Do Accounting Professor Teaching Ratings Reflect Student Learning?

Richard Constand
University of West Florida

This current study provides evidence relating to the question of whether accounting professor teaching ratings in undergraduate classes are related to student learning or if they are just a reflection of the degree of easiness or difficulty of the class and the professors that are being rated. The relationship between average accounting faculty ratings for eight AACSB accredited business schools from RateMyProfessors.com (RMP) and the subsequent average success rate for students from those schools who take and pass the CPA exam on their first attempt are examined. Empirical results indicate that there is a significant positive relationship between students passing the CPA exam and RMP teaching ratings reflecting professor Quality and a significant negative relationship between students passing the CPA exam and RMP ratings reflecting the degree of Easiness for those classes and professors being rated. Additional empirical evidence of the complex relationships between teaching ratings and the personal characteristic of accounting professors who teach either classes for accounting majors or non- majors is also discussed.

Keywords: teaching ratings, accounting, student learning, CPA Exam

INTRODUCTION

One of the most important issues related to the use of teaching ratings by accounting professors is whether ratings reflect student learning or if they reflect other factors such as the difficulty of course content, the demands placed on students, and the grades the professor gives. While many authors suggest ratings do not reflect learning, there are no empirical studies that examine the question of whether accounting professor ratings are related to student learning. The major reason for this is that available data do not allow for the matching of anonymous teaching ratings and the actual grades earned by students providing those ratings. This current study addresses this gap in the literature by focusing on the relationship between average ratings for professors teaching classes for accounting majors at individual schools and an external, objective measure of learning for that same group of students, the success rate for students from those schools who take the CPA exam for the first time.

The ratings data used in this study are from ratemyprofessors.com (RMP), an online platform that allows student to rate individual professors and their classes. Before the website revised their student interface rating page, RMP allowed students to rate their professors at the individual class level using a Likert-type scale on several different dimensions of teaching including professor Helpfulness, professor Clarity, and the perceived Easiness of the professor and the class. RMP also allowed students to rate their Interest in the classes being rated, provide free response comments about the professor and class, and indicate whether the professor is physically attractive by granting "Hot" Chile Pepper awards to professors.1 RMP then calculated professor level average ratings for Helpfulness, Clarity, and Easiness

from the student class ratings and they calculated an overall teaching "Quality" rating that was the average of the professor level averages for Helpfulness and Clarity ratings. These professor level and individual class level ratings were then available for anyone to review online. As of July of 2021, RMP claims their website contains over 15 million class ratings covering over 1.3 million professors at over 7,000 schools.

Many past studies have examined both the RMP numerical data, and the written comments posted by the students on RMP to explore several issues related to teaching ratings. These issues include the comparability of RMP ratings to traditional Student Evaluation of Teaching (SET) ratings administered by schools, differences in faculty ratings across academic disciplines, the relationships between Easiness and attractiveness of the professor on their average teaching quality ratings, and the impact of other class, professor, or student characteristics on those ratings. Few of these studies, however, focus on accounting professor ratings using RMP data, but one does compare ratings for accounting professors to ratings for professors in dozens of different disciplines [Constand, Pace, and Clarke (2016)] and none examine the link between ratings and learning.

The rest of this study is presented as follows. A Literature Review first examines the question of whether RMP data is comparable to SET data and then examines past studies that focus on many of the various factors that have been found to be related to teaching ratings and the relationship between ratings and learning. A Data section describes the RMP ratings data, other sources of data that represent individual professor and class level information, and the CPA pass rate data used as a proxy for learning in this current paper. A Methodology section describes the approaches used to examine the relationship between ratings and student learning, and a Results section discusses the empirical findings. Finally, a Conclusions section discusses the implications of the empirical findings and discusses possible directions for future research.

LITERATURE REVIEW

This section first discusses several topics from the literature including the importance of ratings for accounting professors, the validity of using RMP ratings as a proxy for SET ratings, and the factors that are usually examined as possible factors affecting ratings. A discussion of the literature focusing on the link between ratings and learning is also presented and a discussion of the use of CPA exam results as a measure of the effectiveness of learning is discussed.

Importance of Teaching Ratings for Accounting Professors

Several studies show the importance of SET based teaching ratings in the performance review process in accounting departments. For example, studies show that between 90% and 97% of accounting departments use ratings for annual evaluation purposes and promotion and tenure decisions [Yunker and Sterner (1988), Calderon, Green, and Reider (1994), Green and Wang (2012)]. Another study reports that almost 45% of accounting departments use only student ratings to evaluate` teaching effectiveness [Green, Calderon, Gabbin, and Habegger (1999)]. While the use of ratings for evaluation of teaching effectiveness is widespread and often relied upon by administrators, accounting professors tend to view ratings as a reflection of class difficulty rather than a true measure of teaching effectiveness and student learning [Morgan, Sneed, and Swinney (2003)].

In business schools, many professors consider accounting and finance to be two of the more difficult disciplines taught in the curriculum. One RMP study comparing ratings for accounting professors to ratings in seven other business disciplines and 32 non-business disciplines reports significant differences in both Quality and Easiness ratings between accounting professors and most other professors in the applied and social sciences, humanities, and professional studies areas [Constand, Pace and Clarke (2016)]. That study reports that accounting professors are perceived as being more difficult than in most other disciplines, accounting professors earn significantly lower teaching ratings than professors in other disciplines, and over 50% of the differences across academic disciplines are explained by student's perceived differences in class Easiness. A similar study focusing on finance professor ratings reports similar results [Constand and Pace (2014)]. That study also reports results that indicate differences in Easiness ratings across disciplines explain over 60% of the variation in differences in Quality ratings between finance and non-

finance professors. It should be noted that in both studies discussed above there are no significant differences between the ratings of professors teaching accounting and finance.

Given the widespread use of student evaluations of professors and the importance placed on them by administrators, it is surprising there is not more research focused on ratings of accounting professors. A review of 223 research articles published from 2003 to 2005 in five accounting educational journals resulted in only 6 articles focusing on accounting faculty ratings [Watson, Apostolou, Hassell and Webber (2007)].

