


AN OBJECTIVE EMPIRICAL TEACHING EVALUATION METRIC AT A SOUTHEASTERN US PUBLIC UNIVERSITY

*Métrica objetiva y empírica de evaluación de la enseñanza
en una Universidad Pública del Sureste de EE. UU.*

 **Inessa Korovyakovskaya**¹
korovi@savannahstate.edu

 **Felipe Llaugel**²
fllaugel@udoym.edu.do

 **Dennis Ridley**³
dridley@fsu.edu

¹ College of Business Administration,
Savannah State University, USA

² College of Engineering, Universidad
Dominicana O&M, Dominican Republic

³ School of Business & Industry, Flor-
ida A&M University and Department
of Scientific Computing, Florida State
University, USA

Fecha recepción: 23/08/2019

Fecha aprobación: 09/12/2019



Esta obra está bajo la licencia de Creative
Commons Reconocimiento-NoComercial-Si-
noDerivada 4.0 Internacional.

Abstract

The marginal contribution of faculty to student learning at an AACSB-accredited College of Business Administration in a public university located in a southeastern state in the United States (U.S.) is measured for the first time by an objective quantitative method. Student cumulative Grade Point Average (GPA), centralized to avoid grade inflation relates to the partial amount of credit hours each teacher devotes to students. We proffer that the marginal contribution of the professor to student GPA earned per contact hour of instruction is the regression coefficient associated with the professor. Since the university uses GPA as a measure of progress, contribution to GPA is the professorial teaching contribution to the university objective. Such a teaching contribution is consistent with the professor's assignment of responsibility. The computational results of a five-year empirical data analysis are presented.

Keywords: evaluation, learning outcomes, metric, teaching effectiveness, teaching evaluation, performance evaluation, university.

Resumen

La contribución marginal de los profesores al aprendizaje de los estudiantes en una Facultad de Administración de Empresas acreditada por AACSB en una universidad pública ubicada en un estado del sudeste de los Estados Unidos (EE.UU.), se mide por primera vez mediante un método cuantitativo objetivo. El Promedio Acumulativo de Calificaciones (GPA, por sus siglas en inglés) de los estudiantes, centralizado para evitar la inflación de las calificaciones, se relaciona con la cantidad parcial de horas de crédito que cada profesor dedica a los estudiantes. Consideramos que la contribución marginal del profesor al GPA del estudiante obtenido por hora de contacto de instrucción es el coeficiente de regresión asociado con el profesor. Dado que la universidad usa el GPA como una medida de progreso, la contribución al GPA es la contribución del docente al objetivo de la universidad. Dicha contribución docente es consistente con la asignación de responsabilidad del profesor. Se presentan los resultados computacionales de un análisis de datos empíricos de cinco años.

Palabras clave: eficacia del docente, evaluación, evaluación de la enseñanza, evaluación del desempeño, rendimiento académico, universidad.

1. Introduction

Due to the fact that the decision-making process in academic evaluation has undergone changes in higher education and that there are greater expectations of transparency in tenure and promotion (Bana e Costa & Oliveira, 2012), the evaluation process is being redesigned. Traditionally, academics have been evaluated based on three criteria: teaching, scholarship and service, with different emphasis being placed on one of the individual criteria depending on the type of institution (Fairweather, 2002). Research universities tend to place more emphasis on traditional scholarship while teaching institutions and colleges tend to place more value on teaching and service (Cherry et al., 2017).

The debate about appropriate methods for evaluating teaching has continued in academia for decades. Recently, pressures for accountability are forcing institutions to examine how they value, measure and improve what happens in the classroom. In 2005, US Secretary of Education Margaret Spellings formed the Commission on the Future of Higher Education to examine a national strategy for reforming higher education. Their findings were critical of higher education and the Commission made several suggestions to remedy deficiencies. One of the recommendations was to measure student learning outcomes (Spellings, 2006). Prior to this recommendation, some states established policies that focused on academic productivity in undergraduate teaching (Colbeck, 2002). As a result, there is now increased attention to academic evaluations of the institutions and faculty with the emphasis on outcomes and accountability (Cherry et al., 2017).

This study constructs an objective statistical empirical model for evaluating professorial contribution to student learning in a public university located in a southeastern state in the United States (U.S.). This contribution is one of the three complementary weighted assignments of responsibility that include teaching, research and service (Sharobeam & Howard, 2002). The theory for the model is based

on the Ridley and Collins (2015) professorial evaluation metric (PEM). The PEM is based on student achievement based on GPA. Education theories may argue the pros and cons of how GPA does or does not determine student learning. But the university has stipulated that GPA is its measure of progress within the university and faculty are expected to contribute to said student progress. Therefore, we will consider GPA as a proxy for learning.

The remainder of the paper is organized as follows. Traditional student-based teaching evaluation methods are reviewed in the next section. The empirical case study in a public university located in a southeastern state in the U.S. is presented next. We then discuss the integration of the teaching evaluation metric into a full faculty evaluation metric. Our conclusions include suggestions for future research.

2. Method for Evaluating Teaching

Almost 100 percent of schools/colleges of business “use student evaluation of instruction to measure teaching and classroom performance” (Clayson & Haley, 2011, p.101). It is assumed that students will honestly evaluate professors/instructors and their teaching. Some researchers question the validity of student evaluation of teaching to improve individual instructor performance, modify curriculum, and create comparative scales to evaluate faculty (Clayson & Haley, 2011).

Kozub (2008), Ryan et al. (1980), McNatt (2010) and McPherson et al. (2009) have studied the validity of student evaluations. McNatt (2010) conducted a longitudinal naturally occurring field experiment and concluded that administrators should use caution when interpreting student evaluations with a course and even all courses taught by a given professor if the professor has negative reputation that may result in bias in student evaluations.

