

Student Perception of Effective Teaching – Between Retired Professors and Contract Instructors

Junaid B. Jahangir

The final publication is available at Springer via

<http://dx.doi.org/10.1007/s43545-021-00244-0>

Permanent link to this version <https://hdl.handle.net/20.500.14078/3032>

License All Rights Reserved

Student perception of effective teaching – Between retired professors and contract instructors

Abstract: The objective in this paper is to discern how student comments at the Rate My Professors website distinguish between retired professors and contract instructors in Economics. A qualitative content analysis is used to investigate whether student comments capture effective teaching, as depicted in the academic literature and whether teaching pedagogy has shifted from critical thinking and challenge to easy expectations and easy grades, as part of the corporatization of education.

Keywords: Rate My Professors; effective teaching; easy As and corporate education; teaching Economics

Junaid Jahangir

Assistant Professor of Economics

MacEwan University

7-368

10700 104 Ave

Edmonton AB T5J 4S2

Canada

junaid@ualberta.ca

Student perception of effective teaching – Between retired professors and contract instructors

1. Introduction

Hollywood's depiction of inspirational instructors includes Mr. Hundert in "The Emperor's Club" and Mr. Keating in "Dead Poet's Society." The former is recognized for emphasis on discipline and transfer of values, whereas, the latter is noted for pushing boundaries and thinking outside the box. However, their methods of calling out a student to enforce discipline (Hundert and Bell) or placing stress by pushing one to think on the spot (Keating and Anderson) are not necessarily viewed as effective teaching. The literature on effective teaching includes critical thinking and challenge, clarity and organization, enthusiasm and passion, interaction and active engagement, humour and story-telling, and subject mastery and real life examples. It also includes approachability and avoiding being disrespectful or having unrealistic expectations.

It is interesting to note that based on the Rate My Professors (RMP) website, the average overall rating for retired University of Alberta professors is lower than that received by current contract instructors in Economics (3.58 compared to 3.95). **Apart from bias because of age (Wilson et al. 2014; Joye and Wilson 2015)**, this difference could be due to the former's emphasis on the Keating/Hundert approach or the latter projecting easy expectations and easy grades that arise through the corporatization of education (Marginson and Considine 2000; Mazzarol, Soutar and Seng 2003; Felton, Mitchell and Stinson 2004), as they are more vulnerable to corporatization than current full time tenure

track faculty. If that is the case then, based on the literature, neither approach constitutes effective teaching. Alternatively, retired instructors could have upheld rigorous aspects of effective teaching like critical thinking and challenge, and contract instructors could be favouring palatable aspects like humour and story-telling. In this case, both groups of instructors are noted for effective teaching.

The focus in this paper is on using content analysis to document the differences in teaching pedagogy between retired professors and contract instructors in Economics based on students comments at RMP website. The idea is to let student student perception and experience guide what is viewed as effective teaching and check whether that view is consistent with the academic literature on effective teaching or whether it reflects the corporatization of education. The focus in this paper is also on checking whether retired professors are noted for the rigorous aspects of effective teaching or other factors and whether contract instructors are noted for the palatable aspects of effective teaching or easy expectations and easy grades. More formally, three questions are addressed in this paper, which include checking whether student comments capture effective teaching, whether retired professors retained rigorous academic standards, and whether contract instructors have lowered teaching standards. Addressing these questions allows discerning whether teaching pedagogy has shifted from critical thinking and challenge to easy expectations and easy grades, as part of the corporatization of education.

This paper is divided into seven sections. The next section delves into a select literature review on justifying the use of ratings and comments from the RMP website, the

determinants of teaching effectiveness and issues in teaching evaluations pertaining to the corporatization of education and easy expectations. The third section describes methodology on the data collected and the analytic strategy employed. The fourth section provides a preliminary data analysis. The fifth section delves into the coding process of content analysis. The sixth section uses content analysis and the difference in means method to address the three questions raised in this paper. The final section offers concluding remarks on the results.

2. Literature Review

2.1 Using the Rate My Professors data

For the purposes of this paper, student comments from RMP are used because, unlike official university student evaluations of instructors, they are publicly available. This is not much of a problem, as RMP ratings are strongly correlated with formal university evaluations (Albrecht and Hoopes 2009; Timmerman 2008). Specifically, the strong correlation is based on the question “overall, how would you rate the instructor” (Coladarci and Kornfield 2007). Otto and Sanford (2008) even argue that RMP ratings could be a useful supplement to teaching evaluations. While Legg and Wilson (2012) suggest that students on the RMP website have a negative bias and therefore are not representative of classes, Bleske-Reчек and Michels (2010) indicate that RMP ratings are moderate in tone as opposed to ranting and raving and tend to be more positive than negative. As such, while students who have a strong like or dislike for an instructor are more likely to leave ratings at the RMP website (Sen, Voia and Woolley 2010) thereby biasing the results, this sample selection bias is also true for formal teaching evaluations.

Therefore, using data from the RMP website does not pose an additional problem compared to formal teaching evaluations.

2.2 Determinants of effective teaching

There is no consensus on the definition of effective teaching (Spooren et al., 2013). Based on teaching evaluations, Boex (2000) found that effective teaching includes clarity, presentation skills, the ability to motivate students, and that adopting a demanding stance affects these evaluations. However, Braga et al (2014) conclude that effective teachers require students to exert effort and Becker (2000) suggests avoiding structured tests that fail to challenge students. Effective teaching includes subject mastery (Biggs and Tang, 2011), enthusiasm and approachability (Akerlind 2007). It also includes motivating by real life examples and current issues (Reimann 2004), conveying relevance and telling a good story (Sheridan, Hoyt and Imazeki 2014) and humorous anecdotes and visual illustrations (Dunn and Griggs 2000). In contrast, poor teaching is based on being disrespectful, having unrealistic expectations, and not being knowledgeable (Busler et al. 2017).

There is consensus in the literature on the effectiveness of active learning, which involves listening, writing, reading and reflecting (Meyers and Jonbes 1993). It also includes emphasis on critical skills (Pozo-Munoz et al. 2000), on discovering and constructing knowledge (Barr and Tagg 1995) and on collaborative problem solving (Ongeri 2009). Such interactive teaching is also deemed effective in intermediate and higher-level Economics classes (Hervani and Helms 2004, 267). Similarly, Bonwell (1992) indicates that students must talk, write and make connections with daily lives, as they do

not learn much by memorization. Moreover, students expect to be engaged instead of being passively lectured (Becker 2000) and giving them notes or PowerPoint slides may result in poor attendance and very little engagement (Sheridan, Hoyt and Imazeki 2014; Hamermesh 2019).

However, Andreopoulous and Panayides (2009) found that even the best students, those with a GPA greater than 3.5, liked lecture-based instruction. According to McKeachie (1997) students want organized notes and preparation solely geared towards exams. Similarly, Jordan (2011) stated that student learning-philosophy could be based on memorization and fact acquisition instead of understanding concepts and interpretation and therefore dependent on passive reception for learning instead of active engagement.

In addition to students not valuing effective teaching there are concerns whether students can evaluate instruction quality in the first place, as they are neither trained observers nor privy to instructor pedagogy (Braskamp et al. 1981). However, Lattuca and Domagal-Goldman (2007) mention that a considerable body of research on teaching evaluations finds that students are good judges of clarity, preparation and organization but not content. Likewise, Feldman (2007) and Pan et al. (2009) repudiate claims that students lack maturity to evaluate instruction quality.

In summary, the literature indicates that effective teaching includes critical thinking and challenge, clarity and organization, enthusiasm and passion, interaction and active engagement, humour and story-telling, and subject mastery and real life examples. It also

includes approachability and avoiding being disrespectful or having unrealistic expectations. The literature also indicates that students may not value effective teaching when they prefer passive reception of lectures instead of active engagement, adopt memorization instead of understanding concepts, and want organized notes and preparation solely for exams, all of which indicates that students may not appreciate challenge. Although, the literature indicates that while students are not privy to teaching pedagogy, they are able to discern aspects of effective teaching including clarity and instructor preparation.

