Quality of life among university students in Cagliari. A synthetic indicator

Isabella Sulis, Nicola Tedesco [1]
Dipartimento di Ricerche Economiche e Sociali
Università di Cagliari
Viale S. Ignazio 78, Cagliari
isulis@unica.it, tedesco@unica.it

Abstract: This article reports findings from a survey addressed to measure university students' quality of life in Cagliari. The aim to build up a synthetic indicator of students' 'quality of life at the university' has been pursued by adopting an *ad hoc* modeling approach to scale ordered items (*Item Response Models*) which belongs to the family of the *Generalized Linear and non Linear Mixed Models* . The sensibility of the results has been deeply analyzed by setting up several models with different characteristics. A comparison study with other scaling methods has been made.

Keywords: item response models, students' quality of life, mixed-effects models, synthetic indicator

1 Introduction

In the last years the italian university system faced a phase of overall reorganization. The transition from the old to the new formative system based on two levels (first level- three years - and second level - two years - degree) has represented a break in respect of the past. However, the debate on the redefinition of programme degrees and courses contents in line with the job-

[1] This paper is the result of a joint collaborative research between the two authors. In this joint work I Sulis was responsible for the final editing of §4, 5, 6, 8, 9 and N. Tedesco of sections §1, 2, 3, 7.

market pressures and on the establishment of new efficient administrative-managerial system is still long running. Within this debate the attitude of the institutions towards students' requests and the actions in favor of them have not been modified in meaningful way. In particular an aspect which has been marginally considered is the 'quality of life' of students during their permanence in the university system and its influence on their academic performances. In the last twenty years the efforts of the researchers have been mainly addressed to the effectiveness of university system in terms of 'output'. Specifically, researches have been focused on students' regularity in their curricula, with particular attention to the delay in finishing in time the university study and to the phenomenon of university drop-outs which has deeply characterized the italian university system (Biggeri and Bini, 2001; Porcu and Puggioni, 2003; Bini and Bertaccini, 2007; Chiandotto and Bacci, 2007). Moreover several studies investigated the transition from the university to the job-market (Balbi and Grassia, 2007; Porcu and Tedesco, 2007). A further relevant field of research on the efficiency of the university system focalized on the assessments of 'university quality' (Capursi and Porcu, 2001; Bernardi *et al.*, 2004; Rampichini and Grilli, 2004; Sulis, 2007).

From the other side, in the italian framework the concept of 'students' quality of university life' in terms of students' habits of life during the university studies has never been deeply explored and analyzed. Some attempts advanced in the last decade mainly aimed to measure the 'quality of university life' in terms of adequacy of university structures, infrastructures and facilities; other analyzed students' social environment and its influence on students' well-being than on students'style of life (Gatti and Mandich, 1994; Aureli and Grimaccia, 1999; Maggino and Schifini, 1999).

In this work a modeling approach for the assessment of 'students' quality of life' has been built up and tested on a sample of students' enrolled to three faculties of the University of Cagliari: Economics, Law and Political Sciences. These three faculties have been selected since their students are supposed to be enough homogenous in respect of several characteristics. They have similar formative curricula in terms of contents and teaching methodologies; they are located in the same area of the city and share contiguous buildings and common spaces. This means that they may be considered more similar under a qualitative and quantitative profile The two aims that will be pursed alongside this research are:

1. to determine the level of students' quality of life at the university moving from a bunch of *subjective* indicators;

2. to make a comparisons study between several scaling methods.

2 Measuring quality of life

In social statistics the measurement of the quality of life using a synthetic indicator has been long questioned and the national and international literature produced in the last decades several articles which faced the problem under different prospectives (Cox *et al.*, 1992; Fayers and Hand, 1997; Rampichini and Schifini, 1998; Aureli and Grimaccia, 1999; Fayers and Hand, 2002). In this framework, most of the methodological proposals have been mainly addressed to the identification of indicators apt to monitor from a descriptive point of view the phenomenon in its different aspects. Researchers agree in terms of defining the 'quality of life' as a 'latent' variable that can be measured throughout the use of *objective* and *subjective* indicators of well-being. The latter are addressed to reveal psychological and individual aspects (Larsen *et al.*, 1985; Huebner *et al.*, 2005). The underlying variable need to be firstly operationalized in terms of dimensions and next a set of indicator variables for each dimension are selected according to some rational criteria (Cox *et al.*, 1992; Fayers and Hand, 2002). The definition of the aspects relevant for operationalizing the latent trait is not straightforward and the choice of the components is influenced by many factors: availability of the information; the geographical context in which the university is located; cultural factors; economic conditions of the area etc.. There is a high level of arbitrariness in the phase of definition of many components of the phenomenon that are usually left to single researchers' choices and obviously depend on the framework analyzed. This implies a high level of arbitrariness in the process which strongly influence the final results.

From the other hand, the 'objective approach' moves from the identification of a set of variables supposed to be 'objective' indicators of the standard of life (level of income, the ownership of a house or of particular goods) and defines standard procedures of synthesis of the single components. In the 'subjective approach' the accent in to stress the level of well-being as perceived by single individual (independently from his/her objective standard of life) moving from individuals' responses to a set of indicator items. Mixed strategies use both 'subjective' and 'objective' indicators since they are more apt to highlight the phenomenon under different point of views (Maggino and Schifini, 1999; Aureli and Grimaccia, 1999; Shulz, 1999). One of more interesting aspect of the mixed strategies arises from the analysis of the correlation between the two dimensions (subjective and objective). It should be expected that the higher is the positive correlation between the two dimensions the more they overlap each other. A moderately high negative correlation signals the two indicators are monitoring a phenomenon which is perceived at individual level in a specular way. Whenever the intensity of the correlation suggests that the two components are independent each other both should be taken into account in order to have an overall measure of the 'quality of life'.

This work moves from a *subjective* prospective. Specifically we consider objective indicators (as for instance the *social-economic conditions*) more causal variables which could have influenced the ratings observed for the set of subjective indicator variables .

Therefore, variables in the questionnaire has to be carefully analyzed in order to identify which items are effectively indicators of the 'magnitude' of the latent trait (indicator variable) and which may have influenced the way people perceive it (causal variable). Students' 'quality of life' during the university study is thus considered a latent variable monitored by means of subjective indicator variables. The questionnaire is, in this framework, a tool for measuring the underlying attribute 'students' quality of life', that is operationally defined by its relationships with the indicator variables.