The Validity of RMP Ratings and Comparability to Traditional SET Ratings

There is a growing body of literature that supports the argument that ratings data from RMP is both valid and comparable to ratings obtained from traditional SET rating procedures. Most studies focus on the RMP professor level average ratings and compare them to ratings on similar questions that appear on school administered SET forms. Many large studies across many academic disciplines report that there are significant positive relationships between RMP ratings and similar questions from SET ratings [Coladarci and Kornfield (2007), Brown, Baillie and Fraser (2009), Silva, Silva, Quinn, Draper, Cover and Munoff (2008), Barth (2008)]. One study compares RMP ratings to SET ratings for 1,167 business professors at five different universities and finds the RMP quality ratings to be highly positively correlated with several SET questions that focus on teaching effectiveness, teaching ability, and instructor quality. That study concludes that there is no evidence that the average ratings or the ratings distributions differ across the two ratings platforms [Timmerman (2008)].

Not all authors find unequivocal support for similarities across RMP and SET ratings platforms. For example, one study compares RMP professor level ratings to university SET ratings using a paired t-test for 75 business professors from two schools and reports that quality ratings from the RMP data are significantly lower than the SET ratings [Albrecht and Hoopes (2009)] while another study reports that the online ratings reflect higher teaching effectiveness ratings than the SET ratings [Bruns, Rupert, and Zhang (2011)]. A third argues that web-based ratings forms have a much lower response rate and conclude that web-based and in-class SETs could not be compared to each other [Fogarty, Jonas, and Parker (2013)]. While these studies do raise the question of whether RMP ratings are identical to traditional SET ratings, most of the evidence supports the argument that ratings from the two sources are very similar in many respects and support the use of RMP data for the current study.

There are many factors that are related to teaching ratings outcomes that have been examined in the literature using both SET and RMP data. These factors and their impact on ratings include both professor-level variables and class-level variables. Many early studies using RMP data focus on the relationship between the Quality and the Easiness ratings variables to examine whether ratings are mainly a reflection of class ease or difficulty. Many studies using RMP data provide evidence of a significant positive relationship between average class Easiness and Quality ratings [Marsh (1982, 1984, and 1987)] while others report the strength of these positive relationships vary across disciplines [Felton, Mitchell, and Stinson (2008), Felton, Koper, Mitchell, and Stinson (2008)]. This current study also examines this relationship for accounting professors but also includes an analysis of the relationships between ratings and other factors that may influence those ratings.

Professor Ratings and Student Learning

In 1994, the Teaching and Curriculum Section of the American Accounting Association (AAA) established the Accounting Education Change Commission which issued a report detailing the five dimensions of effective teaching of accounting. Those dimensions include (1) design and course development, (2) the use of well-conceived course materials, (3) presentation skills, (4) well-chosen pedagogical methods, and (5) both guidance and advising [Calderon, Gabbin, and Green (1996)]. The first 4 of these teaching dimensions are reflected in the RMP ratings variable Clarity as all four dimensions contribute to the student's perception of how well the class is organized and how well the professor presents the class content. The fifth dimension, guidance and advising, is reflected in another RMP rating, Helpfulness.

Even though two of the RMP ratings are related to these dimensions of effective teaching identified above, many authors argue RMP ratings do not reflect student learning. For example, some argue that most student comments on RMP focus on the instructor's personality, the amount of coursework, and the instructor's ability to entertain the class rather than on teaching effectiveness and learning [Davison and Price (2009)]. But similar arguments against the validity of traditional SET ratings have been made in the past [Beleche, Fairris, & Marks (2012), Flinn and Crumbley, (2009), Crumbley and Smith (2009), Crumbley, Flinn, and Reichelt (2012), Moore (2009), Moore and Flinn (2009) and Centra (2003)]. And another study reports no significant relationship between ratings and learning in large samples [Uttl, White, and Gonzalez (2017)].

While many opinions relating to the relationship between student learning and teaching ratings have been presented in the literature, there are few actual studies that examine the relationship. This is mainly due to the problems associated with identifying an appropriate variable that reflects student learning. The accounting discipline is fortunate to have one measure of student learning, the ability to pass the Uniform CPA Exam on the first attempt, that perfectly captures the degree of student learning that has taken place. While individual student ratings and data on individual student performance on the CPA exam are not available, it is possible to explore this relationship by focusing on average ratings at different universities and the average CPA pass rates for students from those schools. A similar approach using average CPA pass rates from universities has been used in the past to examine the relationships between pass rates and accounting program characteristics and the point is made that first time pass rates are highly indicative of the quality of the undergraduate accounting program [Nagle, Menk, and Rau (2018)]. This current study examines the relationship between average RMP accounting professor ratings and first time CPA pass rates at eight Florida universities.

DATA

There are several different sources of data used in this study including RMP faculty and class level ratings data, employee data for faculty employed at the universities in the Florida State University System (SUS), individual Florida university Accounting Department websites and publications of the National Association of State Boards of Accountancy. The ratings data represent student ratings and professor data from RMP and additional information from other sources about the faculty who teach the accounting classes that have been rated. Additional information from the Association to Advance Collegiate Schools of Business (AACSB) allows identification of whether the Florida schools are accredited and the types of programs they offer (undergraduate accounting programs, master level accounting and MBA programs, or Ph.D. programs).2 The eight schools examined in the study are Florida Atlantic University (FAU), Florida International University (FIU), Florida State University (FSU), University of Central Florida (UCF), University of Florida (UF), University of North Florida (UNF), University of South Florida (USF), and the University of West Florida (UWF). All are accredited and all offer undergraduate accounting degrees. By restricting the schools to Florida institutions, the process of identifying similar classes operating under different class identifying numbers and names is made much easier allowing a partitioning of the entire sample into smaller subsamples of class ratings with similar characteristics. Finally, publications of the National Association of State Boards of Accountancy, the "Uniform CPA Examination: Candidate Performance Reports" for various years provide information on average pass rates for students graduating from the schools in the study.

Since this study partitions the class ratings into different samples that are the focus of analysis, these average professor level rating provided from RMP are not used.

Professor Data

From the RMP site, professor data representing the professor's name, individual class level ratings for Quality, Helpfulness, Clarity, and student perceptions of the Easiness of the class and their Interest in the class.3 Since this study partitions the class ratings into different samples based on the type of accounting

class being rated, the average professor level ratings provided from RMP are not used, and the individual class level ratings are used to construct averages for particular classes.

