Evaluation of student performance through a multidimensional finite mixture IRT model

Silvia Bacci[1], Francesco Bartolucci[1],
Leonardo Grilli[2], Carla Rampichini[2]

[1] Department of Economics - University of Perugia,
e-mail: silvia.bacci@unipg.it, francesco.bartolucci@unipg.it

[2] Dept. of Statistics, Computer Science, Applications 'G. Parenti' - University of Florence,
e-mail: grilli@disia.unifi.it, rampichini@disia.unifi.it

Final meeting of the research group on 'Mixture and Latent Variable Models for Causal Inference and Analysis of Socio-Economic Data'
February 1-2, 2017, Bologna

Aim

- Analysis of the performance of university students, with reference to 6 compulsory courses of the first year
- The main goal is to classify the students into performance classes on the basis of their performance on the 6 exams; for each exam, the performance is measured by two pieces of information:
  1. enrollment to the exam: whether the student decides to *take the exam* in the observation period (one year)
  2. exam result: conditional on enrollment the student obtain a result (*failed* or *passed with a grade*).
- If the student does not enroll to a given exam, the result is missing: this is informative about the student performance, thus the missingness cannot be ignored

Methods

- We adopt a multidimensional latent class IRT model for the analysis of item responses affected by non-ignorable missingness
- We extend the proposal of Bacci and Bartolucci (2015) to allow for
  - ordinal responses (instead of binary)
  - two kinds of missingness:
    1. structural missing data: the result is missing because the exam is not due (structural missing data are not considered in the model of Bacci and Bartolucci, 2015)
    2. genuine (potentially informative) missing data: the result is missing even if the exam is due because the student did not take it

Bacci S., Bartolucci F. (2015) A multidimensional finite mixture SEM for non-ignorable missing responses to test items, *Structural Equation Modeling*

Data

- We consider the freshmen of A.Y. 2013/2014 in two degree programs (*Economics*; *Business*) of the University of Florence
- Outcome: performance on the compulsory first-year exams in year 2014
- We analyze data about 861 active freshmen (those who enrolled for at least one exam in year 2014 - 89% of the total)
- Exams can be taken in any order (and freely repeated), enrolling via web
  - Courses of semester I: exams in any of the 6 sessions from Jan to Dec
  - Courses of semester II: exams in any of the 4 sessions from June to Dec

Table: *Enrollment rates and exam results for first-year exams (year 2014)*

Course (sem.)		Enroll. rate (%)	Exam grade (%)					Passing rate (%)	
			fail	18-21	22-24	25-27	≥ 28	enroll.	overall
Accounting	(I)	93.5	42.5	15.9	17.3	17.0	7.3	57.5	53.8
Math	(I)	67.8	65.8	16.2	7.3	6.8	4.0	34.2	21.1
Law	(I)	48.3	47.1	14.2	16.1	14.4	8.2	52.9	25.6
Management	(II)	72.5	30.6	8.2	16.7	23.2	21.3	69.4	50.3
MicroEcon	(II)	41.8	41.9	10.6	11.4	18.1	18.1	58.1	24.3
Statistics	(II)	67.0	39.7	16.8	13.5	11.4	18.5	60.3	40.4

Covariates + the course group indicator

		N	Average number of exams	
			enrolled to	passed
All freshmen		**861**	**3.8**	**2.2**
Gender	Male	502	3.8	2.1
	Female	359	3.9	2.2
High school type	Technical	765	3.8	2.1
	Humanities	201	3.7	1.9
	Scientific	321	4.1	2.4
	Other	284	3.6	1.8
High school grade	< 80	596	3.6	1.7
	≥ 80	265	4.4	3.3
Late matriculation	No	759	4.0	2.3
	Yes	102	3.0	1.3
Degree program	Business	588	3.8	2.1
	Economics	273	3.9	2.3
Course group	A-C	257	3.7	2.2
	D-L	240	3.8	2.2
	M-P	204	4.0	2.1
	Q-Z	160	3.9	2.3

Each course has 4 groups (classes), based on the first letter of the student's surname: e.g. Mr. Rossi is assigned to "Statistics Q-Z", therefore the exam "Statistics Q-Z" is due (we can observe a result), while "Statistics A-C" etc. are not due (we cannot observe a result - structural missingness)

Model: basic notation

Remark: in our application 'item' ⇔ 'first-year exam'

- $Y_j = y_j$: response provided by the subject to ordinal item j, $j = 1, ..., m$

  $1, \dots, L$ if item j is observed
  "NA" if item j is skipped

- $R_j = r_j$: item indicator of response

  1 if Y_j is observed
  0 if Y_j is skipped (informative missing)
  "NA" if Y_j is not due (structural missing)

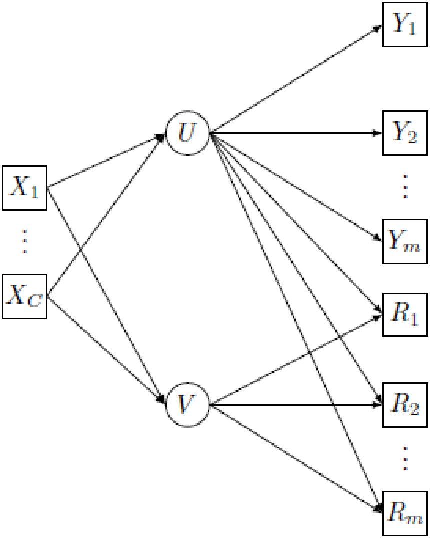
- $X_1, \dots, X_C$: exogenous individual covariates

- U: latent variable denoting the latent trait (performance) measured by the test items

- V: latent variable denoting an individual preference in choosing the test items to answer (determining if an item is observed or missing)

Model: path diagram

Multidimensional LC-IRT model: item binary indicators of answer R_j, item ordinal responses Y_j, latent performance U, latent preference in choosing the test items to answer V



Model: our specification

- The model allows for multidimensional latent variables of each type, but for simplicity we assume
  - a single latent performance U
  - a single latent preference in choosing the exams V
- The exam result is an element of the following set: $\{\text{failed}, 18, 19, \ldots, 29, 30, 30 \text{ with honors}\}$ - this implies 15 ordinal categories, but for simplicity we reduce the categories to 5 as follows:

$$\begin{cases} Y_j = NA & \text{if} \quad R_j = 0 \\ Y_j = 0 & \text{if} \quad R_j = 1 \text{ and } grade = NA \\ Y_j = 1 & \text{if} \quad R_j = 1 \text{ and } 18 \leq grade \leq 21 \\ Y_j = 2 & \text{if} \quad R_j = 1 \text{ and } 22 \leq grade \leq 24 \\ Y_j = 3 & \text{if} \quad R_j = 1 \text{ and } 25 \leq grade \geq 27 \\ Y_j = 4 & \text{if} \quad R_j = 1 \text{ and } grade \geq 28 \end{cases}$$

- Each of the **6 courses** has **4 groups** (i.e. classes with different teachers)
  - any item j refers to a course for a given group $\Rightarrow$ there are m=6×4=24 items ('Accounting A-C', 'Accounting D-L', ..., 'Statistics M-Z')
  - each student is assigned to one group, thus the outcomes for the other groups are missing by construction: $R_j = NA$ (structural missing)

Model: distribution of the latent variables

- We assume that the latent variables U and V have discrete distributions
  - U has support points (latent classes) u_{h_U}
    the number of support points k_U has to be estimated
  - V has support points (latent classes) v_{h_V}
    the number of support points k_V has to be estimated
- Discrete latent variables $\rightarrow$ clustering the individuals into latent classes that are homogeneous with respect to the latent traits
- In the spirit of concomitant variable LC models (Dayton and Macready, 1988), we allow the membership probabilities of the latent classes to depend on observed covariates through a multinomial logit model (Bacci and Bartolucci, 2015):

$$\begin{aligned} \log \frac{\lambda_{h_U}(\boldsymbol{x})}{\lambda_1(\boldsymbol{x})} &= \boldsymbol{x}'\boldsymbol{\phi}_{h_U}, \quad h_U = 2, \ldots, k_U \\ \log \frac{\pi_{h_V}(\boldsymbol{x})}{\pi_1(\boldsymbol{x})} &= \boldsymbol{x}'\boldsymbol{\psi}_{h_V}, \quad h_V = 2, \ldots, k_V \end{aligned}$$