Yunker and Yunker (2003) found a negative relationship between student evaluations and student achievement (see also Coker et al., 1980) and Weinstein (1987). Centra (2003) and Buchert et al.

(2008) found that it is possible for student evaluation of teaching to be influenced by first impressions of instructors and grade expectations. Student evaluations are known to be lower in freshman classes where the students are less mature. Seniors and graduate students are more likely to understand the professor's advocacy for best practices and objectives for high achievement. This can negatively impact young professors who are idealistic with regards to grading standards, quality and intellectual curiosity. They may be accused of being the cause of students failing. Marshall (2005) found that student evaluations were inefficient and ineffective. Highly skewed student evaluations can require the use of percentile rankings of faculty (Clayson & Haley, 2011).

There has been very little research done on the examination of the perceptions and role of the academic administrator in the evaluation process, especially on which factors of the academic evaluation tend to impact classroom instruction and learning outcomes. Academic administrators play a vital role as "the conduit between university policy-makers (board, president and provost) and the academy" (Cherry et al., 2017). They are also the key to hiring and developing new academics and to help professors and instructors meet university standards for promotion and tenure.

Cherry et al. (2017) examined academic administrators' attitudes toward annual faculty evaluation processes and methods. Of the 208 respondents, their findings revealed the following ranking of teaching evaluation methods in their order of importance: Student evaluations had the highest ranking of 39.9% (83) followed by peer evaluations ranking of 27.4% (57). Department head/chair evaluations took the third place in the order of importance with 22.6% (47) of the respondents. Self-evaluation and other methods for evaluating of teaching were in the fourth and fifth places with 6.7% (14) and 3.4% (7), respectively. Time and again, student evaluations continue to play a significant role in evaluating teaching performance in the classroom despite concerns about their validity.

Teaching evaluations that are performed by administrators can be arbitrary. They are based on the administrator's opinion. Administrator evaluations may or may not consider teaching methodology, innovation, currency of syllabus or workload. The administrator may be influenced by student opinions that are no more than popularity contests that are unrelated to learning (Coker, et. al., 1980; Weinstein, 1987). Student complaints to the administrator may lower evaluations when the administrator is more sensitive to student feelings than to upholding standards of academic performance. Empathy for student feelings is desirable. But, overindulgence of students may encourage lack of personal responsibility and less than best study habits. Short term political objectives may supersede lifelong future learning objectives.

Evaluations that are inversely related to learning or progress, or are otherwise unreliable, may cause professors to change their approach to teaching for the worse, discouraging high performance (Coker et. al., 1980; Weinstein 1987). Unreliable evaluations may discourage academic freedom (Dershowitz, 1994; Haskell, 1997; Ryan et. al., 1980). For these reasons better methods for evaluating teaching are required (Ma, 2005; Wolfer & Johnson, 2003). They should be designed to encourage academic rigor, demonstrated academic knowledge and proficiency, critical thinking, understanding and leadership skills.

3. Empirical Case Study

3.1. Teaching Evaluation Score (TES) Data

Grades from 2,194 students in an AACSB-accredited College of Business Administration at a public university located in a southeastern state in the U.S. were collected for the period Fall-2014 to Summer-2018. The majors (Programs) included were a) Accounting; b) Business Computer Information Systems (CIS); c) Business Management Online; d) Business Management; e) Business Marketing; f) Global Logistics and International Business; and g) Master of Business Administration. The data included 348 professors and instructors and 228 courses. Twenty-five (25) of the 228 professors were affiliated

with the College of Business Administration. Given that professors taught in different programs, and courses were repeated during the study period, and included in several programs (majors), the following table is not totalized. Tables 1 and 2 show the composition of the data collected. Figure 1 displays the grade distribution in a histogram.

Table 1. Data Composition for the College of Business Administration

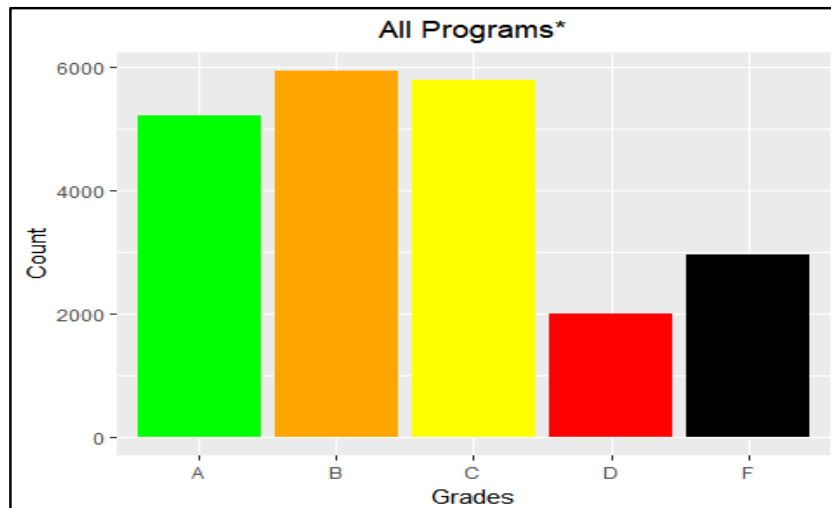
Program	Students	Professors	Courses
Accounting	415	5	123
Bus. Computer Information Systems	205	5	106
Business Management	959	6*	168
Business Marketing	410	4	139
Global Logistics & Int. Business	148	4	104
Business Management Online	78	6*	44
Master of Business Administration	129	10**	31

Note. Some of the professors who teach Management courses in the Business Management major also teach courses in the Business Management Online. Professors who teach at the undergraduate level, also teach at the graduate level.

