Electronic Journal of Applied Statistical Analysis
EJASA, Electron. J. App. Stat. Anal.
http://siba-ese.unisalento.it/index.php/ejasa/index
e-ISSN: 2070-5948
DOI: 10.1285/i20705948v8n3p360

Visualizing multiple results from Nonlinear CUB models with R grid Viewports
By Manisera, Zuccolotto

Published: 20 December 2015

Electronic Journal of Applied Statistical Analysis
Vol. 08, Issue 03, December 2015, 360-373
DOI: 10.1285/i20705948v8n3p360

Visualizing multiple results from Nonlinear CUB models with R grid Viewports

Marica Manisera*and Paola Zuccolotto

University of Brescia, Italy

Published: 20 December 2015

Nonlinear CUB models have been recently introduced in the literature to model rating or ordinal data. They extend the standard CUB (Combination of Uniform and Binomial), which is a mixture model combining a discrete Uniform and a Shifted Binomial random variables. Unlike CUB, Nonlinear CUB models account for the unequal spacing among response categories. Nonlinear CUB can be effectively used in a variety of fields, for example whenever questionnaires with questions having ordered response categories are used to measure human perceptions and attitudes. This paper proposes a new graphical representation, which works with R grid Viewports in order to summarize multiple results from Nonlinear CUB models in a unique plot. A case study on the perceived risk in fraud management is presented.

keywords: rating data, ordinal, feeling, uncertainty, transition probabilities, perceived risk.

1 Introduction

In recent years there has been a growing interest in the statistical models and techniques specifically intended for ordered categorical variables, which typically derive from questions with ordered response categories. For example, to measure customer satisfaction, it is common to ask customers to rate their satisfaction on a given response scale with ordered levels. Such kind of questionnaires are very often used in several fields, not only marketing, to investigate individuals' perceptions, attitudes, behaviors, cognitions and

*Corresponding author:marica.manisera@unibs.it

ISSN: 2070-5948
http://siba-ese.unisalento.it/index.php/ejasa/index

the resulting data, often referred as to rating data, are to be analyzed with appropriate tools.

Among the existing methods addressed to model rating data (for example, Agresti, 2013; Tutz, 2012), an original class of models, called CUB, has been proposed by Piccolo (2003) and D'Elia and Piccolo (2005). In the framework of CUB models, the individuals' decision process is interpreted as the combination of two latent components, called *feeling* and *uncertainty*, related to the model parameters. The feeling component measures the level of agreement with the item being evaluated while the uncertainty component expresses the human indecision that is present in any discrete choice. CUB models can be applied to rating data in any context: the applied statistical papers focus on a wide range of fields, from labour economics (Gambacorta and Iannario, 2013), to happiness economics (Capecchi and Piccolo, 2014), marketing (Iannario et al., 2012), medicine (D'Elia, 2008), sensometrics (Piccolo and D'Elia, 2008) and many others (see the references in Iannario and Piccolo, 2014).

Besides the huge amount of possible applications, several papers developed and extended the methodological aspects of the CUB models (Corduas et al., 2009; Gambacorta et al., 2014; Grilli et al., 2013; Iannario, 2012c,a,b, 2014; Iannario and Piccolo, 2010, 2012; Manisera and Zuccolotto, 2014a,b; Piccolo, 2006, 2014). Among these new developments, in this paper we focus on the Nonlinear CUB model (NLCUB; Manisera and Zuccolotto, 2014b) that has been proposed as a generalization of the standard CUB in order to model the possible "unequal spacing" among the response categories in the respondents' perception. Even when the possible m response categories in a rating scale are coded with m consecutive integers, it is not possible to ensure that the perceived distance between categories, say, 1 and 2 is the same as between 2 and 3: in the Nonlinear CUB philosophy, the respondents can find it more difficult moving from rating 1 to 2 than from rating 2 to 3. This unequal spacing has been formalized by means of the definition of transition probabilities, *i.e.* the probabilities of moving from one rating to the next one during the decision process. When transition probabilities are non-constant, the response categories are unequally spaced in the respondents' mind and the decision process is said to be nonlinear (Manisera and Zuccolotto, 2014b). Unlike standard CUB, NLCUB can be used to model rating data with non-constant transition probabilities.

