Business Education & Accreditation

VOLUME 14 NUMBER 1 2022

CONTENTS

Business Education & Accreditation

Editor in Chief	Managing Editor
Terrance Jalbert	Mercedes Jalbert

Editorial Board

Business Education & Accreditation, ISSN: 1944-5903 (print) and ISSN: 2157-0809 (online) publishes high-quality articles in all areas of business education, accreditation and related fields. Theoretical, empirical and applied manuscripts are welcome for publication consideration. The Journal is published once per year by the Institute for Business and Finance Research, LLC. All papers submitted to the Journal are blind reviewed. The Journal is distributed in print and through EBSCO*Host,* SSRN and RePEc.

The views presented in the Journal represent opinions of the respective authors. The views presented do not necessarily reflect the opinion of the editors, editorial board or staff of the Institute for Business and Finance Research, LLC. The Institute actively reviews articles submitted for possible publication. However, the Institute does not warrant the correctness of information provided in the articles or suitability of information in the articles for any purpose.

This Journal is the result of the collective work of many individuals. The Editors thank the members of the Editorial Board, ad-hoc reviewers and individuals that have submitted their research to the Journal for publication consideration.

ISSN : 1944-5903 (print) and ISSN: 2157-0809 (online)

Business Education & Accreditation
Vol. 14, No. 1, 2022, pp. 1-15
ISSN: 1944-5903 (print)
ISSN: 2157-0809 (online)

www.theIBFR.COM

FACTORS EXPLAINING BUSINESS STUDENTS' SUCCESS IN BUSINESS STATISTICS: A CASE FROM A SCANDINAVIAN BUSINESS SCHOOL

Leiv Opstad, Norwegian University of Science and Technology

ABSTRACT

Statistical skills are strongly linked success in business studies, especially in analyzing risk and in the financial sciences. Therefore, it is useful to acquire more knowledge about factors that can explain the grades achieved in Business Statistics. The objective of this study is to identify variables that are related to performance in Business Statistics among a cohort of business school students in Norway. By using linear regression models, this study tries to identify the relationship between achievement in Business Statistics and several independent variables, including gender, grade point average (GPA) from high school, mathematical background, Big Five personality traits, and attitudes towards statistics (SATS-36). Only attitudes towards statistics were significantly associated with the performance. There is a positive correlation between success in Business Statistics and the two Cognitive Competence and Effort (from SATS-36) dimensions. This is useful knowledge to ensure good results in Business Statistics.

JEL: A20, A22, M20

KEYWORDS: Gender, Big Five, Attitudes Towards Statistics, Mathematical Skills, Regression Model, Success in Business Statistics, Norway, Business School

INTRODUCTION

The introductory Business Statistics course is a major landmark in business courses, especially in the field of finance. Undergraduates need statistical skills to succeed in business subjects; Business Statistics is no different. A strong background in Business Statistics is useful in students' later career (Parker et al., 1999). Despite the usefulness of Business Statistics, many business students have little interest in this field and struggle learning this subject (Nilsson and Hauff, 2018). Business studies appeal to both men and women and there are equal numbers of both undertaking a bachelor's degree in business studies in Norway. Nevertheless, some gender differences remain, and this issue attracts interest among researchers. Why do female students have less interest in statistics and mathematics than the males? (Griffith et al., 2012; Reilly et al., 2019). This might explain why males outperform females in statistics courses (Haley et al., 2007), and to a higher degree choose quantitative economic majors (Worthington and Higgs, 2004). Women tend to prefer accounting, marketing, and management, while a higher percent of male students choose finance. This is in line with findings from Norway (Opstad, 2019).

The choice of educational pathway depends on the students' skills, preferences, and career interests. There is a close relationship between statistics and mathematics (Primi et al., 2020). Students who have any passion and interest in mathematics tend to have the same passion for statistics. The purpose of this study is to identify which factors are linked to grade scores in Business Statistics by using data from a Norwegian University, with a focus on gender, personality traits (Big Five), attitudes towards statistics (SATS-36), and mathematic and academic skills. Since performance in statistics is one of the key factors for success in business studies, it is important to research what determines the achievement in Business Statistics. It is of great value for planning within this field to identify which factors influence the performance in Business

Statistics. The investigation of this issue in this paper will hopefully be a useful contribution. An important contribution of this article is that it simultaneously combines gender, mathematical and academical abilities, and personal characteristics and attitudes towards statistics in the analysis of students' success in Business Statistics. This paper is organized in the following way. First, previous research is presented. On that basis, we will establish a research model as well as postulate some hypotheses. The discussion section focuses on analyzing the various contexts in the research model.

LITERATURE REVIEW

In the first part of this section, it is explained personality traits and attitudes towards statistics. They are key instruments linked to the research model and hypotheses.

The Big Five Personality Traits

The Big Five model for ascertaining personal characteristics (Costa et al., 1992) is widely accepted among researchers. It measures five factors: Agreeableness, Conscientiousness, Neuroticism, Extraversion, and Openness (see Table 1).

Table 1: The Big Five

Trait	**Explanation**
Openness to experience (O)	This person is open to new experiments and ideas
Conscientiousness (C)	This person is well organised, responsible, self-disciplined, effective, and target-oriented
Extraversion (E)	This person is social and oriented towards other people and the world
Agreeableness (A)	This person shows trust and tends to have unselfish manners
Emotional Stability (ES) (Opposite of Neuroticism)	This person tends to be emotionally stable

Openness is linked to intellectually curiosity, Conscientiousness is associated with achieving goals, agreeableness is characterized by the wish to contribute and help others, extraverts are outgoing, and emotional stability relates to not being depressed.