Model: measurement component

The relationships between the latent variables U and V and the manifest variables (item responses $Y_1 \ldots Y_j \ldots Y_m$ and response indicators $R_1 \ldots R_j \ldots R_m$) are described by the measurement part of the model:

- R_j given U and V is a 2PL model (Birnbaum, 1968)

$$\log \frac{q_{h_U h_V,j}}{1 - q_{h_U h_V,j}} = \gamma_{Uj} u_{h_U} + \gamma_{Vj} v_{h_V} - \delta_j$$

where $q_{h_U h_V,j} = Pr(R_j = 1 | U = u_{h_U}, V = v_{h_V})$
identifiability constraints: $\gamma_{Vj} = 1, \delta_j = 0$ for a reference item

- Y_j given U is a Graded Response Model (GRM; Samejima, 1969)

$$\log \frac{p_{h_U,jy}}{1 - p_{h_U,jy}} = \alpha_j u_{h_U} - \beta_{jy}, \qquad y = 2, \ldots, L$$

where $p_{h_U,jy} = Pr(Y_j \geq y | U = u_{h_U})$

identifiability constraints: $\alpha_j = 1, \beta_{j1} = 0$ for a reference item

Model: likelihood inference

- We fit the proposed multidimensional LC-IRT model by maximizing the marginal likelihood using the EM algorithm (Dempster et al., 1977).
- We exploit the R package `MLCIRTwithin` (function `est_multi_poly_within`, which is devoted to the estimation of Multidimensional LC-IRT models in presence of within-item multidimensionality)
  The package has been recently updated to account for the new features of the proposed model, namely ordinal responses and structural missing values.

Bartolucci F., Bacci S., Gnaldi M. (2014). MultiLCIRT: An R package for multidimensional latent class item response models, *Computational Statistics and Data Analysis*, 71, 971–985.

Selection of the number of latent classes

k_U	k_V	$\hat{\ell}$	# par	BIC
2	2	-6520.37	208	14446.41
2	3	-6505.63	217	14477.76
3	2	-6387.18	217	14240.87
3	3	-6364.32	226	14255.96
4	2	-6338.27	226	14203.86
4	3	-6325.72	235	14239.58
5	2	-6323.84	235	14235.84
5	3	-6304.16	244	14257.30

- On the basis of BIC we select $k_U = 4$ latent classes for U and $k_V = 2$ latent classes for V
- In order to check for local maxima, we repeat the model estimation process for different random starting values of the parameters

Testing the ignorability of the missing data mechanism

- If a student decides not to take an exam in the considered year $(R_j = 0)$ then the exam result Y_j is missing: likely this is not ignorable
- In our model the performance U affects both the decision to take an exam R_j and the result Y_j
  $\rightarrow$ the missing data mechanism is not ignorable
- We test the ignorability assumption comparing our multidimensional LC-IRT model with a restricted model where the decision to take an exam R_j does not depend on the performance U, i.e.

$$\gamma_{Uj} = 0, \qquad j = 1, \ldots, 24$$

- $LRT = 2 \times (6533.720 - 6338.268) = 390.904$, with 24 degrees of freedom yielding a very low p-value $\Rightarrow$ we proceed with the proposed multidimensional LC-IRT model accounting for the non-ignorable missing mechanism

Estimated parameters

The selected model with **4** latent classes for the performance U and **2** latent classes for the preference V has 226 parameters:

- discrimination parameters for the effects of the latent variables
  - latent performance U on exam result Y_j ($\hat{\alpha}^*_j$)
  - latent performance U on exam enrollment R_j ($\hat{\gamma}^*_{1j}$)
  - latent preference V on exam enrollment R_j ($\hat{\gamma}^*_{2j}$)
- difficulty parameters shifting the distributions of Y_j and R_j
  - higher $\hat{\delta}^*_j \rightarrow$ lower probability to take the exam
  - higher $\hat{\beta}^*_{jy} \rightarrow$ lower probability of a good result
- latent structure parameters
  - support points of U ($u^*_{h_U}$) and support points of V ($v^*_{h_V}$)
  - estimated coefficients of the multinomial logit model for the probabilities of U ($\boldsymbol{\phi}_{h_U}$) and V ($\boldsymbol{\psi}_{h_V}$)

The asterisk denotes standardization: for ease of interpretation, the support points have been standardized so that the latent variables have mean 0 and standard deviation 1, and the discrimination and difficulty parameters have been transformed accordingly.