Table 2. Grade Distribution by Program

Program	GRADES				
	A	B	C	D	F
Accounting	1155	1097	1008	295	457
Bus Comp Inf. Systems	415	585	617	172	272
Business Management	114	90	76	28	56
Business Marketing	2023	2426	2481	914	1345
Global Logistics & Int. Business	882	1162	1237	471	652
Business Management Online	384	390	330	118	177
Master of Business Administration	262	217	55	7	5

Figure 1. Grade Distribution



3.2. Data format and structure

All students in the College of Business Administration are included in the data. Since any one of these students may take a course from any professor in the

university, all professors must be included. A sample of the student data used in the regression analysis (see Appendix A) is given in Table 3. These names are anonymous for sake of privacy.

Table 3. Sample data taken from the records of 3 students

Semester	Student ID	Grade	Program	Professor	Credits	Course
Fall 2014	915166132	B	Accounting	Professor 54	2	Physical Activity and Stress Mngt
Fall 2014	915166132	C	Accounting	Professor 198	3	Introduction to Anthropology
Fall 2014	915166132	A	Accounting	Professor 241	1	Physical Conditioning
Fall 2014	915166132	A	Accounting	Professor 116	3	Introduction to Anthropology
Fall 2014	915166132	C	Accounting	Professor 268	3	General Biology
Fall 2014	915166132	B	Accounting	Professor 171	2	Physical Fitness for Life
Fall 2014	915138031	A	Accounting	Professor 59	2	Physical Activity and Stress Mngt
Fall 2014	915138031	B	Accounting	Professor 26	3	The Environment of Business
Fall 2014	915138031	B	Accounting	Professor 67	3	Introduction to Anthropology
Fall 2014	915138031	C	Accounting	Professor 5	3	Communicating in Business Env
Fall 2014	915138031	D	Accounting	Professor 177	3	PreCalculus
Fall 2014	915100868	C	Accounting	Professor 11	3	Principles of Managerial Accounting
Fall 2014	915100868	A	Accounting	Professor 285	3	A Survey of US His to Post Civil

3.3. Teaching Evaluation Score (TES) Results

The TES method, explained in Appendix A, was applied to data taken from automated university computer records. The results are shown in Table 4.

Table 4. TES scores for all professors in the university

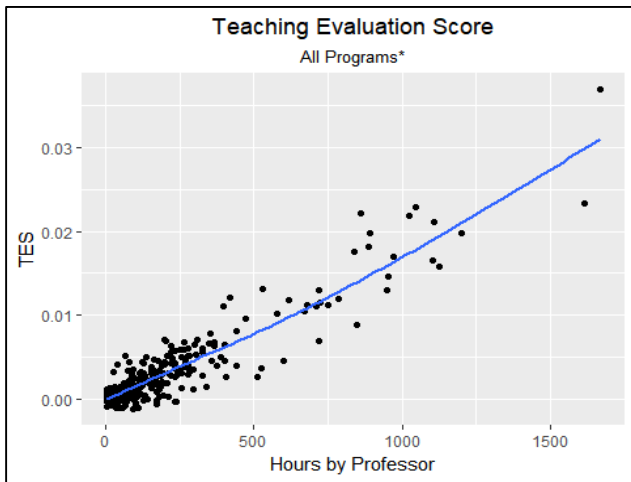
PROFESSOR	b_j	CREDIT HOURS $\sum_{i=1}^n H_{ij}$	$b_j \sum_{i=1}^n H_{ij}$	TES_j
Professor 1	2.580996	1668	4305.102	3.69%
Professor 2	1.691028	1614	2729.319	2.34%
Professor 3	2.555748	1044	2668.202	2.28%
Professor 4	3.003953	861	2586.404	2.21%
Professor 5	2.492423	1023	2549.749	2.18%
Professor 6	2.226235	1107	2464.443	2.11%
Professor 7	2.600446	890	2314.397	1.98%
Professor 8	1.927490	1200	2312.988	1.98%
Professor 9	2.401392	885	2125.232	1.82%
Professor 10	2.452535	837	2052.772	1.76%
...
Professor 339	-4.851096	18	-87.319	-0.07%
Professor 340	-4.218515	21	-88.588	-0.08%
Professor 341	-16.284701	6	-97.708	-0.08%
Professor 342	-2.396708	45	-107.852	-0.09%

Table 4. (Continued)

Professor 343	-3.241247	34	-110.202	-0.09%
Professor 344	-2.832696	39	-110.475	-0.09%
Professor 345	-1.046761	111	-116.191	-0.10%
Professor 346	-2.010987	58	-116.637	-0.10%
Professor 347	-2.827908	43	-121.600	-0.10%
Professor 348	-1.491749	93	-138.733	-0.12%
			$\sum_{j=1}^k b_j \sum_{i=1}^n H_{ij} = 116,823.42$	$\sum_{j=1}^k TES_j = 100.00$

The TES results are plotted in Figure 2. The TES and the grade distribution for each program are included in Appendix C.

Figure 2. Teaching Evaluation Score



The regression model for TES for all programs together has a coefficient of multiple determination R-squared of 0.87 and an adjusted R-squared of 0.8457, representing an excellent goodness of fit indicator. Table 5 shows the detailed breakdown of R-square for the TES model for each of the academic programs. All indicators are appropriate for the case at hand.