Attitudes Towards Statistics (SATS-36)

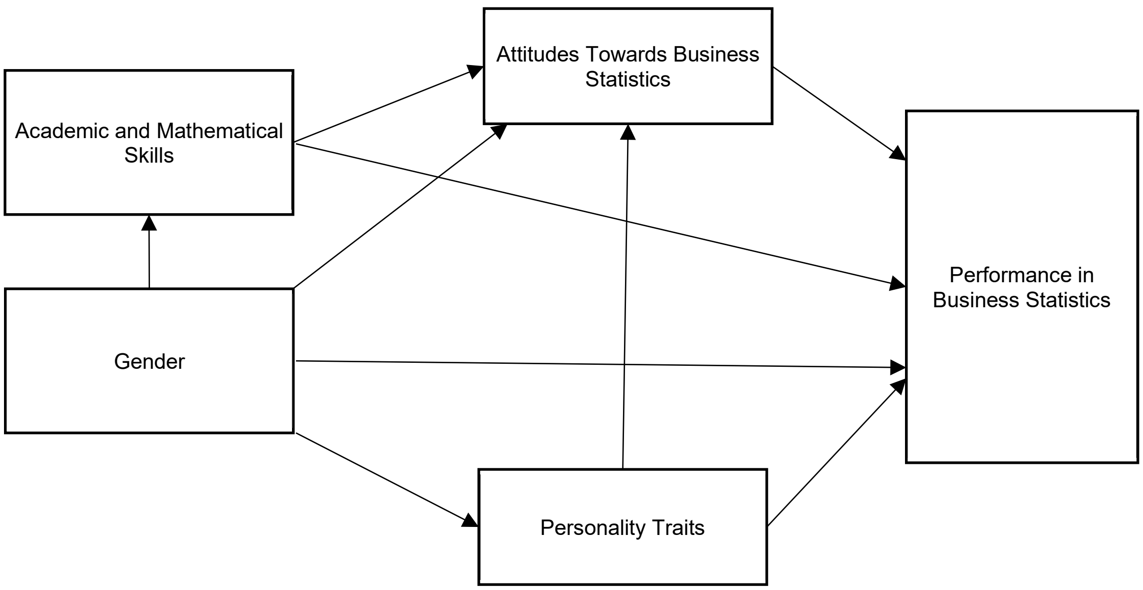
Different methods have been applied for measuring Attitudes towards statistics. This study uses SATS-36, as developed by Schau et al. (1995). It comprises 36 items and six components: Affect (6 items), Cognitive Competence (6 items), Value (9 items), Difficulty (7 items), Interest (4 items), and Effort (4 items). Affect gives an indicator of the person's feelings (positive or negative) about statistics. Cognitive Competence measures intellectual knowledge and skills towards statistics. Value determines the usefulness value of statistics. Difficulty measures if an individual finds it easy or difficult to apply statistics. Interest is an indicator of the degree of interest in statistics. Finally, Effort reveals how much time and effort an individual spends learning statistics. The literature shows that SATS-36 seems to have a high level of reliability and validity (Nolan et al., 2012; Persson et al., 2019).

The Research Model

In line with previous research, this paper introduces a model which analyzes the connection between gender, academic and mathematical skills, personality traits, attitudes towards statistics, and achievement in

Business Statistics. In analyzing the varied factors impact on individual's performance in Business Statistics, one must distinguish between direct and indirect effects (see Figure 1). For example, gender has a direct effect on performance, but also an indirect effect since one can assume there is a link between gender and certain variables like Academic and Mathematical skills, Attitudes towards statistics, and Personality traits. Personality traits and Academic and Mathematical skills can also be divided into a direct and indirect impact; they have a direct influence on performance in Business Statistics, but also indirect via for instance SATS-36 (Attitudes towards Statistics).

Figure 1: Research Model Illustrating Links Between Gender, Personality Traits, Mathematics Skills, Attitudes to Statistics, and Performance in Business Statistics



The figure illustrates the links between gender, personality traits, academic and mathematical skills, attitudes towards statistics and performance in Business Statistics. The model also takes into account that gender is correlated with academic and mathematical skills, personality traits and attitudes towards statistics. Furthermore, academic, and mathematical skills as well as the Big Five influence the attitudes towards statistic. In this way, the research model shows the distinction between direct and indirect effects.

Gender Impact

The gender effect on achievement in statistics is unclear. Some researchers have failed to find any gender impact (Esnard et al., 2021; Lester, 2007; Rabin et al., 2021; VanEs and Weaver, 2018), whereas others state females outperform males in statistics (Lester, 2016), or report higher scores for males (Schram, 1996). For Norwegian business students, Opstad (2018) concluded that male students achieve better grades than female students. Gender matters regarding attitudes towards statistics using SATS-36. Male students tend to have higher values in the Competence, Value, and Interest dimensions according to Hommik and Luik (2017). Rejón-Guardia et al. (2019) report mostly the same result, although they did not find any gender gap regarding Interest. However, they identified a higher score for females for the Effort dimension; specifically, females study harder than males. This is in line with results from a Norwegian Business school (Opstad, 2020). For all the other dimensions, Opstad registered a significant gender gap with the highest value for males and strongest impacts for Difficulty, Value, and Interest. Other researchers suggest the same tendency (Chiesi and Primi, 2015; Tempelaar and Nijhuis, 2007). Additionally, males express more positive attitudes towards statistics, and females are less confident using statistics as a tool. However, some investigators have come to a different conclusion. For instance, Coetzee and Merwe (2010) did not find any gender differences and Mahmud and Zanol (2008) suggested women had more positive attitudes towards statistics than men.

