Students' satisfaction in higher education: the role of practices, needs and beliefs of teachers

Francesca Bassi, Department of Statistical Sciences, University of Padua, Italy, via C. Battisti 241, 35121 Padua Italy, francesca.bassi@unipd.it ph. ++390498274152

Abstract

Purpose

Scope of the paper is the analysis of the evolution of students' satisfaction over time in a large Italian university and the effects on it due to some characteristics of the teachers: their didactic practices, their beliefs and their needs on teaching and learning.

Methodology

The first step of the analysis identifies a latent construct measured with items composing the questionnaire and proposes a reduced set of indicators to measure satisfaction and to model its evolution over time (information collected in three consecutive academic years is available). A second step clusters teachers in homogenous groups with reference to their opinions, beliefs and needs collected with a new survey conducted at the University of Padova with the aim of developing strategies to support academic teachers. Then a mixture conditional latent growth model is estimated with covariates affecting the latent parameters and class membership.

Findings

Model estimation identifies a large group of university courses with a high level of satisfaction which stays constant over time and a small group of problematic courses with low satisfaction, moreover, decreasing over the three considered academic years. Interesting significant effects of covariates related both to the teacher and to the didactic activity are estimated.

Originality

Statistical analyses show that the implementation of innovative didactic practices and commitment to quality of teaching are important factors to be encouraged by the University management. On the contrary, traditionalist way of teaching and low passion for teaching do not improve students' satisfaction.

Keywords: latent class models, latent growth models, quality of the didactics, academic professor

Introduction

Students' participation in university life as well as their perception and evaluation of teaching quality play a major role in higher education. The role of students seems in fact relevant as part of the teaching evaluation process and SETs (students' evaluations of teaching) seem to be an almost universally accepted method of gathering information about the quality of education (Zabaleta, 2007).

The University of Padova was founded in 1222 and is one of the largest in Italy with around 61,000 students and 2,000 professors working in 32 different Departments and 8 Schools: Agricultural Sciences and Veterinary Medicine, Economics and Political Sciences, Law, Engineering, Medicine, Psychology, Science, Human and Social Science and Cultural Heritage. The Italian university system is based on the so called 3+2 reform, which started in academic year 2001-2002 and it is organized in cycles with three consecutive levels: a first-cycle academic degree lasting three years, a second-cycle academic degree of two years and a PhD course. There exists a small group of single-cycle degrees lasting five years, specifically in Medicine, Veterinary, Law and Architecture. (Meggiolaro *et al.*, 2017). At the University of Padova there are 80 first-cycle (bachelor) degree courses, 84 second-cycle (master) courses and 9 single-cycle (5-year-long) degree courses. The University of Padova is ranked at top levels among Italian universities both for the didactic and research activity.

The University of Padova has been collecting information on students' satisfaction for 15 years. The quality assurance system of the University of Padova assigns a considerable role to the students' evaluation of teaching. Great attention and large resources have been allocated since the first 2000s in gathering high quality data for the continuous improvement of the teaching-learning activities. The survey about students' opinions supports the various levels of the internal evaluation process; the survey results are given in a detailed form to individual professors and managers of the various organizational structures (Course Councils, Departments, Athenaeum Schools). Furthermore, some succinct results are published in aggregated form on the university website. Specifically, for each teacher and course, the following indicators are published: the overall level of satisfaction; an indicator related to the organizational aspects of the course (clarity of scopes, examination arrangements, observance of timetable and didactic material); and an indicator related to efficacy of didactics (interest stimulation, clear explanation).

It is truly important that these indicators, and the instrument by which they are collected, have the properties of validity and reliability, which can guarantee its purpose. In Dalla Zuanna *et al.* 2015 the scale has been shown to be valid and reliable following the traditional procedure proposed in the psychometric literature (see, for example, Churchill, 1979) consisting of a number of steps to take in developing a measurement instrument. These steps refer to construct and domain definition, and scale validity, reliability, dimensionality and generalizability (Bassi, 2010). The former analyses show also that the two indicators of satisfaction with organizational aspects and efficacy of didactics are valid and reliable and that the measurement scale is not uni-dimensional: there are four underlying latent factors; only one dimension is strictly related to the teachers and their activity with the students.

Recently a new CAWI (Computer Assisted Web Interviewing) survey was conducted at the University of Padova by means of the research project PRODID (Teacher professional development and academic educational innovation). This study started in 2013 with the aim of developing strategies to support academic teachers and enhance their teaching competence. A specific questionnaire was then developed and addressed to all professors, with regard to their teaching activities adopted in their classes (Dalla Zuanna *et al.*, 2015).

In this paper we use the growth approach (Nagin, 2005) to identify university didactic activities that suffer of a problem of low quality teaching and to evaluate the effect on the dynamics of satisfaction of teachers' attitude towards their academic involvement. We have students' evaluations of a sample of didactic activities over three academic years from 2012 to 2014 and the information collected by the PRODID survey on their teachers.

The paper is organized as follows. Section 1 revises the vast literature on the topic of students' satisfaction. Section 2 describes the survey and the questionnaire and gives some descriptive statistics about the collected data. Section 3 introduces the statistical methodology. Section 4 presents the results obtained estimating the growth models. The paper finishes with some concluding remarks.

1. Literature review

The focus on the service provided and consumer satisfaction in higher education can benefit from benchmarking other services industries that have a long tradition of attention to both the quality of the service offered and customer satisfaction. This may lead to loyalty and create stronger ties with alumni and positive word-of-mouth effects. Additionally, student surveys of their opinion on the university are input for the construction of most rankings in the academic world (see, as an example, The European Teaching Ranking published by Times Higher Education). On the other hand, dissatisfied students may drop out of higher education or move to another institution. Further, this procedure contributes to continuous advancement in terms of programs, campus facilities, and other attributes. Research on SETs addresses relevant issues, including the importance of involving students in evaluation processes as well as the need to obtain significant information that could be

used for improvement (Svinicki and McKeachie 2013; Theall and Franklin 2007; Zabaleta 2007). SETs are considered a valuable tool designed to enhance both students' learning and teaching performance but this is only true if their results are interpreted and used to develop teaching and if students' feedback is transformed into a stimulus for improvement (Zabaleta, 2007). Students' participation and involvement in the higher education assessment processes is fundamental to promote a growing awareness of their role as stakeholders, as underlined in many European documents (European University Association, 2006). Thus, it is important that students' satisfaction with the service provided is measured regularly to evaluate the different attributes involved and assist decision making. However, recent literature has also raised the debate on the potential drawbacks connected with the use of SETs to assess teaching quality. Uttl *et al*. (2016), for example, in their meta-analysis found that in many cases there is no significant association between SETs and achievement in subsequent exams on the same subject. According to these authors, the idea that students learn more by highly rated professors is not always verified, rather students' evaluation is more a measure of satisfaction than of teaching effectiveness. A similar evidence is posed by Boring *et al*. (2016), while Braga *et al*. (2014) found a negative correlation between the measure of teacher effectiveness and students' evaluation. Stroebe (2016) postulates even that the practice of university administration of relying on SETs induces teachers to adopt didactic practices appreciated by students and not necessarily targeted to improve learning.

Despite this ongoing lively debate (Hornstein, 2017), many universities continue to assess the quality of the didactics by means on SETs. In Italy, for example, this practice is made mandatory by the National Agency for the Evaluation of Universities and Research Institutions (ANVUR). This paper wants to give a contribution suggesting a sound procedure to elaborate and then use data collected form the students with reference to their judgment on teaching in university courses. Whether SETs are a measure of teaching effectiveness or are more an expression of students' satisfaction might not still be clear, they are in any case an important piece of information that university management might process in order to monitor and possibly improve the quality of the service.

A general consensus has been reached in the literature about good teaching also on the necessity for SETs instruments to capture the multidimensionality and the complexity of teaching (Roche and Marsh 2000; Rindermann and Schofield, 2001; Saroyan and Amundsen, 2001; Domenech and Descals, 2003; Apodaca and Grad 2005; Burdsal and Harrison 2008; Cheung 2000; Harrison *et al*., 2004; Mortelmans and Spooren 2009). Even if a recent literature recognizes that satisfaction is a binary threshold response (see, for example, Anand and Bansal (2016) and Dolnicar *et al*. (2011), the majority of universities are using ordinal multi-item questionnaires.

For what concerns the extent to which students are capable of providing appropriate teaching evaluations and the identification of possible factors that influence their opinions: Spooren *et al*. 2013 suggest to divide the possible biasing factors in student-related (class attendance, students' effort, expected and final grade, gender, age, pre-course interest and motivation), teacher-related (age, gender, reputation, research productivity, teaching experience, personal traits) and course-related (class size, class attendance rate, class heterogeneity, course difficulty and workload, discipline, level) characteristics that might affect SETs. Numerous are the studies investigating the relationship between SETs and the characteristics of students, courses, and teachers (Aleamoni, 1999; Marsh, 1987 and 2007; Marsh and Roche, 1997; Centra, 1998; Clayson, 2009) and some of these factors continue to provoke discussions among researchers (for example the course workload and the students' grade expectations).

2. The surveys, the questionnaires and preliminary analysis

In the academic years 2012-2013, 2013-2014 and 2014-2015 the questionnaire proposed to the students began with two introductory questions: the first one asked if the student was available

to participate in the survey (if the student was not, no other question was posed), the second one asked what percentage of the lessons of the course under judgement was attended by the student. If the student attended less than 30% of the lessons, he was asked to answer only a number of selected items and a question on why he attended so few classes; otherwise, all items were proposed. The items composing the scale to measure student satisfaction in the case of more than 30% of classes attended are not exactly the same in the three academic years since the instrument was slightly changed in order to obtain always higher quality data. In our analyses, we considered only the 12 items – 11 concerning specific aspects of the didactic activity, and a general one, the gold standard, measuring overall satisfaction - that were submitted in all the considered academic years and that are reported in the following. Students were asked to express their level of satisfaction on a scale from 1 to 10, being 1 the lowest level.

Item S1 At the beginning of the course, were aims and topics clearly outlined?
Item S2 Were examination arrangements clearly stated?
Item S3 Was classes' timetable observed?
Item S4 Is preliminary knowledge sufficient to understand all topics?
Item S5 Independently on how the course was taught, how much are you interested in the topic?
Item S6 Does the teacher stimulate interest towards the topic?
Item S7 Does the teacher clearly explain?
Item S8 Is the suggested material for study adequate?
Item S9 Was the teacher available during office hours?
Item S10 Are laboratories/practical activities/workshops, if included, adequate?
Item S11 Is the requested workload proportionate to the number of credits assigned to the course?
Item S12 How much are you satisfied about this course?

In this paper we consider a sample of 1,854 didactic activities that were evaluated by the students of the University of Padova for the three academic years and did not change teacher nor number of teaching hours in the reference period. Observations with missing data or evident errors were excluded from the analysis. Only questionnaires filled in by regular students who attended at least 50% of the lessons were taken into account. For each course we have also information on the type of degree, bachelor, master or 5-year-long, number of teaching hours and corresponding credits (ECTS), university School where the course is given, role of the teacher, whether assistant, associate or full professor. This sample was analyzed in a previous paper by Bassi *et al.* (2017b), obtaining some important evidences on the quality of the didactics and on its recent evolution over time at the University of Padova. Specifically, the scale was demonstrated to be valid and reliable according to the traditional protocols to evaluate measurement scales. Moreover, factor analysis (Bartholomew and Knott, 1999) showed that one factor explains almost 81% of total variance with very high loadings, greater than 0.84, with all 12 items. This latent factor represents students' satisfaction with university courses. However, there was a problem of fit: the likelihood ratio statistics is equal to 2,401.42 and it rejects the null hypothesis considering a Chi-square distribution with 24 degrees of freedom. This is probably due to the fact that the model does not include co-varying error terms and 11 highly correlated items are too many for only one underlying factor. For this reason, we decided to reduce the 11 items to five indicators in the following way. Items S1, S2, S3 and S8 were aggregated in the indicator that measures satisfaction with reference to organizational aspects (OA); items S6 and S7 were aggregated in the indicator that measures satisfaction with reference to efficacy of didactics (ED). As already said, these two indicators are published by the University of Padova for every didactic activity and have shown to be valid and reliable in a previous work (Bassi *et al.*, 2017a). Items S4 (preliminary knowledge), S9 (teacher availability) and S11 (workload) are kept for the analyses while the other items, S5 (interest in the topic) and S10 (practical activities) are discarded. These last items were those showing the highest residual correlation and also elevated correlation with at least one of the other items still analyzed,

specifically item S5 is highly correlated with S6 and item S10 with items S8 and S9. Performing factor analysis with these five indicators shows the presence of one underlying factor that explains more than 86% of total variance and factor loadings are all greater than 0.88. Moreover, the fit of the model with the reduced number of indicators improves with reference to several indexes (AIC, BIC, likelihood ratio statistics, Root Mean Squared Error). The small loss of information due to the exclusion of items S5 and S10 is highly compensated by a better fit and a lower model complexity. A Cronbach's alpha coefficient of 0.946 ensures internal reliability. In this paper we will use these same indicators as observable measures of students' satisfaction.

In 2013 the University of Padova started a CAWI survey aiming at providing a picture of the teaching experiences developed in the Athenaeum. The PRODID project promoted a research-based approach to creating training programs, faculty learning communities, pilot experimental contexts where teaching innovation could be tested and monitored. Following an evidence-based approach, the project aimed at high lightening teachers' needs, beliefs and practices of teaching and learning , which may constitute a privileged context for the development of innovative teaching activities within the institution. The final questionnaire was developed according to the Framework of teaching (Tigelaar *et al.*, 2004) and it is composed of three sections. The first section, with nine items (Q1 to Q9), focuses on practices developed by the professors of Padova University in their teaching activity. The second section deepens on teachers' beliefs about teaching in higher education with nine items (Q13 to Q21). The third section refers to needs that the professors perceive in order to improve the efficacy of their teaching activity (items Q22 to Q28). Items Q11, Q12 and Q29, Q30 refer to passion and enjoyment for teaching and for research, respectively. Items from Q1 to Q9 generate binary variables (no, yes). All other items require an answer on a seven-point scale where 1 refers to fully disagree and 7 to fully agree (for a more detailed description of the survey, see, Bassi *et al.*, 2017c).

Item Q1 In class I used didactic practices to actively involve students (*e.g.*, case discussion, exercises, problem-based learning).
Item Q2 I invited external contributors during lessons (*e.g.*, stakeholders, experts, testimonials).
Item Q3 I monitored students' learning during the course with specific examinations.
Item Q4 I assessed students' learning integrating different examination modes (written, oral, project work, group assignments, etc.).
Item Q5 I modified my way of teaching according to the advises given by the students in the previous academic year.
Item Q6 In order to teach in English, I changed my methods of teaching.
Item Q7 I used multimedia materials (web, online dictionaries and encyclopedias, audiovisuals) to support my frontal lessons.
Item Q8 I prepared myself or I asked my collaborators to prepare didactic multimedia materials for the lessons.
Item Q9 I used online platforms (Moodle) not only to distribute materials to the students but also to exploit their advanced functions (*e.g.*, online forums participation).
Item Q11 I have a real passion for teaching.
Item Q12 Teaching is for me a really intense and exciting experience.
Item Q13 Theoretical concepts are more important than practical skills.
Item Q14 Active didactic methods (working in group, laboratories, etc.) facilitate the learning process.
Item Q15 Learning is a process that requires students' interaction.
Item Q16 More advanced technologies (e-learning, mobile learning, etc.) facilitate the learning process, involving and motivating students.
Item Q17 Students' opinions on teaching are a very important indicators of quality of the didactics.
Item Q18 A single exam is more appropriate than an integrated one (homework, report, etc.) to evaluate learning.

Item Q19 As students are requested to the evaluate the quality of the didactics, so teachers should be asked to give a self-evaluation.
Item Q20 It is a good practice to personalize didactics on the specific needs of single students.
Item Q21 Teaching in English is a practice that adds value to the academic didactics.
Item Q22 I would like to make students' results more consistent to the topics indicated in the syllabus of the course.
Item Q23 I would like to adapt my didactic proposal to the educational needs of the students.
Item Q24 I would like to be able to use evaluation instruments that could help me in better understanding learning outcomes.
Item Q25 It would be useful to refer to experts in didactics through a confidential mode.
Item Q26 I would be interested in participating to seminars on the didactics (how to build a syllabus, conduct a lesson, evaluate, etc.).
Item Q27 I would be interested in participating to meetings with colleagues to discuss themes relevant to the didactics.
Item Q28 I need a methodological support to integrate in my course innovative technologies for the didactics.
Item Q29 I have a real passion for research.
Item Q30 Doing research is for me an intense and exciting experience.

Information on the courses is available: average number of students attending the lessons, type of the degree, number of hours and credits, School and Department of the Athenaeum. About the teachers we know age, gender and academic role.
We were able to analyze only a small subsample of courses: those among 1,854 that did not change characteristics across the three academic years where the teacher participated also to the PRODID survey. This resulted into 605 didactic activities with the following characteristics: same teacher, same number of hours and credits in the three academic years, the teacher responded to the PRODID questionnaire. Table 1 contains descriptive statistics with reference to our data set of assessed courses.

Table 1. Descriptive statistics of the assessed courses

	Absolute frequency	**%**
Type of degree		
Bachelor	353	58.4
Master	174	28.8
5-year-long	78	12.9
School		
Agricultural Science and Veterinary Medicine	104	17.2
Economics and Political Science	32	5.3
Law	21	3.5
Engineering	144	23.8
Medicine	77	12.7
Psychology	35	5.8
Science	100	16.5
Human and Social Sciences and Cultural Heritage	92	15.3
Role of teacher		
Full	159	26.3
Associate	240	39.7
Assistant	206	34.0
Gender of teacher		
Male	372	61.5

Female	233	38.5
	Mean	**Standard deviation**
Hours	71.2	31.3
Age of teacher	50.7	8.4
Average number of students in class	62.8	48.3

2. Methodology

2.1. Latent class models

Latent class (LC) analysis provides models that consider explicitly the fact that one or more latent variables exist which are not directly observable when studying relationships between observed variables, and takes into account the categorical nature of these variables. Latent class models were introduced by Lazarsfeld and Henry (1968) to express latent attitudinal variables from dichotomous survey items, then they were extended to nominal variables by Goodman (1974a, 1974b), who also developed the maximum likelihood algorithm for estimating latent class models that serves as the basis for many software with this purpose. Later, these models were further extended to include observable variables of mixed scale type, like ordinal, continuous and counts. A traditional latent class cluster model, with one latent variable and two nominal indicators, for example, can be expressed with the following equation:

$$\pi_{abh}^{Y_1Y_2U} = \pi_h^U \pi_{ah}^{Y_1|U} \pi_{bh}^{Y2|U}$$

where $\pi_{abh}^{Y_1Y_2U}$ is the proportion of units in the five-way contingency table; π_h^U is the probability of being in latent class $h = 1,...,H$ of latent variable U; $\pi_{ah}^{Y_1|U}$ is the conditional probability of obtaining the *a-th* categorical response to item Y_1 from members of latent class h; $\pi_{bh}^{Y_2|U}$ is the conditional probability for item Y_2 .An important assumption is that of local independence, that is, given a latent class, the indicators are independent from one another.
Rejection of a traditional H-class latent class cluster model because it does not fit well, means that the local independence assumption does not hold with H classes. In such cases, a model with $H+1$ classes is fitted to the data.

2.2. The growth approach

Growth models study the development of individuals over time capturing the dependence introducing one or more latent variables (Vermunt, 2007). Basically, growth models are regression models for two-level data – time points nested within individuals – in which time enters as a predictor.

Let Y_{it} be an observed variable denoting response for unit i at time t, with $i=1,...,n$, and $t=1,...T$, equally-spaced time points. These repeated observations are regarded as imperfect measures of an underlying latent trajectory. The shape of the growth trajectory (linear, quadratic, etc.) depends on the number of latent variables specified in the model as well as on how the loadings of these latent variables change with respect to time and can be described by the following equation:

$$y_{it} = \alpha_i + \gamma_t \beta_i + \varepsilon_{it}$$

with random effects – intercept and slope - given by:

$$\alpha_i = \mu_\alpha + \zeta_{\alpha_i}$$

and

$$\beta_i = \mu_\beta + \zeta_{\beta_i}$$

The model assumes $\zeta_{\alpha_i} \sim N(0, \Psi_\alpha)$, $\zeta_{\beta_i} \sim N(0, \Psi_\beta)$ and $\varepsilon_{it} \sim N(0, \Theta_t)$. ζ_{α_i}, ζ_{β_i} and ε_{it} are mutually independent for every *i* and *t*. A usual convention for linear growth is that $\gamma_t = t - 1$. The parameters of interest are the means and variances of the random effects and the residual variances over time. This model is defined as unconditional (with no predictors) growth model. Conditional growth models, instead, not only describe but also explain growth, examining predictors of individual change over time. Predictors may be time-constant or time-varying. The random effects are, consequently, specified as

$$\alpha_i = \mu_{0\alpha} + \underline{\mu}_{1\alpha} \underline{x}_i + \zeta_{\alpha_i}$$

and

$$\beta_i = \mu_{0\beta} + \underline{\mu}_{1\beta} \underline{x}_i + \zeta_{\beta_i},$$

where $\underline{x}_i$ is a vector containing the value of time-invariant explanatory variables for individual *i*.

The latent growth mixture model assumes that the population is heterogeneous, and different subpopulations are characterized by different trajectories (Connel and Frey, 2006). The model estimates the intercept and the slope for each class and individual variation around these growth factors:

$$\alpha_{ik} = \mu_{0\alpha k} + \underline{\mu}_{1\alpha k} \underline{x}_i + \zeta_{\alpha_{ik}} \quad (1)$$

$$\beta_{ik} = \mu_{0\beta k} + \underline{\mu}_{1\beta k} \underline{x}_i + \zeta_{\beta_{ik}}, \quad (2)$$

where *k* indicates one of the *K* subpopulations with probability π_k.

A second-order growth model is also known as a growth model with multiple indicators since it allows to model the latent trajectory of a variable that is itself not directly observable (Salguiero *et al*., 2013). Let us now consider, for example, Y_{1it}, Y_{2it}, Y_{3it}, three continuous manifest variables measuring an underlying latent variable η_{it}, the structure of the factors can be considered as invariant over time imposing equality constraints on the factor loadings of the repeated measures. More generally, the value for indicator *j*, *j=1,...,J*, for individual *i* at time *t*, is given by the following expression:

$$y_{jit} = \nu_{jt} + \Lambda_{jt} \eta_{it} + \upsilon_{jit} \quad (3)$$

where Λ_{jt} denotes the factor loadings for indicator *j* at time *t*, and υ_{jit} is an error term. The latent trajectory is then specified for the latent variable η_{it}:

$$\eta_{it} = \alpha_i + \gamma_t \beta_i + \varepsilon_{it}. \quad (4)$$

In the structural equation modeling framework, (3) plays the role of the measurement component, (4) of the structural one.

Starting from the above expressions, conditional and mixed second-order growth models may be specified as in (1) and (2). In this last case, the parameters of special interest are the means and the variances of the random effects $\zeta_{\alpha ik}$ and $\zeta_{\beta ik}$, the dimensions of the mixture components π_k, the factor loadings Λ_{jt}, the regression coefficients of the k latent trajectories, the residual variance over time.

3. Results

In order to summarize the information collected with the PRODID survey, two latent class models was estimated. For the first model, the indictors are the nine binary variables Q1 to Q9, collecting information on the didactic practices adopted by the professors. The best fitting model identifies two latent classes of similar dimension but with different profiles as figures in Table 2 report. Characteristics of the teacher, gender, age and role, and of the course, School, degree, number of hours and average number of students attending, were inserted in the model as active covariates. Only the School has a statistically significant effect. The slightly larger class 1, composed by 55.86% of professors, shows that they are in general more prone to adopt innovative teaching methods. These teachers are more concentrated in the School of Economics and Political Science, Engineering, and Human and Social Sciences and Cultural Heritage. It is interesting to note that there are no significant differences in the two groups for the characteristics of the teacher, nor for degree, the number of hours and the number of students attending the course. We call this binary latent variable didactic practices (DP) since it classifies teachers with reference to their attitude to adopt teaching methods, especially more modern methods than the traditional frontal lesson and new technologies in teaching.

Table 2. Profiles of the two classes of teachers with reference to didactic practices

	Class 1	**Class 2**
Size %	55.86	44.14
Percentage of YES answers		
Q1 active involvement	99.37	68.36
Q2 external contributions	58.79	29.55
Q3 monitoring	62.26	28.23
Q4 integrated examinations	75.75	36.37
Q5 students' suggestions	89.17	72.48
Q6 English	6.61	0.81
Q7 multimedia use	85.13	44.76
Q8 multimedia prepare	58.19	19.64
Q9 online platforms	46.15	9.65
Conditional probabilities		
Agricultural Science and Veterinary Medicine	16.34	18.27
Economics and Political Science	6.63	3.60
Law	2.32	4.93

Engineering	28.55	17.79
Medicine	8.57	17.99
Psychology	5.51	6.13
Science	12.70	21.38
Human and Social Sciences and Cultural Heritage	19.39	9.91

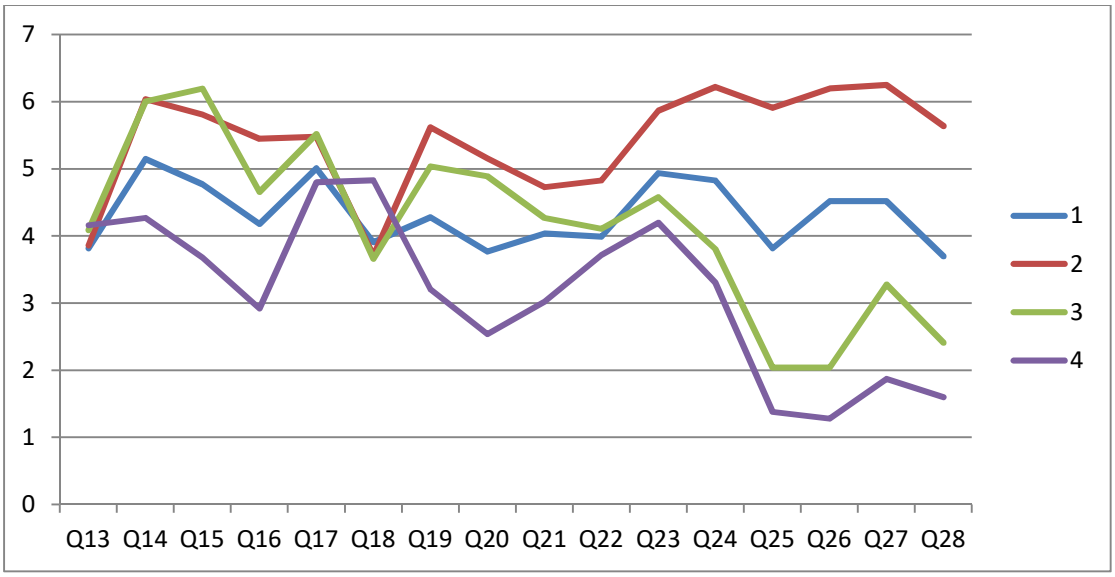
The second latent class model uses as indicators the 16 ordinal variables Q13 to Q28 describing beliefs and needs with reference to didactics of the university professors. Results are listed in Table 3 and represented in Figure 1: four classes of teachers are identified, this is the best fitting model according to the BIC index; School is the only significant covariate at course-level; age, gender and the two items Q11 and Q12 referring to passion and enjoyment for teaching are the significant covariates describing the teacher. In class 1, the largest group of professors (37.04%), almost all the indicators show average values for the items regarding beliefs, and high values for those regarding the needs that the respondents perceive to improve the quality of their didactic activity (we may define these teachers as interested in participating in events and activities related to the quality of the didactics although they do not enjoy teaching too much). With reference to the entire sample, in this class there is a higher proportion of young males, teaching a course in the School of Agricultural Science and Veterinary Medicine, Psychology and Science, the scores on items Q11 and Q12 are the lowest ones. In class 2 (35.22%), teachers are mainly female professors, this is the youngest group, teaching courses in the School of Agricultural Science and Veterinary, Medicine, Psychology, Science and Human and Social Sciences and Cultural Heritage, these professors declare the highest level of passion for teaching, they strongly believe in innovative didactic practices and are very interested in participating in activities regarding this topic (they can be defined as strong believers in the quality of the didactics). Cluster 3 (13.98) identifies professors with high scores on beliefs, medium scores on needs, teaching courses in the School of Engineering, Psychology, and Social Sciences and Cultural Heritage, males (consider the quality of the didactics important but do not feel the need to improve their teaching). Finally in cluster 4, all items received, on average, the lowest scores by the 13.75% of the respondents. This means that these teachers do not believe in the didactic practices involved in the survey nor perceive the need to acquire them. It is interesting to note that in this group we find the highest score on item Q18 that describes the preference for a single exam over an integrated one. In this cluster, we find the oldest respondents, males, teaching courses in the School of Economics and Political Sciences, Law and Engineering (traditionalist professors). This latent variable has been called beliefs and needs (BN) since it classifies teachers with reference to their opinion about teaching in higher education and on the needs they perceive in order to improve the efficacy of their didactics. It is interesting to note that passion and enjoyment for teaching show average values that are statistically different in the four segments; passion and enjoyment for research, instead, are not statistically significant, this implies that they do not have an influence in the attitude towards teaching.

Table 3. Profiles of the four classes of teachers with reference to beliefs and needs and significantly active covariates

	Class 1	**Class 2**	**Class 3**	**Class 4**
Size %	37.04	35.22	13.98	13.75
	Mean scores			
Q13 theoretical concepts	3.82	3.86	4.09	4.16

Q14 active didactic methods	5.15	6.04	6.01	4.27
Q15 students' interaction	4.77	5.81	6.20	3.68
Q16 advanced technologies	4.18	5.45	4.66	2.92
Q17 students' opinion	5.01	5.48	5.52	4.80
Q18 single exam	3.91	3.71	3.66	4.83
Q19 teachers' opinion	4.28	5.62	5.04	3.21
Q20 personalized didactics	3.77	5.16	4.89	2.54
Q21 English	4.04	4.72	4.27	3.02
Q22 consistency with syllabus	3.99	4.83	4.11	3.73
Q23 educational needs	4.94	5.87	4.58	4.26
Q24 evaluation	4.83	6.22	3.81	3.31
Q25 aid from experts	3.82	5.91	2.04	1.38
Q26 seminars on didactics	4.52	6.20	2.04	1.28
Q27 meetings with colleagues	4.51	6.25	3.28	1.87
Q28 support on new technologies	3.70	5.64	2.41	1.60
Teacher's age	49.84	49.24	53.40	53.83
Q11 passion for teaching	5.51	5.92	5.96	6.07
Q12 enjoying teaching	5.45	5.99	5.81	5.85
	Conditional probability			
Agricultural Science and Veterinary Medicine	20.58	20.54	8.95	7.83
Economics and Political Science	5.73	4.25	4.87	7.18
Law	2.95	3.37	2.77	5.84
Engineering	19.70	17.91	28.04	45.69
Medicine	13.89	17.12	28.04	45.69
Psychology	6.10	6.17	6.89	2.83
Science	19.45	13.47	18.39	14.60
Human and Social Sciences and Cultural Heritage	11.58	17.18	23.50	11.47
Male	64.76	46.02	69.38	84.34

Figure 1. Profiles of the four classes of teachers with reference to beliefs and needs.



Different growth models were estimated in order to understand if there was a significant change in the level of students' satisfaction with regards to the courses taught at the University of Padova in the academic years from 2012-2013 to 2014-2015. The best fitting model is a mixture

second-order conditional growth model with a linear latent trajectory and two classes, *i.e.*, a second-order growth model with five indicators (OA, ED, item 4, item 9 and item 11) and one underlying latent factor. Covariates for the intercept and the slope are the two latent variables estimated with the responses given by the professors to the PRODID questionnaire: didactic practices (DP) and beliefs and needs (BF). The mixture component considers the possible existence of more unobservable groups of courses with different latent trajectories describing satisfaction over time and inserting covariates. As covariates for the classes, we considered the characteristics of the course at our disposal: School, type of degree and number of hours. Table 4 reports estimation results.

The two unobservable classes of didactic activities have very different sizes. In the smallest group (4.58%, 28 courses), the average level of satisfaction is lower, 6.17, while in the largest group it is equal to 7.68.

Table 4. Mixture conditional second-order growth model: estimation results

	Class 1		**Class 2**	
	Estimate	**Standard error**	**Estimate**	**Standard error**
Class proportion	0.9542		0.0458	
Intercept				
constant	0.1016*	0.0418	-1.1692*	0.3177
DP = 1	-0.0404	0.0382	-0.5347*	0.2390
DP = 2	0.0404	0.0382	0.5347*	0.2390
BN = 1	-0.1038	0.0598	0.9840*	0.4301
BN = 2	0.0159	0.0618	-1.0374*	0.3626
BN = 3	0.1493	0.0811	-0.3202	0.4797
BN = 4	-0.0524	0.0843	0.3736	0.6778
Slope				
constant	0.0160	0.0247	-0.8797	0.2026
DP = 1	-0.0081	0.0223	0.4661*	0.1558
DP = 2	0.0081	0.0223	0.4661*	0.1558
BN = 1	0.0076	0.0344	-0.4090	0.2463
BN = 2	0.0306	0.0358	0.7447*	0.2089
BN = 3	-0.0658	0.0470	-0.3197	0.2704
BN = 4	0.0277	0.0512	-0.0160	0.3851
Covariance 2nd level				
ψ_{α}	0.3419*	0.0438	0.3419*	0.0438
ψ_{β}	0.000		0.0000	
$\psi_{\alpha\beta}$	-0.0393*	0.0147	-0.0393*	0.0147
BIC	4,801			
AIC	4,695			

* level of significance <0.001

In the smallest group of courses there is a negative and significant effect of belonging to class 1 of variable DP on the intercept, the effect is positive if the teacher belongs to class 1 of the latent variable BN and negative if the professor is classified in group 2 of latent variable BN; this means that having a positive attitude towards didactic practices or belonging to the group of professors who strongly believe in the quality of the didactics diminishes the initial level of satisfaction; while being classified as a professor not very much interested in updating didactics increases the initial level of students' satisfaction. In the largest groups of courses both variables DP

and BN do not have statistically significant effects on the intercept of the latent growth model, moreover the slope is not statistically different from 0, indicating that the average level of satisfaction in these courses has not changed over the three academic years. On the contrary, in the small group the slope is negative and statistically significant, meaning that satisfaction has decreased over time. The slope is, in this group, significantly affected by the positive attitude towards didactic practices and of belonging to class 2 of variable BN with a positive coefficient. When the teacher tends to implement didactic practices or strongly believes in the quality of the didactics, the decrease of satisfaction over time is compensated. The covariance matrix of the latent factors is assumed constant over classes: the variance of the intercept of the model is statistically significant indicating variability among the initial level of satisfaction in the sample.

Estimation of the mixture conditional growth model allows us to isolate a small group of problematic didactic activities with low judgments by students in the academic year 2012-2013 and with no improvement in the quality of the didactics in the subsequent academic years, this dynamics over time, however, might be mitigated by the implementation of innovative didactic practices, by the beliefs on this topic of teachers, their desire to improve, their passion for teaching, as the use of the data collected with the PRODID survey allows to see.

Another interesting result is that the only statistically significant covariates for the variable defining the mixture, are the number of hours of teaching and the number of students attending the lessons. In the smallest group, courses are significantly longer and with less students; there are no differences for what concerns the School, nor the type of degree.

4. Concluding remarks

In the paper we analyze the satisfaction with the didactics of the students of a large Italian university (Padova) using information collected with questionnaires to measure satisfaction over three consecutive academic years and with questionnaires to survey the didactic practices of the teachers and their beliefs and needs on this topic. Only items that were present in the questionnaire for all the three years are considered for courses that maintained the same characteristics as the number of hours of classes and the teacher; some of the teachers of these courses participated also in the survey on didactic practices: Our sample of courses is, for these reasons, not very big (650 units), however, we believe that our analyses reached some interesting and important results that should be considered by the University management.

With factor analysis, one latent construct underlying the 11 considered items, namely satisfaction with the didactic activity was estimated for each of the three consecutive academic years. Latent class analysis identified segments of teachers, homogeneous, on one side, for didactic practices adoption, on the other side, for beliefs and needs on the theme of teaching; moreover, passion for teaching was also measured and revealed to have an important effect on the quality of the didactics. These groups of teachers can be also described in terms of gender, age, role, School where teaching; this means that there are some specific categories of professors and some environments of the Athenaeum more sensible to the goal of raising the quality of the didactics.

We estimated a mixture conditional latent growth model that includes covariates that affect growth factors and class membership. The model reveals as an important instrument to identify a small group of problematic courses with a low initial level of satisfaction and, moreover, with the level of satisfaction decreasing over time in the three considered academic years. As factors affecting growth the classification of teachers in the above segments was used.

The majority of the examined courses seem to go well with an average high satisfaction at the beginning of the observation period (almost 8 in a scale from 1 to 10) that continues with the same level over time. The model manages to separate these many good courses from the small residual group of bad ones. Identifying these courses is very important for the university in order to try to improve the overall quality of the didactics. These courses are characterized by an average greater number of hours of lessons and on average less students attending. The use by the teachers

of didactic practices and their attitude towards the quality of the didactics have an impact on the trajectory over time of the levels of students' satisfaction. Specifically, the implementation of these practices and the commitment to improve the quality of teaching by the professors has an impact that mitigates the decrease in satisfaction over time for the group of problematic courses. This result suggests to the University management not only to continue to monitor the didactic practices implemented by the professors but also to stimulate them to adopt innovative methodologies, especially those based on recent technological instruments.

In general, our model reveals as a very useful instrument to identify the small group of problematic courses in a context of a very large university where, on average, students' satisfaction is high. Implementation of innovative didactic practices and commitment to quality of teaching are important factors to be encouraged by the University management. On the contrary, traditionalist way of teaching and low passion for teaching do not improve student's satisfaction.

In the Italian University system students' opinions are very much considered with reference to the quality of the didactics. Italian universities are judge by ANVUR on their commitment to collect and analyze students' opinions in order to identify critical points, define actions to improve quality and monitor quality over time. This evaluation ends with a score assigned to each Italian Athenaeum and to distribution of public funds. This paper focuses on the data collected at the University of Padova but all Italian universities may benefit from the evidences arising from the analyses.

However, the results presented in the paper may have an interest for the entire university system in Italy but also in many other contexts since the quality of the didactics is seen as a primary goal as the incredible rich literature on the topic, partially quote in this paper, shows. The paper provides some evidences of the factors affecting students' satisfaction and also a suggestion of the type of analysis that can be conducted on this type of data.

References

Aleamoni, L. M. (1999), "Student rating myths versus research facts from 1924 to 1998", *Journal of Personnel Evaluation in Education*, Vol. 13 No.2, pp. 153–166. Anand, A. and Bansan, G. (2016), "Predicting customer's satisfaction (dissatisfaction) using logistic regression", *International Journal of Mathematical, Engineering and Management Sciences,* Vol. 1, No. 2, pp. 77-88.

Apodaca, P. and Grad, H. (2005), "The dimensionality of student ratings of teaching: Integration of uni- and multidimensional models", *Studies in Higher Education*, Vol. 30 No. 6, pp. 723–748.

Bartholomew, D.J. and Knott, M. (1999) *Latent variable models and factor analysis*, Arnold, London.

Bassi, F. (2010), "Experiential goods and customer satisfaction: an application to movies", *Quality Technology & Quantitative Management,* Vol.7 No. 1, pp. 51-67.

Bassi, F., Clerici, R. and Aquario, D. (2017a) "Students' evaluation of teaching at a large Italian University: measurement scale validation", *Electronic Journal of Applied Statistical Analysis*, Vol. 10 No. 1, pp. 93-117.

Bassi F., Dias J.G. and Guerra M. (2017b), "Student's satisfaction in higher education: hoe to identify courses with low quality teaching", in Greselin, F., Mola, F. and Zenga M. (Eds.) *Cladag Book of Short Papers*, Mantova, Universitas Studiorum.

Bassi F., Grilli L., Paccagnella O., Rampichini C. and Varriale R. (2017c), "Analysis of university teaching quality merging student ratings with professor characteristics and opinions", in Greselin, F., Mola, F. and Zenga M. (Eds.) *Cladag Book of Short Papers*, Mantova, Universitas Studiorum.

Boring, A., Ottoboni, K. and Stark, P.B. (2016), "Student evaluations of teaching (mostly) do not measure teaching effectiveness", *Retrieved form Science Open*.

Braga, M., Paccagnella, M. and Pellizzari, M. (2014), "Evaluating students' evaluations of professors", *Economics of Education Review,* Vol. 41(C), 71-88.