RMP also gathered a "rating" on the professor's physical attractiveness by having students identify professors as "Hot" and awarding them "Chili Peppers". A rating of Hot was awarded if a single student had awarded a Chili Pepper in the past. A dummy variable (Hot) takes on a value of 1 if the professor has been given this rating. The RMP data is also used to identify the professor's gender. If the professor's gender is not readily apparent from their first name, the student comments are used to search for references to the professor (him, her, she, he, etc.) or a visit to the professor's school website allows gender identification. A dummy variable (Gen) takes on a value of 1 for males. In addition to gender, the RMP student comments about professors are used to identify those professors who have limited English language skills as students typically complain if understanding the professor's spoken English is a problem.

Another dummy variable (Eng) takes on a value of 1 for professors with strong English language skills and no complaints or zero if there are complaints.

The Florida SUS publishes employee rank, salary, and contract information under the State's "Sunshine Law" government information disclosure requirements. Employee information made public by the SUS provides information about professor rank for full time employees and identifies instructors, permanent lecturers, and other adjunct faculty members. The source also provides information on their contracts including Full Time Equivalent (FTE) employee status allowing for a determination of who is also an administrator (100% FTE on a 12 month contract) or strictly a teacher / researcher (75% FTE on a 9 month contract).

Other teachers who have ratings on RMP but don't appear in the SUS employee data require additional research to identify their rank. Internet searches of Ph.D. program information at the Florida schools allow identification of faculty who are classified as graduate students. Additional internet searches also allow identification of faculty who were graduate students at the time of the ratings, but have now completed their studies, obtained their degrees, and joined faculties at other schools. Other faculty identified as having moved to another school, retired, or having died since their ratings are listed as having the rank they held when the ratings were made. Faculty who are identified as instructors, adjuncts, and other non-tenure track teachers are also included in the study as the control group for the above rank dummy variables. Teachers whose rank, graduate student, or adjunct status were not able to be identified are excluded from the sample. The professor rank data is used to create a series of dummy variables equal to 1 for each rank of full professor (Full) status, associate professor (Assoc) status, assistant professor (Assist) status, or graduate student (GS) status. Emeritus professors and endowed chair holders are classified as full professors.

Class Level Data

The class level data collected from the RMP site contains information on the date the rating was submitted, a class identifying code, a rating (from 1 to 5) for the variables Easiness, Helpfulness, Clarity, and rater Interest, and student comments about the professor and class. It is important to note that the class level data does not contain a quality rating number; RMP calculates the numerical quality rating for the professors by averaging the Helpfulness rating and the Clarity ratings from all their class level ratings. This study calculates an average quality rating for professors using just the classes associated with teaching either undergraduate non- accounting majors or teaching required classes for accounting majors.

For all professors listed on RMP for the schools identified above, class level ratings are collected, and the class codes are reviewed in order to identify the class title. Some students enter class names rather than the correct class codes and if the name can be matched with a standard class title, a correction is made but if no match is found, the rating is dropped. Class codes are also corrected when it is obvious the student has made a typographical error when entering the code transposing two numbers in the code or using an incorrect abbreviation (all Florida schools use either "ACG" or "TAX" as abbreviations in accounting class codes). When class numbers appearing in ratings could not be matched up to actual classes offered at that university (either currently or in the past) that observation is deleted. The corrected class codes are then used to identify ratings for different classes.

The four class level ratings collected from the RMP website are the raw data used in this study: Helpfulness (Help), Clarity (Clarity), Easiness (Ease), and Interest (Interest). As noted earlier, while the RMP professor level data presents a numerical "quality" rating this is not a rating that students actual select when rating a particular class. RMP calculates their faculty quality rating by averaging the Helpfulness and Clarity ratings for all classes for that faculty member.4 Following the RMP method, this study calculates a class rating specific quality variable (Quality) for each class by averaging the Helpfulness and Clarity ratings students enter. Each of the four original class level ratings variables and the Quality variable have a possible value of from 1 to 5 with 5 representing more Helpfulness and Clarity, an easier class and professor, higher rater Interest and higher overall quality. These five variables represent the set of ratings responses for a particular class.

A total of 5,231 individual accounting class level sets of ratings based on RMP published ratings between 2004 and 2013 are used in this study with 2,359 representing sets of ratings for non-major classes and 2,872 representing sets of ratings' required accounting major classes.

CPA Examination Data

One of the major goals of this study is to examine the relationship between faculty ratings and student learning. The CPA Candidate Performance Reports that are published annually provide average performance data on a school by school basis for candidates sitting for the exam. The learning variable used in this study is the school's average passing rate (stated as a percentage of test takers) for the students completing the undergraduate accounting degree who are sitting for the CPA exam for their first time and pass the exam (UGpass). By using data for the first time test takers, the focus of this study is on how well the classwork taken by the students at their schools have prepared them for sitting for the CPA exam and how this learning is related to the teaching ratings earned by professors at those schools. This pass rate data is available from the annual Performance reports for the years 2006 to 2013 and is this study's measure of student learning.5

METHODOLOGY

This study has two main goals; first, to examine the relationships between professor and class characteristics and the ratings students have given to those classes and second, to examine the link between average school ratings for accounting faculty and classes and student learning as measured by average passing rates for CPA candidates.

The first step in the analysis is to use the class code to classify all undergraduate class level ratings into one of two class categories; undergraduate non-major (UGNM) sophomore level classes that are required for all business students, and junior and senior level classes required for undergraduate accounting majors (UGM). The class level data in these two samples are then merged with the professor level data that has been collected from the RMP website and the Florida SUS employee information. All class ratings with incomplete data on the professor level variables are omitted from the study. The final sample for the class level analysis consists of 5,231 total observations containing individual class ratings and the related professor specific data. The UGNM sample contains 2,359 observations and the UGM sample contains 2,872 observations.

Determinants of Professor and Class Ratings

The first stage of the analysis follows the approach used in most other studies that examine the relationships between ratings and explanatory variables by using regressions to explain differences in ratings across professors. This study expands on the existing empirical work, however, in several ways. First, while the explanatory variables examined include variables that were examined by previous authors, several additional explanatory variables are included in this study. These variables have often been discussed as having an impact on teaching ratings, but many have not been examined empirically in past studies focused on accounting professors. The variables include the student's Interest in the class (Interest), the professor's command of the English language (Eng), and the professor's gender (Gen). Other authors

have suggested that the professor's rank also influences ratings, so this study includes dummy variables for both graduate student teachers (GS) and tenure track faculty rank (Assist, Assoc, and Prof).

This study also uses a different focus on the ratings variables than many past studies that use the RMP numerical "Quality" rating that is averaged across all classes taught by a professor. In this study, to allow for an examination of ratings relationships in groups of different level classes, the Quality variable is calculated by averaging the Helpfulness and Clarity on a class by class basis rather than as an average across all classes. This study also examines the ratings responses on the two individual RMP ratings variables, Helpfulness and Clarity, that are used to construct the Quality rating on a class level basis. This allows for a closer examination of how different factors affect both overall quality ratings and the two component parts of that overall RMP rating.

Sample statistics for all observations included in the study and the two subsamples (UGNM and UGM) are presented in Table 1. Correlation coefficients for all variables are calculated and presented in Table 2. These variables are then used in a series of regression to explore how the professor characteristics and student's perceptions of professor / class Easiness and the rater's Interest in the class are related to the ratings given the professor. The regression models used to examine the factors that are related to the three ratings variables are of the following form:

$$\text{Rating} = \alpha + \beta1*(\text{ease}) + \beta2*(\text{Interest}) + \beta3*(\text{Eng}) + \beta4*(\text{Gen}) + \beta5*(\text{Hot}) + \beta6*(\text{GS}) + \beta7*(\text{Assist}) + \beta8*(\text{Assoc}) + \beta9*(\text{Full}) \quad (1)$$

where; Rating = either the Quality, Help or Clarity dependent variable,
Quality = the average of the class level Help and Clarity ratings,
Help= the class level Helpfulness rating,
Clarity = the class level Clarity rating,
Interest = the class level rater Interest rating,
Eng = a dummy variable = 1 if no English language problems reported, 0 otherwise,
Gen = a dummy variable = 1 if professor gender is male, 0 otherwise,
Hot = a dummy variable = 1 if at least one student rated professor as hot, 0 otherwise,
GS= a dummy variable = 1 if teacher is a graduate student, 0 otherwise,
Assist = a dummy variable = 1 if professor is an assistant professor, 0 otherwise,
Assoc = a dummy variable = 1 if professor is an associate professor, 0 otherwise, and
Full = a dummy variable = 1 if a full professor, chairholder, or emeritus professor

These regression models are used to examine the relationship in both subsamples resulting in a total of six regressions reported in Table 3. Results from these analyses will not only allow a comparison to results reported in past studies but provides new evidence of how additional professor level variables and different class levels are related to the ratings accounting professors receive.

The Link Between Professor Ratings and Student Learning

The second part of this study examines the relationship between professor / class ratings and student learning. For accounting students who graduate and then take the CPA exam, their ability to pass the exam on their first attempt is a valid measure of the student learning that taken place during their accounting education. Unfortunately, the data available to match individual student ratings and that student's CPA exam performance is not available. There is, however, average CPA pass rate data available on a school by school basis. By averaging accounting professor ratings from individual schools and examining the relationship between those ratings and how well students from that school perform on their first attempts on the CPA exam, the relationship between accounting professor ratings and student learning becomes apparent. This focus is the major point of this study.

To examine this relationship between professor ratings and student learning, the average annual ratings variables are first calculated for junior and senior level required accounting classes at each of the schools in the study. The average ratings data for these accounting classes in any particular year in each school is

then paired with the CPA pass rates two years in the future. This is because undergraduate accounting programs for majors are approximately two years long and professor teaching performance and student learning is revealed approximately two years later when students take the CPA exam.6 Sample statistics for the lagged undergraduate CPA pass rate variable (UGpass) and for the average annual school ratings variables are presented in Table 4. Correlation tables for the average variables in the UGM sample are presented in Table 5.

Two regression models for each of the samples of average data are examined; the first models focus on the relationships between CPA exam pass rates and the average school RMP Quality ratings variable while the second set of models focus on the relationship between pass rates and the average school Helpfulness (Help) and Clarity ratings that are the two components of the average Quality rating for that school and year. The two models are of the following form:

$$UGpass = \alpha + \beta1*(Quality) + \beta2*(Ease) + \beta3*(Interest), \text{ and} \quad (2)$$

$$UGpass = \alpha + \beta1*(Help) + \beta2*(Clarity) + \beta3*(Ease) + \beta4*(Interest) \quad (3)$$

where; UGpass = the school average lagged pass rate for students taking the CPA exam,
Quality = the annual school average of the class level Helpfulness and Clarity ratings,
Help= the annual school average of class level Helpfulness ratings, and
Clarity = the annual school average of class level Clarity ratings, and Interest = the annual school average of class level student Interest ratings.

The results from these regressions provide insight into the relationship between average school ratings and student learning at those schools as evidenced by the associated CPA exam pass rate. If students rate professors and classes as easier, on average, for a school because the demands placed on students have been lessened over the years as many authors suggest, it is to be expected that they would be less prepared for the CPA exam and their school's average passing rates would be lower. This means the data should show a negative relationship between the Ease variable and the CPA Pass rates. If the RMP Clarity and Helpfulness ratings reflect the dimensions of effective teaching identified by Calderon, Gabbin, and Green (1996), there should be a positive relationship between those average ratings levels and success on the CPA exam.

RESULTS

Sample Statistics

Sample statistics are presented for all undergraduate major class ratings data and for each of the two class level samples in Table 1. The mean Quality, Help and Clarity ratings tend to be very similar across both samples with means between 3.26 and 3.40. The mean the Easiness rating is higher for the UGNM's when compared to UGM's (2.49 verses 2.31 for UGM's). The mean Interest rating is noticeable lower for UGNM's than for UGM's (2.91 verses 3.47), which would be expected given that accounting class for non-majors are less rigorous that classes required for majors.

When the professor characteristic dummy variables (Eng, Gen, Hot) are examined, there are only minor differences across ratings for majors and non-majors. When the professor rank dummy variables are examined the data reveal a surprising fact; only 20% of the UGNM classes are taught by tenure track professors indicating 80% are taught by what are referred to as "contingent faculty", a group that includes graduate students, instructors and other part time adjunct faculty members [Huber, Mafi and Flinn (2009)].7 The data indicate that at the UGNM level, 17% of all classes are taught by graduate students (included in that 80% above), and only 20% of the classes are taught by tenure track or tenured faculty. In contrast, in undergraduate major classes (UGM), 54% are taught by tenure track or tenured faculty and another 14% are taught by graduate students.

TABLE 1
SAMPLE STATISTICS FOR ALL DATA AND TWO CLASS LEVEL SAMPLES

All Data			Undergraduate Non-majors (n=2,359)		Undergraduate Majors (n=2,872)	
	(n=5,231)					
Variable	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
Quality	3.31	1.49	3.29	1.48	3.32	1.5
Help	3.34	1.6	3.29	1.59	3.39	1.61
Clarity	3.27	1.5	3.28	1.5	3.26	1.51
Ease	2.39	1.28	2.49	1.31	2.31	1.26
Interest	3.22	1.35	2.91	1.36	3.47	1.3
Eng	0.92	0.27	0.94	0.25	0.91	0.29
Gen	0.55	0.5	0.5	0.5	0.6	0.49
Hot	0.18	0.38	0.17	0.38	0.18	0.39
GS	0.15	0.36	0.17	0.37	0.14	0.35
Assist	0.1	0.3	0.04	0.21	0.15	0.35
Assoc	0.16	0.36	0.05	0.22	0.24	0.43
Prof	0.14	0.34	0.11	0.32	0.15	0.36

Correlations of All Variables for the Entire Sample

Correlation coefficients between the variables are presented in Table 2.8 Correlations of both Help and Clarity with the Quality variable are almost perfectly positively correlated as is to be expected since the two are used to calculate the Quality variable. The correlations between Ease and each of the two main ratings variables (Help, and Clarity) are both above 50%, the correlations between the rater Interest variable and both ratings are approximately 30%, and the correlations between the Hot dummy variable and the two main ratings are both approximately 25%. All these correlations are significant at the .01 level and all these results are similar to results presented in previous studies.

When the correlations between the Eng variable and the ratings variables are considered, all the correlations are very weak with the absolute value of the correlation coefficients below 6%. The gender dummy also shows some statistically significant correlations with other variables with male professors associated with slightly higher quality, Clarity, and Easiness ratings and negatively correlated with English language skills but, again, the relationships are very weak with the absolute values of these coefficients all below 8%.

Finally, when the correlations of the rank variables are considered, some stronger relationships are apparent. Inclusion in the graduate student (GS) teacher group is positively correlated with the Help and Clarity ratings with correlations of between 21% and 22% and positively correlated with the Ease ratings at about 24% (all significant at the .01 level). GS classification is also negatively correlated (at the .01 level) with the Eng dummy (-15%) and positively correlated with being rated Hot (46%), (at the .01 level). An Interesting pattern is apparent when the correlations between classifications of the four categories of instructor rank are considered relative to one another. While the GS group exhibits positive correlations with the ratings variables (Quality, Help, Clarity, and Ease) as noted above, the correlations disappear for the assistant professor group for Quality, Help, Clarity variables and turn negative for the Ease variable. For the associate and full professors, the correlations with all four of these are negative with the correlations for full professors all approximately -10% and significant at the .01 level.

TABLE 2
CORRELATION COEFFICIENTS - ALL VARIABLES FOR ENTIRE SAMPLE

(n=5,231)

	Quality	Help	Clarity	Ease	Interest	Eng	Gen	Hot	GS	Assist	Assoc
Help	0.9620 ***										
Clarity	0.9566 ***	0.8407 ***									
Ease	0.5714 ***	0.5398 ***	0.5571 ***								
Interest	0.3116 ***	0.2913 ***	0.3072 ***	0.2084 ***							
Eng	-0.0133	-0.0299 **	0.0054	-0.0488 **	-0.0403 **						
Gen	0.0241 *	0.0155	0.0312 **	0.0795 ***	0.0206	-0.0780 ***					
Hot	0.2623 ***	0.2429 ***	0.2610 ***	0.2686 ***	0.0736 ***	-0.0069	-0.0017				
GS	0.2167 ***	0.2188 ***	0.1963 ***	0.2547 ***	0.0600 ***	-0.1472 ***	0.1127 ***	0.4756 ***			
Assist	-0.0178	-0.0133	-0.0211	-0.0325 **	0.0314 **	-0.2137 ***	-0.0645 **	-0.0351 **	-0.1408 ***		
Assoc	-0.0395 **	-0.0422 **	-0.0334 **	-0.0829 **	0.0135	0.0330 ***	0.0254 *	-0.0658 **	-0.1821 ***	-0.1431 ***	
Prof	-0.1034 **	-0.1041 ***	-0.0939 ***	-0.1169 **	0.0129	0.0899 ***	0.2849 ***	-0.1618 **	-0.168 ***	-0.1320 ***	-0.1708 ***

*** Significant at the .01 level
** Significant at the .05 level
* Significant at the .10 level

This pattern may suggest that as accounting educators evolve from graduate students to tenured full professors, they become less vulnerable to the impact of lower ratings and more open to resisting course quality deflation and they begin offering more challenging courses that are considered less easy from the student's point of view. They also receive lower ratings, on average, for Helpfulness and Clarity. Alternatively, the pattern may reflect the fact that more senior professors are assigned harder, more challenging upper level classes, and this impacts their ratings. It's also interesting to note that inclusion in the lower rank instructor groups is negatively correlated with English language skills with the GS variable exhibiting a -15% correlation and the Assist variable exhibiting a -11% correlation with the Eng variable. This may reflect the profession's reliance on more foreign students entering the workforce to make up for the lack of enough English-speaking job candidates.

Professor Characteristics and Teaching Ratings

The results of the six regressions on the three teaching ratings variables (Quality, Help, and Clarity) that explore both student attitudes and the professor characteristic variables are presented in Table 3. Three regressions are reported for each of the two samples: the undergraduate non-major sample (UGNM) and the undergraduate major sample (UGM). The Quality teaching variable is calculated as the average of the other two teaching variables on a class by class basis and the Help and Clarity teaching ratings represent the class by class ratings obtained directly from the RMP website.