The previous literature recognized a gender gap regarding personality traits. Many scientists have reported higher values for men than women for Openness and Conscientiousness, while women attained the highest scores for Extraversion, Agreeableness, and Neuroticism (Costa et al., 2001; Weisberg et al., 2011). Schmitt et al. (2008) concluded that females score higher values in the Extraversion, Conscientiousness, Agreeableness, and Neuroticism dimensions (opposite of Emotional Stability), and lower scores for Openness compared to males across many cultures and countries (55 nations and N=17 637). This in line with findings from Norwegian students (Opstad, 2020). Therefore, the first hypothesis is:

H1: *Gender matters in performance in Business Statistics.*

<u>Academic Skills and Mathematical Background</u>

Some studies have not discovered any association between academic skills, mathematic skills, and success in statistics (Esnard et al., 2021). Others suggest a strong positive relationship exists between mathematical abilities and performance in statistics (Johnson and Kuennen, 2006; Lester, 2007). Quantitative skills are crucial in introductory statistics, and Johnson and Kuennen report a positive link between GPA (Grade Point Average) and achievement in statistics. Moreover, there is a strong positive correlation between mathematical and statistical competence. High mathematical score gives a positive attitude toward statistics (Stanisavljevic et al., 2014), and higher qualifications in mathematics are also positively linked to better grades in statistics (Choudhury and Radakrishnan, 2009; Johnson and Kuennen, 2006).

Among Norwegian business students, Opstad (2018) reported a positive significant correlation between performance in Business Statistics and the academic skills (GPA from high school) and mathematical skills variables; indeed, students skilled in theoretical mathematics in high school tend to get better grades. Mathematics and statistics are connected; this has an impact on students' attitude towards statistics. Put simply, mathematical background matters. Students with mathematical skills have more positive attitudes towards statistics (Carmona, 2004). This is in line with Opstad (2020) for Norwegian business students. For students skilled in theoretical mathematics in high school, there was a significant positive relationship to the Affect, Value, and Difficulty dimensions (SATS-36). The second hypothesis will therefore be:

H2: Academic Skills and Mathematics Background Are Associated with Success in Business Statistics.

<u>Personality Traits</u>

The Conscientiousness dimension helps students to focus on academic tasks and is a good indicator of academic success (Duckworth et al., 2019; Zell and Lesick, 2021). Some argue that Conscientiousness is the only predictor of academic achievement (Buju, 2013). Additionally, Openness tends to be related to academic success, while the result is mixed for the other dimensions.

According to Goldberg (2001), Emotional Stability may be important, while Extraversion and Agreeableness probably have little impact. Opstad (2021b) found a negative correlation between Openness and achievement in mathematics, but when controlling for attitude towards mathematics this impact disappeared. Opstad (2021a) also reported a significant negative connection between performance in macroeconomics and the two dimensions of Openness and Agreeableness. Conscientiousness was positively related to success in macroeconomics. Other researchers have also reported negative associations between Openness and performance (Busato et al., 2000; De Fruyt and Mervielde, 1996). Opstad (2020) suggested a link between personal characteristics and attitude towards statistics. Neuroticism was significantly negatively connected to Cognitive Competence and Affect, while Openness was significantly positively related to Interest, but negatively to Cognitive Competence. Furthermore, he found a significant relationship between Conscientiousness and all dimensions in attitudes towards statistics (SATS-36), although this was not significant for Difficulty; the impact was most strongly linked to Effort. Furnham and Chamorro-Premuzic (2004) confirm that there seems to be a strong link between Conscientiousness and attitudes towards statistics; statistics might apply for Conscientiousness in particular. Hard-working and goal-oriented students use a lot of energy in learning statistics, and they have a positive attitude towards doing so. On the basis of previous findings, the following hypothesis is postulated:

H3: Personality traits are connected to performance in Business Statistics.

Attitudes Towards Statistics

Finney and Schraw (2003) reported a positive link between self-efficacy in statistics and performance. This is in line with Esnard et al., (2021). Several articles have also reported a strong link between performance in statistics and the Affect and Cognitive Competence dimensions (Bechrakis et al., 2011; Nolan et al., 2012). Students with positive attitudes towards statistics tend to perform well in statistics (Lavidas et al., 2020; Sesé et al., 2015; Stanisavljevic et al., 2014). The final hypothesis is thus:

H4: Attitudes towards business statistics are associated with performance in Business Statistics.

DATA AND METHODOLOGY

Sample

The sample consists of 131 students examined in 2019. Students attending the compulsory second-year course in macroeconomics answered the questionnaire. This means that the students have taken the exam in the compulsory course in Business Statistics that runs in the first year. The students answered questions based on the items in Big Five and SATS-36. The participation was voluntary. Around 70 percent of the students attended the course on their chosen day, hence the data might be marginally biased. Even so, the survey gives a good picture of students' attitudes (Bonesrønning and Opstad, 2015). The data are mixed with administrative information about mathematical background, Grade Point Average (GPA) from high school and performance in Business Statistics. Some students did not report personal data. Therefore, regarding information about GPA and grades in statistics, we lacked data for these students (see Table 2).

The average grade in statistics was quite high (close to B). One explanation for this is that there were many applicants to the program and high GPAs from high school are required to be accepted. There is considerable variation in the attitudes towards statistics, with highest scores for Effort. The values for dimensions in the Big Five personality traits vary between 3.3 to 3.9. There are slightly more women than men in the sample, and the values of Skewness, Kurtosis, and Scale Reliability are within the accepted intervals. The Appendix presents the correlations between the variables.