2.3 Issues in teaching evaluations – Corporatization and easy expectations

The corporatization of higher education in many countries by a shift towards a business-oriented model of operation has been well noted (Marginson and Considine 2000; Mazzarol, Soutar and Seng 2003). There is some evidence that increasingly students with modest academic profiles have been admitted (Reimann 2004) and such students, who use a surface approach to education including rote learning and memorization, rationalize that education is a commodity to be bought (Biggs and Tang 2011, p. 4-6). Generally, to the extent that administration wants to retain students, associates teaching effectiveness with higher teaching evaluations and exclusively relies on them, effective teaching may be reduced to satisfying students. The literature shows that teaching evaluations primarily capture student satisfaction (Abrami et al. 2007; Beecham 2009), which does not necessarily imply learning (Beckman and Stirling 2005).

In a corporate model, faculty members, especially contract instructors or those without tenure, who are concerned about potential contract renewal or promotion, would have the incentive and/or pressure to give easier exams, contribute to grade inflation and generally “dumb down” instruction material. Hornstein (2017) notes that teaching evaluations pressurize faculty members to not rock the boat and to not push undergraduate students to maximize their intellectual potential. Therefore, given pressure to win student approval through an easy A, academic standards may have fallen (Felton, Mitchell and Stinson 2004).

The literature suggests that teaching evaluations are problematic if students simply care for grades (Hornstein 2017) or perceived workload (Marsh 1987). Instructors perceived as easier receive higher evaluation scores (McKeachie 1997), indicating student preference for easy classes (Miller 2006). Specifically, in the case of instructors with low ratings, students indicate concerns on workload, grade distribution and teaching practices, whereas, majority of the instructors rated highly are described as “nice,” “easy,” “cool,” “caring,” “laid back,” and “understanding,” which suggests that they are liked for their easy expectations apart from personalities (Felton, Mitchell and Stinson 2004). The literature shows that high teaching evaluations may not be capturing teaching effectiveness but rather dumbing down course material (Becker 2000), personality and entertainment skills (Jordan 2011), easy grading and light course load (Greenwald and Gillmore 1997) and popularity (Aleamoni 1987; Coleman and McKeachie 1981).

However, Feldman (2007) and Pan et al. (2009) suggest that it is untrue that teaching evaluations are popularity contests. Likewise, Marsh and Ware (1982) indicate that entertaining instructors do not necessarily receive higher ratings. Major reviews have consistently shown that teaching evaluations are unaffected by grading leniency and workload, as students learn through challenge and commitment and devalue learning if success is easy due to low workload (Marsh and Roche 2000). According to Theyson (2015), the easiness and quality relationship may be complex as both extremely hard and extremely easy may be viewed as low quality. Therefore, instructors cannot get higher evaluations by offering higher grades and easier courses (Marsh and Roche 2000).

In summary, there are concerns that with the corporatization of education, effective teaching may be reduced to student satisfaction, which is less about learning and more about expected grades and perceived workload. With corporatization, education becomes a commodity to be bought, and instructors feel pressure to give easier exams and higher grades. This suggests that student comments may not capture effective teaching but easy grading and light course load. However, the literature also shows that instructors do not receive higher ratings by offering entertainment, higher grades and easier courses.

In terms of this paper, this literature review justifies using RMP student comments and shows the various aspects of effective teaching. There is, however, difference in the literature on whether student comments capture effective teaching or easy expectations and easy grades. This allows to investigate whether RMP student comments on retired professors and contract instructors capture effective teaching, as identified in the academic

literature, and whether teaching pedagogy has shifted from critical thinking and challenge to easy expectations and easy grades, as part of the corporatization of education.

3. Data and Methods

3.1 Data

This paper is based on student comments on University of Alberta Economics instructors. Information on current full time tenure track Economics faculty is suppressed because their average quality rating of 3.50 and average difficulty level of 3.10 are similar to those of retired professors. Additionally, they are not as pressured by the corporatization of education as contract instructors. Therefore, the focus will be retained on comparing student comments between retired professors and contract instructors to investigate whether they capture effective teaching and whether teaching pedagogy has shifted from critical thinking and challenge towards easy expectations and easy grades. The focus in this paper is on Economics because instruction in this field has been consistently ranked amongst the lowest due to various reasons. These include the requirement of Economics for many students from diverse disciplines (Ongeri 2009), the perception of the subject as boring with little real-life application (Ghosh 2013), fast paced lectures (Reimann 2004), low grades in Economics classes (Cashin 1990), and instructors emphasizing lecture-based teaching instead of active learning methods (Becker and Watts 2001).

The RMP website provides ratings on 14 retired professors and 13 contract instructors at the University of Alberta in the department of Economics. This comprises of a small sample, which makes generalization problematic, but it allows for in-depth

understanding (Bengtsson 2016) to address the questions raised in this paper. Table 1 shows the eight questions that students answer to rate the instructor. The chief question is on instructor quality and is based on the question “How would you rate this professor as an instructor?” This question is answered on a 1 – 5 scale, which is associated with three emoticons - green for awesome (4.0, 5.0), yellow for average (3.0) and red for awful (1.0, 2.0).

The question on difficulty level is also answered on a 1- 5 scale, which gauges the range between an easy A and working hard. The other four questions are answered in a binary fashion and include whether the student would take another class with the instructor, if the course was taken for credit, whether the textbook was used and if attendance was mandatory. Additionally, there are two questions that pertain to the course code and the grade received by the student. Finally, there is space for a comment that is based up to a maximum of 350 characters.

Student ratings and comments were collected from 2002 – 2018 for retired professors, who received a total number of 235 comments. To keep the data manageable and comparable, the ratings and comments for contract instructors were confined within the 2015 – 2019 range, which yielded a total of 370 comments. A slightly later time period is used for contract instructors, **who are a mix of PhD and non PhD instructors with some experienced and other new and who are hired semester to semester or on multi-year contracts**, to allow for the maturation of their teaching pedagogy, otherwise, student

comments end up capturing the mistakes made by instructors in the novice years of their teaching careers.

Data on instructor demographics including ethnicity, accent and gender is based on the author's knowledge based on work experience and is summarized in Table 2 for both instructor categories. This table also includes average instructor quality, average difficulty level and the average grade received by students. Additionally, Table 2 includes data on the percentage of students that would take the course again with the instructor, students in 101/102 introductory level courses, students who deemed attendance mandatory and students who used the text. A preliminary analysis of Table 2 is offered in Section 4.

3.2 Method

The emphasis in this paper is on qualitative analysis because student comments are often not systematically analyzed even as they offer more nuanced and valuable information on improving teaching effectiveness (Santhanam and Jones 2018). These comments reveal what students really feel and think and reveal issues and preferences common to students through word patterns (Jordan 2011). For the purposes of this paper, a content analysis of student comments is undertaken. Generally, content analysis allows making inferences from written data to describe and quantify specific phenomena (Downe-Wambolt 1992). It allows reducing a volume of texts in a systematic fashion by the process of coding, where a code is often an essence capturing word or phrase. These codes are then collapsed by grouping into categories and themes. This form of analysis comprises of both a qualitative and a quantitative methodology, where in the case of the latter, textual facts

are expressed as percentage of key categories (Bengtsson 2016). The idea is to convert written data into quantitative form for systematic drawing out of categories and themes.

In terms of methodology, this paper has some parallels with Silva et al. (2008) that applied content analysis to RMP student comments. They used coding of student comments to obtain 42 categories that were placed in broad category clusters including “instructor characteristics”, “student development” and “course elements”. Their category of “course elements” included items like complexity of exams and fairness of grading. They found that instructor characteristics, based on items like enthusiasm, clarity and organization, generated the most positive and negative comments, whereas, student development, based on items like learning, generated the least comments.

The method employed in this paper is inductive in that student perceptions and experience, evident in the comments, guided the coding process. For this purpose, overall 605 comments, based on 235 comments on retired professors and 370 on contract instructors, were coded. Essentially, the codes were based on keywords and phrases from the comments that depict a certain theme, such as “easy A” or “hard exam”. Hundreds of codes were extracted and a pattern began to emerge that allowed categorizing these codes into 18 category clusters, some of which were larger than others. Overall 1,731 codes were categorized and the percentage of codes that fell in each of the 18 categories was calculated, which allowed using the difference of means tests for subsequent analyses. This coding process is illustrated in Section 5.