3 The survey

This survey considers as reference population 13893 students enrolled at the faculties of Economics, Law and Political Sciences in 2001/2002 a.y.. These three faculties contain 36.9% of students at the university of Cagliari at the moment the survey has taken place. A 'quota' sample has been carried out using a two stage sampling procedure. Specifically, at the first stage units has been stratified according to faculty (Economics,Law and Political Sciences) and a proportional sample has been selected. At the second stage units within each faculty has been stratified according to variable gender. Finally, within each of the six strata a stratified non proportional sample has been selected according to the 'residential status' :'resident', 'non resident', 'back forth traveler'. A constant number of student has been taken for each status. Students have been classified in 'resident' 'non resident' and 'back forth traveler' taking into account of two factors: how far their accommodation was from the university in terms of kilometers and minutes taken by the public transport service. It is not uncommon in the geographical area where the university of Cagliari is to find that the travel time to get to the university is much longer than it could be expected on the bases of the distance in terms of kilometers. The rule which has been adopted defines 'resident' all students who take less than 30 minutes from their houses to the university, 'back forth traveler' who takes between 30 and 60 minutes and 'no resident the others'.

We are aware this criterion does not clearly identified 'back forth' and 'non resident' students; however it has been considered as an enough 'objective' method to stratified the population. The choice of the optimal stratification criterion allowed to have all strata adequately represented in the sample (Table 3). A final sample size of 375 units has been obtained.

Table 3 depicts the number of students in each sub-samples. The rate of sampling is equal to 2,7% of the overall population: 43.2% belongs to Law, 35.2% to Economics and 21,6 to Political Sciences. The distribution of students conditionally upon their age and academic status ('first year student','regular student', 'no regular student') appears to be enough balanced between female and male . It is worth pointing out that student status and age have a specular distribution since the former can be considered a proxy of the latter. It is interesting to have a look at the rate of first year male students who have more than 26 years (2.8%). From Table 3 arises that male students have an average age slightly higher than female and the former usually start

Table 1: *Number of students according to gender, residential status and faculty*

Faculty	Residential status	Gender		Total	
		M	*F*	*a.v.*	*%*
	resident	20	24		
Economics	*back forth traveler*	20	24	162	**35.2**
	non resident	20	24		
	resident	19	35		
Law	*back forth traveler*	19	35	132	**43.2**
	non resident	19	35		
	resident	11	16		
Political Sciences	*back forth traveler*	11	16	81	**21.6**
	non resident	11	16		
Total	*a.v.*	150	225	**375**	
	%	**40.0**	**60.0**		**100.0**

Table 2: *Rate of students in the sample according to gender, academic status and age*

	Academic status			
Age	*first year*	*regular*	*no regular*	**Total**
		Male		
18 – 22	37.5	62.5	-	32.0
23 – 26	-	22.7	77.3	44.0
≥ 27	2.8	-	97.2	24.0
Total	**12.8**	**35.4**	**51.7**	**100.0**
		Female		
18 – 22	32.6	67.4	-	38.2
23 – 26	1.0	28.0	71.0	44.5
≥ 27	-	5.1	94.9	17.3
Total	**12.9**	**39.1**	**48.0**	**100.0**

the university with some year of delay in respect of the latter. However, the differences between genders appear to be irrelevant and of none influence on the results arose from the survey.

4 University students' quality of life in Cagliari: how to measure it?

In this study we move from the hypothesis that 'the quality of life' of a student is mainly determined by his/her style of life and by his/her level of integration in the academic and city environment. We suppose students who have the highest level of 'quality of life' take advantage of all the services of the university, are perfectly integrated in the city and enjoy their students' status by taking part to many external activities. Broadly speaking the indicator items selected provide information on students' habits in their daily life at the university.

The questionnaire used in the survey is structured in sections which provides information on several aspects: students' personal details and the social-economic status of his/her family; students'attitude to use the university facilities and students' style of life in Cagliari. Some of these indicators are applicable just to 'back forth' traveler and 'non resident' students. Other items are addressed to reveal directly or indirectly students' economic conditions (type of accommodation; how much he/ she spends for accommodation, food, etc; type of transport frequently used; if he/she owns a vehicle and type of vehicle; etc).

Items in the questionnaire are of different type: metrical, categorical, ordinal, counts. The main part is composed by questions measured on 'Likert-type' scale addressed to know how often students are involved in specific activities. This set of items are classified as 'subjective' indicators because they move from units' responses in order to define their level of 'quality of life' as students. It may be supposed that the way students' answer is influenced by a subjective component, mainly determined by cultural and social aspects, that need to be explicitly specified in order to sort out a measure of the unobservable attribute. According to the hypothesis followed alongside this research the higher is the number of aspects on which they assert to take part the more they are involved in the social and academic life.

In this phase we bounded the analysis to just those items which are applicable to all students. We selected fourteen 'subjective' indicators concerning how often students attend or use the following services, activities or facilities: *sport, cinema, theaters, cultural activities, bars, clubs, disco, reading, work, lectures, university services as libraries, sport facilities, teacher office hour, canteen*. Excluding 'work' all of them have a positive direction. In our hypothesis to have a 'part-time' o 'full-time' job means to have less time to be involved in the other academic and non academic activities. For this reason we changed the direction of this item before to plug it to the set of indicator variables chosen to measure students' attitude to take part to the academic and non academic activities (university and social life) and students' habits of life. At the same time we use as a control variables all items concerning students' *social-economic conditions* and *academic curricula*. The meaning of the 14 indicators is specified in Table 3.

The measurement of a latent variable throughout a synthetic indicator needs to carefully define the following phases of the process (Bernardi *et al.*, 2004):

1. the *indicator variables* which define the latent variable;
2. the *transformations* to apply to the indicator variables in order to scale themselves;
3. the *weighting scheme* to apply to the re-scaled indicator variables in order to discriminate the relevance of each of them;
4. the *merging function* to summarize the results in a single statement.

The goal is the measurement of *students' quality of life* on the bases of *students's ratings*. The task of building up a synthetic indicator of *university students' quality of life* will be pursued by moving from students' responses to the set of indicator variables. The work will focus on a proper treatment of ordinal variables and subject-specific effects.