Further research on NLCUB as well as real data analyses can be found in Manisera and Zuccolotto (2013, 2014b, 2015b,a,c) and the references therein. In addition, a free program has been developed in the software environment R for NLCUB estimation, graphical representation of a variety of outputs, fit evaluation, along with data simulation according to the NLCUB data generating process (Manisera and Zuccolotto, 2014c).

The aim of this paper is to propose a new graphical representation, which works with R grid Viewports (Murrell, 1999), able to summarize multiple results from NLCUB models in a unique plot. This kind of representation makes it easier to interpret the results from a NLCUB model and compare feeling and uncertainty over several items, or groups of people expressing evaluations on a single item of interest. The paper is organized as follows: Section 2 briefly recalls the main features of CUB and NLCUB models; Section 3 focuses on the NLCUB transition probabilities and their graphical representation, called transition plot (Subsection 3.1), which is one of the ingredients of the new graphical representation proposed in Section 4. Section 5 shows a case study on real data concerned with the fraud risk perceived by the management of a sample of Italian companies. Section 6 concludes the paper.

2 CUB and Nonlinear CUB models

The class of CUB models is built on the basic assumption that, when a subject is asked to express a rating about a given issue on an ordered response scale with m categories, his/her response derives from the combination of a *feeling* attitude towards the evaluated issue and an intrinsic *uncertainty* component surrounding the discrete choice. According to this logic, CUB models fit rating data by means of a mixture of two random variables, namely a Shifted Binomial $V(m, \xi)$ with trial parameter m and success probability $1-\xi$ - modelling the feeling component - and a discrete Uniform $U(m)$ defined over the support $\{1, \ldots, m\}$ - aimed to model the uncertainty component. The resulting random variable R generating the observed ratings r ($r = 1, \ldots, m$) has distribution probability given by

$$p_r(\theta) = \pi Pr(V(m, \xi) = r) + (1 - \pi) Pr(U(m) = r) \tag{1}$$

where $p_r(\theta) = Pr(R = r|\theta)$, $r = 1, \ldots, m$, $\theta = (\pi, \xi)'$, $\pi \in (0, 1]$, $\xi \in [0, 1]$. The parameters $1 - \xi$ and $1 - \pi$ are called *feeling parameter* and *uncertainty parameter*, respectively.

The NLCUB (Manisera and Zuccolotto, 2014a) is a new model in the CUB class, deriving from the attempt to take into account a possible unequal spacing of the rating categories in the respondents' mind. The discrete random variable R is assumed to depend on a new parameter T ($T \geq m-1$) and has the following probability distribution

$$p_r(\theta) = \pi \sum_{y \in l^{-1}(r)} Pr(V(T + 1, \xi) = y) + (1 - \pi) Pr(U(m) = r), \tag{2}$$

whose specific expression is determined by l, a function mapping from $\{1, \ldots, T + 1\}$ into $\{1, \ldots, m\}$. The function $l(\cdot)$ is defined as

$$l(y) = \begin{cases} 1 & \text{if} \quad y \in \{y_{11}, \ldots, y_{g_1 1}\} \\ 2 & \text{if} \quad y \in \{y_{12}, \ldots, y_{g_2 2}\} \\ \vdots & \vdots \quad \vdots \\ m & \text{if} \quad y \in \{y_{1m}, \ldots, y_{g_m m}\} \end{cases} . \tag{3}$$

In expression (3), y_{hs} is the h-th element of $l^{-1}(s)$, and

$$(y_{11}, \ldots, y_{g_1 1}, y_{12}, \ldots, y_{g_2 2}, \ldots, y_{1m}, \ldots, y_{g_m m}) = (1, \ldots, T + 1).$$

Let $g_s = |l^{-1}(s)|$, where $|\cdot|$ denotes the cardinality of a set, the function l is univocally determined by the values $\mathbf{g} = (g_1, \ldots, g_m)'$. In the end, for a given value $\mathbf{g}$, the probability distribution of R can be written as

$$p_r(\theta|\mathbf{g}) = \pi \sum_{i=g_0+\cdots+g_{r-1}}^{g_0+\cdots+g_r-1} \binom{T}{i} (1-\xi)^i \xi^{T-i} + \frac{1-\pi}{m} \tag{4}$$

where $g_0 := 0$ and $T = g_1 + \ldots + g_m - 1$.