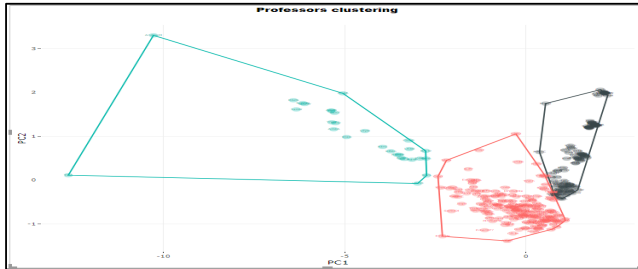
Table 5. Goodness of Fit Indicators for TES models conditioned on all n=348 professors

Program	R-squared	Adjusted R-squared
Accounting	0.9623	0.8779
Bus Comp Inf. Systems	0.9998	0.9645
Business Management	0.8955	0.8397
Business Marketing	0.9664	0.8738
Global Logistics & Int. Business	1.0000	0.9994
Business Management Online	0.9682	0.6898
Master of Business Administration	0.9590	0.9413

With the information (scaled to avoid bias due to dimensions of each variable) included in the TES results, the number of students taught by professor, the count of grades, the count of courses and the credit hours, a principal component analysis was applied to run a clustering model. Figure 3 presents 3 clusters: cluster 1 (left), cluster 2 (center), and cluster 3 (right). Professors in cluster 1 are those with higher TES scores, more students by course, and more courses taught. Professors in cluster 2 have medium TES scores, more reduced courses than those in cluster 1, but similar number of courses taught. Cluster 3 is populated by professors with the lower TES scores, and reduced courses. The clusters show 3 distinct groups of professors whose performance require further analysis to explain the composition of each group. The two first principal components account for more than 90% of data variability. This analysis is meaningful to identify professors that need follow up and to en-

courage those with best practice in the classroom, using several metrics simultaneously.

Figure 3. Clustering Analysis of TES Data



The TES scores for the 25 professors in the College of Business are selected from Table 4 and placed in Table 6. These names are anonymous but in the actual report, the professors will be selected by their real names. This facilitates integration into the comprehensive faculty professorial evaluation metric as discussed below. The College of Business Administration faculty have assignments of responsibility

that are different from other academic units in the university. Therefore, they may be compared only with faculty in their own academic unit.

We notice that the professors in the College of Business Administration occupy the upper echelon of Table 4. This suggests that they are either better teachers or their credits hours taught are more associated with students that are included in the regression analysis. In either case, they do contribute more to the GPA of these particular students. Their total contribution is 34.09% of all professors in the university. For easy interpretation, their TES scores are rescaled so as to add to 100%. The regression analysis was repeated with only College of Business Administration professors. The results are shown in Table 7. The results are similar. To choose between them, we recalculated the adjusted R-squared for the reduced model conditioned on $n=25$ professors. The adjusted R-squared $n=348$ is 0.8457. The adjusted R-squared $n=25$ is 0.8089. Therefore, the full model is considered better.

Table 6. TES scores for all 25 professors in the College of Business Administration conditioned on $n=348$ professors

PROFESSOR	b_j	CREDIT HOURS $\sum_{i=1}^n H_{ij}$	$b_j \sum_{i=1}^n H_{ij}$	TES_j	Rescaled $\frac{TES_j}{\sum_{j=1}^{25} TES_j}$
Professor 10	2.452535	837	2052.7722	1.76%	5.15%
Professor 1	2.580996	1668	4305.102	3.69%	10.81%
Professor 7	2.600446	890	2314.3973	1.98%	5.81%
Professor 12	1.757822	1101	1935.3624	1.66%	4.86%
Professor 115	2.087762	159	331.95422	0.28%	0.83%
Professor 17	2.099878	720	1511.9122	1.29%	3.80%
Professor 211	2.061184	39	80.386167	0.07%	0.20%
Professor 5	2.492424	1023	2549.7493	2.18%	6.40%
Professor 157	3.394052	51	173.09667	0.15%	0.43%
Professor 41	2.034498	366	744.6261	0.64%	1.87%
Professor 14	1.796064	954	1713.4448	1.47%	4.30%
Professor 23	1.917291	681	1305.6751	1.12%	3.28%
Professor 179	11.30157	12	135.61889	0.12%	0.34%
Professor 3	2.555749	1044	2668.2017	2.28%	6.70%
Professor 21	1.875468	723	1355.9632	1.16%	3.40%

Table 6. (Continued)

Professor 19	1.779303	783	1393.1939	1.19%	3.50%
Professor 29	1.217811	846	1030.2685	0.88%	2.59%
Professor 25	1.819648	711	1293.7699	1.11%	3.25%
Professor 22	1.74279	750	1307.0927	1.12%	3.28%
Professor 26	1.835252	669	1227.7835	1.05%	3.08%
Professor 6	2.226235	1107	2464.4425	2.11%	6.19%
Professor 13	1.636574	1125	1841.1457	1.58%	4.62%
Professor 16	1.598537	948	1515.4134	1.30%	3.81%
Professor 11	2.046083	972	1988.7924	1.70%	4.99%
Professor 4	3.003953	861	2586.4039	2.21%	6.49%
			$\sum_{j=1}^k b_j \sum_{i=1}^n H_{ij}=39,826.57$	$\sum_{j=1}^{25} TES_j=34.09\%$	Rescaled $\sum_{j=1}^{25} TES_j=100.00\%$

Table 7. TES scores for all 25 professors in the College of Business Administration conditioned on $n=25$ professors

PROFESSOR	b_j	CREDIT HOURS $\sum_{i=1}^n H_{ij}$	$b_j \sum_{i=1}^n H_{ij}$	TES_j
Professor 10	2.429101	837	2033.158	5.58%
Professor 1	1.869893	1668	3118.981	8.56%
Professor 7	2.508893	890	2232.915	6.13%
Professor 12	1.530892	1101	1685.512	4.63%
Professor 115	2.339626	159	372.001	1.02%
Professor 17	2.326812	720	1675.304	4.60%
Professor 211	1.984656	39	77.402	0.21%
Professor 5	2.126829	1023	2175.746	5.97%
Professor 157	2.401591	51	122.481	0.34%
Professor 41	1.720016	366	629.526	1.73%
Professor 14	1.832542	954	1748.245	4.80%
Professor 23	1.572399	681	1070.804	2.94%
Professor 179	2.393736	12	28.725	0.08%
Professor 3	1.859773	1044	1941.603	5.33%
Professor 21	1.816679	723	1313.459	3.61%
Professor 19	1.767897	783	1384.263	3.80%
Professor 29	1.987247	846	1681.211	4.62%
Professor 25	1.91337	711	1360.406	3.74%
Professor 22	2.027109	750	1520.331	4.17%
Professor 26	1.349424	669	902.765	2.48%
Professor 6	2.083267	1107	2306.177	6.33%
Professor 13	1.434878	1125	1614.237	4.43%
Professor 16	1.738626	948	1648.217	4.53%
Professor 11	1.904503	972	1851.177	5.08%