Table 3: *Indicator items selected to measure students' quality of life*

item	**meaning**
lectures	Attendance at lecturs
library	Use of university library
cus	Use of the university sporting center
canteen	Use of the university refectory
meeting	To have meeting with lecturers during their office hours
work	full time, part time job or full time student
sport	Praticing a sportive activity
cinema	Attendance at cinema
theater	Attendance at theater
cultural	Attendance at cultural events and activities
bar	Attendance at bar
disco	Attendance at disco
reading	Reading non academic books
clubbing	Attendance at clubs

Table 4: *Students' ratings*

item		%	
	never	*sometimes*	*often*
cus	69.60	19.47	10.93
theater	77.07	19.73	3.20
canteen	66.67	21.33	12.00
sport	45.33	25.07	29.60
cultural	60.27	30.93	8.80
reading	33.87	32.80	33.33
work	44.80	35.20	20.00
disco	46.93	37.07	16.00
lectures	11.20	38.13	50.67
bar	18.13	40.80	41.07
meeting	50.93	42.67	6.40
clubbing	22.13	44.00	33.87
library	15.73	51.47	32.80
cinema	25.33	58.13	16.53

5 Choosing a transformation function to re-scale categorical items

Table 5 exhibits the observed rates for each category of the 14 ordinal indicators. It is interesting to stress the unexpected high rate of students who use the category 'never' for the items concerning the attendance at 'cultural events' (60.3%) and 'theater' (77.1%), ' meeting' (50.9%), 'sport' (45.3%) and 'reading' (33.9%).

As a first attempt to explore students' response pattern we dichotomized the indicators collapsing the grade of the scale into two categories 'yes' and 'no' according to two different rules: firstly classifying 'sometimes with yes', after 'sometimes with no'.

Table 5 points out the arbitrariness in the responses arisen postulating questions as binary. The value of the *Spearman* correlation coefficient ($\rho =$

-0.49) between the two rankings highlights the meaningless of using methods for binary data: 'theater' switches from the first to the twelfth rank, 'meeting' from the second to the thirteenth and so forth. Furthermore, the low number of grades (three) of the scale and their asymmetric distribution should suggest to avoid methods which handle ordered categories as metrical.

The present work adopts a modeling approach based on *Item Response Models* in order to assign a metric to the categorical items. *Item response theory approach* is a probabilistic approach for the development of scales which specifies the conditional distribution of the complete response pattern as a function of latent variables and explanatory variables. The aim is to find out the values of the parameters which characterize each item in order to locate them in a continuum on the basis of a 'non subjective' (non arbitrary) criteria. We suppose the intensity of the attribute measured throughout the set of items has a range from $-\infty$ to $+\infty$.

The results obtained by adopting *ad hoc* methods for scaling categorical ordered variables will be compared with more classic methods frequently adopted in literature.

6 Measuring the quality of life of students at Cagliari using Item Response Models

In *Item Response Theory* (*IRT*) the probability to score a category is modeled as a function of a *person parameter*, which measures the intensity at individual level of the attribute, and an 'item parameter', which measure the 'easiness' or the 'difficulty' of the aspect of the attribute measured throughout the item (Bartholomew and Knott, 1999; Moustaki *et al.*, 2004). This framework jointly analyzes individual's response pattern in order to sort out the intensity of the latent variable in each student and sheds some lights on the characteristics of the indicator variables selected.

In *Item Response Theory* the probability for subject p ($p = 1, \ldots, n$) to score category k of item i ($i = 1, \ldots, I$) is function (depending on the link

Table 5: *Rate of positive answers*

item	% rate			
	'often'	*'sometimes-never'*	*'often-sometimes'*	*'never'*
	No	**Yes**	**No**	**Yes**
theater	96.80	3.20	8.80	91.20
meeting	93.60	6.40	6.40	93.60
cultural	91.20	8.80	41.07	58.93
cus	89.07	10.93	10.93	89.07
mensa	88.00	12.00	12.00	88.00
disco	84.00	16.00	33.33	66.67
cinema	83.47	16.53	3.20	96.80
no work	80.00	20.00	29.60	70.40
sport	70.40	29.60	16.53	83.47
library	67.20	32.80	32.80	67.20
reading	66.67	33.33	33.87	66.13
clubbing	66.13	33.87	86.67	13.33
bar	58.93	41.07	16.00	84.00
lectures	49.33	50.67	50.67	49.33

function) of the *item parameters*[2] and *person parameter*; a further parameter, known as *discrimination parameter*, helps to differentiate across items with different discrimination power. *Person parameters* can provide an individual measurement of students' standard of life; *item parameters* can be used to identify a *re-scale* scheme for the set of ordered categorical items; a further parameter, *discrimination parameters* can be considered a kind of weighting scheme for the set of synthetic indicators.

By differentiating the *link* function the framework can be applied to deal with binary, categorical and ordered categorical variables. If a *cumulative logit* link is specified each of the $K-1$ cumulative *logit* expresses the ratio between the probabilities to score for item i category k (for $k=1,\dots,K$) or lower on the probability to score a higher category as function of an item specific threshold and the individual parameter θ_p

$$P(Y_{ip} \leq k|\theta) = \frac{\exp(\tau_{ik} - \alpha_i \theta_p)}{1 + \exp(\tau_{ik} - \alpha_i \theta_p)}. \tag{1}$$

Moving from equation 1 the probability of responding categories lower than k is modeled as follows:

$$\begin{aligned} P(Y_{ip} \leq 1|\theta) &= \frac{\exp(\tau_{i1} - \alpha_i \theta_p)}{1 + \exp(\tau_{i1} - \alpha_i \theta_p)} \\ P(Y_{ip} \leq 2|\theta) &= \frac{\exp(\tau_{i2} - \alpha_i \theta_p)}{1 + \exp(\tau_{i2} - \alpha_i \theta_p)}, \end{aligned} \tag{2}$$

and exact probability of responding categories 'never' (1) 'sometimes' (2) 'often' (3) is

$$\begin{aligned} P(Y_{ip} = 1) &= P(Y_{ip} \leq 1) \\ P(Y_{ip} = 2) &= P(Y_{ip} \leq 2) - P(Y_{ip} \leq 1) \\ P(Y_{ip} = 3) &= 1 - P(Y_{ip} \leq 2). \end{aligned} \tag{3}$$

[2]For polytomous items, we call *item parameters* also the *threshold parameters* which characterize the cut-point of each category (*category parameters*).

The model expresses by equation 1 is known in the psychometric literature and as *Graded Response Model*(Samejima, 1969) [3]. The model allows *threshold parameters* τ_{ik} to differ among items. For each category k, higher values of the thresholds imply greater probabilities of responding in categories lower rather than greater than k. For each category a ranking of the items can be made sorting them according the value of the threshold parameter. For instance, when $k = 1$, the lowest threshold parameter (τ_{i1}) is attached to the item with the greatest probability to score 'often or sometimes' rather than 'never'. Thresholds parameters are also known as cut-point on the logistic scale that mape the range of probability (0-1) onto $(-\infty, +\infty)$. Factor loadings α_i are constrained to be constant across categories.