Estimated discrimination parameters for the exam result Y_j

- The exam results Y_j are affected by the latent performance U through the discrimination parameters $\hat{\alpha}_j^*$
- All those parameters are significantly different from zero $\rightarrow$ All the exams contribute to measure the latent performance U
- Exams with higher discrimination are more sensitive to variations in student performance (Accounting, Mathematics, and Statistics)
- There are differences across groups of the same course, especially for Law and Management

Estimated discrimination parameters for the indicator of taking the exam R_j

- The response indicators R_j (=1 if the student takes the exam in the considered year) are affected
  - by the latent performance U through the discrimination parameters $\hat{\gamma}^*_{Uj}$ and
  - by the latent preference V through the discrimination parameters $\hat{\gamma}^*_{Vj}$

  (this dependence on two latent variables is known as *within-item multidimensionality*, e.g. Adams, Wilson and Wang, 1997)
- The student latent performance U significantly affects the enrollment for most exams: this provides evidence that the enrollment process generating missing exam results is not ignorable (as confirmed by the LRT)
- The student latent preference V has a positive effect for Mathematics and Statistics, and negative for Law; thus V can be interpreted as the preference of the student to take exams in *quantitative subjects* as opposed to exams in qualitative subjects (the effect is significant for about one-third of the items)
- The indicators of taking the exam R_j are affected more by U than by V, namely $|\hat{\gamma}^*_{Uj}| > |\hat{\gamma}^*_{Vj}|$, with the notable exception of Mathematics

From difficulty parameters to predicted probabilities

- The estimates of the difficulty parameters are not easily interpretable
- It is more interesting to look at the predicted probabilities for a student with some hypothetical values of latent performance U and latent preference V (e.g. the mean values $U = V = 0$)
- The predicted probabilities vary with the degree program (*Economics* or *Business*) and the group (A-C, D-L, M-P, Q-Z)
  - We note a large variability among courses and, in some cases, also across groups of the same course for both the enrollment R_j and the exam result Y_j
  - For the majority of courses, the most likely result is a failure and the modal grade of passed exams is 18-21
- To see how the probability to obtain an exam result depends on the student latent performance, we move the value of U (we find that the result is highly influenced by the latent performance)

Estimates for latent performance U & latent preference V

Standardized estimated support points with corresponding average probabilities

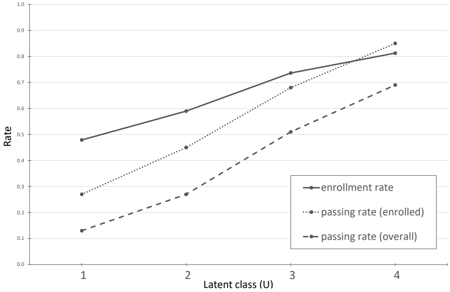
	Performance U latent class h_U				Preference V latent class h_V	
	$h_U = 1$	$h_U = 2$	$h_U = 3$	$h_U = 4$	$h_V = 1$	$h_V = 2$
Support points $(u^*_{h_U}, v^*_{h_V})$	-1.485	-0.129	0.784	1.937	-0.949	1.054
Average probs $(\overline{\lambda}_{h_U}, \overline{\pi}_{h_V})$	0.228	0.395	0.294	0.083	0.526	0.474

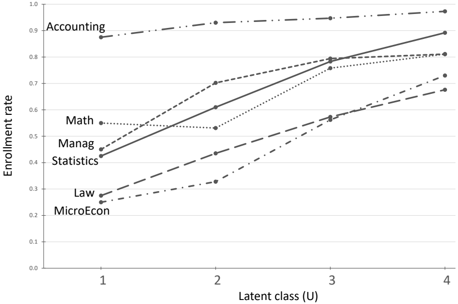
Estimated coefficients of the multinomial logit models for the probabilities of the support points