Table 7. (Continued)

Professor 4	2.236805	861	1925.889	5.29%
		$\sum_{j=1}^k b_j \sum_{i=1}^n H_{ij} = 36,420.54$		$\sum_{j=1}^{25} TES_j = 100.00\%$

4. Integration into the Faculty Professorial Evaluation Metric

The teaching evaluation metric (TEM) may be integrated into an objective professorial evaluation metric (PEM), used to determine a professorial evaluation score (PES). The PEM is designed to incorporate measures of teaching, research and service. It includes the TEM, used to determine a TES; a research evaluation metric (REM), used to determine a research evaluation score (RES); and a service evaluation metric (SEM), used to determine a service evaluation score (SES). The PES is an overall measure a professor’s contribution, expressed as a fraction of the total contribution of all professors in the instructional unit. The PEM accounts for uneven distribution of effort and prior assignment of responsibility between teaching, research and service, between professors, and between different time periods. It is used for annual evaluations, merit reward, tenure and promotion. Professorial contributions require time to take effect. The TEM is discussed in Appendix A; the PEM is defined in Appendix B.

5. Conclusions

The subject institution for this study is a College of Business Administration at a public university located in a southeastern state in the U.S. As in most academic institutions across the U.S., faculty in this AACSB-accredited college are expected to teach, conduct research, provide service to the College, university, profession and community. The faculty of this College is regularly evaluated by students, peer faculty members, and administrators. While there has been some scepticism among researchers in academia about student evaluations being inaccurate and contaminated with bias, this evaluation component is factored in the evaluation formula for faculty tenure and promotion decisions.

From the administrative point of view, student evaluations provide an important insight into the quality of faculty teaching and how much professors’ efforts contribute to learning by students. At the time of the redesign of higher education spearheaded by the Department of Education in the early 2000’s, the student GPA became one of the most important statistics that measures student learning outcomes and the overall quality of instruction for a given institution. Thus, the student GPA affects parent and student decision making to enroll and matriculate in a particular college or university, and administrators’ decision to hire, retain, tenure, and promote their faculty.

We propose that each professor marginally contributes to student GPA in the classes he or she teaches. All of the classes in the data set are 3 credit-hour courses that meet twice a week for 1 hour and 15 minutes, or 3 times a week for 50 minutes. Recent research (Diette & Raghav, 2017, 2018) demonstrates that there is no difference in achieving learning outcomes for classes that meet 2 or 3 times a week. The workloads of professors for the period of 2014-2019 consisted of 3 classes in one semester and 4 classes in another semester of an academic year for a total of at least 120 credit hours per professor per year. In summer sessions, professors taught on average 2 classes each. These teaching loads were designed to provide the faculty with time to conduct research, attend conferences, write and administer grants, and participate in scholarly and professional activities.

Performance standards have been set high by the AACSB accreditation of the College of Business, which stimulated professors’ desire to strive for the high quality instruction in the classroom and to provide students with additional learning and professional opportunities that feed back into their learning and course performance. These include participation in student case competitions, showcases, workshops

and seminars, guest lectures, industry visits, undertaking summer and semester-long internships, conducting undergraduate research and making conference presentations. The faculty motivated and empowered the students to be active in their learning and professional development while still in college. These activities led to achieving learning outcomes as evidenced by mostly “A”, “B”, and “C” grades across the majors of the College of Business Administration in this study. As a result, student enrollment and retention has been high while professors have been receiving kudos, tenure and promotion from the administration. Students have been consistently performing subjective teaching evaluations and professors’ ratings have been high.

6. Recommendations

This study applied an objective statistical empirical model to evaluate the marginal contribution that professors make through their teaching toward student learning. Based on the findings, it is evident that professors do contribute to student success as evidenced by student GPA attribution as a proxy for learning and advancement through the institution. The TES was found to be a reliable evaluation metric that is highly recommended to universities and colleges in the U.S. and around the world for adoption and inclusion into their objective professorial evaluation metric.

References

- Bana e Costa, C. A., & Oliveira, M. D. (2012). A multicriteria decision analysis model for faculty evaluation. *Omega*, 40(4), 424-436 <https://doi.org/10.1016/j.omega.2011.08.006>
- Buchert, S., Laws, E. L., Apperson, J. M., & Bregman, N. J. (2008). First impressions and professor reputation: Influence on student evaluations of instruction. *Social Psychology of Education*, 11(4), 397-408. <https://doi.org/10.1007/s11218-008-9055-1>
- Centra, J. A. (2003). Will teachers receive higher student evaluations by giving higher grades and less coursework? *Research in Higher Education*, 44(5), 495-518. <https://doi.org/10.1023/a:1025492407752>
- Cherry, B., Grasse, N., Kapla, D., & Hamel, B. (2017). Analysis of academic administrators’ attitudes: Annual evaluations and factors that improve teaching. *Journal of Higher Education Policy & Management*, 39(3), 296-306. <https://doi.org/10.1080/1360080x.2017.1298201>
- Clayson, D. E. & Haley, D. A. (2011). Are students telling us the truth? A critical look at the student evaluation of teaching. *Marketing Education Review*, 21(2), 101-112. <https://doi.org/10.2753/mer1052-8008210201>
- Coker, H., Medley, D. M., & Soar, R. S. (1980). How valid are expert opinions about effective teaching? *Phi Delta Kappan*, 62(2), 31-149. <http://bit.ly/2PxAjGT>
- Colbeck, C. L. (2002). State policies to improve undergraduate teaching administrator and faculty responses. *The Journal of Higher Education*, 73(1), 3-25. <https://doi.org/10.1353/jhe.2002.0004>
- Department of Education (2006). A test of leadership: Charting the future of U.S. higher education. Washington, DC.
- Dershowitz, A. (1994). *Contrary to popular opinion*. Berkley Books.
- Diette, T. M. & Raghav, M. (2018). Do GPAs differ between longer classes and more frequent classes at liberal arts colleges? *Research in Higher Education*, 59(4), 519-527. <https://doi.org/10.1007/s11162-017-9478-7>
- Fairweather, J. S. (2002). The ultimate faculty evaluation: Promotion and tenure decisions. *New Directions for Institutional Research*, 2002(114), 97-108. <https://doi.org/10.1002/ir.50>
- Haskell, R. E. (1997). Academic freedom, tenure, and student evaluation of faculty: Galloping Polls in the 21st Century. *Education Policy Analysis Archives*, 5. <https://doi.org/10.14507/epaa.v5n6.1997>
- Llaugel, L. & Ridley, A. D. (2018). A university of Dominican Republic objective empirical fac-