Table 2: Sample Information

Variable	N	Min	Max	Mean	St. Dev.	Skewness	Kurtosis	Scale Reliability Cronbach's Alpha
Performance Statistics (0:F,1:E,2:D,3:C,4:B,5:A)	79	1	5	3.96	0.980	-0.929	0.730	
Cognitive Competence [1)]	131	1.83	7	5.19	1.079	-.516	-0.053	0.84
Value[1)]	131	1.33	6.7	4.44	0.996	-0.220	-0.015	0.85
Difficulty[1)] (Find statistics easy to learn)	131	1.14	5.2	3.52	0.751	.065	.0189	0.65
Interest [1)]	131	1	7	4.50	1.291	-0.222	-0.192	065
Affect [1)]	131	1	7	4.59	1.249	-0.460	0.151	0.84
Effort [1)]	131	1.50	7	5.67	1.086	-1.250	1.847	0.70
N-math [2)] (0:Non N-math,1 :N-math)	131	0	1	0.29	0.456	0.936	-1.142	
Extraversion[3)]	130	1.75	5	3.64	0.783	-0.211	-0.483	0.84
Agreeableness[3)]	130	2.25	5	3.91	0.570	-0.574	0.110	0.49
Conscientiousness[3)]	130	1.50	5	3.67	0.702	-0.569	0.223	0.71
Emotional Stability[3)]	130	1.50	5	3.30	0.791	-0.041	-0.471	0.74
Openness[3)]	130	1.50	5	3.33	0.726	-0.120	-0.482	0.59
Gender (0:F,1:M)	131	0	1	.48	0.502	0.077	-2.025	
GPA (High School)	85	46.9	66.7	51.16	3.20	1.642	5.308	

In statistics, the Likert scale ranged from 1 to 7; 2) Students who have chosen mathematics for natural science at high school; 3) the Likert scale ranged from 1 to 5. Since one relies on students' permission to link the data to GPA and results in Business Statistics, the number of observations for these factors is lower than for the other variables. There are acceptable values on Skewness, Kurtosis and Cronbach's Alfa.

The Regression Models

By using a linear regression model, we can see how different independent variables are linked to the chosen dependent variable (performance in Business Statistics). By using mediation analyses, it is possible to distinguish between direct and indirect effects (Park et al., 2019). Alternatively, one can use different sets of variables in the regression model (Opstad, 2020; Shi et al., 2020). Model 1 (set 1) includes only gender (see Figure 1), whilst Model 2 (Set 2) also includes mathematical and academic skills. Model 3 (set 3) adds personality traits. Finally, Model 4 (set 4) incorporates the complete model. The changes between the models (sets) gives a picture of the indirect impact. For instance, Model 1 (see equation 1) shows the total effect of gender, Model 4 (equation 4) the direct effect, and the difference between (1) and (4) indicates the indirect influence between gender and performance in Business Statistics.

Model 1: $Y_i = a_0 + a_1\mathrm{X1}_i + +\varepsilon_i$ (1)

Model 2: $Y_i = a_0 + a_1\mathrm{X1}_i + a_2\mathrm{X2}_i + a_3\mathrm{X3}_i + +\varepsilon_i$ (2)

Model 3: $Y_i = a_0 + a_1\mathrm{X1}_i + a_2\mathrm{X2}_i + a_3\mathrm{X3}_i + a_4\mathrm{X4}_i + a_5\mathrm{X5}_i + a_6\mathrm{X6}_i + a_7\mathrm{X7}_i + \varepsilon_i$ (3)

Model 4: $Y_i = a_0 + a_1\mathrm{X1}_i + a_2\mathrm{X2}_i + a_3\mathrm{X3}_i + a_4\mathrm{X4}_i + a_5\mathrm{X5}_i + a_6\mathrm{X6}_i + a_7\mathrm{X7}_i + a_8\mathrm{X8}_i + a_9\mathrm{X9}_i + a_{10}\mathrm{X10}_i + a_{11}\mathrm{X11}_i + a_{12}\mathrm{X12}_i + a_{13}\mathrm{X13}_i + a_{14}\mathrm{X14}_i + \varepsilon_i$ (4)

where:

Y = grade attained in Business Statistics (0: F, 1: E, 2: D, 3: C, 4: B, 5: A),
i = student, α0 = constant,

X1 = Gender (0: F, 1: M),
X2 = High school GPA,
X3 = dummy variable for N-maths (0: did not take N-maths, 1: took N-maths),
X4 = Openness, X5 = Extraversion, X6 = Agreeableness, X7 = Conscientiousness,
X8 = Emotional stability, X9 = Cognitive Competence in statistics,
X10 = Perception of the value of statistics, X11 = Difficulty (Stat),
X12 = Interest in statistics X13= Affect in statistics, X14 = Effort in statistics,
ε = stochastic error.

The Big Five personality traits were measured by using 20 items on a 5-point Likert scale where 1 = strongly disagree and 5 = strongly agree. Similarly, SATS-36 were computed on a 7-point Likert scale where 1 = strongly disagree and 7 = strongly agree. In this study, we did not have access to experimental data. Even if there is a correlation between the dependent variable and the independent variables, one must be careful to monitor any causal effects.

RESULTS AND DISCUSSION

Table 3 presents the results from the regression models. Since none of the models show any significant gender effects, there is neither any direct nor indirect gender differences associated with the performance in statistics. Hence, hypothesis 1 (H1) is rejected.

Table 3: Outputs from the Four Linear Regression Models

Variable	**Model 1 (Set 1)**		**Model 2 (Set 2)**		**Model 3 (Set 3)**		**Model 4 (Set 4)**		
	B	***p***	**B**	***p***	**B**	***p***	**B**	***p***	**VIF**
Constant	3.91		2.21		1.40		-1.19		
Gender	0.118	0.60	0.056	0.81	-0.072	0.708	-0.045	0.831	1.52
	(0.224)		(0.227)		(0.259)		(0.208)		
GPA			0.32	0.45	0.054	0.205	0.041	0.222	1.15
			(0.042)		(0.042)		(0.034)		
N-Maths			0.305	0.22	0.248	0.305	0.052	0.792	1.19
			(0.245)		(0.240)		(0.195)		
Openness					-0.396	0.019	-0.168	0.243	1.68
					(0.164)	**	(0.142)		
Extraversion					-0.064	0.682	-0.154	0.220	1.48
					(0.156)		(0.125)		
Agreeableness					0.098	0.665	0.160	0.368	1.19
					(0.225)		(0.176)		
Conscientiousness					0.042	0.829	-0.108	0.490	1.25
					(0.194)		(0.156)		
Emotional Stability					0.238	0.155	0.094	0.519	1.86
					(0.175)		(0.146)		
Affect									
Value							-0.013	0.927	2.43
							(0.139)		
Difficulty							-0.169	0.261	1.73
(find statistics easy)							(0.149)		
Interest							-0.003	0.979	2.58
							(0.107)		
Cognitive Competence							.625	0.000	2.53
							(.123)	***	
Effort							.174	0.087	1.46
							(0.100)	*	
	N=78		N=78		N=77		N=77		
	Adj.R^2=-0.09		Adj.R^2 =-0.02		Adj.R^2 = 0.052		Adj.R^2 = 0.429		
	R^2= 0.04		R^2 = 0.036		R^2= 0.151		R^2 = 0.526		

*Model 1-4 , see equation 1-4. The models show how the different steps influence the estimated variables for identifying direct and indirect impacts. Std. Error in parentheses B= Standardized Coefficients. ***p < 0.01, **p < 0.05 and *p < 0.1, VIF = Variance Inflation Factor . Due to high VIF value (4.5), Affect is not included in the regression models.*

The model specifications (Set 2-Set 4) do not indicate any significant impact on success in Business Statistics related to Mathematical and Academical skills. Therefore, hypothesis 2 (H2) is also not confirmed. Model 3 reveals that only one dimension of personality traits is significantly linked to achievement in Business Statistics; Openness is negatively related to performance in statistics but when controlling for attitudes towards statistics (SATS-36, see Model 4) this effect disappears. The conclusion is that hypothesis 3 (H3) is not confirmed. Two dimensions of SATS-36 are significantly positively correlated with achievements in Business Statistics. Cognitive Competence is strongly related with a high value of the parameter (B= 0.625), whilst the impact of Effort is lower (B=0.174) and only significant at the 10 percent level. The findings confirm hypothesis 4 (H4): Attitudes towards Business Statistics are associated with performance in Business Statistics. In line with other published articles (Nolan et al., 2012), this study reports a strong and significant relationship between Cognitive Competence in statistics and achievement in statistics.

Gender and Performance (Hypothesis 1)

In the field of business administration, there are approximately the same number of males and females. Prior research indicates that a gender difference exists among business and economics students in Norway in performance inclusive Business Statistics (2018) and in choice of major (Opstad, 2019). Despite the gender equalization in Norway, girls tend to a lesser degree to select theoretical mathematics at high school; they prefer more practical mathematics, and this is a disadvantage when studying business subjects (Opstad, 2018; 2019). This is in line with research from other countries (Pritchard et al., 2004). The gender difference in mathematics and attitudes towards mathematics might explain the underrepresentation of women in science, technology, and engineering. Even with a high degree of gender equality, this might explain the existing gender gap (Stoet and Geary, 2018). Opstad and Årethun (2019) report a gender difference in favor of males regarding attitudes towards mathematics. However, it looks like this is changing. Updated figures indicate that the traditional gender divide in attitudes towards statistics is disappearing in Norway. The Pisa test of 2018 reported a higher score in mathematics for females than males in Norway (OECD education, 2020). Furthermore, Utvær (2019) did not find any gender difference in attitudes towards mathematics among Norwegian pupils in primary schools. In recent research among Norwegian business students, Opstad (2020, 2021b) did not notice any gender differences in attitudes towards statistics in either mathematics or in performance in mathematics (Opstad 2021c). The results in this study confirm this tendency. There is no longer any gender difference in performance in business statistics among Norwegian business students.

Mathematical Background and Academic Skills (Hypothesis 2)

According to Opstad (2018), GPAs from high school may not be a good predictor of success at business schools in Norway. The situation seems to be different for mathematical skills since Opstad (2018) suggests background and knowledge in mathematics are a key factor explaining good performance. However, after controlling for attitudes towards mathematics this correlation disappeared in performance in mathematics (Opstad 2021c). This is in line with this study. There is no significant correlation between mathematical abilities and results in Business Statistics in models 3 or 4, even though mathematics and statistics are closely related, and quantitative abilities are important for success in business studies. One explanation may be that N-mathematics focuses on mathematics applied to science areas. This may be less applicable in Business Statistics. Furthermore, it is a requirement to be able to take statistics within a mandatory introductory course in mathematics. This may have contributed to reducing the differences in mathematical abilities from high school.

Personality Traits (Big 5) and Success in Business Statistics (Hypothesis 3)

Previous research provides a mixed picture when it comes to the link between personal characteristics and academic success. Although several researchers point out that there is a correlation between personality traits and attitudes to statistics (Chiesi and Bruno, Opstad, 2020), it is still unclear what the link is between personality traits and successes in statistics. This study does not find any significant correlation in relation to this issue. There is a significant link between Openness and performance in statistics in Model 3, but this effect disappears in Model 4.

Attitudes Towards Statistics (SATS-36) and Achievement in Business Statistics (Hypothesis 4)

In line with other published articles (Nolan et al., 2012), this study reports a strong and significant relationship between Cognitive Competence in statistics and achievement in statistics. This makes sense, as Cognitive Competence is an indicator of knowledge and the ability to use statistics. Furthermore, increased effort in the subject will be rewarded with better grades. This is consistent with previous research results (Dotterweich and Rochelle, 2012). Moreover, dimensions like interest in statistics, value of statistics, and finding statistics easy were not correlated to performance in this subject. However, the overall attitudes towards statistics seem to play an important role in explaining the success therein. To illustrate, adjusted R square increases from 0.052 to 0.429 by including SATS 36. All other variables except attitudes to statistics have little explanatory effect on the results in this investigation.

Limitations and Some Implications

The dataset in this study is only from a single business school in Norway. Subsequently, one must be careful when interpreting these findings in a wider context. In this research, we applied the original version of the Big Five and SATS-36 (translated into Norwegian); an alternative and more robust version might use explanatory factor analysis and present a modified version of the Big Five and SATS-36. Nevertheless, the original version is used in this paper to ensure consistency with previous research. Effort and Cognitive Competence are positively associated with performance in Business Statistics. On the other hand, good grades in Business Statistics will increase the level of Cognitive Competence in statistics, so the causal relationship could go in both directions. Regardless, educators should consider boosting students' attitudes towards statistics

CONCLUDING COMMENTS

The purpose of this article is to identify factors that influence performance in Business Statistics since this subject is an important tool for business students. By using regression models based on data from NTNU Business School, we try to find variables that are significant correlated to achievement in Business Statistics. The data are based on a questionnaire handed out to the students. This information was linked to administrative available data. In order to capture both direct and indirect effects, several regression models are presented with different sets of explanatory variables. Previous research suggests there is a gender gap in performance in Business Statistics. This study does not confirm this. The reason for the lack of identification of the purported gap might be that it has shrunk or no longer exists in business students. GPAs from high school are an indicator of academic skills. This study did identify this variable's relation to performance in Business Statistics. However, the author found no relationship between success in Business Statistics and the two independent variables: mathematical background from high school and personality traits (Big Five). Only attitudes towards statistics were significantly related to performance in Business Statistics. There are also positive relationships between success in this field and two dimensions, specifically Cognitive Competence and Effort. This paper suggests that attitudes towards statistics are a key factor for success in Business Statistics. Further research may be exploring factors that motivate students to learn Business Statistics.

APPENDIX

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	1	-0.011	0.116	-0.083	0.105	-0.052	0.079	0.079	0.120	-0.111	0.053	-0.163	0.024	0.091
2	-0.011	1	0.060	0.172	0.250	0.290	0.274	-0.257	0.046	-0.015	-0.172	-0.114	0.445	0.193
3	0.116	0.060	1	0.637	0.410	0.301	0.314	0.254	-0.268	-0.078	0.065	0.130	0.097	0.166
4	-0.083	0.172	0.637	1	0.574	0.572	0.481	0.071	-0.093	0.035	-0.066	0.249	0.267	0.267
5	0.105	0.250	0.410	0.574	1	0.290	0.698	0.134	0.052	-0.041	-0.088	0.142	0.182	0.305
6	-0.052	0.290	0.301	0.572	0.290	1	0.245	-0.164	-0.084	0.040	-0.106	0.074	0.176	0.210
7	0.079	0.274	0.314	0.481	0.698	0.245	1	0.056	0.169	-0.043	-0.066	0.117	0.172	0.200
8	0.079	-0.257	0.254	0.071	0.134	-0.164	0.056	1	-0.136	0.025	0.272	0.418	-0.213	-0.065
9	0.120	0.046	-0.268	-0.093	0.052	-0.084	0.169	-0.136	1	0.355	-0.105	-0.273	0.271	0.007
10	-0.111	-0.015	-0.078	0.035	-0.041	0.040	-0.043	0.025	0.355	1	0.152	-0.072	0.304	-0.121
11	0.053	-0.172	0.065	-0.066	-0.088	-0.106	-0.066	0.272	-0.105	0.152	1	0.295	-0.025	-0.045
12	-0.163	-0.114	0.130	0.249	0.142	0.074	0.117	0.418	-0.273	-0.072	0.295	1	-0.050	0.049
13	0.024	0.445	0.097	0.267	0.182	0.176	0.172	-0.213	0.271	0.304	-0.025	-0.050	1	0.125
14	0.091	0.193	0.166	0.267	0.305	0.210	0.200	-0.065	0.007	-0.121	-0.045	0.049	0.125	1

1:GPA, 2:Gender, 3: Performance Business Stat, ¤ Stat CogC, 5:Stat Value, 6: Stat Difficult, 7: Stat Interest, 8: Stat Effort, 9: Openness, 10: Extraversion, 11: Agreeableness, 12 : Conscientiousness,13: Emotional Stability, 14: N-mat

BIOGRAPHY

Leiv Opstad is a Professor in Economics at the Norwegian University of Science and Technology (NTNU) Business School in Trondheim, Norway. He teaches economics and public economics for undergraduate and master's degree students. His previous work includes five years at the Ministry of Finance. He has significant experience in teaching at different levels and has written textbooks in macroeconomics and management within the public sector. His research is mainly connected to higher education and business studies.

REFERENCES

Bechrakis, T., Gialamas, V., and Barkatsas, A. N. (2011) ''Survey of Attitudes Toward Statistics (SATS): An investigation of its construct validity and its factor structure invariance by gender'', *International Journal of Theoretical Educational Practice*, vol. 1(1), p.1–15.

Bonesrønning, H. and Opstad, L. (2015) ''Can student effort be manipulated? Does it matter?'', *Applied Economics,* 47(15), p.1511–1524 http:// doi.org/10.1080/00036846.2014.997923

Buju, S. (2013) ''Personality profile of students with technical academic performance'', *Procedia-Social and Behavioral Science*, 78, p.56–60 http://doi.org/10.1016/ j.sbspro.2013.04.250.

Busato, V.V., Prins, F.J., Elshout, J.J. and Hamaker, C. (2000) ''Intellectual ability, learning style, personality, achievement motivation and academic success of psychology students in higher education'', *Personality and Individual Differences*, vol. 29(6). p.1057–1068.

Carmona, J. (2004), ''Mathematical background and attitudes toward statistics in a sample of undergraduate students'' in *10th International Conference on Mathematics Education, Copenhagen* http://www.stat.auckland.ac.nz/~iase/publications/11/Carmona.doc]

Chiesi, F. and Bruno, F. (2021) ‘’Mean differences and individual changes in nursing students'’ attitudes toward statistics: The role of math background and personality traits’, *Nurse Education in Practice*, 52, 103043, p.1-6.] https://doi.org/10.1016/j.nepr.2021.103043

Chiesi, F., & Primi, C. (2015, February) ‘’Gender differences in attitudes toward statistics: Is there a case for a confidence gap?’’ in *CERME 9-Ninth congress of the European society for research in mathematics education,* Pague, Czech Republic, p. 622-628

Choudhury, A. and Radhakrishnan, R. (2009) ‘‘Testing the differential effect of a mathematical background on the statistics course performance: An application of the chow-test’’, *Journal of Economics and Economic Education Research*, vol. 10(3),p.15-26.

Coetzee, S. and Merwe, P. V. D. (2010). ‘’Industrial psychology students' attitudes towards statistics. *SA Journal of Industrial Psychology*, *36*(1), p. 1-8.

Costa, P. T. Jr., Terracciano, A., and McCrae, R. R. (2001) ‘’Gender differences in personality traits across cultures: robust and surprising findings’’*Journal of personality and social psychology,* vol. 81, p.322–331.

De Fruyt, F. and Mervielde, I. (1996) ‘’Personality and interests as predictors of educational streaming and achievement’’, *European Journal of Personality*, vol. 10(5), p.405–425 http://doi.org/10.1002/(sici)1099-0984(199612)10:5.

Dotterweich, D. P., and Rochelle, C. F. (2012) ‘’Online, ITV, and traditional delivery: Student characteristics and success factors in business statistics’’, *American Journal of Business Education*, *5*(2), p.129–138. https://doi.org/10.1108/1836326121128174

Duckworth, A. L., Taxer, J. L., Eskreis-Winkler, L., Galla, B. M., and Gross, J. J. (2019) ‘‘Self-control and academic achievement,’, *Annual Review of Psychology*, 70, 373-399 [online] http://doi.org/10.1146/annurev-psych-010418-103230

Esnard, T. R., Alladin, F. M., and Samlal, K. C. (2021) ‘‘Prior mathematics performance, statistics anxiety, self-efficacy and expectations for performance in statistics: a survey of social sciences students in a Caribbean institution of higher education’’, *Statistics Education Research Journal*, 20(1), p. 4-4 https://doi.org/10.52041/serj.v20i1.98

Finney, S. J., and Schraw, G. (2003) ‘’Self-efficacy beliefs in college statistics courses’’, *Contemporary educational psychology*, *28*(2), p.161-186 https://doi.org/10.1016/S0361-476X(02)00015-2

Furnham, A. and Chamorro-Premuzic, T. (2004) ‘‘Personality and intelligence as predictors of statistics examination grades’’, *Personality and Individual Differences*, vol. 37(5),p.943-955.

Goldberg, L. R. (2001) ‘‘Frozen by success: Why we don’t know nearly enough about the relations between personality attributes and academic performance’’ in *Remarks delivered at the ETS Workshop: Applications to new constructs, Educational Testing Service, Princeton, NJ.*

Griffith, J. D., Adams, L. T., Gu, L. L., Hart, C. L., and Nichols-Whitehead, P. (2012) ‘’Students' attitudes toward statistics across the disciplines: a mixed-methods approach-’, *Statistics Education Research Journal*, vol. 11(2), p.11-30.

Haley, M. R., Johnson, M. F. and Kuennen, E. W. (2007) ‘’Student and professor gender effects in introductory business statistics’’. *Journal of Statistics Education*, vol. 15(3) https://doi.org/10.1080/10691898.2007.11889549

Hommik, C. and Luik, P. (2017) ‘’Adapting the Survey of Attitudes towards Statistics (SATS-36) for Estonian Secondary School Students’’, *Statistics Education Research Journal*, vol. 16(1), p.228-239. https://doi.org/10.52041/serj.v16i1.229

Johnson, M. and Kuennen, E. (2006) ‘‘Basic math skills and performance in an introductory statistics course’’, *Journal of Statistics Education,* 14(2), p.1-15. https://doi.org/10.1080/10691898.2006.11910581

Lavidas, K., Barkatsas, T., Manesis, D. and Gialamas, V. (2020) ‘A structural equation model investigating the impact of tertiary students' attitudes toward statistics, perceived competence at mathematics, and engagement on statistics performance’, *Statistics Education Research Journal*, vol. 19(2),. p.27-41. https://doi.org/10.52041/serj.v19i2.108

Lester, D. (2007) ‘‘Predicting performance in a psychological statistics course’’, *Psychological Reports,* 101(1), p. 334. https://doi.org/10.2466/pr0.101.1.334-334

Lester, D. (2016) ‘’Predicting success in psychological statistics courses’’, *Psychological Reports*, 118(3), p.772–777https://doi.org/10.1177/0033294116647687

Mahmud, Z. and Zainol, M.S. (2008) ‘’Examining postgraduate students’ perceived competency in statistical data analysis and their attitudes toward statistics’’, *International Journal of Education and Information Technologies*, vol. 2(1), p.79–86.

Nilsson, J. and Hauff, J. C. (2018) ‘‘Reducing statistics anxiety using limited teaching resources’’, *Journal of International Education in Business,* vol.11(2), p.312-323.-] https://doi.org/10.1108/JIEB-03-2018-0010

Nolan, M. M., Beran, T. and Hecker, K. G. (2012) ‘’Surveys assessing students’ attitude toward statistics: A systematic review of validity and reliability’’, *Statistics Education Research Journal*, vol. 11(2), p. 103-123. https://doi.org/10.52041/serj.v11i2.333

OECD Education (2020) [https://gpseducation.oecd.org/CountryProfile?primaryCountry=NOR&treshold=10&topic=PI

Opstad, L. (2018) ‘‘Success in business studies and mathematical background: The case of Norway’’, *Journal of Applied Research in Higher Education,* vol. 10(3), p. 399-408. https://doi.org/10.1108/JARHE-11-2017-0136

Opstad, L. (2019) ‘’Different attitudes towards mathematics among economic and business students and choice of business course major in Norway’’, *Social Sciences and Education Research Review*, vol. 6(2),p.6-30.

Opstad L. (2020) ‘’Attitudes towards Statistics among Business Students: Do Gender, Mathematical Skills and Personal Traits Matter?’’, *Sustainability*. 12(15), 6104 https://doi.org/10.3390/su12156104

Opstad, L. (2021a). ‘‘Can we identify the students who have success in macroeconomics depending on exam format by comparing multiple-choice test and constructed response test?’’, *International Journal of Education Economics and Development*, Vol. 12(4), p. 311-328.

Opstad, L. (2021b) ‘‘Factors Explaining Business Student Attitudes towards Mathematics: Does Gender Still Matter?’’, *European Journal of Science and Mathematics Education*, vol. 9(2), p. 13-25.

Opstad L. (2021c) ‘‘Factors Explaining Business Students’ Performance in an Introductory Mathematics Course. What Are the Impacts Of Gender, Academic Ability, Personality Traits, and Attitudes Towards Mathematics?’’, *Advances in Education Sciences*, vol. 3(1) p. 23-43.

Opstad, L. and Årethun. T. (2019) ‘‘Attitude towards mathematics among economics and business students in Norway. Is there any gender difference?’’ in *Proceedings of the 16th International Conference on Enterprises, Systems, Accounting, Logistics and Management* (16th ICESALM 2019). Chania, Crete, Greece

Park, H., Oh, H. and Boo, S. (2019) ‘The Role of Occupational Stress in the Association between Emotional Labor and Mental Health: A Moderated Mediation Model’, *Sustainability,* vol. 11(7), 1886. https://doi.org/10.3390/su11071886

Parker, R. S., Pettijohn, C. E. and Keillor, B. D. (1999) ‘‘The nature and role of statistics in the business school curriculum’’, *Journal of Education for Business*, vol. 75(1), p. 51-54.

Persson, I., Kraus, K., Hansson, L. and Wallentin, F. Y. (2019) ‘‘Confirming the Structure of the Survey of Attitudes toward Statistics (SATS-36) by Swedish Students’’, *Statistics Education Research Journal*, vol. 18(1).. https://doi.org/10.52041/serj.v18i1.151

Primi, C., Bacherini, A., Beccari, C. and Donati, M. A. (2020) ‘’Assessing math attitude through the Attitude Toward Mathematics Inventory–Short form in introductory statistics course students’,, *Studies in Educational Evaluation*, vol. 64, 100838 https://doi.org 10.1016/j.stueduc.2020.100838

Pritchard, R. E., Potter, G. C. andSaccucci, M. S. (2004) ‘‘The selection of a business major: Elements influencing student choice and implications for outcomes assessment’’, *Journal of Education for Business*, vol. 79(3), p. 152-156.

Rabin, L. A., Krishnan, A., Bergdoll, R. and Fogel, J. (2021) ‘’Correlates of exam performance in an introductory statistics course: basic math skills along with self-reported psychological/behavioral and demographic variables’’, *Statistics Education Research Journal*, 20(1), https://doi.org/10.52041/serj.v20i1.98

Rejón-Guardia, F., Vich-I-Martorell, G. A., Juaneda, C. and Cladera, M. (2019) ‘‘Gender differenced in attitudes towards statistics in social science degrees’’ in *EDULEARN19 Proceedings* (p. 2933-2941). IATED. Academy, Palme de Mallorca, p. 2933-3683.

Reilly, D., Neumann, D. L. and Andrews, G. (2019) ‘’Investigating gender differences in mathematics and science: Results from the 2011 Trends in Mathematics and Science Survey’’, *Research in Science Education*, vol. 49(1), p. 25-50.

Schau, C., Stevens, J., