A classification of university courses based on students' satisfaction: an application of a two-level mixture Item Response Theory (IRT) model

SIS 2013

University of Brescia, 19-21 June 2013

M. Gnaldi, S. Bacci

Dapartment of Economics, Finance and Statistics

University of Perugia

FRAMING THE ISSUE

In the educational and evaluation contexts, data usually present several characteristics that need to be taken into account to correctly model the heterogeneity of the analysed phenomenon:

- First, the **variables** of interest are **not directly observable** and we rely on the observed responses to a set of items, which may be differently scored (e.g., binary or ordinal polytomously scored items)
- Second, data usually have a **hierarchical structure** with lower-level units aggregated in higher-level units (i.e., students in courses within departments)

Our aim is to propose a **classication of higher-level units (e.g. university courses) into a smaller number of homogeneous classes** with respect to the quality of teaching on the latent variables measured by the questionnaire (i.e. the Italian questionnaire on students' satisfaction)

- we detect the latent variables measured by the questionnaire items, performing a **model-based hierarchical clustering** and a **factor analysis**
- we classify courses into homogeneous classes by estimating a special case of **multilevel mixture factor model characterized** by (i) an IRT parametrization and (ii) discrete latent variables at all hierarchical levels. It has two-levels taking into account the hierarchical structure of the data: students at the first-level and courses at the second-level.

THE DATA

Questionnaires on students satisfaction with university study courses
Sample of **2145 students** and **77 study courses**

The questionnaire was administered at the end of the 2007 academic year. It included **5 sections**:

(1) Students characteristics;

(2) Satisfaction with the context of study;

(3) Frequency of attendance;

(4) Satisfaction with study courses;

(5) Further information.

For the purposes of the analyses described in this contribution, we accounted for:

- the 12 questions included in section (4) measuring students satisfaction with study courses (5-point Likert scale)
- the 10 covariates (at the student-level)
- 3 additional covariates related to university courses (at the course-level)

Covariates at the students' level

A1 (age): 0 = less than 20, 1= between 21 and 23, 2 = between 24 and 25, 3 = more than 25 years old

A2 (gender): 0 = male, 1= female

A3 (school): 0 = classic and scientific lyceum, 1 = professional institutes, 2 = others

A4 (position): 0 = attendant student, 1 = out of course student

A5 (condition): 0 = student, 1 = worker and student

B5 (preliminary knowledge): 0 = strongly or moderately insufficient, 1= strongly or moderately sufficient

B6 (wish to sit for the exam at the end of the course): 0 = no, 1 = yes

C1 (attendance): 0 = 0,25 max, 1 = 0,50, 2 = at least 0,75

E2 (interest in the course): 0 = strongly or moderately not interested; 1 = strongly or moderately interested

E3 (overall satisfaction): 0 = strongly or moderately dissatisfied; 1 = strongly or moderately satisfied

Covariates at the course level

TM (Course level): 0 = bachelor degree, 1= master degree

CA (Course type): 0 = specialistic (caratterizzante e di base), 1 = non specialistic (affine)

AD (Course disciplinary content): 0 = economics and statistics, 1= law, 2 = history, 3 = psychology and sociology, 4= other contents

THE METHODS AND MODELS

Detection of latent variables

We compare a unidimensional model with a multidimensional counterpart relying on a **hierarchical clustering algorithm** proposed by Bartolucci (2007)

The clustering procedure performs s - 1 steps. At each step, the **Wald test statistic for unidimensionality** is computed for every pair of possible aggregations of items

The aggregation with the minimum value of the statistic is adopted and the corresponding model fitted before going to the next step

The output can be displayed through a **dendrogram** that shows the deviance between the initial (k-dimensional) model and the model selected at each step of the clustering procedure. The results of a cluster analysis depend on the adopted rule to cut the dendrogram (e.g., an information criterion, such as BIC)

Classification of courses into homogeneous classes

Analysis carried out through a special case of **multilevel mixture factor model** characterized by an IRT parametrization and by discrete latent variables at both levels of the hierarchy

This model is equivalent to model "B2" introduced by Vermunt (2008), which represents a special case of the Multilevel Mixture IRT model of Cho and Cohen (2010). However, differently from these authors, we consider (i) ordered **polytomous items** rather than dichotomous items and (ii) **multidimensionality of latent variables**. We also introduce (iii) **covariates at student and course levels** that influence the probability of lower- and upper-level units belonging to the latent classes

The proposed model may be equivalently depicted both as a **three-level model for a univariate response** and as a **two-level model for multivariate responses** (Vermunt, 2008)

Notation

i the generic item ($i = 1, \ldots, s$)

j_k the student or first-level unit j belonging to the latent class k ($k = 1, \ldots, K$)

h_c the course or second-level unit h belonging to the latent class c ($c = 1, \ldots, C$)

$Y_{ij_kh_c}$ the observed response to item i which may assume L ordinal categories ($l = 0, \ldots, L-1$)

$\Theta^{(a)} = (\Theta_1^{(a)}, \ldots, \Theta_D^{(a)})$ the vector of latent variables

$\lambda_i(\theta) = (\lambda_{i0}(\theta), \ldots, \lambda_{i,L-1}(\theta))'$, with $\theta = (\theta^{(1)}\theta^{(2)})$ and $\lambda_{il}(\theta) = p(Y_{ij_kh_c} = l|\theta)$

The proposed model

$$g_y(\lambda_i(\theta)) = \gamma_i \cdot \left(\sum_{d=1}^{D} \delta_{id}(\theta_{j_k h_c d} + \theta_{0 h_c d}) - \beta_{il} \right) \qquad (1)$$

- γ_i: discriminating item parameter (constant among classes)
- β_{il}: difficulty item parameter (constant among classes)
- δ_{id}: binary variable assuming value 1 if item i measures latent variable d ($d = 1, \ldots, D$)

The random effects

$\theta_{j_k h_c d}$

- it represents the first-level random effects
- it has support points $(\theta_{1d}^{(1)}, \ldots, \theta_{Kd}^{(1)})'$ $(d = 1, \ldots, D)$ and corresponding weights $(\pi_1^{(1)}, \ldots, \pi_K^{(1)})$
- it stands for the latent variable d for student j (belonging to latent class k) in course h (belonging to latent class c)

$\theta_{h_c d}$

- it represents the second-level random effects
- it has support points $(\theta_{1d}^{(2)}, \ldots, \theta_{Cd}^{(2)})'$ $(d = 1, \ldots, D)$ and weights $(\pi_1^{(2)}, \ldots, \pi_C^{(2)})$
- it stands for the latent variable d for course h (belonging to latent class c)

Different parameterizations

Item parameterization

- 2PL-type parameterization: when γ_i parameters differ one other for at least a pair of items
- Rasch-type parameterization: when $\gamma_i = 1$ for all items

Link function

- global logit: $g_y(\lambda_i(\theta)) = \log \frac{p(Y_{ij_kh_c} \geq l | \boldsymbol{\theta})}{p(Y_{ij_kh_c} < l | \boldsymbol{\theta})}$
- local logit: $g_y(\lambda_i(\theta)) = \log \frac{p(Y_{ij_kh_c} = l | \boldsymbol{\theta})}{p(Y_{ij_kh_c} = l-1 | \boldsymbol{\theta})}$

The covariates

Note that in the basic version of the proposed model weights $\pi_k^{(1)}$ and $\pi_c^{(2)}$ do not depend on observable covariates, therefore $\pi_k^{(1)} = p(\Theta^{(1)} = \theta_k)$ and $\pi_c^{(2)} = p(\Theta^{(2)} = \theta_c)$

This is equivalent to assuming that individuals (courses) come from K (C) latent classes which are homogenous in terms of the characteristics measured by the questionnaire

The model may be generalized by assuming that weights depend on individual or course covariates, introducing a multinomial logit model for weights

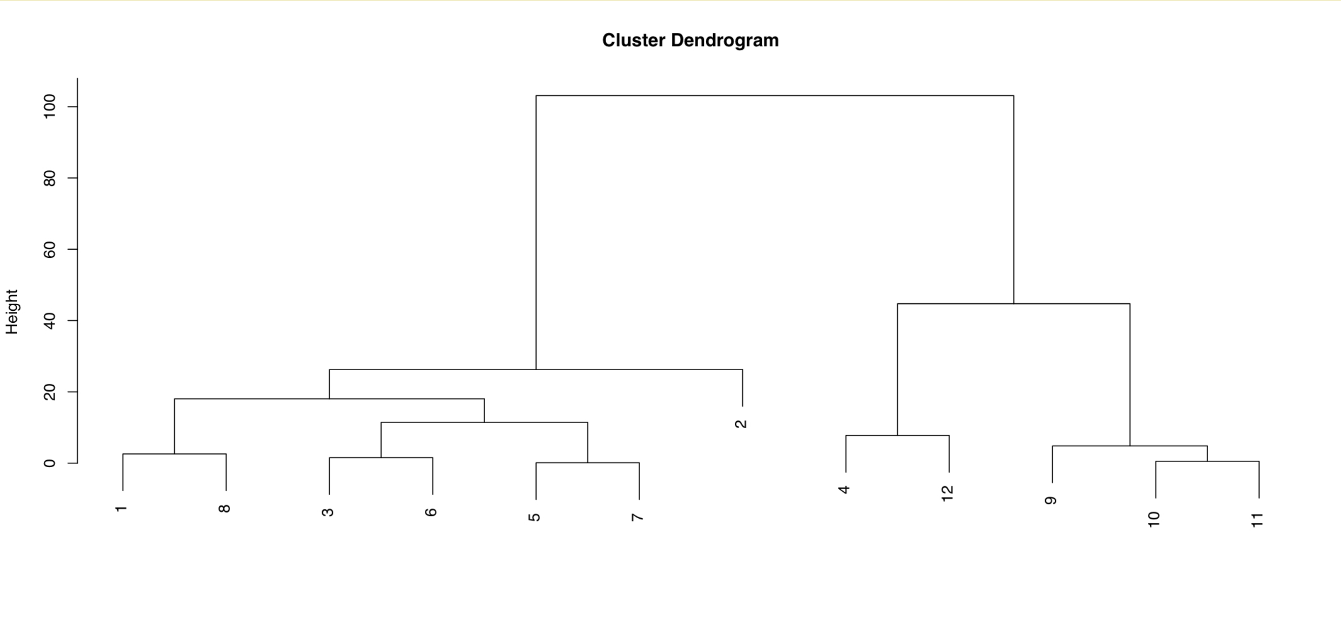
In this way we have a specific weight for each individual j (course h) and latent class k (c)

$$\begin{aligned}
\log \frac{\pi_k^{(1)}(\mathbf{x}_j)}{\pi_1^{(1)}(\mathbf{x}_j)} &= \log \frac{p(\Theta^{(1)} = \theta_k | X_1 = x_{1j}, \ldots, X_p = x_{pj})}{p(\Theta^{(1)} = \theta_1 | X_1 = x_{1j}, \ldots, X_p = x_{pj})} = \\
&= v_{0k}^{(1)} + \sum_{m=1}^{p} v_{mk}^{(1)} x_{mj} \qquad k = 2, \ldots, K; \\
\log \frac{\pi_c^{(2)}(\mathbf{z}_h)}{\pi_1^{(2)}(\mathbf{z}_h)} &= \log \frac{p(\Theta^{(2)} = \theta_c | Z_1 = z_{1h}, \ldots, Z_q = z_{qj})}{p(\Theta^{(2)} = \theta_1 | Z_1 = z_{1h}, \ldots, Z_q = x_{qh})} = \\
&= v_{0c}^{(2)} + \sum_{m=1}^{q} v_{mc}^{(2)} z_{mh} \qquad c = 2, \ldots, C,
\end{aligned} \tag{2}$$

- $\mathbf{x}_j = (x_{1j}, \ldots, x_{pj})$: the observed vector of individual covariates
- $v_{mk}^{(1)}$: the effect of covariate X_m on the logit of $\pi_k^{(1)}(\mathbf{x}_j)$ with respect to $\pi_1^{(1)}(\mathbf{x}_j)$
- $v_{0k}^{(1)}$: the intercept specific for class k
- $\mathbf{z}_h = (z_{1h}, \ldots, z_{qh})$: the observed vector of course covariates
- $v_{mc}^{(2)}$: the effect of covariate Z_m on the logit of $\pi_c^{(2)}(\mathbf{z}_j)$ with respect to $\pi_1^{(2)}(\mathbf{z}_j)$
- $v_{0c}^{(2)}$: the intercept specific for class c

DIMENSIONALITY ASSESSMENT: MAIN RESULTS

Figure 1: Dendrogram resulting from hierarchical clustering: students satisfaction questionnaire.