The NLCUB model includes the standard CUB as a special case, originating with $T = m - 1$ and $g_s = 1$ for all $s = 1, \ldots, m$. Simulation studies and real data analyses with NLCUB can be found in Manisera and Zuccolotto (2013, 2014b), while some deeper theoretical investigations about the identifiability conditions and the use of the EM algorithm for parameter estimation are in Manisera and Zuccolotto (2015b,a).

3 The NLCUB transition probabilities

In Manisera and Zuccolotto (2014b) the idea supporting the definition of the NLCUB model derives from a specific assumption about the psychological mechanism generating the rating expressed by the respondent. Two different approaches, called feeling and uncertainty approach, coexist in the respondents' decision process and the final rating can derive from the feeling or the uncertainty approach with given probabilities. While the uncertainty approach consists of a random judgment, according to the feeling approach the search for the answer is ideally assumed to proceed trough T unconscious consecutive steps. During these steps the respondent makes a screening of all the positive and negative sensations (basic judgments) randomly coming to his/her mind. At each step he/she summarizes the current and the previous basic judgments and transforms them into a rating in the required scale (provisional ratings). The process goes on until all the emotions and the experiences surrounding the issue being evaluated are taken into account.

Thanks to this step-by-step formulation, denoted with R_t the random variable generating the provisional rating at step t, we can define the so-called transition probabilities $\phi_t(s) = Pr(R_{t+1} = s + 1|R_t = s)$, i.e. the probability of moving to provisional rating $s + 1$ at step $t + 1$, given that the provisional rating at step t is s, $s = 1, \ldots, m - 1$. Transition probabilities depend on the parameter ξ and the function l, and describe the respondents' state of mind about the response scale used to express judgments.

Manisera and Zuccolotto (2014b) have derived $\phi_t(s)$ for CUB and NLCUB models. In the CUB models, the transition probabilities are constant over t, s:

$$\phi_t(s) = 1 - \xi \quad \forall t, s \tag{5}$$

with $s = 1, \ldots, m - 1$, $t = 1, \ldots, m - 1$.

In the NLCUB models, the transition probabilities are given by

$$\phi_t(s) = (1-\xi)\frac{\binom{t}{w_{g_s s}}(1-\xi)^{w_{g_s s}}\xi^{t-w_{g_s s}}}{\sum_{h=1}^{g_s}\binom{t}{w_{hs}}(1-\xi)^{w_{hs}}\xi^{t-w_{hs}}} \tag{6}$$

where $w_{hs} = y_{hs} - 1$ and $s = 1, \ldots, m-1$, $w_{1s} \leq t < T$, and are non-constant over t and s, thus allowing to model psychological mechanisms where moving from rating s_1 to $s_1 + 1$ may be easier or harder than moving from s_2 to $s_2 + 1$.

Another parameter of interest is μ, the expected number of one-rating-point increments during the step-by-step reasoning, given by

$$\mu = \phi_0 + (1-\xi)\sum_{t=1}^{T-1}\sum_{s=1}^{m-1}\binom{t}{w_{g_s s}}(1-\xi)^{w_{g_s s}}\xi^{t-w_{g_s s}} \tag{7}$$

with $\mu \in [0, (m-1)]$. With CUB model expression (7) collapses to $\mu = (m-1)(1-\xi)$. With NLCUB models, μ is used as feeling parameter in place of $1-\xi$, because it allows the comparison among NLCUB models with different values **g** (Manisera and Zuccolotto, 2014b).

3.1 The transition plot: a graphical tool to show the perceived distance between consecutive rating categories

Starting from the transition probabilities (6), we can define a very insightful graphical representation, called transition plot, able to show the perceived distance of the rating categories in the respondents' mind.

Let $\phi(s) = av_t(\phi_t(s))$, where $av_t(\cdot)$ denotes averaging over t, be the average transition probability. It can be considered as a measure of the "perceived closeness" between ratings s and $s+1$ and can be transformed into a "perceived distance" $\delta_s = h(\phi(s))$ by means of a proper monotonically decreasing function h (usually $\delta_s = -\log(\phi(s))$).

The perceived distances are the basis for constructing the transition plot, where a broken line joins the points $(s, \tilde{\phi}(s-1))$, $s = 1, \ldots, m$, $\tilde{\phi}(0) = 0$, and $\tilde{\phi}(s-1) = (\delta_1 + \cdots + \delta_{s-1})/(\delta_1 + \cdots + \delta_{m-1})$ for $s = 2, \ldots, m$. In a transition plot the response scale is represented in the x-axis, while the corresponding perceived ratings are displayed in the y-axis. Figure 1 represents two examples of transition plot. The one on the left shows a case of response categories perceived as equally spaced, while the one on the right is an example of unequal spacing, where the ratings 1, 2 and 3 are relatively close each other, differently from ratings 4 and 5. This means that respondents, in their reasoning, find it more difficult moving, say, from rating 3 to 4 than from rating 2 to 3.

