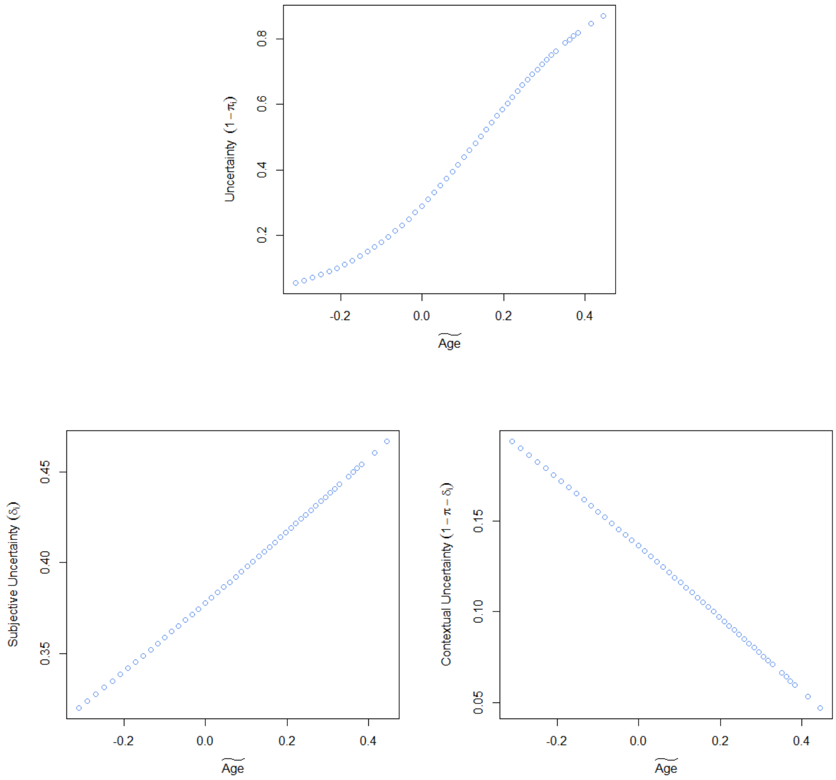

Figure 3.8: Uncertainty estimates of CUB(0,1) and TCUB(0,1) on French data at different values of $\widetilde{age}$.

Even if the sample size for Sweden is not so large as for France, also Swedish behaviours with respect to age are very close to the one observed considering pooled data, letting as suppose that what we observed on the whole sample (with respect to $\widetilde{age}$) was mainly due to the behaviour of these two countries.

3.3.3 Comparing uncertainty of each country with respect to the others

Now we analyse work disability data in order to detect which countries use the response scale in a way different to all the others, having a significant estimated weight of subjective uncertainty. Therefore, we consider the pooled data and compute on them CUB(0,1) and TCUB(0,1) including the living country of the respondent as covariate. The latter was included as a dummy variable for each country which is equal to one if the interviewee belongs to the country of interest, zero otherwise. Results are shown in Table 3.7. The countries that show a significant estimate of the parameter associated with belonging to a specific country are: Germany, the Netherlands, Belgium and Greece.

There are several surprising results.

Differently from the analysis on each country: on the one hand, models for Germany and Greece do now converge; on the other hand, no statistically significant result appears for France and Sweden.

The case of Germany and Greece enlightens the role of the sample size in each answer category: respondents who evaluated themselves as extremely limited are 5 and 10 in Germany and Greece respectively. Looking, for instance, at the gender distribution, the number of males who defined themselves as extremely limited is 2 and 4 in Germany and Greece respectively. It seems that, conditionally to an individual characteristic taking into account, each modality needs to be chosen by at least five respondents to ensure that TCUB model may converge. Results for France and Sweden are not surprising, thinking to the findings highlighted in the previous Section: considering the age of respondents, the behaviours of Swedish and German people are very similar about subjective and contextual uncertainty. However, it seems that the results obtained on the whole sample are mainly driven by the behaviour of just these two countries. In the analyses of this Section, by construction one country is in one group and the other country is in the other group. For this reason, it is reasonable that no statistically significant differences may appear.