- ulty teaching evaluation metric. *Journal of Management and Engineering Integration*, 11(1), 1-10. <http://bit.ly/2sdeZyo>
- Kozub, R. M. (2008). Student evaluations of faculty: Concerns and possible solutions. *Journal of College Teaching & Learning*, 5(11), 35. <https://doi.org/10.19030/tlc.v5i11.1219>
- Ma, X. Y. (2005). Establish internet student-assessing of teaching quality system to make the assessment perfect. *Heilongjiang Researches on Higher Education*, 6, 94-96. <http://bit.ly/2rCQA5t>
- McNatt, D. B. (2010). Negative reputation and biased student evaluations of teaching: longitudinal results from a naturally occurring experiment. *Academy of Management Learning and Education*, 9(2), 225-242. <https://doi.org/10.5465/amle.2010.51428545>
- McPherson, M. A., Jewell, R. T., & Kim, M. (2009). What determines student evaluation scores? A random effects analysis of undergraduate economics classes. *Eastern Economic Journal*, 35(1), 37-51. <https://doi.org/10.1057/palgrave.eej.9050042>
- Ridley, D. & Collins, J. (2015). A suggested evaluation metric instrument for faculty members at colleges and universities. *International Journal of Education Research*, 10(1), 97-114. <http://bit.ly/2E7GThU>
- Ryan, J. J., Anderson, J. A., & Birchler, A. B. (1980). Student evaluations: The faculty responds. *Research in Higher Education*, 12(4), 317-333. <https://doi.org/10.1007/bf00136899>
- Sharobeam, M. H., & Howard, K. (2002). Teaching demands versus research productivity. *Journal of College Science Teaching*, 31, 436-441. <http://bit.ly/2P991rh>
- Spellings, M. (2006). A test of leadership: Charting the future of US higher education. Department of Education.
- Weinstein, L. (1987). Good teachers are needed?. *Bulletin of the Psychometric Society*, 25(4), 273-274. <https://doi.org/10.3758/bf03330353>
- Wolfer, T. A., & McNown, M. (2003). Re-evaluating student evaluation of teaching: The teaching evaluation form. *Journal of Social Work Education*, 39(1), 111-121. <https://doi.org/10.1080/10437797.2003.10779122>
- Yunker, P., & Yunker, J. (2003). Are student evaluations of teaching valid? Evidence from an analytical business core course. *Journal of Education for Business*, 78(6), 313-317. <https://doi.org/10.1080/08832320309598619>

APPENDIX A

THE TEACHING EVALUATION METRIC

The TEM is based on the following regression model:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_j x_{ij} + \dots + \beta_k x_{ik} + \varepsilon_i; \\ i = 1, 2, 3, \dots, n; j = 1, 2, 3, \dots, k \quad (A1)$$

where

y_i = cumulative grade (re-centered around $c=2$) point average of the i^{th} student,

x_{ij} = fraction of total number of semester hours that the i^{th} student was taught by the j^{th} professor,

β_0 = regression parameter representing the extent to which grade point average is unaffected by direct contact hours within the instructional unit,

β_j = regression parameter containing information regarding the impact that the j^{th} professor has on student grade point average, and the errors ε_i are independent and normally distributed with zero mean and variance σ^2 ,

k = number of professors in the instructional unit,

n = number of students.

The marginal rate at which the j^{th} professor contributes to student grade point average, ceteris paribus, is given by β_j grade points per contact hour of instruction. Assuming that grade points measure learning, then β_j represents the institution & instruc-

tional unit context specific teaching effectiveness of the j^{th} professor, in the presence of all contributions by all the other professors. Also, it is assumed that each professor contributes to student learning in some general way through advising or any number of other indirect ways, and that such learning is reflected in β_0 . Therefore, teaching credit is determined from the teaching effectiveness coefficient

$b_j = \beta_0 + \beta_j$ (estimated). It reflects the j^{th} professor's knowledge, proficiency, ability to impart knowledge, contribution to student intellectual development and study habit, ability to leverage the contributions to date made by all other professors, and contribution to student ability to perform in the professor's course, as well as, in other courses taken at the university. In order to correct for grade inflation and differences in grading standards, the grades reported for each class are re-centered around a grade of $c=2$ points before totalling up the grade points. For each class the re-centered grade values are the original grade values minus the average of the grade value for the class plus 2.0. Therefore, this is a professorial peer evaluation of the preparedness of each other's students. It is conducted by the best experts that the university has to offer. Furthermore, the evaluation is kept honest by grade re-centering (the average grade is the same for all professors). Grade re-centering must be explained to all faculty members to avoid any futile temptation to game the system by giving higher grades to their own students.