The traditional CUB model is characterized by "linear" transition plots, as in the left part of Figure 1, while "nonlinear" transition plots like that displayed in the right part of Figure 1 derive from a NLCUB model.

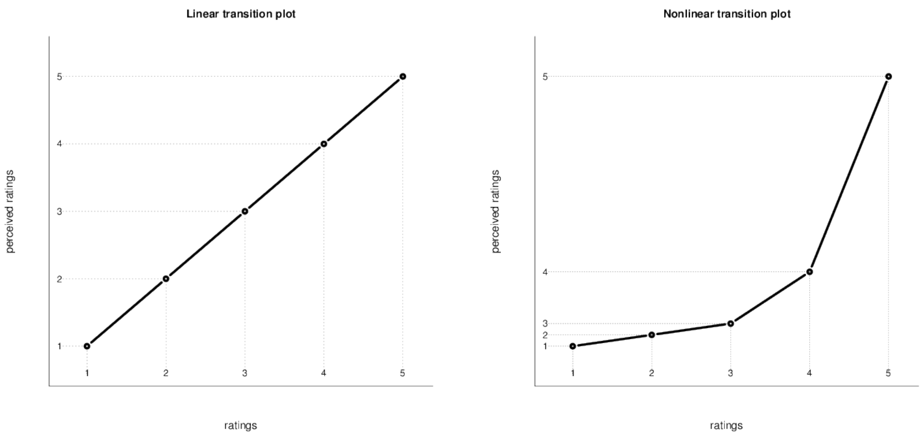

Figure 1: Examples of linear (left) and nonlinear (right) transition plot ($m = 5$)

4 A new graphical representation using the R grid Viewports

When a NLCUB is fitted to rating data, multiple results can be obtained and interpreted to gain insights into the topic under investigation. As detailed in Manisera and Zuccolotto (2014c), the R program `NLCUB` (freely available upon request to the Authors) provides several numerical outputs: estimates for feeling and uncertainty parameters and the **g** vector; the fitted frequencies; a goodness-of-fit index; the full transition probability matrix, containing the estimated $\phi_t(s)$ for all s, t; the estimated transition probabilities $\phi(s)$, computed averaging over t the probabilities in the transition probability matrix; the unconditioned transition probability, that is the estimate of the probability of increasing one rating point in one step of the decision process, independently both on the step and on the rating reached at the previous step; a measure of the nonlinearity of the decision process; the estimate of the uncertainty parameter adjusted for the presence of "don't know" responses, if any (Manisera and Zuccolotto, 2014a, 2015c). A list of several other results are supplied when the chosen estimation procedure is the EM algorithm instead of the Nelder-Mead numerical optimization procedure. Besides those numerical outputs, two plots are automatically provided: the observed vs. fitted frequencies and the transition plot (Subsection 3.1).

In this rich list of outputs, the measures of feeling and uncertainty and the transition probabilities are very important to interpret the phenomenon under study. The main purpose of this paper is to combine such information into a unique graphical representation, in order to obtain an immediate visualization of the results. In detail, two plots are combined:

- first, the plot representing the measures of feeling (μ) and uncertainty ($1 - \pi$) in the parameter square $[0, 1) \times [0, m - 1]$, with uncertainty on x-axis and feeling on y-axis. In this plot, the location of the point representing the NLCUB model has an immediate meaning in terms of feeling and uncertainty of the item being analyzed;

- secondly, the transition plot (Subsection 3.1), allowing a quick interpretation on how respondents perceive the response scale and, in particular, the spacing among the response categories.

The idea is to represent, for each item under study, the corresponding NLCUB in the parameter square not by a point, but by a a very small and stylized transition plot. The location of such stylized transition plots gives information about feeling and uncertainty, while the transition plot itself gives information about the nonconstantness of the transition probabilities and finally about the nonlinearity of the item.