Beeing the aim of this work showing in which terms a TCUB model can be useful, we further consider the models computed including Germany, Belgium, the Netherlands and Greece as covariates for subjective uncertainty. To this aim, we drawn these estimated TCUB models with significant estimates of the parameters associated to the specific country as points in a Simplex Plot in Figure 3.9, exploiting the relation: $\pi + \delta_i + \nu_i = 1$.

Table 3.7: CUB(0,1) and TCUB(0,1) with dummy country covariate for uncertainty and subjective uncertainty respectively.
***p-value<0.001, **p-value<0.01, *p-value<0.05, ·p-value<0.1

	Par	**DE**		**NL**		**BE**		**GR**		**IT**		**FR**		**SW**		**ES**	
CUB(0,1)	ξ	0.878	***	0.878	***	0.878	***	0.844	***	0.877	***	0.877	***	0.876	***	0.878	***
	β_0	1.085	***	0.977	***	1.079	***	0.814	***	1.072	***	1.088	***	1.136	***	1.158	***
	β_1	-0.323	·	0.725	***	-0.198		1.189	***	-0.102		-0.137		-0.541	*	-0.880	***
	$L(\hat{\boldsymbol{\theta}})$	-5711.995		-5706.339		-5712.821		-5688.676		-5713.240		-5712.946		-5708.986		-5701.298	
	AIC	11429.990		11418.680		11431.640		11383.350		11432.480		11431.890		11423.970		11408.600	
	BIC	11449.230		11437.920		11450.880		11402.600		11451.720		11451.130		11443.210		11427.840	
TCUB(0,1)	ξ	0.901	***	0.901	***	0.902	***	0.901	***	0.901	***	0.901	***	0.901	***	0.902	***
	π	0.685	***	0.686	***	0.685	***	0.677	***	0.687	***	0.686	***	0.686	***	0.686	***
	ω_0	-1.480	***	-1.346	***	-1.451	***	-1.281	***	-1.352	***	-1.406	***	-1.352	***	-1.427	***
	ω_1	0.548	**	-0.642	*	0.379	*	-4.170	*	-0.3664		0.083		-0.395		0.247	
	$L(\hat{\boldsymbol{\theta}})$	-5584.205		-5590.093		-5589.145		-5577.851		-5591.330		-5592.707		-5591.370		-5591.538	
	AIC	11176.410		11188.190		11186.290		11163.700		11190.660		11193.410		11190.740		11191.080	
	BIC	11202.070		11213.840		11211.950		11189.360		11216.320		11219.070		11216.400		11216.730	

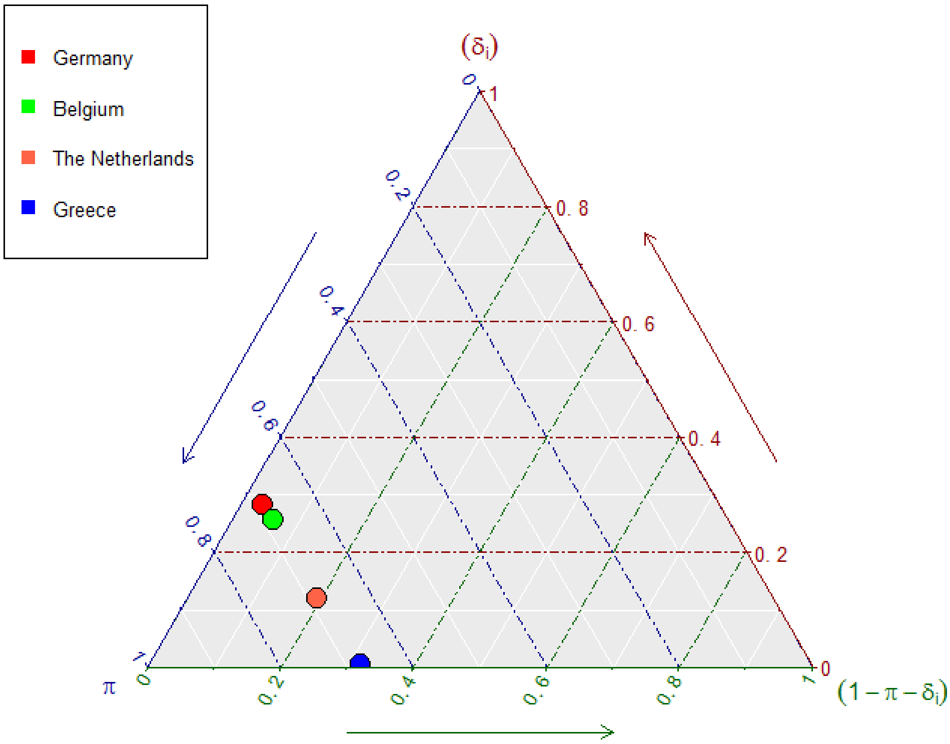

Figure 3.9: Estimated TCUB models parameter of uncertainty for Germany, Belgium, the Netherlands and Greece (Table 3.7) are shown as points in a simplex plot with a size proportional to the estimate of the feeling parameter.

Coordinates of each point show the weights of these components, while point size is proportional to the estimate of the feeling parameter (even if it is basically the same for each one in these cases). This plot allows to compare the different estimates with respect to the covariates, since the latter are dummy variables. It let also us interpreting appropriately stability, relative weight and shifting of the components among the specified subgroups of respondents. Therefore, this plot points out that the estimates of ξ (the size of the points) and π are effectively the same in the different models. What changes is the estimate of the parameter related to the subjective uncertainty (δ_i) and, as a consequence, also the contextual one. We decide to insert no estimate of the baseline of the country's dummy: "other respondents expect the ones living in the country in object" to clarify the interpretation. Looking at this Simplex Plot it is easy to notice how and how much the weight of the two components of the uncertainty vary with respect to the characteristic of the country that we consider. For example, it is evident that Greeks (with respect to the others) behave the opposite of Germans. In particular, comparing the behaviour of Germans to the others' one, we can observe that they are fuzzier according to estimates of CUB(0,1) (Table 3.7). Rather, the implementation of the TCUB(0,1) allows to notice that German people have indeed a heavy weight of the uncertainty component, but of the subjective one. They are therefore more inclined to use the response scale according to our assumed triangular distribution than people of the other countries. The same phenomenon can be detected in the case of Belgian people, although the corresponding estimate of δ_i is smaller than the German one. Moreover, the uncertainty parameters computed by the CUB(0,1) are not significant. The opposite behaviour can be founded for the Netherlands and Greece. Their respondents seem, indeed, to be less inclined to choose central categories with respect to the others. For these countries, it is observed a higher uniform use of the response scale available, particularly for Greece, whose weight of the subjective component (δ_i) is much lower than the others' one. This result means that Greek respondents have a large estimated weight of the contextual uncertainty, letting us suppose that these people were more influenced by some kinds of external factors than by individual response styles or subjective features. This might point out the need to pay more attention on the questionnaire administration for these two countries, but for Greece in particular, where it is evident that there is something in the context (as for instance, the general culture of this country or something specific of its labour market features) that might lead to this evidence.

Conclusions

This work aims at describing a new approach useful to analyse ordinal data, in particular the ones coming from rating surveys, where people are requested to evaluate objects, items, services, by choosing among a list of ordered categories. This model is an extension of a standard CUB model, with the goal of separating not only the feeling of the respondent toward the item to the uncertainty component that naturally characterises human choices, but to disentangle this uncertainty component in a subjective and a contextual one. Indeed, looking at the definition of uncertainty according to a standard CUB model, we may identify two fundamental traits. One is related to characteristics and/or background of respondents, who choose a response category irrespective of its content but following their particular response style (say, subjective uncertainty). The other is due to time pressure, equivocal wording and questionnaire administration in general, that might lead respondent to choose indifferently one category among the m available (say, contextual uncertainty). In order to take into account of these aspects, we extended the structure of a standard CUB model which consists, in brief, of the mixture of two components: a shifted Binomial and a discrete Uniform distribution. The first deals with feeling and the second with uncertainty. The extended model adds to these two elements of the mixture a third one: a Triangular distribution, specified as in Kokonendji and Zocchi (2010), whose weight can be related to one or more subjective covariates by means of a logistic link (as in standard CUB models with covariates). In such way, the feeling component is yet captured by the shifted Binomial distribution, while the Uniform distribution, representing the totally indifferent choice (since it assigns the same probability to each response category) is now suitable to model the contextual uncertainty component. The Triangular distribution (more flexible than the Uniform one) aims at taking into account the subjective uncertainty and in particular response styles such as *response contraction bias*, *resoluteness in the extremes* and "no choice" tendency. This model is therefore called TCUB (**CUB** models with a **T**riangular uncertainty component). In order to see the potential of this new extended CUB model in some empirical examples, we applied it to a dataset of the SHARE survey, regarding self-reported work disability level in eight European countries. These results were then compared to the ones of other CUB models.

To summarize the findings of TCUB application, we can state that, as far as components are concerned, it is noticeable a substantial homogeneity among the estimated feeling parameters of all computed models. In this context of study, this means that the work incapability *level* is estimated basically in the same way by all different models. The same happens to the estimated *weight* of this component (π) in TCUB models and in the CUB(0,0) ones. Thus, we can affirm that the basic structure of the model is not radically changed by the introduction of this extension. This might confirm that TCUB allows to investigate uncertainty in a more effective way, without affecting the other parameter estimates. This is in line with the aim of TCUB models which is to disentangle the uncertainty dimension in the two aforementioned parts.

We had the chance to see how the TCUB model might be a solution if the aim is to better understand the nature of the uncertainty dimension and to interpret and visualize its estimates. In the reported cases (Chapter 3, Section 3.3), these models achieve a better fitting than the already existing ones. Their performances in fact are preferable in terms of log-likelihood's improvement and automatic selection criteria AIC and BIC. Concerning interpretation, the TCUB model, with gender as covariate for explaining subjective uncertainty, allows to point out that women has an higher weight of subjective component with respect to men. Their uncertainty in choosing the answer was mainly due to the tendency to prefer central categories, than to a totally indifferent choice.

Furthermore, we observe that if a subgroup of people has a higher level of estimated uncertainty in CUB(0,1), for such subgroup is observed also a higher level of subjective uncertainty in TCUB(0,1). We can therefore suppose that, at least for this empirical example, the uncertainty captured by the standard CUB model with covariates in these subgroups might be mostly explicated by subjective characteristics. Consequently, large uncertainty estimates might depend mostly on individual's characteristics, background and response styles and not so much on external factors. Such a evidence might suggest that there is no heavy fuzziness in response process due to time pressure, number of questions and questionnaire administration in general.

Moreover, in our analysis, the subjective component becomes more relevant as the age of respondents increases. The reverse trend can be observed for the contextual uncertainty. This finding is worthier if we take into account that a standard CUB model with the same covariate for uncertainty can detect only an increasing trend. Analysing this data per country, the trends observed for age applying a TCUB on the whole dataset are the same of the ones observed for Sweden and, in particular, for France. Therefore, we suppose that the trends of the pooled data are primarily driven by the respondents' behaviour of these two countries (which count together about 29% of the sample size).

The analysis of TCUB models separately for each country involved in the survey arose some relevant questions: which sample size ensures the accuracy of standard error estimates? Which is the minimum size in each category that ensures model identification? The latter question is particularly relevant. Indeed for the two countries with the smallest number of frequencies in a response modality (less then 1.3% of respondents of these countries chosen the fifth category), the TCUB model does not converge. Since the specification of the TCUB model needs the introduction of at least one covariate, the number of respondents having that characteristic (i.e. beeing male or female) and choosing the fifth answer category may be extremely low. This explains why the estimations for Greece and Germany do not converge when these countries are analysed separately, while no convergence problems appear with TCUB models specifying the country as a covariate. Some findings suggest a minimum number of 5 answers in each modality, conditionally to the individual characteristic, to reach the convergence. The sample size might also affect, in this case, the estimate of standard errors and therefore the significance of the estimated parameters. These issues are currently under investigation and need to be further studied.

From the cross-country comparisons, German and Belgian people show a significant large estimated weight of the subjective component with respect to the other countries. This means that their decision making process mainly consists in feeling component and subjective uncertainty, i.e. due to response styles for example. The contrary can be observed in the Netherlands and especially in Greece. The latter, indeed, has a large estimated weight of the contextual uncertainty, letting us suppose that Greek people may be more affected by some kind of external factors than response styles or subjective features. This findings might point out the need to pay more attention on the questionnaire administration, for these two countries but in particular for Greece, where uncertainty is almost totally contextual. We can argue that there is something in the context, as for instance the general culture of this country or something specific of its labour market characteristics that might lead to this evidence.

The possibility to draw these considerations is very relevant. For example, let us think at a researcher who has to analyse the results of a pilot survey. In such a context these findings can provide the necessary information to understand if uncertainty observed is due to subject's peculiarities, with the possibility to take them into account in the analyses, or to the questionnaire administration, suggesting for which countries or subgroups of respondents it is convenient to adjust and enhance it.

In this work we provide only some results, useful to see the potential and the richness of this extension of CUB models, also with its limits due to the fact that this proposal is under scrutiny yet. More can be observed, by including more subjective covariates (if available), but a lot of its potential can be already seen with the reported analysis.

Bibliography

Agresti, A. (2010). *Analysis of ordinal categorical data*, volume 656. John Wiley & Sons.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE transactions on automatic control*, 19(6):716–723.

Banks, J., Kapteyn, A., Smith, J. P., and van Soest, A. (2009). Work Disability is a Pain in the****, Especially in England, the Netherlands, and the United States. In *Health at Older Ages: The Causes and Consequences of Declining Disability among the Elderly*, pages 251–293. University of Chicago Press.

Baumgartner, H. and Steenkamp, J.-B. E. (2001). Response styles in marketing research: A cross-national investigation. *Journal of marketing research*, 38(2):143–156.

Börsch-Supan, A. (2005). The survey of health, ageing, and retirement in europe. *Methodology, Mannheim Research Institute for the Economics of Aging (MEA).*

Börsch-Supan, A., Brandt, M., Hunkler, C., Kneip, T., Korbmacher, J., Malter, F., Schaan, B., Stuck, S., and Zuber, S. (2013). Data resource profile: the survey of health, ageing and retirement in europe (share). *International journal of epidemiology.*

Börsch-Supan, A., Brugiavini, A., Jürges, H., Mackenbach, J., Siegrist, J., and Weber, G. (2005). *Health, Aging and retirement in Europe: First results from the survey of health, ageing and retirement in Europe.* Mannheim Research Institute for the Economics of Aging (MEA) Mannheim.

Bound, J. (1989). Self-reported vs. objective measures of health in retirement models. Technical report, National Bureau of Economic Research.

Bozdogan, H. (1990). On the information-based measure of covariance complexity and its application to the evaluation of multivariate linear models. *Communications in Statistics-Theory and Methods*, 19(1):221–278.

Capecchi, S. and Piccolo, D. (2015). Dealing with large heterogeneity/uncertainty in sample survey with ordinal data. *IFCS Proceedings, University of Bologna.*

Corduas, M., Iannario, M., and Piccolo, D. (2009). A class of statistical models for evaluating services and performances. In *Statistical methods for the evaluation of educational services and quality of products*, pages 99–117. Springer.

D'Elia, A. (2000). Il meccanismo dei confronti appaiati nella modellistica per graduatorie: sviluppi statistici ed aspetti critici. *Quaderni di Statistica*, 2:173–203.

D'Elia, A. (2003). Modelling ranks using the inverse hypergeometric distribution. *Statistical Modelling*, 3(1):65–78.

D'Elia, A. and Piccolo, D. (2005). A mixture model for preferences data analysis. *Computational Statistics & Data Analysis*, 49(3):917–934.

Di Iorio, F. and Iannario, M. (2012). Residual diagnostics for interpreting cub models. *Statistica*, 72(2):163.

Fligner, M. A. and Verducci, J. S. (1993). *Probability models and statistical analyses for ranking data*, volume 80. Springer.

Gottard, A., Iannario, M., and Piccolo, D. (2016). Varying uncertainty in cub. *Advances in Data Analysis and Classification*, 10(2):225–244.

Holland, P. W. and Wainer, H. (2012). *Differential item functioning*. Routledge.

Iannario, M. (2008). A class of models for ordinal variables with covariates effects. *Quaderni di Statistica*, 10:53–72.

Iannario, M. (2009a). Fitting measures for ordinal data models. *Quaderni di statistica*, 11:46–79.

Iannario, M. (2009b). Selecting feeling covariates in rating surveys. *Italian Journal of Applied Statistics*, 20(1):21–1.

Iannario, M. (2010). On the identifiability of a mixture model for ordinal data. *Metron*, 68(1):87–94.

Iannario, M. (2012a). Hierarchical cub models for ordinal variables. *Communications in Statistics-Theory and Methods*, 41(16-17):3110–3125.

Iannario, M. (2012b). Modelling shelter choices in a class of mixture models for ordinal responses. *Statistical Methods & Applications*, 21(1):1–22.

Iannario, M. (2012c). Preliminary estimators for a mixture model of ordinal data. *Advances in Data Analysis and Classification*, 6(3):163–184.

Iannario, M. (2014). Modelling uncertainty and overdispersion in ordinal data. *Communications in Statistics-Theory and Methods*, 43(4):771–786.

Iannario, M. and Piccolo, D. (2009). A program in r for cub models inference. *Available via Internet, URL http://www. dipstat. unina. it/CUBmodels/, Version*, 2.

Iannario, M. and Piccolo, D. (2010). A new statistical model for the analysis of customer satisfaction. *Quality Technology & Quantitative Management*, 7(2):149–168.

Iannario, M. and Piccolo, D. (2011). University teaching and students' perception: Models of the evaluation process. In *Statistical Methods for the Evaluation of University Systems*, pages 93–112. Springer.

Iannario, M. and Piccolo, D. (2012a). Cub models: Statistical methods and empirical evidence. *Modern Analysis of Customer Surveys*, pages 231–258.

Iannario, M. and Piccolo, D. (2012b). A framework for modelling ordinal data in rating surveys. *Proc Joint Stat Meet, Section on Statistics in Marketing, San Diego, California*, pages 3308–3322.

Iannario, M. and Piccolo, D. (2014). Inference for CUB models: a program in R. *Stat & Appl*, 12:177–204.

Iannario, M. and Piccolo, D. (2015). A generalized framework for modelling ordinal data. *Statistical Methods & Applications*, pages 1–27.

Iannario, M., Piccolo, D., and Simone, R. (2016). *CUB: A Class of Mixture Models for Ordinal Data*. R package version 0.1.

Jöreskog, K. G. and Moustaki, I. (2001). Factor analysis of ordinal variables: A comparison of three approaches. *Multivariate Behavioral Research,*, 36(3):347–387.

Kokonendji, C. C. and Zocchi, S. S. (2010). Extensions of discrete triangular distributions and boundary bias in kernel estimation for discrete functions. *Statistics & probability letters*, 80(21):1655–1662.

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of psychology*.

Marden, J. I. (1996). *Analyzing and modeling rank data*. CRC Press.

McCullagh, P. (1980). Regression models for ordinal data. *Journal of the royal statistical society. Series B (Methodological)*, pages 109–142.

McCullagh, P. and Nelder, J. A. (1989). *Generalized linear models*, volume 37. CRC press.

McKelvey, R. D. and Zavoina, W. (1975). A statistical model for the analysis of ordinal level dependent variables. *Journal of mathematical sociology*, 4(1):103–120.

McLachlan, G. and Krishnan, T. (2007). *The EM algorithm and extensions*, volume 382. John Wiley & Sons.

McLachlan, G. and Peel, D. (2004). *Finite mixture models*. John Wiley & Sons.

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(2):337–357.

Paulhus, D. L. (1991). Measurement and control of response bias. *Measures of personality and social psychological attitudes*, pages 17–59.

Piccolo, D. (2003). On the moments of a mixture of uniform and shifted binomial random variables. *Quaderni di Statistica*, 5(1):85–104.

Piccolo, D. (2006). Observed information matrix for mub models. *Quaderni di Statistica*, 8:33–78.

Piccolo, D. and D'Elia, A. (2008). A new approach for modelling consumers' preferences. *Food Quality and Preference*, 19(3):247–259.

R Core Team (2016). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Schwarz, G. et al. (1978). Estimating the dimension of a model. *The annals of statistics*, 6(2):461–464.

Stevens, S. S. (1951). Mathematics, measurement, and psychophysics. *Handook of Experimental Psychology*.

Stevens, S. S. (1958). Problems and methods of psychophysics. *Psychological Bulletin*, 55(4):177.

Stevens, S. S. (1968). Measurement, statistics, and the schemapiric view. *Science*, 161(3844):849–856.