Hot and easy in Florida: The case of economics professors

Oliver Schnusenberg
The University of North Florida

Cheryl Froehlich
The University of North Florida

Abstract

We investigate whether overall quality ratings recorded by students on RateMyProfessors.com are related to the perceived easiness and hotness RMP ratings for a sample of 110 economics professors in Florida. Results indicate that students who utilize RMP reward easy and hot economics professors with higher RMP overall quality ratings. Second, the correlation between RMP overall quality ratings and overall in-class instructor ratings is about 0.53 for a subsample of instructor ratings. Third, only clarity is significant when regressing the overall in-class instructor ratings on RMP clarity, helpfulness, easiness, and hotness ratings.

Keywords: Student evaluations, RateMyProfessors, clarity of instruction, online evaluation

I. PURPOSE AND MOTIVATION

The concept that students should evaluate their professors was first introduced in the mid-1920s at the University of Washington. Since then, however, years of confusion, discontent, and concern about the usefulness of student evaluations have ensued. Furthermore, the pragmatic use of student evaluations by administrators to determine a faculty member's tenure, promotion, and merit pay raises only added fuel to the myriad of faculty concerns with the concept of student evaluations (Algozzine, Beattie, et al. [2004]).

With the advent of the Internet, the ease of securing information has increased exponentially and some university websites now post university-sanctioned student evaluation results. Additionally, there are several privately-owned sites that allow students to anonymously enter their evaluation of an instructor and/or course at a university using that site's questionnaire, which is typically different from the university-sanctioned questionnaire. RateMyProfessors.com (henceforth RMP) is one of the online services that allow students to anonymously rate their professors. The service has been available to students since 1999. As of March 2005, the site boasted an impressive 3,132,184 ratings for 4,467 universities, with an average of 4,000 ratings per day added to the site.

Students visiting RMP can rate their professors in four major categories: helpfulness, clarity, easiness, and hotness. The helpfulness and clarity factors are combined into an "overall quality" rating, such that the overall RMP quality rating is a simple average of the helpfulness and clarity factors. The RMP helpfulness, clarity, and easiness ratings have a 1 to 5 scale. RMP only defines ratings of 1 and 5 for the helpfulness and clarity criteria. A rating of 1 is defined as "very unhelpful" and "very unclear," respectively for helpfulness and clarity. A rating of 5 is defined as "very helpful" and "very clear," respectively. Easiness is also rated on a 1 to 5 scale, with a rating of 1 being defined as "very hard" and a rating of 5 as "very easy." The hotness criterion is indicative of a given professor's appearance. "Hotness" is a score on RMP that is a function of "hot" votes and "not hot" votes; students have to rate a professor as either "hot" or "not." For every "hot" vote, the score increases by 1. For every "not hot" vote, the score decreases by 1. A chili pepper is displayed next to the teacher's name if "hotness" is positive. Negative "hotness" scores are not displayed, but they do exist. For example, if a professor has ten hot votes and twelve not hot votes, no chili pepper is displayed, but the fact that the not hot votes exceed the hot votes is also not mentioned. Furthermore, if the next student rates this professor as hot, such that there are eleven hot and twelve not hot votes, a chili pepper would still not be displayed.

The objective of this study is threefold. First, we investigate whether a professor's RMP overall quality rating, which is the average of the students' rating of the professor's clarity and helpfulness, is positively related to the easiness and hotness criteria on RMP for a sample of 110 economics professors at public universities in Florida. A finding indicating that easiness and/or hotness are relevant factors in determining a professor's overall quality rating on RMP would indicate that the self-selected sample of students using RMP reward easier and more attractive professors with higher RMP overall quality ratings.

Second, for a subset of 38 economics professors at public universities in Florida, the RMP overall quality ratings are examined to determine whether they are positively correlated to university-sanctioned in-class questionnaires' overall instructor ratings. If the RMP overall quality ratings for the professors are not highly correlated to their university-sanctioned student questionnaires' overall instructor rating, then it stands to reason that either students use different

criteria to evaluate professors in the classroom versus on RMP or that only a particular subset of students from any given class sees the need to evaluate professors on RMP, which could be indicative of a response bias (Sorensen and Johnson, 2003)

Third, the overall in-class instructor ratings for the subset of economics professors are examined to determine if they are significantly related to the RMP clarity, helpfulness, easiness, and hotness ratings. Since many universities use in-class teaching evaluations as their primary tool to evaluate teaching effectiveness and in turn use teaching effectiveness as a criterion to award merit pay raises, promotion, and/or tenure, an investigation into the factors underlying ratings recorded on RMP and their relationship to overall in-class ratings is warranted. Particularly, RMP provides a unique opportunity to investigate whether the perceived helpfulness, clarity, easiness, and hotness of a professor are related to university-sanctioned overall in-class instructor ratings.

All three objectives will be accomplished by investigating instructors in the economics department at public universities in Florida. In alphabetical order, these universities are: Florida Agricultural and Mechanical University, Florida Atlantic University, Florida Gulf Coast University, Florida International University, Florida State University, New College of Florida, The University of Central Florida, The University of Florida, The University of North Florida, The University of South Florida, and The University of West Florida.

The remainder of this paper is organized as follows. Section II provides a general description of RMP. Section III provides a review of related literature. Section IV presents the data and methodology. Results are presented in Section V. Section VI concludes and provides some implications.

II. A DESCRIPTION OF RATEMYPROFESSORS.COM

The stated purpose of RMP is to be a free resource for students. To rate a professor on RMP, students select their state and their university in that state. If the professor that a student is trying to rate is already listed on the site, students simply click on the name of the professor to rate them. If the professor is not listed, students can add the professor and rate them. Consequently, students can anonymously rate a given professor at their university along the criteria of clarity, helpfulness, easiness, and hotness. Students using the site can leave ratings without creating an account, but the site allows students to create either a free account or a gold account, which allows users to contact other registered users to obtain their opinion about a professor or course. In addition, students can leave written comments for professors on RMP.

In addition to leaving anonymous ratings for professors, RMP users also have the ability to suggest changes. For example, if a professor's name is spelled wrong, users have the ability to click on a link to report the error. Furthermore, if a rating is inappropriate, users have the ability to flag the rating to have it reviewed. These suggested changes are handled by "School Administrators." These administrators also accept new teachers, and are part of the RMP staff, working with the site owners and operators. To qualify as an administrator, individuals must be students at that school and have an RMP account.

General guidelines for ratings are provided by RMP, and a direct link to these guidelines is presented to RMP users when they enter their ratings for a professor. These are reproduced from the website in Table 1 (Appendix). RMP promises to enforce the "do not's" with the help of the administrators. Violations of the guidelines will result in the rating's comment being removed or the entire rating being deleted.

Given the anonymity students using RMP enjoy, there are some general potential pitfalls of the site. For instance, "The Valley Star," the independent student newspaper at Los Angeles Valley College, featured an article about RMP on February 23, 2005, in which a student pointed out that some students may use the site to get back at a teacher for giving a bad grade, while at the same time acknowledging that students can use the site to provide each other with suggestions for good teachers. Similarly, on September 7, 2004, Professor Kenneth Westhues from the University of Waterloo wrote an article in "The Record" warning against the pitfalls of the site. Specifically, he points out that a certain number of evaluations published on the site are noise. Furthermore, while cookies and filters attempt to prevent multiple ratings by the same user or by users who have never taken the professor being rated, it is impossible to prevent such abuses entirely. Another point of concern is that institutions vary in terms of how carefully they monitor comments, due to the nature of the school administrators. These points are also raised in the September 2001 "Teaching Matters Newsletter" at the University of Waterloo.

Although it is possible that some students will use RMP to post inaccurate or inappropriate comments, there are several reasons to believe students will not do so. First, RMP states on its website that over 65 percent of the comments left by users are positive in nature. Furthermore, while RMP acknowledges that the ratings recorded by users are simply a collection of opinions and, as such, not statistically valid, the administrators of the site state that they receive a large amount of emails telling them that the ratings are highly accurate. Second, the authors of the September 2001 University of Waterloo "Teaching Matters Newsletter" attempt to assess the statistical validity of the site by examining the rating of all Distinguished Teacher Award winners at the University of Waterloo. Of the 16 award winners, 15 were rated in the high-quality category. Moreover, the total number of instructors rated in the high-quality category was 124, followed by 72 professors rated in the middle category, and 54 in the low category. The authors mention that this is consistent with the way instructors are rated on course evaluations, since university instructors are generally rated above average on course-evaluation instruments. However, no rigorous statistical analysis was conducted in the newsletter to verify this observation empirically. Third, Giesey, Chen, and Hoshower (2004) find that a sample of engineering students generally consider the improvement of teaching to be the most important outcome of teaching evaluations. Therefore, although RMP ratings of professors are not university-sanctioned, if students believe that professors will attempt to improve their teaching performance based on ratings left at RMP, they may use the site to provide constructive criticism of professors rather than using the site to leave inappropriate or inaccurate ratings for professors.

If students utilize RMP to provide constructive criticism of professors, while acknowledging that they are not necessarily representative of all the students taking classes with a given professors, ratings left on RMP can be useful to professors. Insofar RMP helps professors identify their own weaknesses early in the semester, it can be a useful tool to both support teacher growth and enhance teacher professionalism, two common themes in research on teaching evaluations (Bernstein, 2004). For instance, Andrews (2004) argues that teachers will support an evaluation program if there is adequate time to allow one to remediate defects and deficiencies. Similarly, Algozzine, Beattie, et al. (2004) suggest the use of interim evaluations to provide formative information about ways to improve teaching. Both of these studies indicate that teachers can use early feedback to improve their teaching performance. RMP can be particularly beneficial, since students can leave ratings and comments at any time during the semester, and not only when the semester ends. Thus, professors are provided with feedback

throughout the semester, which may allow them to improve. The ultimate result could be improved results on university-sanctioned in-class evaluations.

While RMP ratings have the potential to provide professors with early feedback, they also allow us to investigate whether students, at least the self-selected sample of students using RMP, value the easiness and attractiveness of a professor. Specifically, RMP ratings are used to investigate three objectives: (1) how the RMP overall quality rating for economics professors at public universities in Florida relates to the perceived easiness and hotness of a professor, (2) how the site's overall quality rating for a professor correlates to the professor's overall in-class instructor rating, and (3) how the overall in-class instructor rating relates to the perceived clarity, helpfulness, easiness, and hotness of the professor on RMP.

III. REVIEW OF RELATED LITERATURE

The literature on student evaluations is large and dates back to the 1920s. This section focuses on the literature investigating the perception of student evaluations by various parties and professor characteristics influencing student evaluations.

There is some evidence that student evaluations are viewed differently by university administrators and faculty. Calderon and Green (1997), for example, find that 95 percent of administrator responses to their survey view student evaluations as a valid criterion of teaching effectiveness. Similarly, Morgan, Sneed, and Swinney (2003) find that administrators believe student evaluations measure teaching effectiveness more effectively than faculty. Conversely, faculty members believe that their personality is the primary determinant of ratings on student evaluations, although course work load, the type of course, and the grade distribution also have an impact on evaluations. Other papers showing these relationships are by Kemp and Kumar (1990); Yunker and Sterner (1988); Simpson (1995); and Marsh and Overall (1981).
Since attractiveness is one aspect of a teacher's personality, an investigation of RMP hotness ratings and their impact on both RMP overall quality and overall in-class instructor ratings affords us an opportunity to investigate whether personality affects student evaluations.

A variety of articles have investigated teacher characteristics and their impact on student evaluations. For example, gender (Marsh and Roche [1997]) is found not to have a significant influence on evaluations, while the experience of the instructor may bias student ratings (Cohen [1981, 1982]). Instructor enthusiasm and expressiveness were also found to significantly influence student evaluations. See, for example, Shevlin, Banyard, Davies, and Griffiths (2000) and Abrami, Perry, and Leventhal (1982).
Certain student characteristics, such as student interest in the course and student/teacher attitude similarities can also influence evaluations. Also see Granzin and Painter (1973); Marsh (1980); and Marsh and Roche (1997).

Several articles in the literature also investigate the evaluation factors considered important by students. For example, Giesey, Chen, and Hoshower (2004) find that students consider the improvement of teaching, the improvement of course content and format, making evaluation results available for students' decisions, and tenure and promotion decisions the most important outcomes of teaching evaluations. Landrum and Dillinger (2004) find that the three most important factors predicting course evaluations are recommendations among students to take a course, the expected grade, and the overall instructor rating. Wilhelm (2004) provides additional insight and finds that students are twice as likely to choose a course with an instructor who receives excellent course evaluations. Moreover, students are willing to put up with poor

course evaluations or a heavy workload if they believe that they will gain a great deal of useful knowledge. Nevertheless, Grimes, Millea and Woodruff (2004) find that those students with an external locus of control, who are more likely to blame factors other than themselves for their failure, tend to provide lower evaluations and blame the instructor for their performance. The authors also point out that teaching strategies and advice that result in students assuming control for their own learning lead to both academic success and higher teaching evaluations.

Combined, these studies indicate that students are more likely to take professors who receive good evaluations, and that these evaluations are at least partly based on the expected, especially for students with an external locus of control. Since RMP provides students with an ability to rate easiness, and since RMP allows students to post ratings prior to the end of the semester (i.e., based on expected grades), this study may reveal additional insight into factors considered important by students.

Several articles in the empirical literature investigate the effect of looks and/or beauty on salary and/or performance. For instance, Hamermesh and Biddle (1994) examine the impact of looks on earning using interviewers' ratings of respondents' physical appearance. They find that plain people earn less than average-looking people, which they label a plainness penalty. This penalty is about 5 to 10 percent. They also find a slightly less pronounced beauty premium, which is the difference in salaries between good-looking people and average-looking people. Furthermore, Hamermesh and Biddle find no significant differences in looks across genders or occupational classes. In a student evaluation setting, this indicates that students may award more attractive professors with higher evaluations.

The study by Hamermesh and Parker (2003) is most closely related to the present study. The authors directly investigate whether beauty leads to differences in productivity that may generate earnings differences. To accomplish their objective, the authors investigate student instructional ratings for a group of university professors and acquire six independent measures of their beauty. Hamermesh and Parker find that instructors who are viewed as better-looking receive higher instructional ratings. Moreover, the impact exists within university departments and even within particular courses, and is larger for male than for female instructors. The authors acknowledge that their measure of beauty could merely be a proxy for a variety of related unmeasured characteristics that might positively affect instructional ratings. Furthermore, it may be that students pay more attention to good-looking professors and learn more, which would be a productivity effect.

Given the increased exposure and publicity of RMP and the documented relationship between appearance and salary in both teaching and non-teaching professions, an investigation of both the nature of the overall quality ratings posted on RMP and their relationship to overall in-class ratings of professors is warranted. For example, if the RMP ratings are significantly affected by easiness and hotness ratings on RMP, this would indicate that students (at least the self-selected sample using RMP) value a professor's level of easiness and a professor's appearance. Furthermore, if the university-sanctioned overall in-class ratings are significantly related to helpfulness, clarity, easiness, or hotness, then we gain some insight into the factors used by students to assign the "overall instructor rating," even though this is only true for the subsample of students completing both the in-class questionnaires and using RMP. Since many universities use in-class student evaluations as their primary indicator of teaching effectiveness, this is an important research topic, particularly since at least two of the RMP factors (easiness and hotness) are not available on university-sanctioned course evaluations.

The present study contributes to the existing literature in at least three ways. First, we contribute to the ongoing discussion about whether "hotness" or "easiness" directly leads to differences in productivity as measured by in-class instructor evaluations by regressing overall in-class instructor ratings on measures of "hotness" and "easiness" recorded by students on RMP. Second, by using student ratings of "hotness" and "easiness," we are able to investigate directly whether it is these ratings themselves that cause differences in the overall ratings, both those provided by students on RMP and in the classroom. That is, unlike Hamermesh and Parker, we do not have to obtain exogenous measures of attractiveness (or beauty), since it is a measure on RMP. Nonetheless, we acknowledge that the RMP hotness rating could proxy for another unmeasured factor. Third, based on Hamermesh and Parker's findings, we isolate economics professors.

IV. DATA AND METHODOLOGY

IV.1. Data

To investigate our first objective, whether RMP overall quality ratings are related to the perceived easiness and hotness, we obtained our data from RMP. Specifically, our sample consists of all ratings posted on RMP for the previously identified eleven public universities in the state of Florida. The ratings were sorted by department and the overall quality, helpfulness, clarity, easiness, and hotness ratings for each course, the total number of ratings for each professor, and the total number of economics professors rated from each university were tabulated for each course taught in the economics department at these universities. Additionally, we determined the gender of each professor using the professor's first name. The total number of economics professors for which all data items were available was 173.

To accomplish our second and third objectives of investigating whether the overall quality ratings from RMP are positively correlated with the overall in-class instructor ratings for a subset of public Florida universities and whether the RMP variables of clarity, helpfulness, easiness, and hotness are related to the professors' in-class evaluations, we obtained overall in-class instructor evaluations for the sample universities. Results for the most recent available semester in-class instructor ratings were obtained from those universities that made these results publicly available. Results for in-class evaluations were available for Florida Atlantic University, Florida International University, the University of Florida, The University of North Florida, and The University of South Florida. The last evaluations for Florida International University and The University of South Florida were available for the Spring 2004 semester. Evaluations for the other three universities were available for the Fall 2004 semester. Of the five institutions in Florida that published their instructor ratings online all except for Florida Atlantic University (FAU) used a 1 to 5 scale, with "1" being poor and "5" being excellent. FAU uses a 1 to 5 scale, with "1" being excellent and "5" being poor. Consequently, the FAU scale was converted to be synchronous with the other universities' scales. The total number of professors for which the overall in-class ratings were available on the university website was 60.