The factor loading α_i describes the effect of the person parameter (which measures 'student's quality of life') on the cumulative probability of responding up of a category. If the *discrimination parameters* are constant, e.g. $\alpha_i = 1$ for $i = 1, \ldots, I$, all items discriminate in the same way across individuals of different *person parameters*. The negative sign on the discrimination parameter α_i indicates that as 'student's quality of university life' increases the response for the observed item is more likely to fall at the high end of the scale (Moustaki, 2003). The effect of the person and item parameter is additive: for any item i the higher the value of an individual on the latent trait (θ_p), the higher is the probability to score higher categories.

A more parsimonious model for ordered variables is the *Proportional Odds Models* with *threshold parameters* constant across items and an item parameter β_i which shift the cut points towards the low end of the scale

$$P(Y_{ip} \leq k|\theta) = \frac{\exp(\tau_k - \beta_i - \alpha_i\theta_p)}{1 + \exp(\tau_k - \beta_i - \alpha_i\theta_p)}. \tag{4}$$

The greater is the positive value of β_i the bigger is the probability to score high categories.

The measurement problem can be alternatively formulated within the framework of the *hierarchical models* where level-1 units are students' answers to each single item (repeated measurements on the same student) and students are the level-2 units. In doing this person parameters θ_p are spec-

[3]The random-effects version of a *Proportional Odds Model* (Agresti, 2002).

ified to be random effects which vary among subjects by following a $\theta_p \sim \mathcal{N}(0, \sigma_\theta^2)$ distribution (Bartholomew, 1998; Boeck and Wilson, 2004).

Bayes' theorem is used to get the posterior distribution of θ given the vector of observed responses $\boldsymbol{y}$

$$f(\theta|\mathbf{y}) \propto f(\theta) f(\mathbf{y}|\theta).$$

Moving from the posterior distribution, the individual parameter may be predicted on the scale of θ using as measures of location and accuracy of θ_p the $E(\theta|\mathbf{y})$ and its $Var(\theta|\mathbf{y})$.

7 A comparison study among models with different characteristics

In the following four one-dimensional models are set up and results compared in terms of goodness of fit, weighting schemes, ranking of the items according to their 'easiness' and interpretability. The characteristics of the four models are depicted in Table 6. Model 4 shows the best goodness-of-fit measure, followed in terms of AIC by Model 2, Model 3 and Model 1. The simplest model (M_1) load all items on the latent trait with the same weight and thresholds τ_k are specified to be constant across the items; model (M_2) differs from model (M_1) since allows factor loadings to vary and fixes the factor loading of item *reading* equal to 1; model (M_3) has thresholds which vary across items whereas factor loadings are fixed equal to 1; the most complex models M_4 leaves free both *threshold parameters* and factor loadings.

The four models specify person parameters as random-effects which follow a normal distribution and assume one latent factor. The estimates and their standard errors for models with constant thresholds (M_1 and M_2) are shown in Table 7 and 7. Both models agree in indicating exactly the same ranking of the items in terms of 'how easy is to score higher categories'. The 'easiness' of the item are measured throughout the *item parameter* β_i. The value of 'reading' has been set to 0 in order to estimate freely the others *threshold* and *item parameters* and to collocate them in a scale $(-\infty, +\infty)$. 'Reading' has been selected as reference item since it shows the most heterogeneous rate of responses across the three categories (0.34,0.33,0.33). From

Table 6: *Comparisons between Graded Response Models in terms of AIC*

M	*L-1 units*	*L-2 units*	$Var(\theta)$	τ^*_{ik}	α_i	*n° param.*	*AIC*
M_1	5250	375	.62 (.07)	$=$	$=1$	16	9880.9
M_2	5250	375	1.34 (.34)	$=$	$\neq 1$	29	9652.4
M_3	5250	375	.65(.07)	$\neq$	$=1$	29	9733.0
M_4	5250	375	.42 (.15)	$\neq$	$\neq 1$	42	9381.5

* *Threshold parameters* are constant across items $=$ or differ among them $\neq$

both table arise that 'attendance at lecture' is the easiest item, followed by 'no work activity' and ' clubbing'. The most difficult are the items at the bottom of the ranking: 'the use of university canteen', 'the use of sportive facilities' and ' attendance at theater'. The value of *item parameters* agree with the results arose from the descriptive analysis where items 'attendance at theater' and 'attendance at lecture' exhibits respectively the lowest and the highest rate of responses in category *never*.

Fixing the load of 'reading' equal to 1 and leaving the others free to vary arises that the three aspects 'attendance at theater' (0.99), 'the use of university sport facilities'(0.89) and 'attendance at disco' (0.90) have a discrimination power close to 1, whereas items 'sport' (1.97), 'bar' (1.54) and 'clubbing' (1.37) discriminate more between subjects with different level of 'quality of life'. The lowest factor loadings are attached to aspects strictly linked to academic activities:'use of university canteen' (-0.04), 'no work'(0.006), 'attendance at lecture' (0.126) and 'use of the university library'(0.25).

The analysis continues by comparing results of M_3 and M_4. The estimated cut-points for both models are given in Table 7 and 7. The *proportional odds model* with different *threshold parameters* across the items makes the ranking of the aspects in terms of 'easiness' not unique since they can be sorted according to the values of the cut-points on the first or on the second category. To make easier the comparison of the rankings (according to their level of easiness obtained) sorted out using the four modeling approaches we evaluate the value of the *Spearman correlation coefficient*.

Table 7: *One dimension Graded Response Model* M_1

items	*thres. par.*		*item par.*	*discr. par.*
	τ_1	τ_2	β_i	α_i
lectures	-.977(.110)	1.061 (.111)	1.123 (0.145)	1
no work	-.977(.110)	1.061 (.111)	0.742 (0.144)	1
bar	-.977(.110)	1.061 (.111)	0.637 (0.141)	1
library	-.977(.110)	1.061 (.111)	0.484 (0.140)	1
clubbing	-.977(.110)	1.061 (.111)	0.349 (0.140)	1
reading	-.977(.110)	1.061 (.111)	0.000 (0.000)	1
cinema	-.977(.110)	1.061 (.111)	−0.206 (0.138)	1
sport	-.977(.110)	1.061 (.111)	−0.435 (0.145)	1
disco	-.977(.110)	1.061 (.111)	−0.839 (0.144)	1
meeting	-.977(.110)	1.061 (.111)	−1.177 (0.144)	1
cultural	-.977(.110)	1.061 (.111)	−1.475 (0.148)	1
canteen	-.977(.110)	1.061 (.111)	−1.673 (0.154)	1
cus	-.977(.110)	1.061 (.111)	−1.817 (0.156)	1
theater	-.977(.110)	1.061 (.111)	−2.366 (0.164)	1

Table 8: *One dimension Graded Response Model M_2*

items	*thres. par.*		*item par.*	*discr. par.*
	τ_1	τ_2	β_i	α_i
lectures	-1.025 (.123)	1.136 (.123)	1.140 (.154)	.126 (.102)
no work	-1.025 (.123)	1.136 (.123)	.766 (.157)	.006(.104)
bar	-1.025 (.123)	1.136 (.123)	.750 (.154)	1.549 (.253)
library	-1.025 (.123)	1.136 (.123)	.499 (.149)	.251 (.099)
clubbing	-1.025 (.123)	1.136 (.123)	.390 (.149)	1.370 (.231)
reading	-1.025 (.123)	1.136 (.123)	.000 (.000)	1.000 (.000)
cinema	-1.025 (.123)	1.136 (.123)	-.196 (.142)	.635 (.121)
sport	-1.025 (.123)	1.136 (.123)	-.642 (.177)	1.971 (.327)
disco	-1.025 (.123)	1.136 (.123)	-.889 (.150)	.905 (.171)
meeting	-1.025 (.123)	1.136 (.123)	-1.138 (.151)	.276 (.102)
cultural	-1.025 (.123)	1.136 (.123)	-1.518 (.156)	.754 (.155)
canteen	-1.025 (.123)	1.136 (.123)	-1.617 (.165)	-.046 (.110)
cus	-1.025 (.123)	1.136 (.123)	-1.915 (.171)	.889 (.201)
theater	-1.025 (.123)	1.136 (.123)	-2.574 (.198)	.995 (.206)

Table 9: *One dimension Graded Response Model M_3*

items	*thres. par.*		*discr. par.*
	$\tau_{i(1)}$	$\tau_{i(2)}$	α_i
lectures	- 2.26 (.175)	-0.02 (.117)	1
library	-1.866 (.155)	.804 (.123)	1
bar	-1.715 (.147)	.475 (.118)	1
no work	-1.510 (.142)	.220 (.118)	1
clubbing	-1.446 (.138)	.796 (.122)	1
cinema	-1.211 (.133)	1.840 (.150)	1
reading	-.742 (.123)	.834 (.123)	1
sport	-.205 (.118)	1.019 (.126)	1
disco	-.119 (.118)	1.861 (.152)	1
meeting	.058 .(118)	2.919 (.219)	1
cultural	.491 (.120)	2.595 (.192)	1
canteen	.803(.123)	2.162 (.169)	1
cus	.944 (.126)	2.314 (.175)	1
theater	1.377 (.136)	3.702 (.300)	1

Looking at the first cut-point a level of agreement equal to 0.95 is detected between the ranking provided by M_3 and M_4, 0.94 between (M_1,M_2) and M_3 and 0.97 between (M_1,M_2) and M_4. On the second cut point the ranking provided by M_3 still shows a good agreement ($\rho = 0.95$) with the results arisen under (M_1,M_2). The agreement with M_4 is sensible weaker ($\rho = 0.73$). The model to scale the set of item has been chosen also by considering the uncertainty related to its estimates. The high values of the standard errors for several *item* and *discrimination parameters* which characterize M_4 signal that their values are poorly determined. This make the estimates of these parameters strongly unreliable (Barholomew *et al.*, 2002). On the basis of the uncertainty of its estimates, we preferred to define the latent variable scaling the items using the parameters provide by M_2 [4].

Person parameters provide estimates of students' position on the latent trait 'quality of life'. It is interesting to stress the high values of the *Pearson correlation coefficient* between the expected values of person parameters estimated using the three models (all pairs of indexes show values greater than 0.90). Since the three models provide similar rankings of the students we selected the one which shows the best goodness of fit and smaller variability of the expected values of the person parameters. The estimates of the fix and random parameters provided by Table 7 are used in order to get the posterior estimates of the person that will be used in further analysis as indicators of 'students' quality of life' analysis. Figure 1 shows the estimate of the person parameter for each student and its 95% confidence interval ordered according to the expected value from the lowest to the highest. Students are judged to have significantly different parameters of quality of university life if and only if their respective intervals do not overlap (Goldstein and Spiegelhalter, 1996).

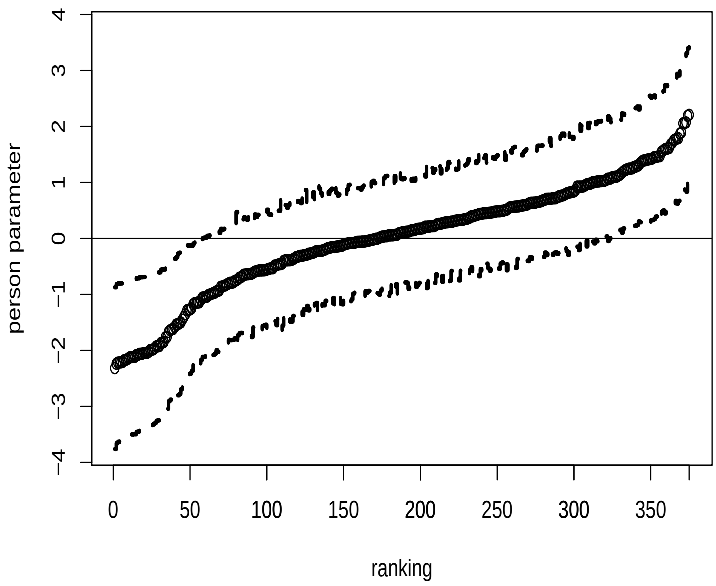
Figure 1 does not allow to make a clear ranking across students but highlights the existence of clusters of students which are characterized by different levels of 'quality of university life'. Three main groups can be detected: the first is composed by students who have the overall confidence interval

[4] Barholomew *et al.* (2002) recommend when standard errors are fairly large in relation to the difference in estimates to be aware against placing undue weight on small inequalities among the loadings.

Table 10: *One dimension Graded Response Model M_4*

items	*thres. par.*		*discr. par.*
	$\tau_{i(1)}$	$\tau_{i(2)}$	α_i
bar	−7.466 (2.346)	2.135 (0.900)	12.870 (4.621)
clubbing	−3.658 (0.570)	2.139 (0.532)	7.187 (1.686)
lectures	−2.073 (0.164)	−0.026 (0.104)	.144 (.169)
library	−1.700 (0.144)	0.725 (0.112)	.379 (.182)
no work	−1.387 (0.129)	0.209 (0.104)	.078 (.161)
cinema	−1.257 (0.152)	1.878 (0.170)	1.442 (.315)
reading	−0.708 (0.125)	0.778 (0.125)	1.000 (fixed)
sport	−0.210 (0.143)	1.078 (0.154)	1.689 (.364)
disco	−0.133 (0.143)	2.010 (0.186)	1.703 (.377)
meeting	0.036 (0.106)	2.707 (0.213)	.386 (.182)
cultural	0.461 (0.123)	2.509 (0.200)	1.043 (.268)
canteen	0.694 (0.110)	1.996 (0.159)	-.181 (.180)
cus	0.860 (0.120)	2.159 (0.173)	.633 (.224)
theater	1.437 (0.167)	3.804 (0.329)	1.455 (.357)

Figure 1: *Expected value of person parameter 'student's quality of life' and pairwise 95% overlap intervals*



under the 0; the second highlights a fairly large cluster whose confidence intervals cross the 0; finally a third group shows confidence intervals which lie completely in the range of positive values.

8 A comparison with more 'classic' scaling methods for ordered variables

In the following the ranking of the individual measures of 'quality of students' life' obtained by using the specific modeling approach for categorical ordered data is compared with the values that we would observed by adopting two 'classic' scaling methods. In this attempt to make a comparison we are not considering the uncertainty on the final score associated to the mean value of the person parameters. The first method assigns numbers in arithmetic progression to contiguous modalities, making implicitly the strong assumption of constant distance between adjacent categories. In our analysis each of the three categories of responses 'never', 'sometimes', 'often' are replaced with numerical values '1','2','3'. The set of items are merged in a synthetic indicator by adopting a linear function with equal weights (all equal to 1). The final score has been divided with its maximum: the maximum value it could assume under the assumption that all responses were equal to 3 (max. 42). In the forth we will refer to this method as M_5. The second method supposes that each ordinal item i is generated by a latent continuous variable z^*. The observed ordinal variable is linked to the latent z^* (Torgerson, 1958; Jöreskog, 2002) through the following relationship:

$$Y_{pi} = k \Leftrightarrow \gamma_{K-1} < z^* \leq \gamma_K, \quad k\text{=}1,\ldots,\, K \tag{5}$$

where Y_{ip} is the rating given by unit p to item i, $k = 1,\ldots,K$ is the number of categories and γ are called thresholds parameters

$$-\infty \leq \gamma_0 \leq \gamma_1 \leq \gamma_2 \leq \cdots \leq \gamma_{K-1} \leq \gamma_K \leq +\infty.$$

The distribution of the underlying latent variable is supposed to be standard normal, with density function $\phi(u)$ and distribution function $\Phi(u)$. Thus, the probability to score a category k is given by

$$\pi_k = Pr[\gamma_{k-1} < z^* < \gamma_k] = \int_{k-1}^{k} \phi(u)du = \Phi(\gamma_k) - \Phi(\gamma_{k-1}), \qquad (6)$$

and any threshold parameter γ_k is determined using the following relationship

$$\gamma_k = \Phi^{-1}(\pi_1, \pi_2, \ldots, \pi_k). \qquad (7)$$

In this way, the underlying variable assigns a metric to the ordinal categories. The percentage of responses in category k is used in order to sort out consistent estimates of π_k, for $k=1, \ldots, K$. By using relation (7) the lower and upper bound of each category are consistently estimated. The final score assigned to each ordinal category corresponds to the median value of $\hat{\gamma}$ between the two extremes $\hat{\gamma}_{k-1}$, $\hat{\gamma}_k$ (Table 11).

The set of 14 items, scaled using estimates provided by Table 11 have been summarized in a single indicator (M_6) using a linear function with weights set equal to 1 and the final score has been divided by its maximum. The position indexes of the three indicators are depicted in Table 8, instead *Pearson Correlation Coefficient* between pairs of them are shown in Table 8. From the matrix of correlation arises a high level of agreement between the values calculated by using the three different scaling methods (the value of ρ within pairs of indicators is always greater than 0.90). The maximum value of the coefficient is observed between M_5 and M_6. From Figure 2, arises the existence of some clusters of units that make the difference between the synthetic indicators obtained by using 'classic' scaling methods and M_2.

In an attempt to better understand the reasons of these differences, the pattern of responses of some 'anomalous units'[5] has been deeply investigated. From the exploratory analysis of these patterns arises some interesting features that motivates the use of such a modeling approach. We examined several subject profiles, examples of which are given in Table 13. From Figure 3, we selected some of the units whose position in the scale is sensible to the choice of the scaling method (obviously taking into account that M_2 and M_5 have different range). Specifically the cluster (a) of those units which

[5]Units which show a remarkable different behavior adopting different scaling methods.

Table 11: *Estimates of threshold parameters*

item	$\hat{\gamma}_{1.Med}$	$\hat{\gamma}_{2.Med}$	$\hat{\gamma}_{3.Med}$
	never	*sometimes*	*often*
lectures	-1.589	-0.517	0.664
library	-1.414	-0.215	0.978
cus	-0.391	0.818	1.601
canteen	-0.431	0.750	1.555
meeting	-0.660	0.591	1.852
no work	-1.282	-0.316	0.759
sport	-0.750	0.198	1.045
cinema	-1.142	0.111	1.387
theater	-0.292	1.123	2.144
cultural events	-0.521	0.698	1.706
bar	-1.337	-0.291	0.823
disco	-0.724	0.398	1.405
reading	-0.957	0.007	0.967
clubbing	-1.223	-0.148	0.957

Table 12: *Position Indexes for the three indicators M_2,M_5, M_6*

Position Index	M_2	M_5	M_6
Min.	-2.315	0.3333	-0.712
1^st Q	0.574	0.5476	-0.153
Median	0.108	0.6190	0.041
Mean	0.000	0.6135	0.012
3^rd Q	0.698	0.6905	0.209
Max.	2.208	0.8333	0.619

Table 13: *Pearson Correlation Coefficient between pairs of synthetic indicators*

Methods	M_2	M_5	M_6
M_3	1.000	0.923	0.901
M_5	0.923	1.000	0.973
M_6	0.902	0.973	1.000

show a better position when scaling method M_5 is adopted: 134, 126, 188, 348 and 324.