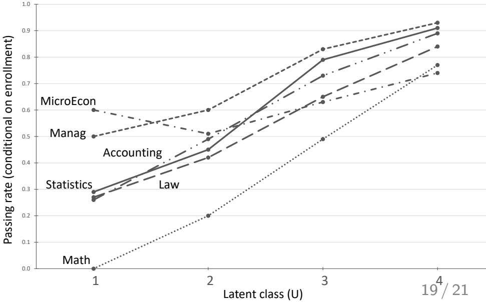
	Model for performance U			Model for preference V
	$\hat{\phi}_{11j}$	$\hat{\phi}_{12j}$	$\hat{\phi}_{13j}$	$\hat{\psi}_{11j}$
Constant	0.748	0.568	-1.956	-0.869
Degree Economics $(j = 1)$	-0.434	-0.030	0.045	0.716
Female $(j = 2)$	0.434	0.059	-0.487	-0.147
HS grade $(j = 3)$	0.014	0.118	0.265	-0.020
HS technical	–	–	–	–
HS humanities $(j = 4)$	0.138	0.148	0.691	-0.240
HS scientific $(j = 5)$	-0.061	1.021	2.219	1.963
HS other $(j = 6)$	-0.163	-0.163	-0.282	-0.336
Late enrollment $(j = 7)$	-0.191	-1.415	-1.834	-0.744

Parameters in red have p-value < 0.05.

Enrollment rates and passing rates for the latent classes







Final remarks

- Student performance measured by exam results of taken exams AND by indicators of taking the exams, thus accounting for the decisions to take or not to take any given exam in the considered year (non-ignorable missingness) - this is relevant in the Italian university system, where many students do not take all the compulsory exams in the expected period (revealing the patterns is essential for corrective actions).
- Meaning of the latent variable U: higher latent class $\rightarrow$ higher probability of taking exams AND passing exams, thus U is interpreted as overall performance.
- Noteworthy differences among classes held by different teachers $\rightarrow$ teacher effect on both the probability of taking the exam and the result (this raises a fairness issue).
- The proposed LC-IRT model is suitable for a wide range of applications characterized by ordinal items with non-ignorable missing item responses, e.g. in achievement tests, customer satisfaction and quality of life.

Thanks for your attention!

Silvia Bacci, Francesco Bartolucci, Leonardo Grilli, Carla Rampichini

Estimated item discrimination parameters

latent performance U on exam result Y (scaled parameters $\hat{\alpha}_j^*$)
latent performance U on exam enrollment R (scaled parameters $\hat{\gamma}_{1j}^*$)
latent preference V on exam enrollment R (scaled parameters $\hat{\gamma}_{2j}^*$)