4. Preliminary Data Analysis

Table 2 provides information of 14 retired professors and 13 contract instructors in Economics. It indicates that the average instructor quality is lower (3.58/5) and the average difficulty level is higher (3.32/5) for retired instructors compared to the average quality (3.95/5) and difficulty level (2.69/5) for contract instructors. It also indicates that retired professors are mostly white and native speakers of English (92.86%). In contrast, while 69.23% of contract instructors are white, only 53.85% are native English speakers. Another distinguishing feature is that only 7.51% of the student ratings are for 101/102 level classes for retired professors, compared to 32.87% for contract instructors.

All of this indicates that the lower quality of retired professors may be driven by higher difficulty level, even as they are predominantly white, native English speakers, who didn't teach many introductory Economics classes; factors that are usually associated with higher instructor ratings. Similarly, the higher quality of contract instructors seems to be driven by lower difficulty level, even as many are ethnic, whose native language is not English and who have taught relatively more introductory classes; factors that are usually associated with lower instructor ratings. The bias in student rating based on ethnicity (Boatright-Horowitz and Soeung 2009; McPherson and Jewell 2007; Smith 2007), accent (Ogier 2005), and large introductory classes (Pienta 2017) is well documented.

The above findings motivate the analysis of student comments to discern whether the lower average quality of retired professors is explained by critical thinking and

challenge or other factors and whether the higher average quality of contract instructors is explained by easy expectations and easy grades or effective teaching.

5. The Coding process

The coding process for content analysis proceeded in two steps. In the first step, codes were extracted based on keywords, phrases and themes that emerged from student comments. In the interest of retaining focus on teaching style, course complexity and student learning, extraneous information was ignored. This focus was shaped and confirmed as the coding process continued for the 605 comments and a pattern on key themes began to emerge from the student comments. Sometimes, previous coding were revisited to ensure consistency with the coding process. In the second step, the codes were grouped into 18 larger category clusters, which were based on the overall themes that emerged from the student comments. This coding process can be illustrated by the following two comments, respectively for a retired professor and a contract instructor, who were coded as R10 (from R1 – R14) and C11 (from C1-C13).

“Unbelievably boring and very inarticulate in a lecture setting. It's somewhat frustrating to listen to him talk in class. However, the group project is easy and interesting; mid term and final are both easy if you read the readings and attend class. If you have an interest in the world outside North America, I would recommend this class”

“He kept the class interesting and made sure everyone understood the concepts before moving on. You need to attend class because he does not post any notes online. His exams are ok as long as you take good notes in class and do his practice exams.”

The first comment indicates that the student found the instructor boring and inarticulate. The comment then indicates that the project and exams were easy contingent on reading the material and attending class. Retaining focus on teaching style, course complexity and student learning for the purpose of coding, the keywords/phrases allowed for the codes of ‘boring’, ‘inarticulate’, ‘easy exams’, ‘attend class’, and ‘read material’. Any extraneous information in this comment was ignored.

Similarly, the second comment indicates that the student found the class interesting and that the instructor ensured understanding. The comment then indicates that it was important to attend class because notes were not provided. Additionally, the student emphasizes to take notes and to do the sample exams. There is less extraneous information in this comment compared to the first. For the purpose of coding, the keywords/phrases allowed for the codes of ‘interesting’, ‘ensured understanding’, ‘attend class’, ‘take notes’, ‘do sample exams’.

The above is based on the first step in the coding process. In the second step, these codes were grouped into 18 category clusters that were based on the overall themes that arose from coding the 605 comments. The codes ‘boring’ and ‘inarticulate’, in the first comment and ‘interesting’ and ‘ensured understanding’ in the second comment comprise

of distinct categories. They were respectively assigned to the broad categories of 'Boring/Slide reader', 'Unclear/Tangents/Accent', 'Interaction/Interest/Passion', and 'Critical skills/Learning'. However, the codes 'attend class', and 'read material' in the first comment and 'attend class' and 'take notes', in the second comment seem as if they could be further collapsed.

Nonetheless, based on the iterative process of revisiting coding, it was decided to retain **each of** these codes **instead of combining them**, as this allows to capture student intensity for these factors. What this means is that the student in the first comment emphasizes that you have to 'attend class', and 'read material', which are capturing two different facets of the broad category 'Attend/Take notes/Work hard'. Collapsing both of them into one code would fail to capture both facets and therefore fail to capture student intensity of preference against having to put in effort for the class. Similarly, the student in the second comment underscores that you have to 'attend class' and 'take notes.' Both these facets were included in the broad category 'Attend/Take notes/Work hard'. Thus, assigning both codes to this category captures student intensity of preference against having to put in effort for the class.

Finally, the code 'easy exams' in the first comment and 'do sample exams' in the second comment were both assigned to the category 'Easy A/Pattern/Fair'. A value judgment was made in the construction of this category as student comments seemed to suggest that exams based on sample exams were easy, fair or offered a boost to the GPA. As an illustration of this association between patterned exams and fairness or easiness,

consider for instance the following comments for contract instructors coded as C1, C4, and C5 and retired professor R8.

“All the exams were very similar to his previous exams ... Overall the course is a GPA booster”

“his exams were super fair, almost identical to practice exams.”

“Super easy exams. Very much like the practice questions and previous exams.”

“If you like to be pampered, then take this class ... tests directly from her notes.”

This two step method of coding and categorizing involved iterations of revisiting coding and included value judgment in grouping codes for categorizing, a process that was undertaken for all 605 comments. While the 18 categories, which can be seen in Table 3, were induced from the preponderance of key themes in student comments, several of these categories were also identified in the literature review on effective teaching and in the context of issues in Economics instruction.

To recapitulate, the literature review above indicated that effective teaching includes critical thinking and challenge, clarity and organization, enthusiasm and passion, interaction and active engagement, humour and story-telling, subject mastery and real life examples, and approachability and avoiding being disrespectful or having unrealistic expectations. It also showed that students may not value effective teaching when they prefer passive reception of lectures, want organized notes and preparation solely for exams, thereby avoiding challenge. The literature also highlighted that issues in Economics

instruction include the perception of the subject as boring with little real-life application, fast paced lectures, low grades and instructors avoiding active learning methods.

As such, the categories that are substantiated from the literature review on effective teaching include ‘Mastery’, ‘Critical skills/Learning’, ‘Interaction/Interest/Passion’, ‘Fun/story’ and its converse ‘Boring/Slide reader’, ‘Clarity/Organization’ and its converse ‘Unclear/Tangents/Accent’, ‘Personality/Helpful/Available’ and its converse ‘Rude/Personality’, ‘Easy A/Pattern/Fair’ and its converse ‘Tough/Grading/Exams’, ‘Notes given/Skip/No text,’ and its converse ‘Attend/Take notes/Work hard’, and ‘Tech’ for technology use. The literature on issues in Economics instruction offered the coding categories of ‘Fast pace’, ‘Math’ and ‘Real world relevance’. Finally, student comments also allowed for an additional category of ‘Text expensive/bad’.

Five categories ‘Easy A/Pattern/Fair’, ‘Attend/Take notes/Work hard’, ‘Personality/Helpful/Available’, ‘Clarity/Organization’ and ‘Fun/story’ comprised of the largest amount of codes and each had a converse category so that the positive and negative codes were separated for each category pair. On the other hand, the other categories are much smaller and therefore, in the case of the categories of ‘Mastery’, ‘Critical skills/Learning’, and ‘Real world relevance’ the positive and negative codes were included in the same category to simply focus on whether student comments captured these categories. While there are many codes, the key codes assigned to each of the 18 categories and the percentage of codes in each category are presented in Table 3.

6. Content Analysis

Overall, 1,731 words/phrases extracted from the 605 student comments as codes were grouped into 18 categories. Based on this assignment, the percentage of codes in each category was calculated to determine the relative importance of that category in student comments. For instance, there are 261 codes out of a total of 1,731 (15.08%) in the ‘Easy A/Pattern/Fair’ category, but only 69 (3.99%) codes in the ‘Critical skills/Learning’ category. These percentages facilitate the use of the difference in means method, which allow answering the three questions raised in this paper as follows.

6.1 Do student comments capture effective teaching?

The first question is motivated from the literature on the corporatization of education, which suggests that effective teaching may be reduced to student satisfaction that is less about learning and more about expected grades and perceived workload. This necessitates investigating whether students comments capture effective teaching as depicted in the academic literature or whether they capture easy expectations and easy grades that are part of the corporatization of education.

Table 3 sheds some light in answering this question. It shows that the largest categories include ‘Easy A/Pattern/Fair’ (15.08%) and ‘Attend/Take notes/Work hard’ (13.63%). If these categories are combined with their converse categories ‘Tough/Grading/Exams’ (10.86%) and ‘Notes given/Skip/No text’ (6.12%) then this indicates that the main concerns in student comments revolve around grades (25.94%) and work load (19.76%). Combining the other three categories with their converses indicates

that after grades and work load, student concerns are about clarity (13.23%), instructor personality (11.67%) and a fun approach to instruction (11.44%). Similarly, combining the categories of 'Mastery' and 'Critical skills/Learning' to focus on student learning (7.11%) and combining 'Interaction/Interest/Passion' and 'Real world relevance' to focus on active learning (6.99%) indicates that concerns on student development are much smaller compared to concerns on grades and workload. Other categories are negligibly small.

Overall, the findings of this paper indicate that student comments revolve around grades and workload (45.70%), then instructor personality and approach (36.34%), followed by student development (14.11%). These findings are somewhat similar to the results of Silva et al. (2008), which found that instructor personality and approach generated more positive and negative comments compared to student development. The findings of this paper indicate that student comments capture palatable aspects of effective teaching that include clarity and organization, enthusiasm and passion, interaction and active engagement, humour and story-telling, relatively more than the rigorous aspects of effective teaching that include critical thinking, challenge, and subject mastery. However, the largest categories capture a preoccupation with grades and workload, which reflects the disposition of easy expectations and avoiding challenge. In short, student comments capture concerns of easy expectations and easy grades that are part of the corporatization of education, but they also capture many aspects of effective teaching, although relatively more of the palatable aspects than the rigorous aspects of effective teaching.

6.2 Did retired professors maintain rigorous standards?

The second question is motivated by the observation that the average quality rating and the average difficulty level for retired professors are respectively 3.58 and 3.32 compared to 3.95 and 2.69 for contract instructors. This necessitates investigating whether the lower average quality rating and the higher average difficulty level for retired professors are explained by the rigorous aspects of effective teaching or other factors including the Keating/Hundert approach of placing students under stress or enforcing strict discipline. Rigorous in this context does not mean using a teacher centred approach where the student is expected to do all the work but an emphasis on critical thinking and challenge that arise through subject mastery, and where emphasis is on concept-based teaching instead of rote learning and on exams that test learning instead of the ability to regurgitate information.

Table 4 provides some focus in answering this question. It shows the application of a difference of means test on the percentage of codes between the 14 retired professors and 13 contract instructors for each of the 18 category clusters. It indicates that retired professors receive statistically significantly more codes in the ‘Mastery’ and ‘Rude/Personality’ categories and contract instructors receive statistically significantly more codes in the ‘Easy A/Pattern/Fair’ and ‘Notes given/Skip/No text’ categories. This suggests that where contract instructors are distinguished for the provision of lecture notes and easy or lecture based patterned exams, retired instructors are noted both for subject mastery and unfavourable personalities.

To conduct a sensitivity analysis for this result, the difference of means test was conducted multiple times. To do so, the 6 top rated retired professors, those who received a quality rating greater than the mean rating of 3.58, were separated from the 8 lower rated professors. Similarly, the 5 top rated contract instructors, those who received a quality rating greater than the mean rating of 3.95, were separated from the 8 lower rated instructors. The results of the difference of means test, which are suppressed for space considerations, indicate that the Top 6 retired professors are significantly noted for their subject mastery compared to the Bottom 8 contract instructors, and that the Bottom 8 retired professors are noted for unfavourable personalities compared to the Top 5 contract instructors.

Table 5 continues with this sensitivity analysis. The results respectively indicate that the Bottom 8 retired professors are noted for boring instruction methods and unfavourable personalities when compared to the Top 6 retired professors, that there is no statistically significant difference between the Top 6 retired professors and the Top 5 contract instructors, and that the Bottom 8 retired professors are noted for unfavourable personalities when compared to the Top 5 contract instructors. The results of the sensitivity analysis can be presented as follows.

Top 6 retired professors/ Bottom 8 contract instructors: Subject Mastery

Top 6 retired professors/ Top 5 contract instructors: No statistically significant difference

Bottom 8 retired instructors/ Top 5 contract instructors: Unfavourable personalities

Bottom 8 retired instructors/ Bottom 8 contract instructors: Unfavourable personalities

Bottom 8 retired instructors/ Top 6 retired professors: Boring instruction and Unfavourable personalities

Collectively, this analysis indicates that top retired instructors are noted for subject mastery and lower rated retired instructors are noted for boring instruction and unfavourable personalities. However, sometimes the same retired professor is noted for both aspects, as evident through the following snippets of comments on retired professor R9.

“one of those memorable profs that truly challenges you but makes you know the material entirely. It was a difficult class ... but I never once regretted taking it.”

“Spare yourself and your GPA ... This guy was a bit obnoxious and rude to the students.”

Retired professor R9 is also noted for the Keating/Hundert approach as evident through the comment that he “throws chalk at you if you sleep in his class. If you did not read the case before coming to class (he'll ask you a question) then you have to leave.”

Comments on other retired instructors substantiate the finding that retired instructors are noted both for the rigorous aspects of effective teaching and for other factors like unfavourable personalities that sometimes include the Keating/Hundert approach of placing students under stress or enforcing discipline. Table 6 shows a snapshot of

comments, based on high and low ratings of retired instructors, that capture rigorous aspects of effective teaching as well as unfavourable aspects of retired professors.

The rigorous aspects of effective teaching include having a “bear-trap grip on the topic”, that “if you take a genuine interest”, the instructor would “bend over backwards for you” despite the course being challenging, and that the the instructor is able to spark interest by showing “the beauty of economics” through “currents events to explain economic theories” and by teaching and pushing “the class to think”. The unfavourable aspects of retired professors include acknowledging that while the instructor is “knowledgeable” “it was hard to take anything away from the lectures”, that the instructor is “unbelievably boring and very inarticulate” even if the “mid term and final are both easy”, and that the instructor “can lose students in his meticulous notation”, thereby making “things more complicated than they have to be”.

To recapitulate, the difference in means tests and the salient comments on retired professors indicate that they are distinguished for both the rigorous aspects of effective teaching and other unfavourable aspects including the Hundert/Keating style of placing stress or enforcing discipline. While a distinction was made between top rated professors, noted for subject mastery and challenge, and low rated professors, noted for boring pedagogy, sometimes the same instructor is noted for both aspects in teaching. This means that while retired professors were noted to have upheld rigorous standards, they were also noted to have unfavourable aspects that did not align with effective teaching as depicted in the literature.

6.3 Do contract instructors lower teaching standards?

The third question is motivated by the observation, as noted earlier, that the average quality and average difficulty level are respectively higher and lower for contract instructors compared to retired instructors. This necessitates investigating whether contract instructors are noted for the palatable aspects of effective teaching or whether they lower teaching standards through easy expectations and easy exams. As in the previous section, a difference in means test in Table 4 and the sensitivity analysis in Table 5 helps provide a focus to addressing this question.