As mentioned in Section 3, the transition probabilities result from the step-by-step reasoning and, therefore, are only related to the feeling part of the NLCUB model, while are not connected with the uncertainty latent component. When two items are closely located in the parameter square, it is interesting to see if their transition plots are similar or not: the same expected number μ of one-rating-point increments during the step-by-step reasoning can be associated with different perceptions of the response scale. On the contrary, similar shapes of the transition plot can be observed for two items differently positioned in the parameter square: this means that the spacing between response categories is perceived in the same way in the two items, but the final result in terms of feeling μ is different. For example, consider the case of two items both with linear transition plot but with a very different feeling μ: consecutive ratings are considered as equally spaced in both items, but to reach high ratings has to be interpreted as equally easy when the feeling is high or equally difficult when the feeling is low. This kind of interpretation can be directly obtained by the proposed graphical representation and makes comparisons among different items very easy and immediate.

From a computing point of view, in order to represent a small stylized transition plot in the square representing the parameter space, we use the grid Viewports developed in the R environment (Murrell, 1999). Viewports are very flexible tools that describe rectangular regions on a graphics device (in our case, the parameter square) and define a number of coordinate systems within those regions, where the stylized transition plots are located.

In the following, we provide an example code that users can employ to obtain the proposed graphical representation. Preliminarily, the data set must be read. It must be stored in a data frame, called `data`, like in the example, with n rows and p columns containing the ratings expressed by n subjects to p items of interest; in order to enhance the plot, we recommend assigning the items' names to the columns and create the object labels (`labels <- names(data)`). Also, the number of ordinal categories of the response scale must be stored in `cats` and the p-dimensional arrays `Mu` and `Pai` should be initialized; they will be filled up with the p estimates of μ and π, respectively.

The focus in the following example code is on the proposed diagram; it is worth noting that the part concerning the NLCUB estimation can be customized: the user can, for example, choose between two different optimization algorithms (Nelder-Mead optimization or the EM algorithm) or decide to give the frequency table instead of the vector of

n ratings as input in the `NLCUB` function (in this case, `freq.table` must be set to `TRUE`, which is the default option). All the details about the R program `NLCUB`, which is freely available, are in Manisera and Zuccolotto (2014c). For details on the usage of the R grid Viewport see Murrell (1999).

Example code

```
plot.new()
vp <- viewport(x=x1,y=y1,width=w1, height=h1,
just =c("left","bottom"))
pushViewport(vp)
grid.rect(gp=gpar(col=gray(0.6)))
vp <- viewport(x=x2,y=y2,width=w2, height=h2,
just = c("left","bottom"))
pushViewport(vp)
for (col in 1:ncol(data)){
source(NLCUB.R)
NLCUB(r=data[,col],m=cats,freq.table=FALSE)
Pai[col] <- pai
Mu[col] <- mu

vp <- viewport(x=(1-pai),y=mu/(cats-1),width=w3 height=h3,
just = c("centre","centre"))
pushViewport(vp)
stytrplot(tp=transprob,lab=labels,log.scale=TRUE)
upViewport()
}
par(new=TRUE, fig=gridFIG())
par(mar=c(0,0,0,0))
plot(1-Pai, Mu, xlim=c(0,1), ylim=c(0,(cats-1)),
type="n",xaxt="n",yaxt="n", bty="n")
axis(1,line=lin1)
axis(2,line=lin2)
upViewport()
upViewport()
par(new=TRUE, fig=gridFIG())
par(mar=c(0,0,0,0))
plot(1-Pai, Mu, xlim=c(0,1), ylim=c(0,(cats-1)),
type="n",xaxt="n",yaxt="n")
text(0.5,0,"Uncertainty",cex=1.4)
text(0,(cats-1)/2,"Feeling",srt=90,cex=c1)
text(0.5,(cats-1)-0.1,"Main title",cex=c2)
dev.off()
```

Several arguments can be set as needed. To obtain the diagrams in Figures 2 and 3 we set `x1`=`y1`=0.12, `w1`=`h1`=0.76, `x2`=`y2`=0.08, `w2`=`h2`=0.84, `w3`=`h3`=0.20, `lin1`=`lin2`=2.5, `c1`=1.4, `c2`=1.8 and changed the `"Main title"` appropriately.