Item		$U \to Y_j$			$U \to R_j$			$V \to R_j$		
Course	Group	$\hat{\alpha}_j^*$	$\hat{\text{se}}_{\alpha_j^*}$	p-value	$\hat{\gamma}_{Uj}^*$	$\hat{\text{se}}_{\gamma_{Uj}^*}$	p-value	$\hat{\gamma}_{Vj}^*$	$\hat{\text{se}}_{\gamma_{Vj}^*}$	p-value
Account	A-C	2.127	0.285	< 0.001	0.904	0.375	0.016	0.401	0.274	0.144
	D-L	**1.795**	0.259	< 0.001	0.535	0.349	0.125	0.594	0.433	0.170
	M-P	2.527	0.380	< 0.001	1.172	0.533	0.028	-0.503	0.525	0.338
	Q-Z	2.337	0.357	< 0.001	0.433	0.318	0.173	0.522	0.311	0.093
Math	A-C	2.241	0.410	< 0.001	1.918	0.893	0.032	2.448	1.176	0.037
	D-L	2.134	0.386	< 0.001	1.278	0.440	0.004	2.251	0.772	0.004
	M-P	**1.731**	0.402	< 0.001	1.700	0.503	0.001	1.782	0.587	0.002
	Q-Z	2.963	0.707	< 0.001	5.050	4.605	0.273	6.783	5.266	0.198
Law	A-C	1.849	0.427	< 0.001	0.903	0.196	< 0.001	-0.380	0.221	0.085
	D-L	3.016	0.506	< 0.001	1.303	0.254	< 0.001	-0.390	0.267	0.144
	M-P	**1.391**	0.306	< 0.001	1.144	0.267	< 0.001	-0.804	0.314	0.011
	Q-Z	1.783	0.425	< 0.001	1.033	0.241	< 0.001	-0.279	0.214	0.192
Manag	Bus A-L	2.990	0.530	< 0.001	2.287	0.389	< 0.001	0.675	0.315	0.032
	Bus M-Z	**3.163**	0.484	< 0.001	1.339	0.249	< 0.001	-0.304	0.221	0.169
	Eco A-L	1.976	0.389	< 0.001	1.106	0.250	< 0.001	0.539	0.296	0.068
	Eco M-Z	**0.662**	0.306	0.030	2.212	0.500	< 0.001	0.605	0.455	0.184
MicroEcon	A-C	1.249	0.325	< 0.001	1.429	0.247	< 0.001	0.310	0.238	0.193
	D-L	1.130	0.402	0.005	3.114	0.598	< 0.001	0.450	0.319	0.159
	M-P	1.822	0.366	< 0.001	1.889	0.353	< 0.001	-0.447	0.294	0.128
	Q-Z	**2.350**	0.588	< 0.001	2.202	0.440	< 0.001	0.926	0.282	0.001
Statistics	A-C	2.787	0.445	< 0.001	2.333	0.466	< 0.001	0.946	0.387	0.014
	D-L	2.496	0.389	< 0.001	1.567	0.290	< 0.001	0.808	0.295	0.006
	M-P	2.867	0.497	< 0.001	1.772	0.319	< 0.001	0.104	0.292	0.722
	Q-Z	2.258	0.468	< 0.001	1.998	0.469	< 0.001	1.020	0.347	0.003

Predicted probabilities of enrollment R

Probabilities at some values of latent performance U and latent preference V.

Item		$P(R = 1 \mid U = u, V = v)$							
Course	Class	u	0	$-\sigma_U$	$+\sigma_U$	0	0	Range	
		v	0	0	0	$-\sigma_V$	$+\sigma_V$	$\pm\sigma_U$	$\pm\sigma_V$
Accounting	A-C		0.96	0.91	0.98	0.94	0.98	0.07	0.03
	D-L		0.95	0.92	0.97	0.92	0.97	0.05	0.06
	M-P		0.97	0.92	0.99	0.98	0.96	0.07	-0.03
	Q-Z		0.91	0.87	0.94	0.86	0.95	0.07	0.09
Mathematics	A-C		0.78	0.37	0.96	0.27	0.97	0.58	0.72
	D-L		0.78	0.50	0.93	0.28	0.97	0.42	0.69
	M-P		0.77	0.38	0.95	0.36	0.95	0.57	0.59
	Q-Z		0.87	0.10	1.00	0.02	1.00	0.89	0.99
Law	A-C		0.32	0.16	0.54	0.41	0.25	0.38	-0.17
	D-L		0.51	0.22	0.79	0.60	0.41	0.57	-0.19
	M-P		0.58	0.31	0.81	0.76	0.38	0.51	-0.37
	Q-Z		0.56	0.31	0.78	0.63	0.49	0.47	-0.14
Management	Bus A-L		0.88	0.43	0.99	0.79	0.94	0.56	0.15
	Bus M-Z		0.79	0.50	0.94	0.84	0.74	0.44	-0.10
	Eco A-L		0.64	0.37	0.85	0.51	0.76	0.47	0.25
	Eco M-Z		0.88	0.45	0.99	0.81	0.93	0.53	0.13
MicroEcon	A-C		0.27	0.08	0.60	0.21	0.34	0.52	0.13
	D-L		0.14	0.01	0.79	0.10	0.21	0.78	0.11
	M-P		0.60	0.18	0.91	0.70	0.49	0.72	-0.21
	Q-Z		0.53	0.11	0.91	0.31	0.74	0.80	0.43
Statistics	A-C		0.85	0.36	0.98	0.69	0.94	0.63	0.25
	D-L		0.68	0.31	0.91	0.49	0.83	0.60	0.34
	M-P		0.69	0.27	0.93	0.67	0.71	0.66	0.04
	Q-Z		0.87	0.47	0.98	0.70	0.95	0.51	0.25