Table 4 indicates that contract instructors receive statistically significantly more codes in the ‘Easy A/Pattern/Fair’ and ‘Notes given/Skip/No text’ categories. This indicates that contract instructors are distinguished for the provision of lecture notes and easy or lecture based patterned exams. For sensitivity analysis, a difference of means test, whose results are suppressed for space considerations, revealed that the Top 5 contract instructors are noted for easy or patterned exams, a fun approach to teaching and the provision of notes when compared to the Bottom 8 retired professors.

Table 5 continues with this sensitivity analysis. The results indicate that the Top 5 contract instructors are noted a fun approach to teaching when compared to the Bottom 8 contract instructors, and that there is no statistically significant difference between the Top 5 contract instructors and the Top 6 retired professors. Other comparisons do not specifically distinguish contract instructors. The results of the sensitivity analysis can be presented as follows.

Top 5 contract instructors/ Bottom 8 retired professors: Fun approach to teaching, Easy or patterned exams, and Provision of notes

Top 5 contract instructors/ Bottom 8 contract instructors: Fun approach to teaching

Top 5 contract instructors/ Top 6 retired professors: No statistically significant difference

Collectively, this analysis indicates that while contract instructors are noted for easy or patterned exams and provision of notes, the top rated contract instructors are also distinguished for a fun approach to teaching. This means that contract instructors are known for the palatable aspects of effective teaching like fun and story telling but also for easy expectations and easy or patterned exams. The salient comments on contract instructors, based on high and low ratings in Table 7, reveal both these aspects. As an example, the following comment received by the highest rated contract instructor C1 captures both easy grades and a fun approach to teaching/learning.

“This was the only lecture I used to look forward to every MWF. The prof was funny and the connections between course material and real life made understanding content easy. All the exams were very similar to his previous exams ... Overall the course is a GPA booster and the prof made class exciting.”

In another comment, instructor C2 is noted for effective teaching but that again is coupled with an emphasis on easy exams.

“My stats/calc background was pretty weak but he basically retaught me everything and made it easy to understand. Exams are surprisingly easy and he'll give you plenty of help ... it's so easy with Professor.”

Sometimes this emphasis on easy exams or grades comes at the expense of a fun approach, which is indicated by the comment that even as the “lecture is not interesting ... he would like to give student high grades.” This predominate concern with grades is also manifest in the comment that the instructor “wants his students to succeed to the point where he is shifting the curve over 10% to benefit everyone's GPA”. It is also clear that low rated instructors are noted for failing to “prepare you for the actual tests” and that “some questions feel unfair, as you can't solve them even if you did study off the notes.” Comments on other low rated instructors indicate that the instructor “gave us a super easy practice midterm, and made the actual exam hard”. Moreover, low rated contract instructors are noted for being “boring”, an issue which get compounded when the “lecture is mandatory”.

Overall, the comments on both highly rated and low rated contract instructors indicate a predominate concern with easy or patterned exams and easy grades. However, comments on highly rated contract instructors also highlight a fun approach to teaching and sometimes these comments allude to learning. This substantiates the findings from the difference of means tests that distinguish contract instructors on the basis of a fun approach to teaching, easy or patterned exams, and provision of notes. This means that while contract instructors may have lowered rigorous standards through easy exams, they are also noted

to have upheld teaching standards through palatable aspects of effective teaching as depicted in the literature.

7. Conclusions

This paper was motivated by the observation of lower average quality ratings and higher average difficulty levels of retired professors compared to contract instructors, based on the RMP ratings in Economics at the University of Alberta. It was also motivated by the observation in the literature on the corporatization of education, which affects contract instructors more than current full time tenure track faculty. The focus, therefore, in this paper was retained on comparing retired professors to contract instructors, who are starkly different from them, than to current full time tenure track faculty that are similar in terms of average quality and average difficulty level.

A quantitative linear regression based on the RMP data does not provide as much of an indepth scrutiny that is possible with a qualitative analysis of student comments. This facilitated investigating whether student comments capture effective teaching or easy expectations, whether retired professors upheld rigorous standards, and whether contract instructors have lowered teaching standards. The objective was to determine whether teaching pedagogy has shifted from critical thinking and challenge towards easy expectations and easy grades, as part of the corporatization of education.

To this end, a content analysis was undertaken in two steps, where student comments were first coded and then the codes were assigned to 18 category clusters. These

category clusters were inductively determined from the data and they were also supported by the literature review on effective teaching. Based on this coding and categorizing, the percentage of codes in each of the 18 categories were computed for retired professors and contract instructors. This allowed for using the difference of means tests, which helped to provide focus in answering the three questions raised in this paper.

For the first question, the findings of this paper indicate that while student comments predominately capture concerns of easy expectations and easy grades that are part of the corporatization of education, they also capture many aspects of effective teaching, although relatively more of the palatable aspects (instructor personality and fun approach) than the rigorous aspects (critical thinking and challenge).

For the second question, the findings indicate that while retired professors were noted to have upheld rigorous standards through subject mastery, they were also noted to have unfavourable aspects in terms of personality and the Keating/Hundert approach that does not align with effective teaching as depicted in the literature. For the third question, the findings of this paper indicate that while contract instructors may have lowered rigorous standards through easy exams, they are also noted to have upheld teaching standards through palatable aspects of effective teaching like a fun approach to teaching/learning, as depicted in the literature.

Overall, the results show that there is no statistically significant difference between the top retired professors and the top contract instructors based on any of the 18 category

clusters that captured aspects of effective teaching and easy expectations. However, the top retired professors are distinguished for subject mastery and the top contract instructors are noted for a fun approach to education. Although, retired professors are also noted for unfavourable personalities and the Keating/Hundert approach of say throwing a chalk or calling out a student, and contract instructors are also noted for easy expectations and easy grades by “shifting the curve” or by giving out “surprisingly easy” exams.

This means that academic standards have not necessarily fallen. They have changed form, where rigorous aspects of effective teaching are replaced by palatable aspects and where unfavourable personalities and approaches are replaced by easy expectations and easy grades. In other words, the old positive and negative aspects of teaching are replaced by a new set of positive and negative aspects. These results, however, are limited to the study based on Economics instructors at the University of Alberta and future extensions of this study may determine whether they hold across other disciplines and teaching institutions. Returning though to Mr. Keating and Mr. Hundert, it does seem that some approaches look inspirational only in Hollywood movies.

Declarations

Funding: No funding was received to assist with the preparation of this manuscript.

Conflicts of interest/Competing interests: The author has no relevant financial or non-financial interests to disclose.

Availability of data and material: The dataset generated during and/or analysed during the current study are available from the corresponding author on reasonable request. (The data is public anyway at the RMP website)

Code availability: Not applicable

References

- Abrami, PC, d'Apollonia S, Rosenfield S (2007) The dimensionality of student ratings of instruction: What we know and what we do not. In: Perry RP, Smart JC (ed) *The scholarship of teaching and learning in higher education: An evidence-based perspective*. Springer, New York, pp 385–456.
- Akerlind GS (2007) Constraints on academics' potential for developing as a teacher. *Studies in Higher Education* 32(1): 21-37.
- Albrecht S, Hoopes J (2009) An empirical assessment of commercial web-based professor evaluation services. *Journal of Accounting Education* 27(3): 125-132.
- Aleamoni L (1987) Student rating myths versus research facts. *Journal of Personnel Evaluation in Education* 1(1): 111-119.
- Anderson H, Siegfried JJ (1997) Gender Differences in Rating the Teaching of Economics. *Eastern Economic Journal* 23(3): 347-357.
- Andreopoulos GC, Panayides A (2009) Teaching Economics to the best undergraduates: What are the problems? *American Journal of Business Education* 2(6): 117-121.
- Baek SG, Shin HJ (2008) Multilevel analysis of the effects of student and course characteristics on satisfaction in undergraduate liberal arts courses. *Asia Pacific Education Review* 9(4): 475-486.
- Barr RB, Tagg J (1995) From teaching to learning: A new paradigm for undergraduate education. *Change* 27 (November/December): 13-25.
- Becker WE (2000) Teaching Economics in the 21st century. *The Journal of Economics Perspectives* 14(1): 109-119.
- Becker WE, Watts M (2001) Teaching Methods in U.S. Undergraduate Economics Courses. *The Journal of Economic Education* 32(3): 269-279.
- Beckman M, Stirling K (2005) Promoting critical thinking in Economics Education. Working Paper.
- Beecham R (2009) Teaching quality and student satisfaction: Nexus or simulacrum? *London Review of Education* 7: 135–146.
- Bengtsson M (2016) How to plan and perform a qualitative study using content analysis. *NursingPlus Open* 2: 8-14.
- Biggs J, Tang C (2011) *Teaching for quality learning at university*, 4th edition. Open University Press.