The response patterns in Table 13 show that the higher positions observed for M_5 in respect of M_2 are mainly determined by responses provided to items 'no work','canteen','library' and 'lecture'. These items have an extremely low discrimination power (0.006, -0.046, 0.251, 0.126) and their overall influence in determining the final factor score of the synthetic indicator is marginal. The same consideration can be argued for observations 308 and 140, which score high categories (sometimes and never) for the same set of items. The higher position of M_2 observed for these two units is mainly due to responses provided to item 'cus' (0.889) and 'sport' (unit 140) and 'cinema'(unit 308) that greatly influence the final score. In particular, the high factor loading of 'sport' pushed unit 140 in a better position in respect of unit 308 on the scale of M_2.

Figure 2: *Scatter plot between pairs of synthetic indicators*

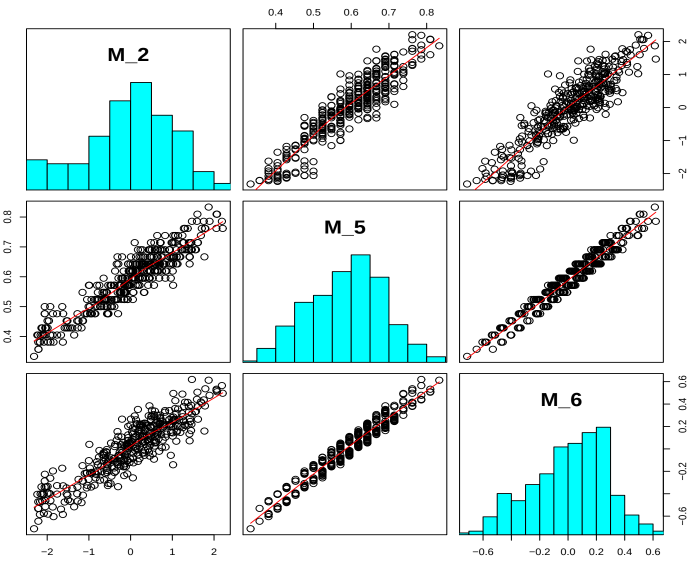
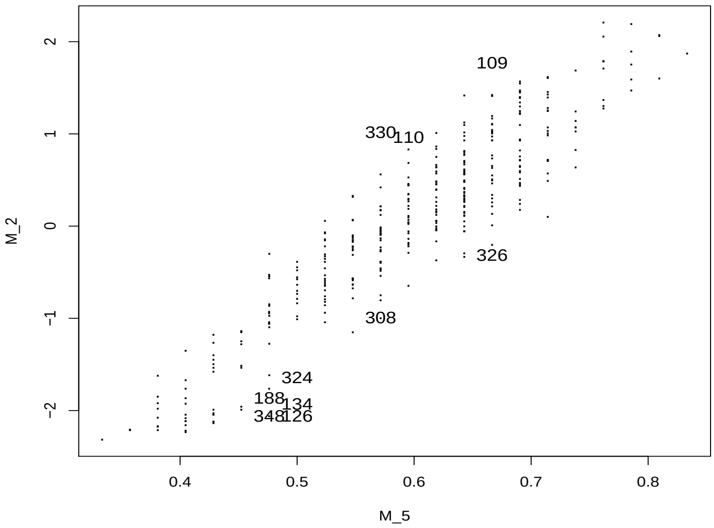


Figure 3: *Scatter plot between M_5 and M_2*



Other differences are detectable having a look at two response profiles which provides exactly the same score under method M_5 (b): the difference between the response profiles of unit 109 and 326 are strongly highlighted by M_2 which assumes a value equal to 1.77 (max 2.20) for unit 109 and equal to -0.315 for unit 326 . Subject 109 is involved in activities which have greater discrimination power ('sport', 'bar', 'clubbing', 'reading', 'theater') in respect of subject 326. This characteristic arises using the scale method M_2 which attaches higher levels of 'quality of life' to unit 109, instead method M_5 does not allow to distinguish the items, and thus the response pattern of the two subjects, in terms of 'quality of life' from a qualitative point of view. The last five units at the right end of Table 13 (c) show better positions on the scale of the synthetic indicator M_2 than M_5. Units 330 and 110 provide most of the responses in the right end of the scale for items with particular high factor loadings: 'sport' (1.971), 'club' (1.370), 'bar' (1.594),'disco' (0.905). The higher value of unit 109 in respect of 330 and 110 is determined by responses provided to item 'theater' and 'cinema'.

Table 14: *Response pattern of 'anomalous' units in respect of the value of M_2 and M_6*

item	**a** *b. p.*[1] *under* M_5							**b** *e. p.*[2] *under* M_5		**c** *b. p. under* M_2		
	134	*126*	*188*	*348*	*324*	*308*	*140*	*109*	*326*	*330*	*110*	*109*
lectures	3	3	3	3	3	3	3	1	3	1	1	1
library	3	3	3	3	3	2	3	1	3	1	2	1
cus	1	1	1	1	1	2	3	2	3	1	1	2
canteen	2	3	1	3	3	3	3	1	3	1	1	1
meeting	2	1	3	1	1	2	1	1	3	1	1	1
no work	2	2	3	3	3	3	3	1	1	1	1	1
sport	1	1	1	1	1	1	3	3	1	3	3	3
cinema	1	1	1	1	1	3	1	3	2	1	1	3
theater	1	1	1	1	1	1	1	2	1	1	1	2
cultural	1	1	1	1	1	1	1	2	1	1	3	2
bar	1	1	1	1	1	1	1	3	2	3	3	3
disco	1	1	1	1	1	1	1	2	1	3	3	2
reading	1	1	1	1	2	2	1	3	2	3	1	3
clubbing	1	1	1	1	1	1	1	3	2	3	3	3

[1]b.p=better performance

[2]e.p=equali performance

9 Conclusion

In this work a modeling approach for the assessment of ‘students’ quality of life’ has been built up and tested on a sample of students’ enrolled to three faculties of the University of Cagliari: Economics, Law and Political Sciences. Results arose from the use of different scaling methods based on *Item Response Models* highlight the aspects which allow to differentiate more across subjects who have different habits of life. The method attaches the greatest discrimination power to all the activities not directly linked to the university life; specifically to attend frequently pubs, clubs, cultural events, theaters and sporting centers is what make the difference between those students who are just involved in their academic studies and those who are perfectly integrated in the city, enjoy their students’ status by taking part to many external activities and try to take the greatest advantage from the city environment. An interesting point which has not been faced in this work concerns the association between the ‘students’ quality of life’ and students’ academic success; unfortunately, the unavailability of information did not allow us to carry on further investigation on this aspect.

Another issue left unexplored is the dimensionality of the set of indicator variables. The determination of the number of the factors underling the latent trait ‘students’ quality of life’ is a delicate issue that we overcame by allowing the 14 indicators to load on one factor. Further researches on this topic should deeply explore the behavior of two clusters of indicators which define a two-dimensional latent trait.

The main advantage of this methodological approach is that none assumption is made on the distances between adjacent categories and the ordinal scale of the items is specifically taken into account in the estimation of threshold parameters. Nevertheless, leaving the factor loadings free to vary across the items implicitly provides a weight schemes for the set of indicators and allow to overcome the assumption that all items load in the same way on the latent trait.

Comparisons between this methods and two ‘classic’ scaling methods frequently adopted in the pass to score ordered variables reveal that even if there is a high level of agrement between rankings sorted out using the three approaches, the latter poorly discriminate between subjects who are more or less involved in non academic activities.

By means of these comparisons we showed as these models turn out to be useful research tools in the phase of definition of different components of a composite indicators. Clearly, more investigations and more information are necessary to assess the potentiality and the pitfalls of these scaling methods in the context of ‘quality of life’ of students.

References

Agresti A. (2002) *Categorical Data Analysis*, Wiley-Intersciencie, Hoboken, New Jersey.

Aureli E. and Grimaccia E. (1999) Un percorso metodologico per lo studio della qualità della vita universitaria degli studenti, in: *Contesti di Qualità della vita. Problemi e misure*, Franco Angeli.

Balbi S. and Grassia M.G. (2007) Profiling and labour market accessibility for the graduates in economics at naples university, in: *Effectiveness of University Education in Italy*, Fabbris L., ed., Physica-Verlag.

Barholomew D., Steele F. and I. Moustaki J.G. (2002) *The Analysis and Interpretation of Multivariate Analysis for Social Scientists*, Chapman & All.

Bartholomew D.J. (1998) Scaling unobservable constructs in social science, *Applied Statistics*, 47, 1–13.

Bartholomew D.J. and Knott M. (1999) *Latent Variable Models and Factor Analysis*, Kendall.

Bernardi L., Capursi V. and Librizzi L. (2004) Measurement awareness: the use of indicators between expectations and opportunities, in: *Atti XLII Convegno della Società Italiana di Statistica. Bari, 9-11 Giugno 2004*, Società Italiana di Statistica.

Biggeri A. and Bini M. (2001) Evaluation at university and state level in italy: need for a system of evaluation and indicators, *Tertiary Education and Management*, 7, 149–162.

Bini M. and Bertaccini B. (2007) Evaluating the university educational process. a robust approach to the drop-out problem, in: *Effectiveness of University Education in Italy*, Fabbris L., ed., Physica-Verlag.

Boeck P.D. and Wilson M. (2004) *Item Response Models: a Generalized Linear and non Linear Approach*, New York: Springer.

Capursi V. and Porcu M. (2001) La didattica universitaria valutata dagli studenti: un indicatore basato su misure di distanza fra distribuzioni di giudizi, in: *Atti Convegno Intermedio della Società Italiana di Statistica 'Processi e Metodi Statistici di Valuatzione', Roma 4-6 giugno 2001*, Società Italiana di Statistica.

Chiandotto B. and Bacci S. (2007) Measurement of university external effectiveness based on the use of acquired skills, in: *Effectiveness of University Education in Italy*, Fabbris L., ed., Physica-Verlag.

Cox D., Fitzpatrick R., Fletcher A.E., Gore S.M., Spiegelhalter D.J. and Jones D.R. (1992) Quality of life assessment: can we keep it simple?, *Journal of the Royal Statistical Society, A*, 155, 353–393.

Fayers P.M. and Hand D.J. (1997) Factor analysis, casual indicators and quality of life, *Quality of Life Research*, 6, 139–150.

Fayers P.M. and Hand D.J. (2002) Casual variables, indicator variables and measurement scales: an example from quality of life., *Journal of the Royal Statistical Society, B*, 165, 233–261.

Gatti A. and Mandich G. (1994) *Vivere da studente*, ERSU.

Goldstein H.G. and Spiegelhalter D.J. (1996) League tables and their limitations: Statistical issues in comparisons of istitutional performance, *Journal of the Royal Statistical Society, A*, 159, 385–443.

Huebner E., Valois R., Paxton R.J. and Drane J. (2005) Middle school students' perceptions of quality of life, *Journal of Happiness Studies*, 6, 15–24(10).

Jöreskog K.G. (2002) Sctructural equation modeling with ordinal variables using lisrel, Technical report.

Larsen R., Diener E. and Emmons R.A. (1985) An evaluation of subjective well-being measures, *Social Indicators Reaserch*, 46, 4.

Maggino F. and Schifini S. (1999) Qualità della vita universitaria validazione di strumenti soggettivi, in: *Contesti di Qualità della vita. Problemi e misure*, Franco Angeli.

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 Statistical and Psychology*, 56, 337–357.

Moustaki I., Joreskog K. and Mavridis D. (2004) Factor models for ordinal variables with covariate effects on the manifest and latent variables: A comparison of LISREL and IRT approaches, *Structural Equation Modeling Journal*, 11(4), 487–513.

Porcu M. and Puggioni G. (2003) Laurea e abbandono universitario. Uno studio comparativo dei due evnti su un campione di immatricolati dell'università di cagliari, in: *LAID-OUT: scoprire i rischi con l'analisi di segmentazione*, Fabbris L., ed., CLEUP,Padova.

Porcu M. and Tedesco N. (2007) Transition from university to the job market. a time analysis of university of cagliari graduates, in: *Effectiveness of University Education in Italy*, Fabbris L., ed., Physica-Verlag.

Rampichini C. and Grilli L. (2004) Analysis of university course evaluations: from descriptive measures to multilevel models, *Statistical Methods & Applications*, 13.3, 357–371.

Rampichini C. and Schifini S. (1998) A hierarchical ordinal probit model for the analysis of life satisfaction in italy, *Social Indicators Research*, 44, 41–69.

Samejima F. (1969) Estimation of ability using a response pattern of graded scores, *Psychometrika Monograph*, 17.

Shulz W. (1999) Explaining quality of life:the controversy between subjective and objective variables, EuReporting Working Paper 10, European Commission.

Sulis I. (2007) *Measuring students' assessments of 'university course quality' using mixed-effects models*, Ph.D. thesis, Università di Palermo.

Torgerson W.S. (1958) *Theory and Methods of Scaling*, Wiley and Son, New York.