For each university, the size of the university, as measured by student enrollment as reported on the last available fact sheet published on the university website, was also obtained. In addition to the variables listed above, each professor on RMP was required to have at least three overall quality ratings. This additional control was imposed to prevent lone ratings left by students who were "testing out" the site or student who left ratings for nonexistent professors

that have not yet been removed by the universities' RMP administrators. In addition to three ratings, the models were also repeated for a minimum number of four and two ratings. Although the sample sizes were smaller, the results are very similar to the results reported here and are available from the authors upon request. After including the descriptive variables and imposing the additional restriction of at least three overall quality ratings per professor, the final sample size was 110 total quality ratings from RMP and 38 overall in-class ratings.

Summary statistics for the 110 RMP ratings in our sample are provided in Table 2 (Appendix). The average enrollment (including graduate and undergraduate students) of the eleven public universities in Florida was 23,881 students. The smallest public university in Florida is New College of Florida with an enrollment of 692 students. The largest public university in Florida is the University of Florida with an enrollment of 48,765 students. Across the eleven public universities in Florida, an average of 28.45 economics instructors were rated on RMP, with a minimum of three economics professor for New College of Florida and a maximum of 43 economics professors for the University of Central Florida. We were able to obtain the total department size for nine of the eleven universities. The average department size for those nine universities was 20.6 faculty members. We did not obtain department sizes for Florida Agricultural & Mechanical University and for Florida Gulf Coast University. As shown in Table 2, the average value for the male dummy variable was 0.79, indicating that 79 percent of the 110 economics professors rated were male.

The average "hotness" rating for an economics professor at a public university in Florida is 7.89 percent. This means that, on average, 7.89 percent of the ratings recorded on RMP indicate that the professor is "hot." Moreover, for six professors all of the ratings indicated that the professor was "hot."

Table 2 also shows that the RMP overall quality rating for economics professors at public universities in Florida was 3.36 (median = 3.35), with a standard deviation of 0.88. As discussed earlier, the total quality rating is an average of the helpfulness and clarity ratings students leave on RMP. The respective averages for these categories are 3.43 and 3.28, with respective standard deviations of 0.92 and 0.91. The average easiness rating on RMP was 3.00 (median = 3.00), with a standard deviation of 0.81.

The nature of the RMP data is subject to self-selection bias on part of the students participating in those ratings. This possible negative response bias present in online evaluations is mentioned by Sorenson and Johnson (2003), who point out that response rates may also be lower in an online setting, since students can complete evaluations on their own time or not at all. Additionally, the participation rates for the overall in-class ratings are not readily available, since the majority of the universities investigated display only summary results for a given professor instead of the complete sheet providing the response rate for each class and professor.

Despite the self-selection bias inherent in using RMP ratings and the lack of response rates, we believe that analyzing the RMP ratings by themselves provides some valuable information about how the group of students utilizing the RMP site, considered as a population, views the professors. Furthermore, investigating the subsample of overall in-class ratings may provide some valuable information about the relationship between the "overall instructor" ratings and their relationship to criteria available on RMP, such as easiness and hotness. We acknowledge, however, the nature of the self-selection bias of the RMP data and the limited size of the overall in-class ratings, which may significantly bias our results.

In order to get a better sense of how and why students utilize RMP, a brief survey was conducted using a sample of 238 students at The University of North Florida (UNF). The survey

is reproduced in Table 3. As an incentive to complete the survey, extra credit equal to 2 percentage points on the final exam in the Fall 2005 semester was offered. The survey had to be completed on the Blackboard System. The survey administrator had no way of determining which specific answers a student provided, only whether or not a student completed the survey, assuring anonymity. This was clearly communicated to students both in class and in the survey instructions. A brief description of the site was also provided at the beginning of the survey. The survey was refined with the help of two students to avoid ambiguity.

The survey was conducted in two undergraduate Financial Management classes and one graduate Advanced Financial Management class at UNF. In the three classes combined, 377 students were eligible to complete the survey. Thus, the response rate was 63.13 percent. While students in the surveyed classes are probably not overlapping with the students conducting the ratings in our sample, the survey provides some indication how and why students use RMP in general, and whether the students using RMP are different from those not using RMP.

The purpose of conducting this survey was to determine 1) if students have used RMP to either rate a professor or obtain information about a professor; 2) if students are more likely to rate a very good or a very bad professor; 3) how students define a very good and a very bad professor.

The results from the survey are presented in Table 3 (Appendix). Panel A of Table 3 presents answers to some general questions about the students; Panel B presents answers to questions for those students who have used RMP to rate a professor; Panel C presents answers to questions for those students who have not used RMP to rate a professor. In each panel, the answers with the most frequent responses are presented in bold print for easier readability.

As shown in Panel A, the vast majority of students in the survey (87 percent) have heard about RMP, and 65 percent have used the site to find information about their professors. However, only about 31 percent of the students (or 73 students) participating in the survey have used the website to actually rate a professor.

Panel B asks questions specifically directed at those students who have used RMP to rate one or more of their past or present professors. Students were asked to select “Not applicable” if they have not used the site to rate a professor. For each question, exactly 165 students responded “Not applicable.” Consequently, the percentages shown in Panel B are for the sample of 73 students that have used RMP to rate a professor. Panel B shows that 1) 97.27 percent of students who have used RMP to rate a professor have a GPA between 2.00 and 4.00; 2) 86.30 percent of these students are juniors or seniors; 3) these students are most commonly management, finance, or accounting majors.

The remaining questions in Panel B attempt to determine the most common reason why students have used RMP to rate a professor. Interestingly, 39.73 percent of students who have used RMP used it to rate “excellent” professors as opposed to 27.40 percent who thought rated professors were “extremely bad”. Insofar as this sample reflects the typical students using RMP to rate a professor, it does not appear that students simply utilize the site to complain about their professors, which substantiates the arguments inherent in Giesey, Chen, and Hoshower (2004) and Landrum and Dillinger (2003) that students may use the site to communicate with each other and provide constructive criticism. Also interesting is the fact that 21.92 percent of students in the sample who have used RMP did so for their own personal satisfaction. The next question in Panel B that students who have used RMP consider an excellent professor as one who is very clear in explaining difficult concepts (56.16 percent of responses) and is very helpful and approachable (20.55 percent of responses). Notably, not even one student considers a professor

excellent just because the professor is good-looking. Responses to the last question in Panel B of Table 3 reveal that students who have used RMP consider a professor "extremely bad" most commonly if the professor is inconsistent in course expectations (35.62 percent) or is unable to explain difficult concepts (32.88 percent). This agrees with the findings of Wilhelm (2004) that students will put up with poor evaluations if they gain a great deal of knowledge, which is undoubtedly not the case if a professor cannot clearly explain difficult concepts. Interestingly, no students consider a professor "extremely bad" if the professor is either unattractive or because the professor gave them a bad grade.

Panel C asks questions directed at those students who *have not used* RMP to rate one or more of their past or present professors. Questions in this panel are designed to reveal any differences between those students who have used RMP to rate their professors and those that have not. In each question, students were asked to select "Not applicable" if they have used the site to rate a professor. For each question, exactly 73 students responded "Not applicable." The general questions in Panel C indicate that 1) 97.58 percent of students who have used RMP to rate a professor have a GPA between 2.00 and 4.00; 2) 87.27 percent of these students are juniors or seniors; 3) these students are most commonly management, finance, or accounting majors. In summary, the general characteristics of the students who have not used RMP to rate a past or present professor do not appear to be different from those who have used RMP.

The remaining questions in Panel C attempt to determine the most common reason why students would use RMP *in the future* to rate a professor, even if they have never used RMP before. 35.15 percent of students who have not used RMP indicated that they may use it to rate an "excellent" professor, while 35.76 percent indicate they may use it to rate an "extremely bad" professor. This is an interesting response pattern, since the majority of students who have used RMP used it to rate "excellent" professors. Nevertheless, the percentage responses again indicate that students do not only use RMP to complain about professors, even if they have not yet used the site. Students who have not used RMP consider an "excellent" professor as someone who is very clear in explaining difficult concepts (58.79 percent of responses) or who is very helpful and approachable (29.70 percent of responses). This is also very similar to students who have used RMP. The last question in Panel C also reveals that students who have not used RMP define an "extremely bad" professor very similarly to students who have used RMP; the most common responses were for professors who are inconsistent in course expectations (37.58 percent of responses) and who are unable to explain difficult concepts (33.33 percent of responses). Again, no students consider a professor "extremely bad" if the professor is either unattractive or because the professor gave them a bad grade.

Overall, the survey results reported in Table 3 indicate that: 1) students who have used RMP most commonly used it to rate an "excellent" professor; 2) students who have not used RMP are equally likely to use the site to rate "excellent" or "extremely bad" professors; 3) students, whether they have used RMP to rate professors or not, think that an "excellent" professor is one who is very clear in explaining difficult concepts and who is very helpful and approachable; 4) students, whether they have used RMP to rate professors or not, think that an "extremely bad" professor is one who is either inconsistent in course expectations or who is unable to explain difficult concepts.

It should be noted that our sample contains very few economics majors. However, we conducted this survey to gain some general insights into the reasons students utilize RMP. Since the students participating in the survey are from a variety of majors, we feel that the survey results provide a good general sense of the students using RMP. Overall, the survey results based

on our modest sample at UNF do not reveal any significant differences between students who have used RMP versus those that have not.

IV.2. Methodology

Our first objective is to investigate whether overall quality ratings recorded by students on RMP are positively related to the easiness and hotness RMP ratings. To accomplish that objective, we utilized the following regression model:

$$TOT_i = \alpha_0 + \alpha_1 RAT_i + \alpha_2 EAS_i + \alpha_3 NUM_i + \alpha_4 HOT_i + \alpha_5 MAL + \alpha_6 SIZ_i + \varepsilon_i , \qquad (1)$$

where

TOT_i = the total quality rating for professor *i* recorded on RMP;

RAT_i = the number of ratings for professor *i* recorded on RMP;

EAS_i = the easiness rating for professor *i* recorded on RMP;

NUM_i = the number of economics professors rated on RMP that are at the same university as professor *i*;

HOT_i = the total hotness rating for professor *i* on RMP, scaled by the total number of ratings for professor *i*;

MAL_i = a dummy variable equal to unity if professor *i* is male and zero otherwise; and

SIZ_i = the size of the university of professor *i* according to the last published fact sheet on the university website.

Since the RMP overall quality rating is computed as the average of the clarity and helpfulness ratings, the latter two variables are not included as regressors in Equation (1).

The variables NUM, SIZ, RAT, and MAL deserve additional discussion. These variables are included as control variables in the regression. NUM and SIZ, the number of economics professors at a given school rated on RMP and the total enrollment of the university, respectively, are used as proxies for size. Weis (1991) argues that professors at larger universities may have less time to devote to students due to higher research demands. Moreover, evaluation criteria used differ between AACSB-accredited and on-accredited institutions, which tend to be smaller. Lein and Merz (1977-78) find that accredited institutions place a much higher emphasis on research. This could affect the overall performance of professors in the classroom.

RAT, the number of ratings for a given professor, is also included as a control variable. Due to self-selection bias and based on our survey results, professors who are perceived as particularly good or bad by students may receive more RMP ratings than professors perceived to be of "average" quality, resulting in an upward or downward bias. Moreover, while we require at least three ratings per professor to be included in the sample, the number of ratings as a regressor serves as an additional safeguard against possible biases resulting from only a few ratings per professor.

MAL, the male dummy variable, is also included as a control variable to control for the gender of the professor being rated. Hamermesh and Parker (2003) find that instructors who are viewed as better-looking receive higher instructional ratings. This effect is larger for male than for female instructors. The gender of each professor was identified based on the professor's first name. In cases where the first name did not clearly indicate the gender of the professor, we visited the university website and obtained a picture of that professor.

Our second objective, investigating the correlation between RMP's overall quality rating and overall instructor in-class evaluations, is accomplished by computing the Pearson correlation coefficient between the overall RMP rating and the overall in-class instructor evaluation for the subsample of 38 ratings.

To accomplish our third objective, investigating whether the RMP factors such as easiness and hotness are related to the overall in-class instructor evaluation, Equation (1) was modified as follows:

$$ICE_i = \alpha_0 + \alpha_1 CLA_i + \alpha_2 HEL_i + \alpha_3 RAT_i + \alpha_4 EAS_i + \alpha_5 NUM_i + \alpha_6 HOT_i + \alpha_7 MAL + \alpha_8 SIZ_i + \varepsilon_i, \quad (2)$$

where

ICE_i = the "overall quality of instructor" rating for professor *i* obtained from in-class evaluations;

CLA_i = the clarity rating for professor *i* recorded on RMP;

HEL_i = the helpfulness rating for professor *i* recorded on RMP;

All other variables are as defined previously.

Importantly, Equation (2) includes the clarity and helpfulness ratings from RMP. The in-class evaluations obtained from the universities' websites provide a unique opportunity to investigate their relationship to the clarity and helpfulness criteria provided by students on RMP. Equation (1), which utilizes the RMP overall quality rating as the dependent variable, did not include these two independent variables, because the overall quality ratings on RMP are computed as the average of the helpfulness and clarity scores on RMP. Inclusion of these two variables in Equation (2) will indicate whether overall in-class instructor evaluations are influenced by the perceived clarity and helpfulness of the professor.

Two additional methodological considerations are noteworthy. First, in both Equations (1) and (2), forward stepwise regression was utilized. In these regressions, we required a variable to have a p-value of at least 0.50 to enter the model. Forward stepwise regression was utilized to investigate which specific factors are most significantly related to the dependent variable. Although this p-value is much higher than traditional significance levels, we required it to see if the adjusted r-square of the model changes when additional variables are included. The significant variables in each model can still be observed in the lower-step models. Second, since we use aggregate student responses (instead of individual student responses), heteroscedasticity in the model is likely. We first tested for heterscedasticity using White's (1980) test. Unable to detect any heteroscedasticity using this test, we also employed the Breusch-Pagan (1979) test for heteroscedasticity. The latter test tests for heteroscedasticity attributable to specific regressors. Specifying the aggregated RMP ratings for easiness, hotness, clarity, and helpfulness as the regressors most likely causing heteroscedasticity, we were able to detect heteroscedasticity in some models of the stepwise regression. This was corrected for using the generalized method of moments (GMM) procedure.

V. RESULTS

This section discusses four separate results. First, the correlation coefficients for the variables presented in Table 2 are discussed. Second, results from estimating Equation (1) are presented. Third, the correlation coefficients between the RMP overall quality rating and overall in-class instructor evaluations are discussed. Next, results from estimating Equation (2) are presented.

The correlation coefficients for the variables in Table 2 are displayed in Table 4. Most of the variables exhibit the correlation coefficients that would be expected. For example, since the RMP overall quality rating is a simple average of the helpfulness and clarity ratings, we would expect perfect correlation between each of these variables and the overall quality rating. Table 4 (Appendix) indicates that helpfulness and clarity are very highly positively correlated with the overall quality rating, with correlation coefficients of 0.96 and 0.96, respectively. The reason these correlation coefficients are not exactly 1.0 is due to rounding errors.

Some of the relationships in Table 4 deserve further discussion. First, it seems that the number of RMP ratings per professor is positively related to the total number of economics professors rated from a given school (r = 0.36). Also, school size is positively correlated with both the number of economics professors at a given school and the ratings per professor on RMP, with correlation coefficients of 0.73 and 0.26, respectively. It appears that students from larger schools rate more professors and record more ratings for each professor. Interestingly, hotness is negatively correlated to the number of economics professors rated, the ratings per professor on RMP, to male professors, and to the size of the school, although the correlation coefficients are relatively small in absolute value.

From Table 4, it seems that easiness, helpfulness, and clarity are all positively correlated with hotness, with correlation coefficients of 0.30, 0.44, and 0.47, respectively. As Hamermesh and Parker (2003) point out, however, it is possible that higher instructional ratings for better-looking professors are simply due to a third factor that is unmeasurable. Finally, more helpful economics professors are also perceived as clearer (r = 0.86), clearer professors are perceived as easier (r = 0.50), and more helpful professors are perceived as easier (r = 0.59). Furthermore, the correlation coefficients between overall quality and both hotness (r = 0.47) and easiness (r = 0.56) are relatively high.

These correlations are interesting in light of the survey results reported in Table 3. Recall that the survey indicated that students perceive “excellent” professors as those that are very helpful and that are very clear in explaining difficult concepts. Consequently, it is not surprising that the overall quality rating is significantly related to easiness, when easiness is also highly correlated with clarity and helpfulness indicates an

To investigate whether the overall quality ratings recorded on RMP are related to easiness and hotness, Equation (1) was applied to the 110 overall in-class ratings of economics professors at public universities in Florida. The results from estimating equation (1) are displayed in Table 5 (Appendix).

Before applying the forward stepwise regression, the full model was estimated. In the full model, the number of ratings per professor (RAT), the perceived easiness of the professor (EAS), the hotness of the professor (HOT), and the male dummy variable (MAL) are all positive and significant at traditional levels. Although we require a minimum of three ratings per professor to be included in the sample, it does appear that professors that receive more RMP ratings have a higher overall quality RMP rating, on average, as indicated by the significant coefficient for the variable RAT. The coefficient for EAS, 0.50, indicates that the overall quality rating increases by about 0.50 points for every additional point of perceived easiness. The coefficient for HOT, 1.77, indicates that, for every additional ten percent of total ratings that indicate a professor is hot, the overall quality rating increases by about 0.18 points. The coefficient for MAL indicates that male professors receive a total quality rating that is about 0.32 points higher than the overall quality rating female professors receive, which agrees with the findings reported by Hamermesh and

Parker (2003). The adjusted r-square for the full model indicates that 42.10 percent of the variation in the overall quality rating is explained by the regressors included in the model.

The additional models in Table 5 report the results from implementing the forward stepwise regression. The first variable entering the model in the stepwise regression is EAS (Model 2), followed by HOT (Model 3). Interestingly, Model 3 has an adjusted r-square of 40.29 percent, very close to the r-square reported for the full model; it appears that most of the variation in the overall quality rating is explained by EAS and HOT. The third variable entering the model is the male dummy variable, MAL (Model 4). Recall that Hamermesh and Parker (2003) find that instructors, particularly male instructors, who are viewed as better-looking receive higher instructional ratings. The results reported here similarly indicate that male professors receive higher overall quality ratings than female instructors on RMP. We did investigate whether male professors are more likely to be perceived as easy or hot by utilizing an interaction terms between the male dummy variable and EAS and HOT. Neither of the interaction terms was statistically significant at conventional levels when included in the regression together with EAS and HOT. This is consistent with the findings of Hamermesh and Biddle (1994), who find no significant gender differences across occupations, but is inconsistent with the findings of Hamermesh and Parker (2003), who find a more pronounced beauty effect for male instructors.

The next two variables entering the model are the number of ratings per professor, RAT (Model 5), and the size of the university, SIZ (Model 6). Although the coefficient for SIZ is only marginally negative and significant, it indicates that professors from larger universities, in terms of total enrollment, receive, on average, lower overall quality ratings on RMP. If larger universities proxy for a higher degree of research emphasis, then this finding agrees with Weis (1991) that research emphasis leads to less time for students. It is noteworthy to mention that Model 6, which includes all significant variables in the model with a p-value less than 0.50, has the highest adjusted r-squared of 42.57 percent among the models reported.

Overall, the results presented in Table 5 indicate that students rating economics professors at public universities in Florida on RMP heavily utilize both the perceived easiness and hotness of the professors in assigning their ratings. Both the coefficients and the explanatory power of these two variables is fairly consistent across the different models, although the number of ratings per professor, the gender of the professor being rated, and the size of the university are also significant.

To accomplish the second objective, the correlation coefficient between RMP overall quality ratings and overall in-class instructor ratings for the subsample of 38 economics professors was computed. The correlation coefficient was 0.53. While this coefficient definitely indicates a positive relationship between the two types of ratings, it also indicates that students use different criteria when evaluating their professors in class versus on RMP. For example, even though the survey results did not reveal this, it could be that students using RMP have an external locus of control, as argued by Grimes, Millea, and Woodruff (2004). Moreover, the correlation coefficient may be low because some factors considered important by students, such as tenure and promotion decisions (Giesey, Chen, and Hoshower [2004]) are not applicable to RMP ratings or because of the small sample size.

The third objective of the present study is to examine the overall in-class ratings for the subset of 38 economics professors in order to determine if they are significantly related to the clarity, helpfulness, easiness, and hotness ratings recorded on RMP. To accomplish the third objective, Equation (2) was estimated.

Results from estimating equation (2) are displayed in Table 6 (Appendix). As in Table 5, the full model was estimated first, then forward stepwise regression was utilized. In the full model in Table 5, the only significant variable is the RMP clarity rating for the professor, with a coefficient of 0.27. This indicates that the in-class overall instructor ratings are, on average, 0.27 points higher for every additional point of perceived clarity as indicated on RMP. Interestingly, neither the perceived RMP easiness and hotness (EAS and HOT) are significant in the full model. In fact, when the forward stepwise regression procedure is utilized, both HOT and EAS enter the model in Models 3 and 4, respectively, but neither variable is significant at traditional levels. However, Model 3, which includes both clarity and hotness, has the highest adjusted r-square of 33.28 percent, indicating that 33.28 percent of the variability in in-class overall instructor ratings is explained by the perceived clarity and hotness of the professor.

The findings reported here contradict the findings by Hamermesh and Parker (2003) that there is a performance beauty premium. We note again, however, that the sample utilized here is small and that the subsample of students using RMP may differ from the total sample of students in the class. The results reported here are nevertheless interesting because in-class instructor evaluations do not ask students to assess the easiness and hotness of a professor. Consequently, this is the first study to investigate whether there is a direct relationship between the perceived easiness and hotness of a professor and the in-class overall instructor ratings. While only 38 economics professors at public Florida universities are investigated here, the results reported here for this small sample indicate that students do not seem to reward easier or more attractive professors with higher in-class evaluations.

Thus, while it appears that the ratings students leave on RMP are driven primarily by the perceived easiness and hotness of the professor in question, overall in-class ratings are related only to the perceived clarity of the professor as indicated by students on RMP. Easiness and hotness of the professor, as indicated by students on the website, do not seem to drive the "overall rating of instructor" category for economics professors at public universities in Florida. It should also be mentioned here that the RMP easiness of a professor is very highly correlated with helpfulness and clarity. Thus, if students think a professor is easier because the professor is more clear in his or her explanations and/or because the professor is helpful, then we should expect the positive relationship between the RMP overall quality rating and easiness. This conjecture is further solidified by the survey results reported in Table 3, since students perceive "excellent" professors to be those that are very clear and helpful and not those that are easy graders.

From a productivity perspective, these results indicate that students, when they fill out the paper in-class evaluations, actually assign higher ratings to those professors that are the most clear in their explanations. Arguably, this is a trait that is indicative of a productive teacher, as being clear and concise in the classroom presumably increases student learning. This is suggested both by Giesey, Chen, and Hoshower (2004) in that students consider the improvement of teaching and course content important, and by Shevlin, Bynard, Davies, and Griffiths (2000) and Abrami, Perry, and Leventhal (1982), who find that instructor expressiveness influences student evaluations. It is still possible, however, that factors found to influence teaching evaluations in previous research, such as professor experience (Cohen [1981, 1982]) or student/teacher attitude similarities (Abrami and Mizener [1983]), drive the results reported here.

VI. CONCLUSION AND IMPLICATIONS

The purpose of this study is to investigate student ratings recorded on RateMyProfessors.com (RMP), a public website that allows students to evaluate their professors along various criteria. The objective of this study is threefold. First, we seek to investigate whether RMP overall quality ratings are positively related to the perceived easiness and hotness criteria of economics professors at public universities in Florida. Second, we investigate whether the overall quality ratings on RMP are positively correlated with the overall in-class instructor ratings for a subset of public Florida universities. Third, we seek to determine whether the overall in-class instructor ratings at public Florida universities are significantly related to the clarity, helpfulness, easiness, and hotness ratings recorded on RMP.

In an attempt to identify the reasons why students use RMP, a brief survey was conducted for a sample of 238 business students at The University of North Florida. The results from this survey indicate that: 1) students who have used RMP most commonly used it to rate an "excellent" professor; 2) students who have not used RMP are equally likely to use the site to rate "excellent" or "extremely bad" professors; 3) students, whether they have used RMP to rate professors or not, think that an "excellent" professor is one who is very clear in explaining difficult concepts and who is very helpful and approachable; 4) students, whether they have used RMP to rate professors or not, think that an "extremely bad" professor is one who is either inconsistent in course expectations or who is unable to explain difficult concepts.

For a sample of 110 ratings on RMP, our results indicate that students who utilize RMP reward easy and hot professors with a higher rating. On average, for every additional 10 percent of total ratings that indicate a professor is hot, that professor's total quality rating increases by approximately 0.18 points on a 1 to 5 scale. For every additional point of perceived easiness on a 1 to 5 scale, the total quality rating increases by about 0.50 points.

Our results also indicate that the correlation coefficient between total quality ratings on RMP and the overall in-class ratings for a subsample of 38 economics professors at five public universities in Florida is 0.53. Furthermore, when regressing overall in-class instructor ratings on clarity, helpfulness, easiness, and hotness ratings from RMP, only the professor's clarity enters the model in a forward stepwise regression procedure – neither easiness, hotness, nor helpfulness appear important considerations for students when they evaluate their instructor using paper and pencil.

These findings contradict the finding by Hamermesh and Parker (2003), who find that instructors who are viewed as better-looking receive higher instructional ratings. However, our results are interesting from an instructional standpoint; Wilhelm (2004) argues that students are willing to put up with instructors who receive poor course evaluations if they believe they will gain a great deal of knowledge. This indicates that students may assign low course evaluations for other reasons, but that they reward professors from whom they learn a lot. Arguably, clearer professors are better able to impart knowledge on their students, so again the importance of clarity is not surprising.

The findings reported here support the view that students reward professors based on a criterion that arguable increases teaching productivity – clarity. Consequently, it does not appear that economics professors at public Florida universities are hot and easy; at a minimum, their hotness and easiness does not seem to be rewarded by students when they conduct their in-class evaluations.

Appendix

Table 1. Suggested Rating Guidelines Posted on RMP.

- **Do's**
  - Be honest.
  - Poor spelling WILL NOT cause your rating to be removed; however, poor spelling may result in your rating being discredited by those who read it.
  - Limit your comments to the professor's professional abilities. Try not to get personal.
  - Try to be objective in your assessment of the professor.

- **Do Not's**
  - Threaten to kill or harm your professor. Not only will the rating be deleted, but we will notify the authorities of your IP address and the time you rated. *This is enough information to identify you.*
  - Talk about your professor's sex life. This includes:
    - Claiming that the professor sleeps with students, even if he or she has slept with you.
    - Claiming that he or she is homosexual.
  - Direct racist, sexist or homophobic remarks at your professor.
  - Post ratings for people who do not teach classes at your school.
  - Criticize the way a professor looks or dresses. Appearances have little to do with a professor's ability to teach the material.
  - Use the comment area to talk about irrelevant subjects like the football team; comments should be pertinent to the class and/or the professor who teaches the class.
  - Include sexual innuendo in your comment.
  - Sign your comment with your name, initials, pseudo name, or any sort of identifying mark.
  - Put an e-mail address into your comment.
  - Include a URL to a webpage or website, *unless it is relevant to the class.*
  - Claim that the professor *has been* or *will be* fired.
  - Rate a professor more than once for the same class.
  - Write your comments in any language other than English (unless you attend a French-Canadian school).

Table 2. Summary Statistics for A Sample of 110 RMP Ratings for Economics Professors at Public Universities in Florida.

	Econ Profs Rated[a]	**Ratings Per Prof**	**Male = 1**	**Total Enrollment[b]**
Average	28.45	17.98	0.79	23,880.91
Median	33.00	8.00	1.00	25,383.00
StDev	13.19	25.99	0.41	16,496.12
Min	3.00	3.00	0.00	692.00
Max	43.00	129.00	1.00	48,765.00

	Hotness %[c]	**Easiness[d]**	**Helpfulness[d]**	**Clarity[d]**	**Overall Quality[d,e]**
Average	7.89	3.00	3.43	3.28	3.36
Median	0.00	3.00	3.55	3.30	3.35
StDev	18.35	0.81	0.92	0.91	0.88
Min	0.00	1.30	1.00	1.00	1.20
Max	1.00	4.70	5.00	5.00	4.90

Notes to Table 2:

a Represents statistics for the number of economics professors at the eleven public universities in Florida.

b Total student enrollment (undergraduate and graduate) at the eleven public universities in Florida during the Fall Semester of 2004 as obtained from the university websites. The most recent enrollment data available for The University of West Florida was for the 2002/2003 academic year.

c Computed by dividing the hotness rating from RMP by the total number of ratings per professor.

d Helpfulness, clarity, and total quality ratings are on a scale from 1 (worst) to 5 (best). Professors with a total quality rating of 5 receive a smiley face on the RMP website. Easiness is on a scale from 1 (hardest) to 5 (easiest).

e The average of the helpfulness and clarity ratings as posted by RMP

Table 3. Survey Results for 238 Students at The University of North Florida.[a,b]

Panel A – General Questions (n = 238)
I have heard about the website called www.RateMyProfessors.com.
Yes: 208 (87.40%)
No: 30 (12.60%)
I have used the website www.RateMyProfessors.com to find out how professors at a university are viewed by students.
Yes: 154 (64.71%)
No: 84 (35.29%)
I have used the website www.RateMyProfessors.com to rate one or more of my professors.
Yes: 73 (30.67%)
No: 165 (69.33%)
How did you first learn about the site www.RateMyProfessors.com?
Found it on the Internet: 8 (3.36%)
Heard about it from a friend: 173 (72.69%)
Read about it in a magazine or student publication: 5 (2.10%)
My professor mentioned it in class: 31 (13.03%)
Through this survey: 21 (8.82%)
Panel B – Questions for Students Who Have Used RMP (n = 73)
If you HAVE USED www.RateMyProfessors.com to rate one or more of your past or present professors, what is your current cumulative GPA? (Remember that this survey is anonymous, and honesty is of the utmost importance to the validity of the results). If you HAVE NOT USED www.RateMyProfessors.com, select "Not applicable" as your answer.
0.00 – 0.99: 0 (0.00%)
1.00 – 1.99: 2 (2.74%)
2.00 – 2.99: 31 (42.47%)
3.00 – 4.00: 40 (54.80%)
If you HAVE USED www.RateMyProfessors.com to rate one or more of your past or present professors, which of the following best describes your status? If you HAVE NOT USED www.RateMyProfessors.com, select "Not applicable" as your answer.
Freshmen: 0 (0.00%)
Sophomore: 1 (1.37%)
Junior: 41 (56.16%)
Senior: 22 (30.14%)
Graduate: 9 (12.33%)

If you HAVE USED www.RateMyProfessors.com to rate one or more of your past or present professors, which of the following best describes your major? If you HAVE NOT USED www.RateMyProfessors.com, select "Not applicable" as your answer.
Economics: 1 (1.37%)
Management: 16 (21.92%)
Accounting: 10 (13.70%)
Finance: 17 (23.29%)
Financial Services: 1 (1.37%)
Transportation and Logistics: 0 (0.00%)
International Business: 6 (8.22%)
Marketing: 8 (10.96%)
MBA: 9 (12.33%)
MACC: 0 (0.00%)
Other: 5 (6.85%)
If you HAVE USED www.RateMyProfessors.com to rate one or more of your past or present professors, which of the following best describes the reason you rated the professor? If you have used the site for more than one reason, select the answer that reflects your most significant reason. If you HAVE NOT USED www.RateMyProfessors.com to rate a professor, please select "Not applicable" as your answer to this question.
I thought a professor was an excellent professor, and I wanted to provide other students who may take this professor with information about him or her: 29 (39.73%)
I thought a professor was extremely bad, and I wanted to provide other students who may take this professor with information about him or her: 20 (27.40%)
I thought a professor was of average quality, and I wanted to provide other students who may take this professor with information about him or her: 5 (6.85%)
I wanted to leave a favorable or unfavorable rating for my own personal satisfaction, whether or not anyone else reads it: 16 (21.92%)
None of the above: 3 (4.11%)
If you HAVE USED www.RateMyProfessors.com to rate one or more of your past or present professors, what makes a professor "excellent" in your opinion? Answer this question even if you did not rate your professor(s) as "excellent." Even if you agree with more than one statement, select only the answer that best describes your definition of "excellent." If you HAVE NOT USED www.RateMyProfessors.com, select "Not applicable" as your answer.
The professor is very clear in explaining difficult concepts: 41 (56.16%)
The professor is an easy grader: 5 (6.85%)
The professor is very helpful in- and outside of the classroom and is very approachable: 15 (20.55%)
The professor is very good-looking: 0 (0.00%)
The professor is very fair in grading: 10 (13.70%)
The professor is very respectful towards students: 0 (0.00%)
Other: 2 (2.74%)

If you HAVE USED www.RateMyProfessors.com to rate one or more of your past or present professors, what makes a professor "extremely bad" in your opinion? Answer this question even if you did not rate your professor(s) as "extremely bad." Even if you agree with more than one statement, select only the answer that best describes your definition of "extremely bad." If you HAVE NOT USED www.RateMyProfessors.com, select "Not applicable" as your answer.
The professor is inconsistent in expectations for the course, and course requirements change frequently throughout the semester: 26 (35.62%)
The professor is unable to explain difficult concepts: 24 (32.88%)
The professor is unattractive: 0 (0.00%)
The professor is disrespectful towards students: 7 (9.59%)
The professor gave me a bad grade: 0 (0.00%)
The professor is a very hard grader, even though the grading criteria are fair: 2 (2.74%)
The professor is not helpful and not very approachable: 10 (13.70%)
The professor does not treat all students equally: 2 (2.74%)
Other: 2 (2.74%)
Panel C – Questions for Students Who Have Not Used RMP (n = 165)
If you HAVE NOT USED www.RateMyProfessors.com to rate one or more of your past or present professors, what is your current cumulative GPA? (Remember that this survey is anonymous, and honesty is of the utmost importance to the validity of the results). If you HAVE USED www.RateMyProfessors.com, select "Not applicable" as your answer.
0.00 – 0.99: 0 (0.00%)
1.00 – 1.99: 4 (2.42%)
2.00 – 2.99: 64 (38.79%)
3.00 – 4.00: 97 (58.79%)
If you HAVE NOT USED www.RateMyProfessors.com to rate one or more of your past or present professors, which of the following best describes your status? If you HAVEUSED www.RateMyProfessors.com, select "Not applicable" as your answer.
Freshmen: 0 (0.00%)
Sophomore: 1 (0.61%)
Junior: 99 (60.00%)
Senior: 45 (27.27%)
Graduate: 20 (12.12%)

If you HAVE NOT USED www.RateMyProfessors.com to rate one or more of your past or present professors, which of the following best describes your major? If you HAVE USED www.RateMyProfessors.com, select "Not applicable" as your answer.
Economics: 3 (1.82%)
Management: 46 (27.88%)
Accounting: 35 (21.21%)
Finance: 24 (14.55%)
Financial Services: 2 (1.21%)
Transportation and Logistics: 6 (3.64%)
International Business: 8 (4.85%)
Marketing: 15 (9.09%)
MBA: 20 (12.12%)
MACC: 0 (0.00%)
Other: 6 (3.64%)
If you HAVE NOT USED www.RateMyProfessors.com to rate one or more of your past or present professors, which of the following would cause you to leave ratings for one more of your professors? If you HAVE USED www.RateMyProfessors.com to rate a professor, please select "Not applicable" as your answer to this question.
I think a professor is an excellent professor, and I would want to provide other students who may take this professor with information about him or her: 58 (35.15%)
I think a professor is an extremely bad professor, and I would want to provide other students who may take this professor with information about him or her: 59 (35.76%)
I think a professor is of average quality, and I would want to provide other students who may take this professor with information about him or her: 4 (2.42%)
I would leave a favorable or unfavorable rating for my own personal satisfaction, whether or not anyone else reads it: 18 (10.91%)
None of the above: 26 (15.76%)
If you HAVE NOT USED www.RateMyProfessors.com to rate one or more of your past or present professors, what makes a professor "excellent" in your opinion? Answer this question even if you did not rate your professor(s) as "excellent." Even if you agree with more than one statement, select only the answer that best describes your definition of "excellent." If you HAVE USED www.RateMyProfessors.com, select "Not applicable" as your answer.
The professor is very clear in explaining difficult concepts: 97 (58.79%)
The professor is an easy grader: 7 (4.24%)
The professor is very helpful in- and outside of the classroom and is very approachable: 49 (29.70%)
The professor is very good-looking: 0 (0.00%)
The professor is very fair in grading: 5 (3.03%)
The professor is very respectful towards students: 0 (0.00%)
Other: 6 (3.64%)

If you HAVE NOT USED www.RateMyProfessors.com to rate one or more of your past or present professors, what makes a professor "extremely bad" in your opinion? Answer this question even if you did not rate your professor(s) as "extremely bad." Even if you agree with more than one statement, select only the answer that best describes your definition of "extremely bad." If you HAVE USED www.RateMyProfessors.com, select "Not applicable" as your answer.
The professor is inconsistent in expectations for the course, and course requirements change frequently throughout the semester: 62 (37.58%)
The professor is unable to explain difficult concepts: 55 (33.33%)
The professor is unattractive: 0 (0.00%)
The professor is disrespectful towards students: 0 (0.00%)
The professor gave me a bad grade: 0 (0.00%)
The professor is a very hard grader, even though the grading criteria are fair: 7 (4.24%)
The professor is not helpful and not very approachable: 26 (15.76%)
The professor does not treat all students equally: 0 (0.00%)
Other: 5 (3.03%)

Notes to Table 3:

[a] Students were given the following RMP description and instructions to complete the survey:

Description: RateMyProfessors.com is a website that allows students to anonymously rate their professors along criteria such as helpfulness, clarity, easiness, and attractiveness of a professor. Students simply access the website without providing any personal information, select the professor they would like to rate at their university (or add the professor if he or she is not already listed on the site), and select the ratings they believe to be appropriate. In addition, students can write comments about the professor(s).

Instructions: This extra credit survey will provide students with two percentage points extra credit on the final exam in the course. For example, if there are 30 questions on the final exam, and you answer 21 questions correctly, that would be 70%. With this extra credit, you would have a 72%.

Anonymity in this survey is extremely important, as the results will be used for research purposes. I will be able to determine everyone who completed the survey, without knowing which student provided what answers. Please DO NOT email me to tell me you completed the survey and the answers you provided, as this will render the results useless.

Please answer the following questions as honestly as possible. If none of the provided answers describe the answer you want to provide, please select "None of the above."

Make sure that you read each question VERY CAREFULLY before selecting an answer. Some questions are very similar to one another and may differ by only one or two words.

[b] A total of 377 students were enrolled in the three classes to which the survey was administered. Thus, the response rate is 238/377 = 63.13%.

Table 4. Correlation Coefficients for Descriptive RMP Variables and University-Specific Variables Used in the Regression Analysis for A Sample of 110 RMP Ratings for Economics Professors at Public Universities in Florida.

	Econ Profs Rated[a]	**Ratings Per Prof**	**Male = 1**	**Total Enrollment[b]**
Econ Profs	1.00	0.36	0.05	0.73
Ratings		1.00	-0.09	0.26
M=1			1.00	0.05
Total Enrollment				1.00

	Hotness %[c]	**Easiness[d]**	**Helpfulness[d]**	**Clarity[d]**	**Overall Quality[d,e]**
Econ Profs	-0.05	0.01	-0.01	-0.10	-0.05
Ratings	-0.13	0.01	0.04	0.06	0.06
M=1	-0.21	-0.01	0.09	-0.00	0.05
Total Enrollment	-0.12	0.08	-0.07	-0.07	-0.07
Hotness %	1.00	0.30	0.44	0.47	0.47
Easiness		1.00	0.59	0.50	0.56
Helpfulness			1.00	0.86	0.96
Clarity				1.00	0.96
Overall Quality					1.00

Notes to Table 4:

a Represents statistics for the number of economics professors at the eleven public universities in Florida.

b Total student enrollment (undergraduate and graduate) at the eleven public universities in Florida during the Fall Semester of 2004 as obtained from the university websites. The most recent enrollment data available for The University of West Florida was for the 2002/2003 academic year.

c Computed by dividing the hotness rating from RMP by the total number of ratings per professor.

d Helpfulness, clarity, and total quality ratings are on a scale from 1 (worst) to 5 (best). Professors with a total quality rating of 5 receive a smiley face on the RMP website. Easiness is on a scale from 1 (hardest) to 5 (easiest).

e The average of the helpfulness and clarity ratings as posted by RMP.

Table 5. Regression Results from Estimating Equation (1)[a] for a Sample of 110 RMP Ratings for Economics Professors at Eleven Public Universities in Florida. (*t*-statistics in parentheses)[b,c]

	Intercept	RAT	EAS	NUM	HOT	MAL	SIZ
Model 1	1.65 (5.22)***	0.01 (2.18)**	0.50 (4.67)***	-0.00 (-0.38)	1.77 (4.09)***	0.32 (1.95)*	-0.00 (-0.68)
Model 2	1.62 (4.90)***	--	0.61 (5.76)***	--	--	--	--
Model 3	1.72 (5.61)***	--	0.50 (4.63)***	--	1.56 (3.43)***	--	--
Model 4	1.52 (4.88)***	--	0.49 (4.55)***	--	1.70 (3.93)***	0.27 (1.73)*	--
Model 5	1.44 (4.64)***	0.00 (1.93)*	0.49 (4.58)***	--	1.80 (4.07)***	0.30 (1.92)*	--
Model 6	1.64 (5.19)***	0.00 (2.13)**	0.50 (4.75)***	--	1.75 (4.12)***	0.32 (1.97)*	-0.00 (-1.74)*

Notes to Table 5:

[a] $TOT_i = \alpha_0 + \alpha_1 RAT_i + \alpha_2 EAS_i + \alpha_3 NUM_i + \alpha_4 HOT_i + \alpha_5 MAL + \alpha_6 SIZ_i$, where (1)

TOT_i = the total quality rating for professor *i* recorded on RMP;

RAT_i = the number of ratings for professor *i* recorded on RMP;

EAS_i = the easiness rating for professor *i* recorded on RMP;

NUM_i = the number of economics professors rated on RMP that are at the same university as professor *i*;

HOT_i = the total hotness rating for professor *i* on RMP, scaled by the total number of ratings for professor *i*;

MAL_i = a dummy variable equal to unity if professor *i* is male and zero otherwise; and

SIZ_i = the size of the university of professor *i* according to the last published fact sheet on the university website.

[b] The full model is complemented by additional models based on forward stepwise regression, where the most significant variable is identified and added to the model first. The procedure terminates when the remaining variables have significance at less than 50%.

[c] The model was tested first tested for heteroskedasticity using White's (1980) test. Although White's test was unable to detect any heteroskedasticity in the model, the Breusch-Pagan (1979) test was also conducted. It was assumed that the error variance varies with the set of regressors that are aggregate student responses (i.e., EAS and HOT). Generalized method of moments estimation (GMM) was subsequently used to correct for the heteroscedasticity.

*,**,*** Significant at the 10%, 5%, and 1% level, respectively

Table 6. Regression Results from Estimating Equation (2)[a] for a Sample of 38 "Overall Rating of Instructor" In-Class Ratings for Economics Professors at Eleven Public Universities in Florida. (*t*-statistics in parentheses)[b,c]

	Intercept	CLA	HEL	RAT	EAS	NUM	HOT	MAL	SIZ
Model 1	2.61 (5.66)***	0.27 (2.06)**	0.08 (0.60)	-0.00 (-0.59)	0.04 (0.38)	0.01 (0.29)	-0.92 (-1.31)	0.02 (0.08)	0.00 (0.51)
Model 2	3.05 (13.45)***	0.30 (4.74)***	--	--	--	--	--	--	--
Model 3	2.93 (12.32)***	0.34 (4.89)***	--	--	--	--	-0.83 (-1.44)	--	--
Model 4	2.82 (9.66)***	0.31 (3.99)***	--	--	0.07 (0.78)	--	-0.82 (-1.43)	--	--

Notes to Table 6:

[a] $$ICE_i = \alpha_0 + \alpha_1 CLA + \alpha_2 HEL + \alpha_3 RAT_i + \alpha_4 EAS_i + \alpha_5 NUM_i + \alpha_6 HOT_i + \alpha_7 MAL + \alpha_8 SIZ_i, \quad (2)$$

where

ICE_i = the "overall quality of instructor" rating for professor *i* obtained from in-class evaluations;

CLA_i = the clarity rating for professor *i* recorded on RMP;

HEL_i = the helpfulness rating for professor *i* recorded on RMP;

RAT_i = the number of ratings for professor *i* recorded on RMP;

EAS_i = the easiness rating for professor *i* recorded on RMP;

NUM_i = the number of economics professors rated on RMP that are at the same university as professor *i*;

HOT_i = the total hotness rating for professor *i* on RMP, scaled by the total number of ratings for professor *i*;

MAL_i = a dummy variable equal to unity if professor *i* is male and zero otherwise; and

SIZ_i = the size of the university of professor *i* according to the last published fact sheet on the university website.

[b] The full model is complemented by additional models based on forward stepwise regression, where the most significant variable is identified and added to the model first. The procedure terminates when the remaining variables have significance at less than 50%.

c The model was first tested for heteroskedasticity using White's (1980) test. Although White's test was unable to detect any heteroskedasticity in the model, the Breusch-Pagan (1979) test was also conducted. It was assumed that the error variance varies with the set of regressors that are aggregate student responses (i.e., CLA, HEL, EAS, and HOT). Generalized method of moments estimation (GMM) was subsequently used to correct for the heteroscedasticity.

*,**,*** Significant at the 10%, 5%, and 1% level, respectively

References

Abrami, P. and D. Mizener. 1983. "Does the Attitude Similarity of College Professors and Their Students Produce 'Bias' in Course Evaluations?" *American Educational Research Journal* 20: 123-136.

Abrami, P., P. Perry, and L. Leventhal. 1982. "The Relationship Between Student Personality Characteristics, Teacher Ratings, and Student Achievement." *Journal of Educational Psychology* 74: 111-125.

Algozzine, B. J. Beattie, M. Bray, C. Flowers, J. Gretes, L. Howley, G. Mohanty, and F. Spooner. 2004. "Student Evaluation of College Teaching; A Practice in Search of Principles." *College Teaching* 52: 134-141.

Andrews, H. 2004. *Accountable Teacher Evaluation! Toward Highly Qualified and Competent Teachers.* Stillwater, OK: New Forums Press.

Bernstein, E. 2004. "What Teacher Evaluation Should Know and Be Able to Do: A Commentary." *National Association of Secondary School Principals Bulletin* 88: 80-88.

Biddle, J.E. and D.S. Hamermesh. 1998. "Beauty, Productivity, and Discrimination: Lawyers' Looks and Lucre." *Journal of Labor Economics* 16: 172-201.

Calderon, T. and B. Green. 1997. "Use of Multiple Information Types in Assessing Accounting Faculty Teaching Performance." *Journal of Accounting Education* 15: 221-237.

Centra, J. 1993. "Reflective Faculty Evaluation." *Student Evaluations of Teaching: What Research Tells Us*. San Francisco: Jossey-Bass.

Campos, S. "Grading Tables Have Been Turned." *The Valley Star*. Los Angeles Valley College. February 23, 2005.

Cohen, P. 1981. "Student Ratings of Instruction and Student Achievement: A Meta-Analysis of Multisection Validity Studies." *Review of Educational Research* 51: 281-309.

Cohen, P. 1982. "Validity of Student Ratings in Psychology Courses: A Meta-Analysis of Multisection Validity Studies." *Teaching of Psychology* 9: 78-82.

Ellis, D. and Bulman-Fleming, B. "RMP.ca." *Teaching Matters Newsletter*. The University of Waterloo. September 2001.

Giesey, J., Y. Chen, and L. Hoshower. 2004. "Motivation of Engineering Students to Participate in Teaching Evaluations." *Journal of Engineering Education* 93: 303-312.

Granzin, K. and J. Painter. 1973. "A New Explanation for Students' Course Evaluation Tendencies." *American Educational Research Journal* 10: 115-124.

Grimes, P., M. Millea, and T. Woodruff. Spring 2004. "Grades – Who's to Blame? Student Evaluation of Teaching and Locus of Control." *Journal of Economic Education* 35: 129-147.
Hamermesh, D.S. and A.M. Parker. 2003. "Beauty in the Classroom: Professors' Pulchritude and Putative Pedagogical Productivity." NBER Working Paper No. W9853.
Hamermesh, D.S. and J.E. Biddle. 1994. "Beauty and the Labor Market." *The American Economic Review* 84: 1174-1194.
Justice, M. "Find Spring Classes With Web Services." *Central Florida Future*. The University of Central Florida. December 6, 2004.
Kemp, B. and G. Kumar. 1990. "Student Evaluations: Are we Using Them Correctly?" *Journal of Education for Business* 65: 106-111.
Landrum, R. and R. Dillinger. 2004. "The Relationship Between Student Performance and Instructor Evaluations Revisited." *The Journal of Classroom Interaction* 39: 5-9.
Lein, C. and C. Merz. 1977-78. "Faculty Evaluation in Schools of Business: The Impact of AACSB Accrediation on Promotion and Tenure Decisions." *Collegiate News and Views* Winter: 21-24.
Marsh, H. 1980. "The Influence of Student, Course, and Instructor Characteristics in Evaluations of University Teaching." *American Educational Research Journal* 17: 219-237.
Marsh, H. and J. Overall. 1981. "Validity of Students' Evaluations of Teaching Effectiveness." *Journal of Educational Psychology* 72: 468-475.
Marsh, H. and L. Roche. 1997. "Making Students' Evaluations of Teaching Effectiveness Effective: The Central Issues of Validity, Bias, and Utility." *American Psychologist* 52: 1187-1197.
Morgan, D., J. Sneed, and L. Swinney. 2003. "Are Student Evaluations A Valid Measure of Teaching Effectiveness: Perceptions of Accounting Faculty Members and Administrators." *Management Research News* 26: 17-32.
Shevlin, M., P. Baynard, M. Davies, and M. Griffiths. 2000. "The Validity of Student Evaluation of Teaching in Higher Education: Love Me, Love My Lectures?" *Assessment & Evaluation in Higher Education* 25: 397-405.
Simpson, R. 1995. "Uses and Misuses of Student Evaluations of Teaching Effectiveness." *Innovative Higher Education* 20: 3-5.
Sorenson, D. and T. Johnson. 2003. *Online Student Ratings of Instruction*. San Francisco: Jossey-Bass.
Weis, W. 1991. "What's Going on in Business Schools?" *Decision Line* 22: 3-4.
Westhues, K. "RMP." *The Record*. Kitchener, Ontario, Canada. September 7, 2001.
White, H. 1980. "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test of Heteroskedasticity." *Econometrica* 48: 817-838.
Wilhelm, W. 2004. "The Relative Influence of Published Teaching Evaluations and Other Instructor Attributes on Course Choice." *Journal of Marketing Education* 26: 17-30.
Yunker, P. and J. Sterner. 1988. "A Survey of Faculty Performance Evaluation in Accounting." *Accounting Educators' Journal* 26: 63-71.

Hot and Easy in Florida