The function `stytrplot` in the example code is aimed at representing the transition plot in its stylized form and has inputs given by the transition probabilities (tp), the items' labels (lab) and the `log.scale` logical flag, referring to the scale used to transform the transition probabilities into perceived distances to be plotted. In detail, the function `stytrplot` is defined as follows (the graphical parameters are set to the values used in the following Figures 2-3 and can be modified):

```
stytrplot <- function{tp,lab,log.scale=TRUE}{
m <- length(tp)+1
if (log.scale==TRUE){tp <- -log(tp)}
trp <- cumsum(rbind(0,as.matrix(tp)))/max(cumsum(rbind(0,as.matrix(tp))))
par(mar=c(0.9,0.9,0.9,0.9))
par(new=TRUE, fig=gridFIG())
plot(trp,type="o", cex=0.5,ylim=c(-0.1,max(trp)+0.1),lty=1,lwd=1.7,
xlim=c(0.7,m+0.2),xlab="",ylab="", main=labels[col],
xaxt="n",yaxt="n", cex.lab=0.8,bty="n",cex.main=1)
}
```

In both the example code and the stytrplot function, a number of graphical parameters can be customized, for example the size for text and symbols, by resorting to the usual graphical commands in R (for example, see the `par()` function or some well-known high-level plot functions, like `axis`, `title`, etc.).

5 Case study

In this section, we show how the new graphical representation proposed in Section 4 can be effectively used to interpret the results from a survey investigating the management's perceveid risk of being victim of frauds.

All around the world, companies are increasingly under attack from cyber-criminals, hacktivists and spies with the final aim to defraud companies. Nowadays, a huge number of computers, networks, wireless links, smartphones, and other devices connect billions of people to the Internet. Moreover, in troubled economic times, the risk of fraud increases making traditional protections not appropriate for meeting security needs. Fraud management aims at detecting and preventing fraudulent activity and applies measures to ensure security of data. These measures include, among many others, risk assessment. When making risk assessment, besides the objective risk level, it is also important to assess the management's attitude towards fraud, that is their perceived risk.

In this scenario, with the aim to assess the management's perceived risk, NetConsulting (Milano, Italy) carried out a survey in 2013 in Italy. The collected data involve $n = 116$ managers of small (24%), mid-sized (31%) and large (45%) companies.

We performed two separate analyses. In the first, we considered data about a general perception of fraud risk: the 116 managers were asked to express a judgment about (i) their personal degree of perceived fraud risk (Individual Perceived Risk, IPR) and (ii) the company's degree of perceived fraud risk (Business Perceived Risk, BPR), both measured on a 5-point Likert scale (from 1=very low to 5=very high). We fitted a NLCUB model to the resulting data and, using the code described in Section 4, we obtained Figure 2. It clearly shows that managers perceive a medium-high level of risk (μ equals 2.23 and 2.63 for IPR and BPR, respectively), with a slightly higher uncertainty when the Business risk is considered ($1 - \pi = 0.17$ for BPR vs. 0.02 for IPR). The stylized transition plot is linear for IPR (meaning that respondents considered the 5 response ratings as equally spaced and the NLCUB reduces to a standard CUB model) while is nonlinear s-shaped for BPR: respondents consider that moving from rating 2 to 3 is the most difficult transition in the response scale.

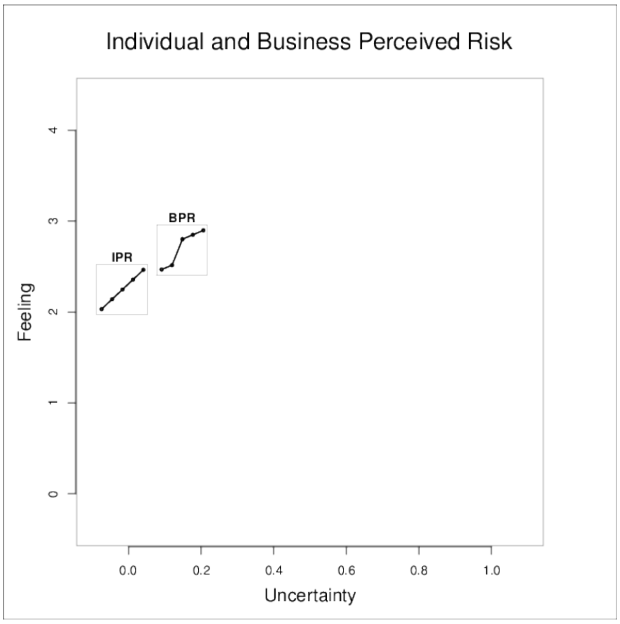

Figure 2: Individual Perceived Risk (IPR) and Business Perceived Risk (BPR): measures of uncertainty (x-axis) and feeling (y-axis), together with stylized transition plots.