TABLE 3
REGRESSIONS ON RATINGS VARIABLES FOR TWO SUB-SAMPLES

	Undergraduate Non-Majors			Undergraduate Majors		
	Quality (n=2,359)	Help (n=2,359)	Clarity (n=2,359)	Quality (n=2,872)	Help (n=2,872)	Clarity (n=2,872)
Intercept	0.7731 ***	0.9009 ***	0.6453 ***	1.0209 ***	1.1585 ***	0.8832 ***
	6.21	6.47	5.03	8.95	9.17	7.60
Ease	0.61776 ***	0.6254 ***	0.6102 ***	0.5332 ***	0.5482 ***	0.5181 ***
	32.64	29.57	31.30	26.84	24.93	25.61
Interest	0.1564 ***	0.1444 ***	0.1684 ***	0.2675 ***	0.2659 ***	0.2691 ***
	8.85	7.32	9.26	15.07	13.52	14.89
Hot	0.30434 ***	0.2041 **	0.4046 ***	0.3561 ***	0.3180 ***	0.3942 ***
	4.07	2.44	5.25	5.27	4.25	5.73
Eng	0.53629 ***	0.4558 ***	0.6168 ***	0.0362	-0.0410	0.1134
	5.02	3.82	5.61	0.44	-0.45	1.37
Gen	-0.1746 ***	-0.2282 ***	-0.1210 **	0.0586	0.0257	0.0915 *
	-3.49	-4.09	-2.35	1.16	0.46	1.78
GS	0.27859 ***	0.4229 ***	0.1343 *	0.1388 *	0.1804 **	0.0972
	3.56	4.84	1.67	1.70	2.00	1.17
Assist	0.6070 ***	0.6570 ***	0.5570 ***	-0.0794	-0.1147	-0.0442
	4.8	4.65	4.28	-1.09	-1.43	-0.60
Assoc	0.1448	0.0472	0.2424 **	0.0064	-0.0240	0.0369
	1.38	0.40	2.25	0.10	-0.35	0.58
Prof	-0.2486 ***	-0.2842 ***	-0.2129 ***	0.0068	-0.0062	0.0197
	-3.15	-3.23	-2.62	0.09	-0.08	0.26
F-statistic	210.01 ***	173.25 ***	193.97 ***	166.27 ***	140.79 ***	154.92 ***
Adjusted R-square	0.4437	0.3967	0.4241	0.3413	0.3047	0.3255

*** Significant at the .01 level
** Significant at the .05 level
* Significant at the .10 level

All the regression models are significant at the .01 level and with adjusted r-square values of between 30% and 45%. The results also indicate that both samples often exhibit different patterns of significant coefficients for some explanatory variables. There are also, however, important similarities across both samples with all regressions exhibiting strong positive relationships (significant at the .01 level) between the teaching ratings and the two class-oriented student ratings (Ease and Interest). Classes that students consider to be easier or classes in which students have a stronger Interest tend to receive higher teaching ratings. Past studies that focus on the RMP average quality variable calculated from all a professor's class ratings report this same pattern of significant relationships. When the relationship between student attitudes towards professor physical attractiveness (Hot) and teaching ratings is considered, the relationships are strong in both undergraduate samples with the positive relationships significant at the .01 level in five of the six regressions and significant at the .05 level in the sixth regression.

When the other explanatory variables are considered, some differences across samples and across teaching ratings variables do become apparent. When the English language proficiency variable (Eng) is considered, this professor characteristic appears to be very important to students in the non-major classes with strong positive relationships (significant at the .01 level) apparent in all three teaching ratings regressions. In contrast, there are no significant relationships between professor English language proficiency and teaching ratings for the undergraduate major account classes. The differences in the undergraduate sample results may be because classes for non-majors are often taught by graduate students and Table 2 indicates that there is a negative correlation (-15.09%) between the GS dummy variable and the Eng dummy variable.

Differences across the undergraduate samples are also apparent when the gender variable (Gen) is examined with strong negative relationships apparent for the non-major class sample (at the .01 level) but no negative relationships apparent in the accounting major sample. For the non- major classes, male faculty are considered significantly less helpful (at the .01 level) and are considered to lack Clarity (at the .05 level) in their presentations. These significant negative relationships are not apparent in the classes for majors.

The results for the professor rank dummy variables also reveal differences across the different samples. In the non-major classes, graduate students and assistant professors are given significantly higher Help and Clarity ratings (three of four coefficients significant at the .01 level) while full professors are given significantly lower ratings on these variables (at the .01 level). All these significant relationships disappear for both assistant and full professors when the major classes are examined.

The results of the regressions of professors and class characteristics on the three ratings variables provide evidence in support of the following conclusions. First, students' attitudes towards class Easiness and their Interest in those classes are positively related to higher ratings for instructor Helpfulness, Clarity, and overall Quality in both samples. Second, the positive relationship between students' perceptions of instructor attractiveness and ratings is strong.

Third, the relationship between ratings and English language proficiency is positive and significant in the non-major classes, but not in the major classes. Fourth, the gender variable is negatively related to ratings in non-major classes indicating male professors receive lower ratings but is not significant in major classes. Finally, professor rank is related to professor ratings in non-major classes with graduate students and assistant professors receiving higher ratings and full professors earning lower ratings, but this pattern is not apparent in the classes for majors.

The Links Between Professor Ratings and Student Learning

The sample statistics for the UGM sample are presented in Table 4. The observations in the samples represent school year averages for the instructor ratings variables (Quality, Help and Clarity), the student attitude ratings (Ease and Interest), and for the lagged CPA pass rates (UGpass). Note the very large ranges for some variables. The annual CPA pass rates range from a low of 26% to a high of 74% across the data for different schools and different years. When the ratings variables are considered, there are also wide ranges for the three ratings variables. The ranges are smaller for the Ease and Interest variables.