Predicted probabilities of exam result Y

Predicted probabilities at some values of latent performance U

Item		$P(Y = y_k \mid U = 0)$					Success rate $P(Y > 0 \mid U)$			
Course	Class	0	1	2	3	4	$U = u$			
		Failed	18-21	22-24	25-27	28-30	$-\sigma_U$	0	$+\sigma_U$	range
Accounting	A-C	0.35	0.33	0.17	0.12	0.03	0.18	0.65	0.94	0.76
	D-L	0.65	0.17	0.13	0.05	0.01	0.08	0.35	0.77	0.68
	M-P	0.38	0.22	0.30	0.09	0.01	0.12	0.62	0.95	0.84
	Q-Z	0.14	0.29	0.35	0.18	0.03	0.37	0.86	0.98	0.61
Mathematics	A-C	0.84	0.11	0.03	0.01	0.01	0.02	0.16	0.64	0.62
	D-L	0.80	0.14	0.04	0.02	0.00	0.03	0.20	0.67	0.65
	M-P	0.87	0.07	0.04	0.02	0.01	0.03	0.13	0.45	0.43
	Q-Z	0.91	0.09	0.00	0.00	0.00	0.01	0.09	0.67	0.66
Law	A-C	0.82	0.11	0.06	0.01	0.00	0.03	0.18	0.58	0.54
	D-L	0.53	0.17	0.24	0.06	0.01	0.04	0.47	0.95	0.91
	M-P	0.71	0.14	0.04	0.09	0.02	0.09	0.29	0.62	0.53
	Q-Z	0.48	0.26	0.18	0.07	0.01	0.15	0.52	0.87	0.71
Management	Bus A-L	0.22	0.19	0.35	0.18	0.06	0.15	0.78	0.99	0.83
	Bus M-Z	0.76	0.06	0.13	0.04	0.01	0.01	0.24	0.88	0.87
	Eco A-L	0.38	0.20	0.21	0.16	0.04	0.18	0.62	0.92	0.74
	Eco M-Z	0.10	0.03	0.09	0.57	0.21	0.82	0.90	0.95	0.12
MicroEcon	A-C	0.71	0.12	0.10	0.05	0.02	0.11	0.29	0.59	0.48
	D-L	0.81	0.05	0.05	0.06	0.04	0.07	0.19	0.42	0.35
	M-P	0.36	0.16	0.16	0.25	0.08	0.23	0.64	0.92	0.69
	Q-Z	0.65	0.12	0.07	0.08	0.09	0.05	0.35	0.85	0.80
Statistics	A-C	0.50	0.29	0.12	0.06	0.03	0.06	0.50	0.94	0.88
	D-L	0.55	0.23	0.16	0.04	0.02	0.06	0.45	0.91	0.84
	M-P	0.67	0.20	0.08	0.04	0.02	0.03	0.33	0.90	0.87
	Q-Z	0.62	0.20	0.06	0.07	0.05	0.06	0.38	0.85	0.79

Testing differences among the groups of a course

Likelihood-ratio tests comparing, separately for each course, the *full model* with the *restricted model with groups collapsed for the course under consideration*

Model	$\hat{\ell}$	# par	Deviance	df	p-value
Full model	-6336.114	226	–	–	–
Collapsing Business	-6389.587	202	106.947	24	0.000
Collapsing Mathematics	-6349.616	202	27.004	24	0.304
Collapsing Law	-6397.061	202	121.894	24	0.000
Collapsing Management	-6464.912	202	257.596	24	0.000
Collapsing Economics	-6411.623	202	151.017	24	0.000
Collapsing Statistics	-6358.218	202	44.208	24	0.007

- The likelihood ratio tests reveal significant differences among the groups for nearly all the courses (teacher effect)
- The only exception is represented by Mathematics