Burdsal, C. A. and Harrison, P. D. (2008), "Further evidence supporting the validity of both a multidimensional profile and an overall evaluation of teaching effectiveness", *Assessment & Evaluation in Higher Education*, Vol. 33 No. 1, pp. 567–576.
Centra, J. A. (1998). *Development of The Student Instructional Report II*, Princeton, NJ: Educational Testing Service.
Cheung, D. (2000). "Evidence of a single second-order factor in student ratings of teaching", *Structural Equation Modeling*, Vol. 7 No. 3, pp. 442–460.
Clayson, D. E. (2009). "Student evaluations of teaching: Are they related to what students learn? A meta-analysis and review of the literature", *Journal of Marketing Education*, Vol. 31 No. 1, pp. 16–30.
Connel A.M. and Frey A.A. (2006), "Growth mixture modelling in developmental psychology: overview and demonstration of heterogeneity in developmental trajectories of adolescent antisocial behaviour", *Infant and Child Development,* Vol. 15 No.6, pp. 609-621.
Cronbach, L.J. (1951), "Coefficient Alpha and the internal structure of tests", *Psychometrika,* Vol. 16 No. 3, pp. 297-334.
Churchill, G.A. (1979), "A paradigm for developing better measures of marketing constructs", *Journal of Marketing Research,* Vol.16 No. 1, pp. 64-73.
Dalla Zuanna, G., Bassi, F., Clerici, R., Paccagnella, O., Paggiaro, A., Aquario D., Mazzuco C., Martinoia, S., Stocco, C. and Pierobon, S. (2015), *Tools for teaching assessment at Padua University: role, development and validation. PRODID Project (Teacher professional development and academic educational innovation)* (Report of Research Unit n.3). Padua: Department of Statistical Sciences, University of Padua.
Dolnicar, S., Grün, B., and Leisch, F. (2011), "Quick, simple and reliable : forced binary survey questions",*International Journal of Market Research*, Vol. 53 No. 2, pp. 231-252
Domenech, F. and Descals, A. (2003), "Evaluation of the University teaching / learning process for the improvement of quality in higher education", *Assessment and Evaluation in Higher Education,* Vol. 28 No. 2, pp. 165-178.
European University Association (2016), *Quality Culture in European Universities: a Bottom up Approach*. Brussels: EUA.
Goodman, L. A. (1974a), "The analysis of systems of qualitative variables when some of the variables are unobservable: Part I. A modified latent structure approach", *American Journal of Sociology*, Vol. 79 No. 5, pp. 1179-1259.
Goodman, L. A. (1974b), "Exploratory latent structure analysis using both identifiable and unidentifiable models", *Biometrika*, Vol. 61 No. 2, pp. 215-231.
Harrison, P., Douglas, D. and Burdsal, C. (2004), "The relative merits of different types of overall evaluations of teaching effectiveness", *Research in Higher Education*, Vol. 45 No. 3, pp. 311–323.
Hornstein, H.A. (2017), Student evaluations of teaching are an inadequate assessment tool for evaluating faculty performance, *Cogent Education,* Vol. 4 No. 1, pp. 1-8.
Lazarsfeld, P. F. & Henry, N. W. (1968), *Latent Structure Analysis*. Boston: Houghton Mifflin.
Marsh, H. W. (1987), "Student's evaluations of university teaching: Research findings, methodological issues, and directions for further research", *International Journal of Educational Research*, Vol. 11 No. 4, pp. 253–388.
Marsh, H. W. (2007), "Students' evaluations of university teaching: Dimensionality, reliability, validity, potential biases and usefulness", in Perry, R.P. & Smart, J.C. (Eds.) *The scholarship of teaching and learning in higher education: An evidence-based perspective*. New York, Springer.
Marsh, H. W. and Roche, L A. (1997), "Making students' evaluations of teaching effectiveness effective: The critical issues of validity, bias and utility", *American Psychologist*, Vol. 52 No. 11, pp. 1187–1197.
Meggiolaro, S., Giraldo, A. and Clerici, R. (2017), "A multilevel competing risks model for analysis of university students' carriers in Italy", *Studies in Higher Education*, Vol. 42 No. 7, pp. 1259-1274.

Mortelmans, D. and Spooren, P. (2009), "A revalidation of the SET37-questionnaire for student evaluations of teaching", *Educational Studies*, Vol. 35 No 5., pp. 547–552.
Nagin, D. (2005), *Group-based modeling of development*, Harvard University Press.
Rindermann, H. and Schofield, N. (2001), "Generalizability of multidimensional student ratings of university instruction across courses and teachers", *Research in Higher Education,* Vol. 42 No. 4, pp. 377-399.
Roche, L. and Marsh, W. H. (2000), "Multiple dimensions of University teacher self-concept: construct validation and the influence of students' evaluations of teaching", *Instructional Science,* Vol. 28 No. 5/6, pp. 439-468.
Saroyan, A. and Amundsen, C. (2001), "Evaluating University teaching: time to take stock", *Assessment and Evaluation in Higher Education,* Vol. 26 No. 4, pp. 341-353.Salguiero, M.F., Smith, P.W.F. and Vieira, M.T.D. (2013), "A multi-process second-order latent growth curve model for subjective well-being", *Quality and Quantity,* Vol. 47 No. 2, pp. 735-752.
Spooren, P., Brockx, B. and Mortelmans, D. (2013), "On the validity of student evaluation of teaching: The state of the art", *Review of Educational Research,* Vol. 83 No. 4, pp. 598-642.
Stroebe, W. (2016). Why good teaching evaluations may reward bad teaching: on grade inflation and other unintended consequences of student evaluations. *Perspectives on Psychological Science,* , Vol. 11 No. 6, pp. 800-816.
Svinicki, M. and McKeachie, W. J. (2013), *McKeachie's Teaching Tips: Strategies, Research, and Theory for College and University Teachers*. (13th ed.). Belmont, CA: Wadsworth.
Theall, M. and Franklin, J. (2007). (Eds.) *Student Ratings of Instruction: Issues for Improving Practice: New Directions for Teaching and Learning*. San Francisco, Jossey-Bass.
Tigelaar, D.E.H., Dolmand, D.H.J.M., Wolfhagen, I.H.A.P. and var der Fleuten, C.P.M. (2004), "The development and validation of a framework for teaching competencies in higher education", *Higher Education,* Vol. 48 No. 2, pp. 253-268.
Uttl, B., White, C.A. and Gonzalez, D.W. (2016), Meta-analysis of faculty's teaching effectiveness: student evaluation of teaching ratings and student learning are not related. *Studies in Educational Evaluation,* Vol. 54 No. 1, pp. 22-42.
Vermunt, J.K. (2007), "Growth models for categorical response variables: standard, latent-class, and hybrid approaches", in van Montfort K., Oud H. and Satorra A. (Eds.) *Longitudinal Models in the Behavioural and Related Sciences*, Mahwah, Erlbaum, pp. 139-158.
Zabaleta, F. (2007), "The use and misuse of student evaluations of teaching", *Teaching in Higher Education,* Vol. 12 No. 1, pp. 55–76.