TABLE 4
SAMPLE STATISTICS FOR SCHOOL YEAR AVERAGES

Undergraduate Accounting Majors (n=64)

Variable	Mean	Std Dev	Min	Max
UGpass	0.56	0.11	0.26	0.74
Quality	3.27	0.43	2.11	3.98
Help	3.34	0.51	2.09	5.00
Clarity	3.20	0.44	2.00	4.00
Ease	2.20	0.39	1.40	3.13
Interest	3.49	0.35	2.89	4.75

Correlations between Teaching Ratings and CPA Exam Average Pass Rates

Table 5 presents the correlations for the variables in the UGM sample. While the correlations between the Quality, Help, and Clarity variables reflect the pattern of relationships seen earlier in Table 2, the interesting results are those representing the relationships between the average student ratings at a school in any year and the CPA passing rate for that school two years later. Professor Clarity has a significant positive correlation with the lagged passing rate of approximate 30% (significant at the .05 level) and since Clarity is used to calculate the overall quality ratings, quality is also positively correlated to the lagged pass rates (at the .10 level). The average class Easiness rating, however, is negatively correlated to the lagged average school passing rate (significant at the .10 level) while the student Interest measure is not significantly correlated with pass rates.

TABLE 5
CORRELATION COEFFICIENTS - AVERAGE RATINGS AND LAGGED CPA PASS RATES

(n=64)

	UGpass	Quality	Help	Clarity	Ease
Quality	0.2319 *				
Help	0.1310	0.9121 ***			
Clarity	0.2968 **	0.8834 ***	0.6136 ***		
Ease	-0.2245 *	0.4517 ***	0.5381 ***	0.2549 *	
Interest	-0.0538	0.1099	0.1444	0.0465	0.0867

*** Significant at the .01 level
** Significant at the .05 level
* Significant at the .10 level

The Relationship Between Teaching Ratings and Student Learning

The results of the regressions examining the relationships between RMP ratings and teaching effectiveness, as measured by lagged CPA pass rates, are presented in Table 6. The two regressions on the undergraduate major class ratings are significant (at the .05 level or better) with adjusted r-square values of 14% to 16%. In the first regression that includes the school wide average Quality ratings, the Quality coefficient is positively rated to the lagged pass rate while the average Easiness rating is negatively related to the passing rate (both relationships significant at the .01 level). The second regression that replaces the Quality rating with its two components, Help and Clarity, show that a school's average instructor Clarity is positively related to that school's average pass rates (at the .10 level) and the average perceived Easiness of classes is negatively related (at the .01 significance level) to the pass rate. These regression results suggest that for undergraduate accounting classes, higher professor teaching ratings are positively

associated with greater student learning as measured by CPA exam pass rates and that class and professor ratings of Easiness are negatively related to CPA pass rates.

TABLE 6
REGRESSIONS ON CPA PASS RATES

	UGpass (n=64)	UGpass (n=64)
Intercept	0.5258 **	0.5057 **
	3.52	3.28
Quality	0.1050 ***	
	3.23	
Help		0.0340
		0.94
Clarity		0.0707 *
		1.98
Ease	-0.1109 ***	-0.1034 ***
	-3.13	-2.73
Interest	-0.0194	-0.0173
	-0.55	-0.49
F-statistic	4.74 ***	3.59 **
Adjusted R-square	0.1531	0.1433

*** Significant at the .01 level
** Significant at the .05 level
* Significant at the .10 level

SUMMARY AND CONCLUSIONS

This study examines both the relationships between accounting professor teaching ratings and characteristics of those professors and the relationship between average student ratings of accounting professors at different schools and student learning as measured by the average pass rates on the CPA exam at those schools. The major findings are as follows. First, for classes taken by both majors and non-majors, there are strong positive relationships between ratings and student perceptions of the Easiness of the professor and class, their Interest in the class, and their perceptions of the professor's physical attractiveness. These results are similar to results reported in many different studies covering different disciplines, but they do not provide any insight into whether or not ratings reflect student learning.

The more important findings reported in this study is the significant positive relationship between average Quality ratings for professors teaching accounting major classes at individual schools and the average pass rate for students from those schools when they sit for the CPA exam two year later and the significant negative relationship between student's perceptions of the Easiness of their classes and their schools average passing rates on the exam. These results suggest that overall quality teaching ratings do, on average, reflect student learning and that classes that are perceived as being harder and more demanding better prepare students for taking and passing the CPA exam.

ENDNOTES

1. Since the data was collected for this research paper, RMP has changed the format of their ratings' collections interface and no longer collect data on student's perceptions of professor Helpfulness or Clarity, on student Interest in the class, or on students' perceptions of professor attractiveness.
2. Individual school program information is obtained from the interactive database containing information from the 2013-14 AACSB Business School Questionnaire (BSQ) available at https://datadirect.aacsb.edu/public/profiles/search.cfm.
3. RMP data include ratings for all types of faculty including tenured and tenure track professors, instructors, graduate teaching assistants and other adjunct faculty but this paper will use the generic term "professors" to refer to all these instructor categories.
4. See Appendix I for the RMP information on how the ratings variables were defined or calculated when this study's data were collected.
5. Because of changes in the data RMP collects and publishes, it is impossible to expand this study to include more recent data.
6. In addition to the results reported here the analyses were repeated with one year lags and the pattern of significant results was identical.
7. Huber, Mafi and Flinn (2009) provide an extensive discussion of the reasons for the heavy reliance on contingent faculty in business schools and the implications for accounting education programs.
8. While there are many highly significant (at the .01 level) correlation coefficients shown in the table not all the statistically significant correlations are large enough to be of interest since significance may be driven, in part, by the large sample size.

REFERENCES

Barth, M.M. (2008). Deciphering Student Evaluations of Teaching: A Factor Analysis Approach. *Journal of Education for Business*, *84*, 40–46.

Beleche, T., Fairris, D., & Marks, M. (2012). Do course evaluations truly reflect student learning? Evidence from an objectively graded post-test. *Economics of Education Review*, *31*(5), 709–719.

Brown, M.J., Baillie, M., & Fraser, S. (2009). Rating Ratemyprofessors.com: A Comparison of Online and Official Student Evaluations of Teaching. *College Teaching*, *57*(2), 89–92. DOI:10.3200/CTCH.57.2.89-92

Bruns, S.M., Rupert, T.J., & (May) Zhang, Y. (2011). Effects of converting student evaluations of teaching from paper to online administration. *Advances in Accounting Education*, *12*, 167–192.