To perform the second analysis, we used the data from a set of questions asking the 116 managers to rate, on a 4-point Likert scale, their perceived fraud risk when using some different Information and Communication Technologies: Web 2.0 and Social Networks, cloud storage and computing, Bring Your Own Device and Legacy technologies. Figure

3 shows that the perceived risk is higher when using Web 2.0 and Social Networks (SOCNET, $\mu = 2.20$) and Bring Your Own Device (BYOD, $\mu = 2.14$) technologies, intermediate when using cloud storage and computing (CLOUD, $\mu = 1.72$), while the lowest perceived risk refers to the use of the traditional legacy technologies (LEG, $\mu = 1.03$). All the perceptions are expressed with a low level of uncertainty: $1-\pi$ ranges from 0.01 (BYOD) to 0.19 (SOCNET). It seems that there is a direct relationship between the degree of perceived risk, on one hand, and the number of possible connections involved in the use of each ICT, on the other hand. At the same time, we are able to analyze in Figure 3 the four transition plots. All of them are nonlinear, with different shapes: s-shaped for CLOUD (the most difficult transition in the respondents' perceptions is between ratings 2 and 3), convex for SOCNET and LEG (proceeding from one rating to the next one is more and more difficult) and concave for BYOD (proceeding from one rating to the next one is easier and easier).

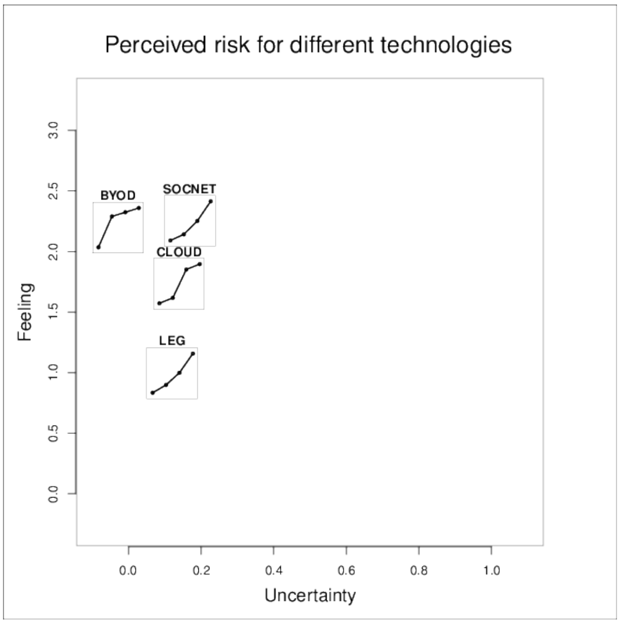

Figure 3: Perceived fraud risk when using Web 2.0 and Social Networks (SOCNET), cloud storage and computing (CLOUD), Bring Your Own Device (BYOD) and Legacy technologies (LEG): measures of uncertainty (x-axis) and feeling (y-axis), together with stylized transition plots.

Thanks to the proposed graphical representation, comparisons among items become more effective respect to the possible interpretation made relying only on parameter space and transition plots separately considered. Here, for example, we immediately see that respondents perceive the response scale of SOCNET and LEG in the same way; however, the difference in the feeling measure makes it possible to say that proceeding from rating 3 to rating 4 is the most difficult transition for both items, but it is more difficult to reach high ratings for LEG than SOCNET, because of the difference in the feeling of the two items. In addition, BYOD and SOCNET have comparable measures of feeling (with higher uncertainty for SOCNET) but different spacing of the response categories in the respondents' mind.

6 Conclusions

Nonlinear CUB models, recently introduced in the literature to model rating or ordinal data, extend the standard CUB because is able to account for the unequal spacing among response categories perceived by respondents. Among the various outputs from a NLCUB, we focus on two graphical representations: (i) the plot, in the parameter space, of the measures of feeling (agreement with the item) and uncertainty (indecision surrounding the discrete choice) and (ii) the transition plot, representing the perceived spacing of the rating categories in the respondents' mind.

Combining these two plots, we proposed an insightful graphical representation, which works with R grid Viewports, able to summarize multiple results from Nonlinear CUB models in a unique plot. The R code is given in detail, so that users can take advantage of this paper as a guide to effectively make this plot when applying NLCUB models.