Bleske-Rechek A, Michels K (2010) RateMyProfessors.com: Testing Assumptions about Student Use and Misuse. *Practical Assessment, Research & Evaluation* 15(5): 1-12.

Boatright-Horowitz S, Soeung S (2009) Teaching white privilege to white students can mean saying good-bye to positive student evaluations. *American Psychologist* 64(6): 574–575.

Boex JFL (2000) Attributes of effective Economics instructors: An analysis of student evaluation. *Journal of Economic Education* 31(3): 211-227.

Bonwell CC (1992) Risky business: Making active learning a reality. *Teaching Excellence. POD Network in Higher Education. 1992/93 Essay Series: 1-3.*

Braga M, Paccagnella M, Pellizzari M (2014) Evaluating students' evaluations of professors. *Economics of Education Review* 41: 71–88.

Braskamp L, Ory J, Pieper D (1981) Student written comments: Dimensions of instructional quality. *Journal of Educational Psychology* 73(1): 65-70.

Busler J, Kirk C, Keeley J, Buskist W (2017) What constitutes poor teaching? A preliminary inquiry into the misbehaviours of not-so-good instructors. *Teaching of Psychology* 44(4): 330-334.

Cashin W (1990) Students do rate different academic fields differently. In: Theall M, Franklin J (ed) *Student ratings of instruction: Issues for improving practice*, New Directions for Teaching and Learning. 43 (Fall). Jossey-Bass, San Francisco, pp 113-21.

Coladarci T, Kornfield I (2007) RateMyProfessors.com versus formal in-class student evaluations of teaching. *Practical Assessment, Research & Evaluation* 12(6): 1- 15.

Coleman J, McKeachie WJ (1981) Effects of instructor/course evaluations on student course selection. *Journal of Educational Psychology* 73(2): 224-26.

Downe-Wamblot B (1992) Content analysis: method, applications and issues. *Health Care for Women International* 13: 313-321.

Dunn R, Griggs SA (2000) *Practical Approaches to Using Learning Styles in Higher Education*. Bergin & Garvey, Westport, CT.

Feldman KA (2007) Identifying exemplary teachers and teaching: Evidence from student ratings. In: Perry RP, Smart JC (ed) *The scholarship of teaching and learning in higher education: An evidence-based perspective*. Springer, New York, pp 93–143.

Felton J, Mitchell J, Stinson M (2004) Web-based student evaluations of professors: the relations between perceived quality, easiness and sexiness. *Assessment & Evaluation in Higher Education* 29(1): 91-108.

Ghosh IK (2013) Learning by doing models to teach undergraduate Economics. *Journal of Economics and Economic Education Research* 14(1): 105-119.

Greenwald, AG, Gillmore GM (1997) Grading leniency is a removable contaminant of student ratings. *American Psychologist* 52(11): 1209–1217.

Hamermesh DS (2019) 50 years of teaching introductory Economics. *The Journal of Economic Education* 50(3): 273-283.

Hamermesh D, Parker A (2005) Beauty in the Classroom: Instructors' Pulchritude and Putative Pedagogical Productivity. *Economics of Education Review* 24(4): 369-376.

Hervani A, Helms MM (2004) Increasing Creativity in Economics: The Service Learning Project. *Journal of Education for Business* 79(5): 267-274.

Hornstein HA. (2017) Student evaluations of teaching are an inadequate assessment tool for evaluating faculty performance. *Cogent Education* 4(1) 1304016.

Jordan, DW (2011) Re-thinking student written comments in course evaluations: Text mining unstructured data for program and institutional assessment. Doctoral dissertation. California State University.

Joye SW, Wilson JH (2015) Professor age and gender affect student perception and grades. *Journal of the Scholarship of Teaching and Learning* 15(4):126-138.

Lattuca LR, Domagal-Goldman JM (2007) Using qualitative methods to assess teaching effectiveness. *New Directions for Institutional Research* 136: 81-93.

Legg,AM, Wilson JH (2012) RateMyProfessors.com Offers Biased Evaluations. *Assessment & Evaluation in Higher Education* 37(1): 89-97.

Marginson S, Considine M (2000) *The enterprise university: Power, governance, and reinvention in Australia*. Cambridge University Press, New York.

Marsh HW (1987) Students' Evaluations of University Teaching: Research Findings, Methodological Issues, and Directions for Future Research. *International Journal of Educational Research* 11(3): 253-388.

Marsh HW, Ware JE (1982) Effects of expressiveness, content coverage and incentive on multidimensional student rating scales: New interpretations of the Dr. Fox effect. *Journal of Educational Psychology* 74(1): 126-134.

Marsh HW, Roche LA (2000) Effects of grading leniency and low workload on students' evaluations of teaching: Popular myth, bias, validity, or innocent bystanders? *Journal of Educational Psychology* 92(1): 202-228.

Mazzarol T, Soutar GN, Seng MSY (2003) The third wave: Future trends in international education. *International Journal of Educational Management* 17(3): 90-99.

McKeachie WJ (1997) Student ratings: The validity of use. *American Psychologist* 52(11): 1218-1225.

McPherson MA, Jewell RT (2007) Leveling the playing field: Should student evaluation scores be adjusted? *Social Science Quarterly* 88(3): 868-881.

Meyers C, Jones TB (1993) *Promoting active learning: Strategies for the college classroom*. Jossey Bass, San Francisco, CA.

Miller JD (2006) How to fight ratemyprofessors.com. *Inside Higher Ed*, January 31.

Otto J, Sanford DAJ, Ross DN (2008) Does ratemyprofessor.com really rate my professor? *Assessment & Evaluation in Higher Education* 33(4): 355-368.

Mobius MM, Rosenblat TS (2006) Why Beauty Matters. *American Economic Review* 96(1): 222-235.

Niu SC (2005) Comparing Two Populations. December 22, University of Texas at Dallas. https://www.utdallas.edu/~scniu/OPRE-6301/documents/Two_Populations.pdf Accessed 26 March 2021

Ogier J (2005) Evaluating the effect of a lecturer's language background on a student rating of teaching form. *Assessment & Evaluation in Higher Education* 30(5): 477-488.

Ongeri JD (2009) Poor student evaluation of teaching in economics: a critical survey of the literature. *Australasian Journal of Economics Education* 6(2): 1-24.

Pan D, Tan G, Ragupathi K, Booluck K, Roop R, Ip Y (2009) Profiling teacher/teaching using descriptors derived from qualitative feedback: Formative and summative applications. *Research in Higher Education* 50(1): 73-100.

Pienta NJ (2017) The slippery slope of student evaluations. *Journal of Chemical Education* 94(2): 131-132.

Pozo-Munoz C, Reboloso-Pacheco E, Fernandez-Ramirez B. (2000) The 'ideal teacher' implications for student evaluation of teacher effectiveness. *Assesment & Evaluation in Higher Education* 25(3): 253-263.

Reimann N (2004) First-year teaching-learning environments in Economics. *International Review of Economics Education* 3(1): 9-38.

Santhanam E, Lynch B, Jones J (2018) Making sense of student feedback using text analysis – adapting and expanding a common lexicon. *Quality Assurance in Education* 26(1): 60-69.

Sen A, Voia M, Woolley F (2010) Hot or Not: How appearance affects earnings and productivity in academia. *Carleton Economic Papers* 10-07.

Sheridan BJ, Hoyt G, Imazeki J (2014) Targeting Teaching: A primer for new teachers of Economics. *Southern Economic Journal* 80(3): 839-854.

Smith BP (2007) Student ratings of teacher effectiveness: An analysis of end-of-course faculty evaluations. *College Student Journal* 41(4): 788–800.

Spooren P, Brockx B, Mortelmans D (2013) On the validity of student evaluation of teaching: The state of the art. *Review of Educational Research* 83: 598–642.

Theyson KC (2015) Hot or not: The role of instructor quality and gender on the formation of positive illusions among students using ratemyprofessors.com. *Practical Assessment, Research & Evaluation* 20(4): 1 -12.

Timmerman T (2008) On the Validity of RateMyProfessors.com. *The Journal of Education for Business*. 84(1): 55-61.

Weeden J, Sabini J (2005) Physical Attractiveness and Health in Western Societies: A Review. *Psychological Bulletin* 131(5): 635–653.

Wilson JH, Beyer D, Monteiro H (2014) Professor age affects student ratings: Halo effect for younger teachers. *College Teaching* 62(1): 20-24.

Table 1: Rate My Professor questions

Question	Response
Course Code	Selection of course
How would you rate this professor as an instructor?	1 – 5
How hard did you have to work for this class?	1 – 5 (the scale on easy A versus working hard)
Would you take this Prof again?	Yeah - Um, No
Was this class taken for credit?	Yeah - Um, No
Textbook use	Yeah - Um, No
Attendance	Mandatory – Non Mandatory
Grade Received	Selection of grade
Here's your chance to be more specific	<p>Comment (350 characters)</p> <p>The site gives suggestions as:</p> <ul style="list-style-type: none"> your unique experience writing / reading intensity attendance policy availability outside of class required participation

Table 2: Summary of the Rate My Professors data

Instructors	# of instructors	# of Comments	White	Native English	Male	Quality	Difficulty level	Would take again (%)	100 level	Attendance	Text used	Average grade
UfA Retired	14	235	92.86%	92.86%	85.71%	3.58	3.32		7.51%		42.12%	
UfA Contract Teaching	13	370	69.23%	53.85%	84.62%	3.95	2.69	65.30%	32.87%	47.59%	42.62%	3.72
Sum	27	605										

**Empty cells indicate missing data*

Table 3: Codes and Categories in Content analysis

Categories	Percentages	Codes (based on keywords/phrases/themes)
Personality/Helpful/Available	8.78%	accessible outside class, cares, helpful office hours, inspirational, kind, nice, professional, emails quickly, respectful, encourages, goes out of the way, hot, humble, respects opinions, works hard
Rude/Personality	2.89%	disrespectful, does not care, rude, shames, unhelpful, arrogant, calls you out, condescending, delayed grading, late responses, not helpful
Clarity/Organization	6.53%	clear expectations, explains well, organized, cares about the subject, clear exam guidelines, clear grading, clear lectures, clear notes, concise, easy to understand, good feedback, simplifies, repeats
Unclear/Tangents/Accent	6.70%	accent, auditory class, bad notes, confusing, difficult to understand, doesn't answer questions, scattered, self-study, tangents, bad examples, bad writing, disorganized, doesn't explain well, makes it hard, makes mistakes, no feedback, poor wording, vague instructions, unstructured
Critical skills/Learning	3.99%	concepts not memorization, learned lots, analyze and understand, care about learning, challenging, changed perception, critical skills, debates, ensures understanding, in-depth, pushed to think, multiple perspectives, learned nothing, don't learn much, memorization
Interaction/Interest/Passion	4.91%	CP matters, engages, interesting, passionate, discussions, sparked interest, CP optional, doesn't like disagreement
Easy A/Pattern/Fair	15.08%	easy A, easy exams, easy HW, easy marking, follow chart, gives away exams, hints, GPA booster, helps succeed, lecture based exams, similar exams, fair exams, fair marking, homework based exams, many sample exams, prepares for exams, sample based exams, review questions are exams, shifts curve, straight forward exams, text based exams
Tough/Grading/Exams	10.86%	different exams, difficult exams, hard quizzes, high expectations, no homework, pop quizzes, tough grader, work hard, tough online course, hard A, hard final, hard homework, long exams, low grades, not an easy A, not a GPA booster, study for exams, tricky exams, worked but scored low
Fun/Story	6.99%	enjoyable, fun, funny, entertaining, stories, makes dull fun, time flies
Boring/Slide reader	4.45%	boring, dry, disengaging, reads PowerPoint slides, dull, monotone, not interesting, slow paced
Real world relevance	2.08%	good examples, real world applications, relevant, real life connections
Fast Pace	0.75%	fast paced, talks fast, throws information, type fast,
Mastery	3.12%	knows stuff, smart, brilliant, smart, thorough, doesn't know stuff, gets stuck, unable to answer, reads notes
Attend/Take notes/Work hard	13.63%	attend class, discipline, lots to read, no notes, pay attention, puts on spot, read text, study hard, take notes, bonus marks in class, CP marks, don't skip class, lots of work, lots to read, old school, work hard,
Notes given/Skip/No text	6.12%	can skip class, online notes provided, gives notes, don't need notes, everything posted, great notes, text not used, skip class, text not needed
Math	1.16%	lots of graphs, math heavy, calculus based
Text expensive/bad	0.46%	useless text, expensive text, text required for quizzes,
Tech	1.50%	bonus marks on e-polls, iclicker is a waste of time, video recording of lectures, can book online exam dates

Table 4: Determinants of effective teaching based on student comments

Categories	RETIRED (\bar{X}_1)	CONTRACT (\bar{X}_2)	n1	n2	s1 ²	s2 ²	n1+n2-2	(n1-1)s1 ²	(n2-1)s2 ²	Sp ²	Sqrt	$\bar{X}_1 - \bar{X}_2$	t
Personality/Helpful/Available	8.66%	8.84%	14	13	0.0104	0.0031	25	0.14	0.04	0.0069	0.0320	-0.18%	-0.06
Rude/Personality	5.12%	1.80%	14	13	0.0017	0.0009	25	0.02	0.01	0.0013	0.0140	3.32%	2.38*
Clarity/Organization	6.01%	6.78%	14	13	0.0026	0.0056	25	0.03	0.07	0.0040	0.0244	-0.77%	-0.32
Unclear/Tangents/Accent	7.77%	6.18%	14	13	0.0191	0.0095	25	0.25	0.11	0.0145	0.0464	1.59%	0.34
Critical skills/Learning	6.36%	2.83%	14	13	0.0051	0.0019	25	0.07	0.02	0.0036	0.0230	3.53%	1.54
Interaction/Interest/Passion	4.42%	5.15%	14	13	0.0021	0.0015	25	0.03	0.02	0.0018	0.0164	-0.73%	-0.45
Easy A/Pattern/Fair	7.60%	18.71%	14	13	0.0064	0.0117	25	0.08	0.14	0.0089	0.0364	-11.12%	-3.05**
Tough/Grading/Exams	14.31%	9.18%	14	13	0.0089	0.0044	25	0.12	0.05	0.0068	0.0317	5.13%	1.62
Fun/Story	4.77%	8.07%	14	13	0.0041	0.0055	25	0.05	0.07	0.0048	0.0266	-3.30%	-1.24
Boring/Slide reader	5.83%	3.78%	14	13	0.0081	0.0011	25	0.11	0.01	0.0048	0.0266	2.05%	0.77
Real world relevance	1.59%	2.32%	14	13	0.0010	0.0015	25	0.01	0.02	0.0013	0.0136	-0.73%	-0.53
Fast Pace	0.88%	0.69%	14	13	0.0006	0.0002	25	0.01	0.00	0.0004	0.0080	0.20%	0.25
Mastery	6.54%	1.46%	14	13	0.0074	0.0001	25	0.10	0.00	0.0039	0.0240	5.08%	2.11*
Attend/Take notes/Work hard	15.90%	12.53%	14	13	0.0110	0.0072	25	0.14	0.09	0.0092	0.0370	3.37%	0.91
Notes given/skip/no text	2.65%	7.81%	14	13	0.0012	0.0035	25	0.02	0.04	0.0023	0.0184	-5.16%	-2.80**
Math	0.71%	1.37%	14	13	0.0003	0.0012	25	0.00	0.01	0.0007	0.0102	-0.67%	-0.65
Text expensive/bad	0.88%	0.26%	14	13	0.0003	0.0000	25	0.00	0.00	0.0002	0.0052	0.63%	1.21
Tech	8.66%	8.84%	14	13	0.0104	0.0031	25	0.14	0.04	0.0069	0.0320	-0.18%	-0.06

** , * denote significance at the 1% and 5% levels

Notes: The second and third columns indicate the average % of comments for highly and poorly rated instructors in each of the 18 categories. The fourth and fifth columns indicate the number of instructors in the specific categories. The next two columns indicate the variances of the coding categories. The four columns after that use the methods described in Niu (2005) to determine the pooled variance “S²” for the coding categories. The pooled variances allow computing the t statistic to test for the significance of the difference between the coding categories for the highly rated and poorly rated instructors.

Table 5: Sensitivity analysis

Categories	Top Retired (6)	Low Retired (8)	t	Top Retired (6)	Top Contract (5)	t	Low Retired (8)	Low Contract (8)	t	Top Contract (5)	Low Contract (8)	t
Personality/Helpful/Available	9.63%	8.18%	0.34	9.63%	8.23%	0.27	8.18%	10.06%	-0.46	8.23%	10.06%	-0.42
Rude/Personality	1.07%	7.12%	-3.20**	1.07%	1.11%	-0.02	7.12%	2.76%	2.42*	1.11%	2.76%	-0.84
Clarity/Organization	10.70%	3.69%	1.97	10.70%	7.44%	0.86	3.69%	6.31%	-0.82	7.44%	6.31%	0.29
Unclear/Tangents/Accent	3.21%	10.03%	-1.08	3.21%	1.74%	0.20	10.03%	12.03%	-0.33	1.74%	12.03%	-1.58
Critical skills/Learning	4.28%	7.39%	-1.01	4.28%	1.90%	0.65	7.39%	4.14%	1.10	1.90%	4.14%	-0.71
Interaction/Interest/Passion	6.42%	3.43%	1.32	6.42%	5.70%	0.28	3.43%	4.73%	-0.61	5.70%	4.73%	0.41
Easy A/Pattern/Fair	11.23%	5.80%	1.03	11.23%	22.63%	-2.01	5.80%	14.79%	-1.89	22.63%	14.79%	1.39
Tough/Grading/Exams	9.09%	16.89%	-1.82	9.09%	8.54%	0.11	16.89%	10.45%	1.58	8.54%	10.45%	-0.43
Fun/Story	8.56%	2.90%	1.49	8.56%	12.97%	-1.06	2.90%	2.37%	0.15	12.97%	2.37%	2.63*
Boring/Slide reader	0.53%	8.44%	-2.31*	0.53%	3.16%	-0.61	8.44%	4.73%	1.09	3.16%	4.73%	-0.46
Real world relevance	1.60%	1.58%	0.01	1.60%	2.53%	-0.43	1.58%	2.17%	-0.33	2.53%	2.17%	0.17
Fast Pace	2.67%	0.00%	2.50*	2.67%	0.16%	1.96	0.00%	1.38%	-1.34	0.16%	1.38%	-1.11
Mastery	9.09%	5.28%	1.26	9.09%	1.27%	2.00	5.28%	1.78%	1.14	1.27%	1.78%	-0.17
Attend/Take notes/Work hard	15.51%	16.09%	-0.12	15.51%	12.66%	0.49	16.09%	13.02%	0.64	12.66%	13.02%	-0.07
Notes given/skip/no text	4.28%	1.85%	0.90	4.28%	9.34%	-1.78	1.85%	6.31%	-1.85	9.34%	6.31%	1.03
Math	1.07%	0.53%	0.36	1.07%	0.63%	0.28	0.53%	2.37%	-1.37	0.63%	2.37%	-1.05
Text expensive/bad	1.07%	0.79%	0.42	1.07%	0.00%	1.27	0.79%	0.59%	0.30	0.00%	0.59%	-0.90

Note: The numbers in parentheses in the first row refer to the number of instructors in each of the instructor categories.

Table 6: Select comments for Retired professors

Instructor	High rating	Instructor	Low rating
R1	Bear-trap grip on the topic, unconventional style, impossible to be bored, and that's saying something for an econ class. Posts everything to the website immediately after class, very reliable on getting marks back. I'm a Political Science major, and I would take more econ classes if he were teaching them, just because he was teaching them.	R6	Knows his stuff very well but can lose students in his meticulous notation, makes things more complicated than they have to be. Exams generally were tough and often had a very low class average. I don't see the point of having such hard exams because you don't learn anything from an exam that less than half is able to pass. Study past exams!
R9	one of those memorable profs that truly challenges you but makes you know the material entirely. It was a difficult class (lots of readings), but I never once regretted taking it. It's also not so hard that I'd recommend ditching it for an easy-A class. This is truly an unforgettable class that you'll want to come to.	R9	Spare yourself and your GPA because there are tons of other econ classes taught by professors that actually take their job seriously. This guy was a bit obnoxious and rude to the students. The tests were harder than you'd expect from a prof who spent all his time telling stories.
R12	Been a couple of years since I had him, but I still remember him well. His classes will give you hand cramps, and is one of the tougher classes in undergrad. His ... course would be considered a grad level at some Universities. He's very smart and if you take a genuine interest, he will bend over backwards for you.	R10	Unbelievably boring and very inarticulate in a lecture setting. It's somewhat frustrating to listen to him talk in class. However, the group project is easy and interesting; mid term and final are both easy if you read the readings and attend class. If you have an interest in the world outside North America, I would recommend this class
R14	He's inspirational. He showed me the beauty of economics. He used current events to explain economic theories. He pushed the class to think. He taught me how to think. Isn't that what a great professor does - to inspire the students? His style is unique. Chat with him in his office, it's lovely. I learnt so much more chatting with him. I miss him	R11	Although I have no doubt that he's knowledgeable ... I found his teaching style pretty disappointing. It was hard to take anything away from the lectures. Very vague with concepts he is trying to illustrate. His diagrams are incomplete and almost useless. Also he is very boring. Had to learn everything from the textbook. Would not take again

Table 7: Select comments for Contract instructors

Instructor	High rating	Instructor	Low rating
C1	This was the only lecture I used to look forward to every MWF. The prof was funny and the connections between course material and real life made understanding content easy. All the exams were very similar to his previous exams and online homework took about an hour a week. Overall the course is a GPA booster and the prof made class exciting.	C5	Granted, prof ... is a nice guy by recording his class and he is probably look handsome for some people. However, I should say the material he taught is too easy compared with his exam. I still remember he gave us a super easy practice midterm, and made the actual exam hard. Personally, I think he basically just read the ppp during the term.
C2	My stats/calc background was pretty weak but he basically retaught me everything and made it easy to understand. Exams are surprisingly easy and he'll give you plenty of help with assignments in office hours. He likes to scare students into thinking this class is impossible, but it's so easy with Professor ...	C8	Poorly run class. Often gets stuck on his own problems, has students correct him, and he doesn't provide answers for his practice exams.
C4	One of the best profs I had. Extremely nice guy and genuinely wants his students to succeed to the point where he is shifting the curve over 10% to benefit everyone's GPA. His personality is out there and he has a few phrases that can drive you insane after listening to them for a semester.	C4	If you want to pass, you better do the practice exams. However, his notes and practice exams fails to prepare you for the actual tests, as a lot of questions on the exams are tested on material that isn't covered in nearly enough detail in class. Consequently, some questions feel unfair, as you can't solve them even if you did study off the notes.
C12	He is a very nice professor, and I personally think he is the best Econ professor in U of A. His lecture is not interesting but very useful. He would like to give student high grades and very helpful. He really care about if student can understand his content or not. I strongly recommend to take his course.	C12	Boring lecture, full of ERROR on slides. Illogic and disorganized Useless practice question Not testing on the major model, but those fine negligible details If I have a chance I will NEVER take any class with him. You won't get anything from his lecture, however, his lecture is mandatory You will have a really hard time if you choose him.