Calderon, T.G., Gabbin, A.L., & Green, B.P. (1996). Summary of promoting and evaluating effective teaching: American Accounting Association teaching and curriculum section. *Journal of Accounting Education*, *15*, 367–383.

Calderon, T.G., Green, B.P., & Reider, B.P. (1994). Extent of use of multiple information sources in assessing accounting faculty teaching performance. In *American Accounting Association Ohio Regional Meeting Proceeding*.

Centra, J.A. (2003). Will Teachers Receive Higher Student Evaluations by Giving Higher Grades and Less Course Work? *Research in Higher Education*, *44*, 495–518. https://doi.org/10.1023/A:1025492407752

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

Constand, R.L., & Pace, R.D. (2014a). Student Evaluations of Finance Faculty: Perceived Difficulty Means Lower Faculty Evaluations. *Journal of Financial Education*, *40*, 14–43.

Constand, R.L., & Pace, R.D. (2014b). Addendum - Student Evaluations of Finance Faculty: Perceived difficulty means lower faculty evaluations. *Journal of Financial Education*, *40*, 177–178.

Constand, R.L., Pace, R.D., & Clarke, N. (2016). Accounting faculty teaching ratings: Are they lower because accounting classes are more difficult? *Journal of Accounting and Finance*, *16*(4), 70–86.

Crumbley, D.L., & Smith, S. (2009). The Games Professors Play in the Dysfunctional Performance Evaluation System Used in Higher Education: Brainstorming Some Recommendations. *Measure Learning Rather Than Satisfaction in Higher Education*, pp. 41– 58.

Davison, E., & Price, J. (2009). How do we rate? An evaluation of online student evaluations. *Assessment & Evaluation in Higher Education*, *34*(1), 51–65. DOI: 10.1080/02602930801895695

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

Felton, J.F., Koper, P.T., Mitchell, J.B., & Stinson, M. (2008). Attractiveness, Easiness, and other issues: Student evaluations of professors on RateMyProfessors.com. *Assessment & Evaluation in Higher Education*, *33*, 45–61.

Flinn, R.E., & Crumbley, D.L. (2009). Multiple teaching performance measures needed: SET management similar to earnings management. *Measure Learning Rather than Satisfaction in Higher Education*.

Fogarty, T.J., Jonas, G.A., & Parker, L.M. (2013), The medium is the message: Comparing paper-based and web-based course evaluation modalities. *Journal of Accounting Education*, *31*(2), 177–193.

Green, B.P., & Wang, G. (2012). Identifying sources of evidence used to assess faculty teaching performance. *Advances in Accounting Education*, *13*, 109–126.

Green, B.P., Calderon, T.G., Gabbin, A.L., & Habegger, J.W. (1999). Perspectives on implementing a framework for evaluating effective teaching. *Journal of Accounting Education, 17*, 71–98.

Huber, M., Mafi, S., & Flinn, R. (2009). *A Balanced Look at the use of Contingent Faculty in Accounting Education*. Accounting Student Learning, American Accounting Association.

Larry Crumbley, D., Flinn, R., & Reichelt, K.J. (2012). Unethical and deadly symbiosis in higher education. *Accounting Education, 21*(3), 307–318.

Marsh, H.W. (1982). Factors Affecting Student's Evaluations of the Same Course Taught by the Same Instructor on Different Occasions. *American Education Research Journal*, *19*, 485–497.

Marsh, H.W. (1984). Students' Evaluations of University Teaching: Dimensionality, Reliability, and Validity, Potential Biases and Utility. *Journal of Educational Psychology*, *76*, 707–754.

Marsh, H.W. (1987). Students' Evaluations of University Teaching: Research Findings, Methodological Issues, and Directions for Future Research. *International Journal of Educational Research*, *11*, 253–387.

Moore, P. (2009). Why We Should Measure Student Learning: A Glossary of Collegiate Corruption. *Measure Learning Rather than Satisfaction in Higher Education*, pp. 17–28.

Moore, P., & Flinn, R.E. (2009). The Limitations of Measuring Student Learning. *Measure Learning Rather than Satisfaction in Higher Education*, pp. 91–110.

Morgan, D.A., Sneed, J., & Swinney, L. (2003). Are Student Evaluations a Valid Measure of Teaching Effectiveness: Perceptions of Accounting Faculty Members and Administrators. *Management Research News*, *26*, 17–32.

Nagle, B.M., Menk, K.B., & Rau, S.E. (2018). Which accounting program characteristics contribute to CPA exam success? A study of institutional factors and graduate education. *Journal of Accounting Education*, *45*, 20–31.

Public Accounting Report. (2008). *27th Annual Professor's Survey – 2008, Top 25 undergraduate Programs*. Public Accounting Report - The Independent Newsletter of the Accounting Profession Since 1978, October 7-8.

RateMyProfessor. (n.d.). *Rating description*. Retrieved December 2014, from http://www.ratemyprofessors.com/help.jsp

Silva, K.M., Silva, F.J., Quinn, M.A., Draper, J.N., Cover, K.R., & Munoff, A.A. (2008). Rate my professor: Online evaluations of psychology instructors. *Teaching of Psychology*, *35*, 71–80.

Timmerman, T. (2008). On the Validity of RateMyProfessor.com. *Journal of Education for Business*, *84*, 55–61.

UCLA: Academic Technology Services, Statistical Consulting Group. (n.d.). *Introduction to SAS*. Retrieved September 2010, from http://www.ats.ucla.edu/stat/sas/output/ttest.htm

Uttl, B., White, C.A., & Gonzalez, D.W. (2017). Meta-analysis of faculty's teaching effectiveness: Student evaluation of teaching ratings and student learning are not related. *Studies in Educational Evaluation, 54*, 22–42.

Watson, S.F., Apostolou, B., Hassell, J.M., & Webber, S.A. (2007). Accounting education literature review (2003–2005). *Journal of Accounting Education, 25*, 1–58.

Yunker, P.J., & Sterner, J. (1988, Fall). A Survey of Faculty Performance Evaluation in Accounting. *Accounting Educators Journal*, pp. 63–71.