Results from a survey investigating the management's perceived fraud risk show that the proposed plot is easy to make and interpret and gives interesting insights in the understanding the phenomenon under study.

Acknowledgement

The Authors wish to thank NetConsulting (Milano, Italy) for making the data available. This research was partially funded by STAR project (University of Naples Federico II - CUP: E68C13000020003) and partially by a grant from the Italian Ministry of Education, University and Research, MISURA: 'Multivariate statistical models for risks evaluation', PRIN project, prot. 2010RHAHPL_005.

References

Agresti, A. (2013). *Categorical Data Analysis.* Wiley, New York, 3^{rd} edition.

Capecchi, S. and Piccolo, D. (2014). Modelling the latent components of personal happiness. In Perna, C. and Sibillo, M., editors, *Mathematical and Statistical Methods for Actuarial Sciences and Finance*, pages 49–52. Springer.

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

D'Elia, A. (2008). A statistical modelling approach for the analysis of tmd chronic pain data. *Stat Methods Med Res*, 17:389–403.

D'Elia, A. and Piccolo, D. (2005). A mixture model for preference data analysis. *Comput Stat Data An*, 49:917–934.

Gambacorta, R. and Iannario, M. (2013). Measuring job satisfaction with CUB models. *Labour*, 27:198–224.

Gambacorta, R., Iannario, M., and Vallian, R. (2014). Design-based inference in a mixture model for ordinal variables for a two stage stratied design. *Aust NZ J Stat*, 56:125–143.

Grilli, L., Iannario, M., Piccolo, D., and Rampichini, C. (2013). Latent class CUB models. *Adv Data Anal Classif*, 8:105–119.

Iannario, M. (2012a). Hierarchical CUB models for ordinal variables. *Commun Stat-Theor M*, 41:3110–3125.

Iannario, M. (2012b). Modelling shelter choices in a class of mixture models for ordinal responses. *Stat Method Appl*, 20:1–22.

Iannario, M. (2012c). CUBE models for interpreting ordered categorical data with overdispersion. *Quad Stat*, 14:137–140.

Iannario, M. (2014). Modelling uncertainty and overdispersion in ordinal data. *Commun Stat-Theor M*, 43:771–786.

Iannario, M., Manisera, M., Piccolo, D., and Zuccolotto, P. (2012). Sensory analysis in the food industry as a tool for marketing decisions. *Adv Data Anal Classif*, 6:303–321.

Iannario, M. and Piccolo, D. (2010). A new statistical model for the analysis of customer satisfaction. *Qual Technol Quantit Manag*, 7:149–168.

Iannario, M. and Piccolo, D. (2012). CUB models: Statistical methods and empirical evidence. In Kenett, R. S. and Salini, S., editors, *Modern Analysis of Customer Surveys*, pages 231–258. Wiley, New York.

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

Manisera, M. and Zuccolotto, P. (2013). Nonlinear CUB models: some stylized facts. *QdS - J Methodol Appl Statist*, 1–2.

Manisera, M. and Zuccolotto, P. (2014a). Modeling "don't know" responses in rating scales. *Pattern Recogn Lett*, 45:226–234.

Manisera, M. and Zuccolotto, P. (2014b). Modeling rating data with Nonlinear CUB models. *Comput Stat Data An*, 78:100–118.

Manisera, M. and Zuccolotto, P. (2014c). Nonlinear CUB models: the R program. *Stat & Appl*, XII:205–223.

Manisera, M. and Zuccolotto, P. (2015a). Numerical optimization and EM algorithm in a mixture model for human perceptions analysis. *Working paper*.

Manisera, M. and Zuccolotto, P. (2015b). On the identifiability of Nonlinear CUB models. *J Multivar Anal*, 140:302–316.

Manisera, M. and Zuccolotto, P. (2015c). Treatment of 'don't know' responses in a mixture model for rating data. *Metron*, forthcoming.

Murrell, P. (1999). Layouts: A mechanism for arranging plots on a page. *J Comput Graph Stat*, 8:121–134.

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

Piccolo, D. (2006). Observed information matrix for MUB models. *Quad Stat*, 8:33–78.

Piccolo, D. (2014). Inferential issues on CUBE models with covariates. *Commun Stat-Theor M*, 43.

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

Tutz, G. (2012). *Regression for categorical data*. Cambridge University Press, Cambridge.