Main results for dimensionality assessment

The two groups of items (i.e. items on students' satisfaction: D1-D12) correspond to different dimensions which may be identified as the satisfaction with the:

Course content and the ways it is taught (e.g. clarity, content consistency etc.);

Course organisation (e.g. assessment system, punctuality, teacher's helpfulness etc.)

CLASSIFICATION OF COURSES INTO HOMOGENEOUS CLASSES: MAIN RESULTS

Rating Scale Model choice

Table 1: GRSM and RSM: number of parameters, log-likelihood, deviance, BIC index.

	GRSM	RSM
num.par.	92	82
Log-lik.	-23264	-23513
BIC	47223	47646

Effect of covariates

Table 2: Effect of covariates.

	coef	s.e.	z-value	p-value
Latent trait 1				
B5	1.255	0.284	-4.426	<0.0001
B6	1.704	0.665	-2.561	0.0100
E2	2.281	0.438	-5.212	<0.0001
E3	11.043	0.976	-11.319	<0.0001
TM	-1.801	0.555	3.244	0.0012
CA	1.941	0.874	-2.221	0.0260
Latent trait 2				
A1	1.052	0.131	-8.064	<0.0001
A2	0.636	0.212	-3.000	0.0027
C1	1.438	0.258	-5.582	<0.0001
TM	1.058	0.498	-2.124	0.0340
CA	-1.737	0.768	2.262	0.0240

Effect of covariates at the students level

The odds of students' satisfaction with the COURSE CONTENT is higher for students with sufficient preliminary knowledges, who intend to take the exam at the end of the course, and who are interested in the content of the course

Satisfaction with the COURSE ORGANISATION is positively affected by age of students, gender and attendance: older and female students that assiduously attend lectures tend to have higher odds of satisfaction

Effect of covariates at the course level

Satisfaction with the CONTENT of MSc courses is lower that of Bachelor degree courses

Satisfaction with the CONTENT of non specialistic (affini) courses is higher than that of specialistic courses.

Vice versa:

Satisfaction with the ORGANISATION of MSc courses is higher than that of Bachelor degree courses

Satisfaction with the ORGANISATION of non specialistic (affini) courses is lower than

Support point estimates

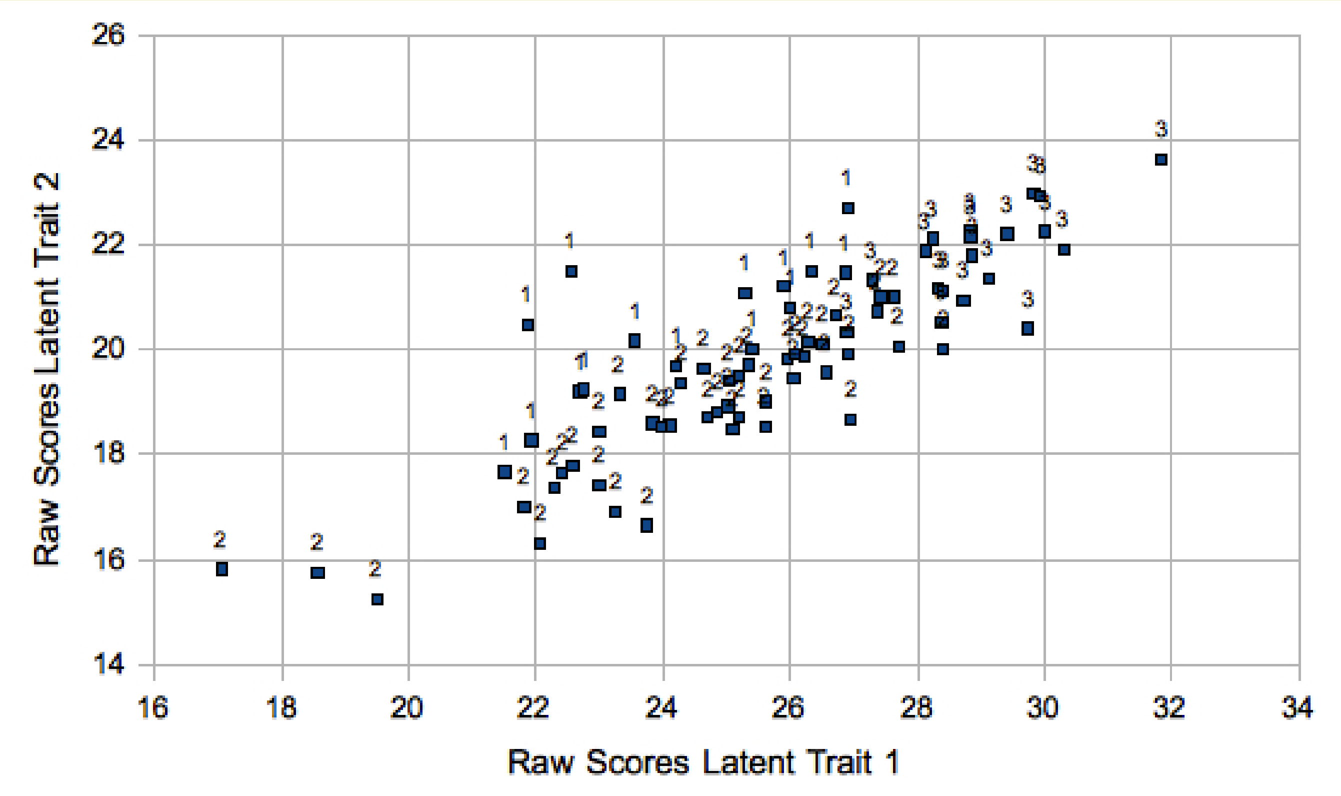
Table 3: Support point estimates for latent classes at course level.

	c=1	c=2	c=3
Latent trait 1	1.916	2.375	3.792
Latent trait 2	0.780	-0.215	0.592
Average weight	0.187	0.562	0.251

Table 3: Percentiles of raw scores on latent trait 1 and latent trait 2, by latent class at course level ($c = 1, 2, 3$).

	Latent trait 1			Latent trait 2		
	c=1	c=2	c=3	c=1	c=2	c=3
10%	17	19	23	16	15	17
25%	20	21	26	18	17	20
50%	24	25	29	20	19	22
75%	28	28	32	23	21	24
90%	31	31	34	24	23	25

Figure 2: Average raw score on latent trait 2 by average raw score on latent trait 1, by courses (labels denoting the second-level latent class.



References

Andrich, D. (1978). A rating formulation for ordered response categories. Psychometrika, 43, 561-573

Bartolucci, F. (2007). A class of multidimensional IRT models for testing unidimensionality and clustering items. Psychometrika, 72:141-157.

Birnbaum, A. (1968). Some latent trait models and their use in inferring an examinees ability. In Lord, I. F. M. and M.R. Novick (eds.), Reading, M. A.-W., editors, Statistical theories of mental test scores, pages 395-479.

Bacci, S. and Caviezel, V. (2011). Multilevel IRT models for the university teaching evaluation. Journal of Applied Statistics, 38, 2775-2791 (2011)

Cho, S.-J. and Cohen, A. S. (2010). A multilevel mixture IRT model with an application to DIF. Journal of Educational and Behavioral Statistics, 35, 336-370

Grilli, L. and Rampichini, C. (2007). Multilevel factor models for ordinal variables. Structural Equation Modeling, 14(1), 1-25.

Vermunt, J.K. (2003). Multilevel latent class models. Sociological Methodology, 33, 213-239

Vermunt, J.K. (2008). Multilevel latent variable modeling: an application in education testing. Austrian Journal of Statistics, 37, 285-299