The teaching evaluation score (TES_j) for the j^{th} professor is based on a combination of the teaching effectiveness coefficient (b_j) and the teaching workload. It is measured by the total contribution to the number of student credit hours earned by students who were taught by the j^{th} professor, expressed as a fraction of the contribution to the grand total number of student credit hours made by all professors.

$$TES_j = \left(\sum_{i=1}^n b_j H_{ij} \right) / \left(\sum_{j=1}^k \sum_{i=1}^n b_j H_{ij} \right) = (b_j \sum_{i=1}^n H_{ij}) / \left(\sum_{j=1}^k b_j \sum_{i=1}^n H_{ij} \right) \quad (A2)$$

where H_{ij} represents the number of contact hours

that the j^{th} professor taught the i^{th} student in the evaluation year. If the i^{th} student was not taught by the j^{th} professor during the evaluation year, then $Hg = 0$. Since the intercept α_0 represents the extent to which GPA is unaffected by direct contact hours within the instructional unit, its contribution to cumulative GPA must be distributed equally to all professors in the unit. That is, $b_j = \beta_0 + \beta_j$ (*estimated*). If the regression is run through the origin the estimated regression coefficients will change from β_j (*estimated*) to $\beta_0 + \beta_j$ (*estimated*). Therefore, we choose to run the regression without an intercept. Either way, the value of TES_j is the same since TES_j is calculated from $b_j = \beta_0 + \beta_j$ (*estimated*).

If it were true that large class sizes lower teaching effectiveness, then the teaching effectiveness coefficient will be lowered. However, multiplying the teaching effectiveness coefficient by the number of student credit hours will increase the TES, and thereby offset the effect of class size on the TES. In order to assist in maximizing teaching effectiveness, the university should attempt to equalize and reduce class sizes. Where large class sizes are unavoidable, technology may be used to mitigate reductions in teaching effectiveness.

APPENDIX B

THE PROFESSORIAL EVALUATION METRIC

For the purpose of giving the most general possible description of the PEM model, assume that the number of professors in an instructional unit is k . Then let the professor code be j where $j=1, 2, 3, \dots k$. The professorial evaluation score for the j^{th} professor is determined as follows:

$$PES_j = T_j TES_j + R_j RES_j + S_j SES_j \quad (B.1)$$

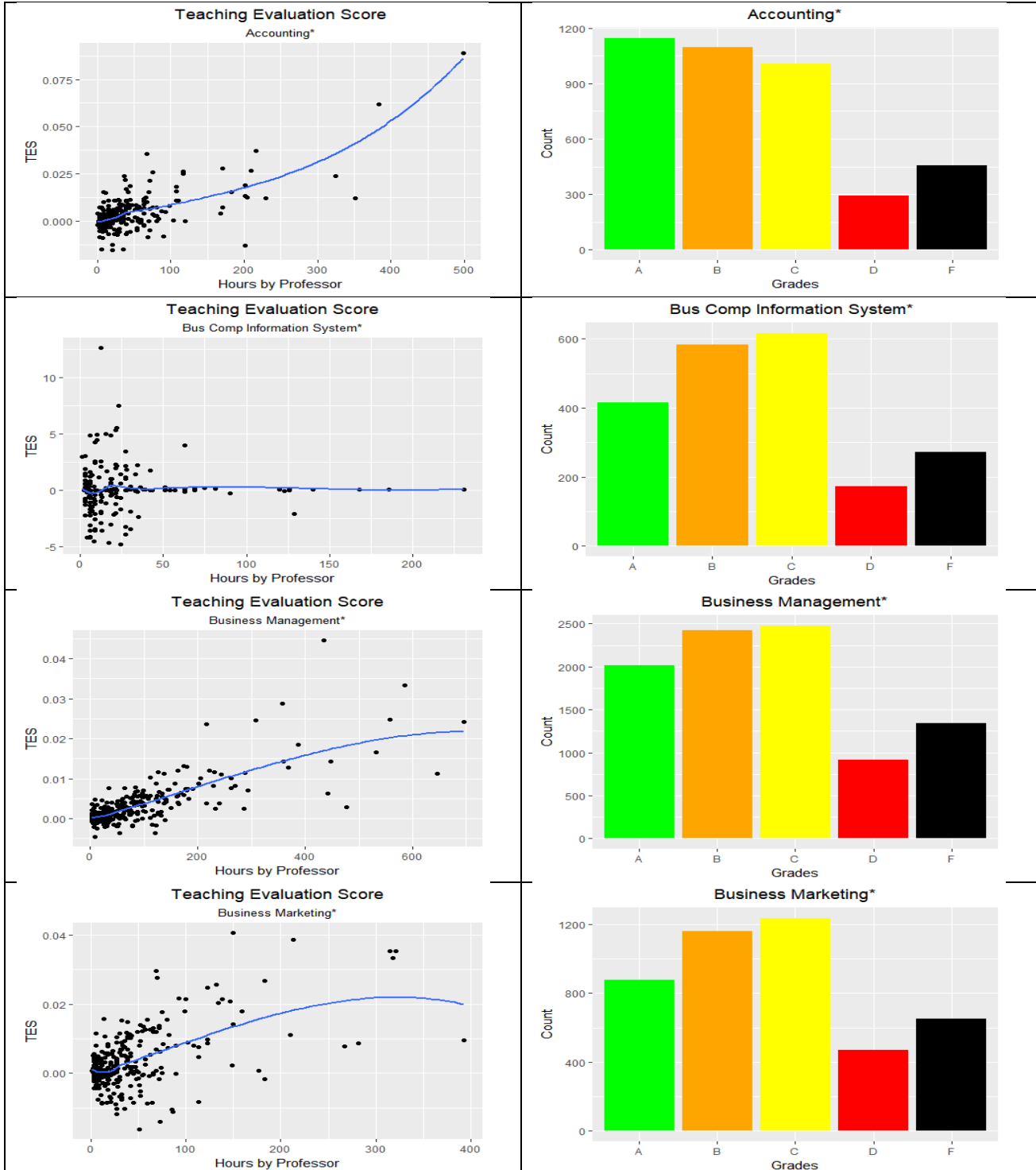
where

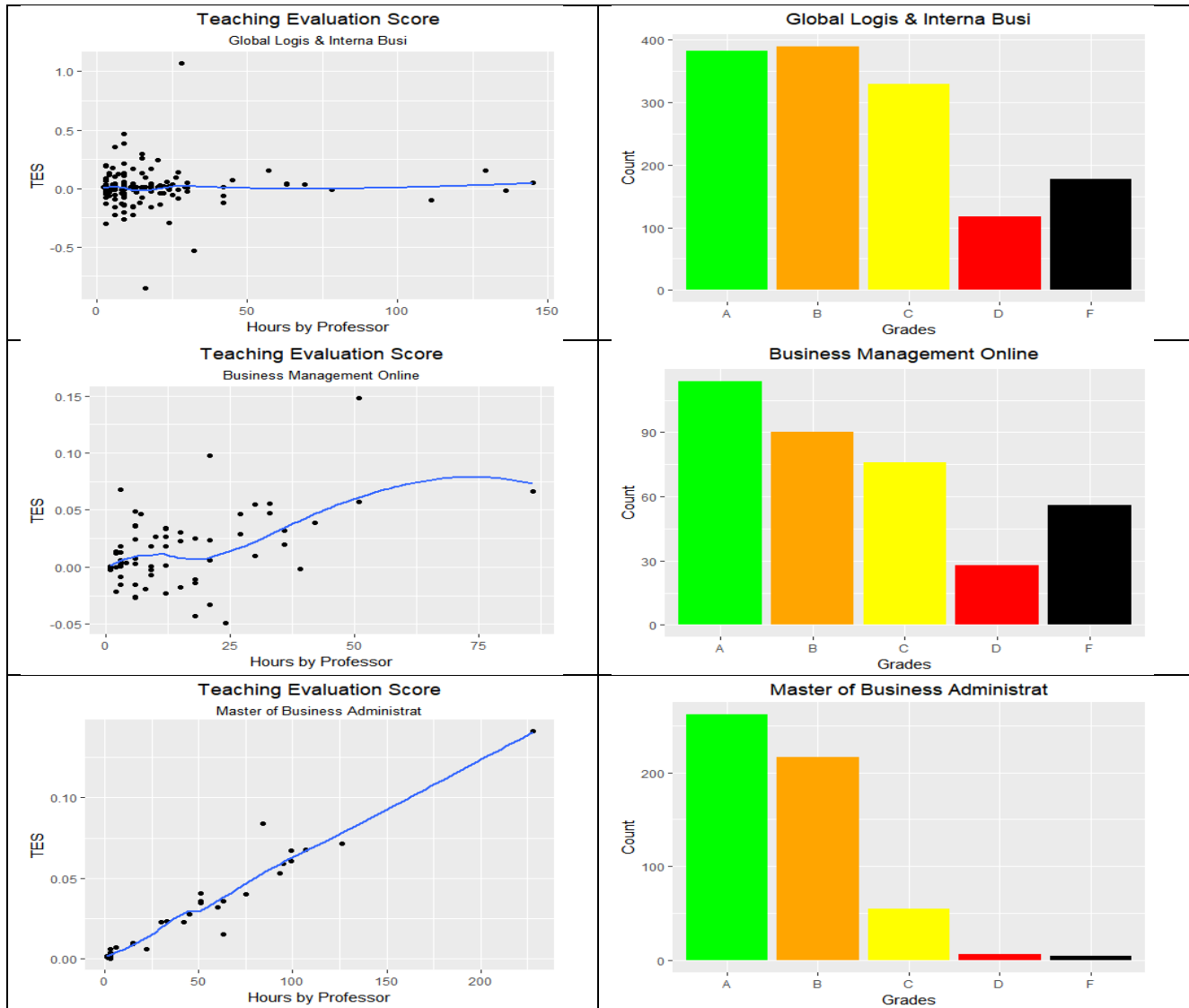
- TES_j = Fraction of total professorial teaching contribution made by the j^{th} professor,
- RES_j = Fraction of total research contribution made by the j^{th} professor,
- SES_j = Fraction of total service contribution made by the j^{th} professor,
- T_j = Fraction of the j^{th} professor's assignment of responsibility given to teaching (≥ 0.25 i.e., average at least one 3hr. course per semester to maximize the TES contribution to the PES).
- R_j = Fraction of the j^{th} professor's assignment of responsibility given to research (≥ 0.20),
- S_j = Fraction of the j^{th} professor's assignment of responsibility given to service ($\geq 0.05 \leq 0.1$),
- $T_j + R_j + S_j = 1$ (assigned prior to the evaluation period then revised later to maximize PES),
- $j = 1, 2, 3 \dots k$, and $k =$ number of professors in the instructional unit.

A 5-year total PES will measure long, continuous and productive contributions. These and additional details of the models for determining TES, RES and SES are given in Ridley and Collins (2015).

APPENDIX C

TES RESULTS FOR INDIVIDUAL PROGRAMS





CÓMO CITAR:

Korovyakovskaya, I., Llaugel, F., & Ridley, D. (2020). An Objctive Empirical Teaching Evaluation Metric at a Southeastern US Public University. *Revista Caribeña de Investigación Educativa (RECIE)*, 4(1), 34-48. <https://doi.org/10.32541/recie.2020.v4i1.pp34-48>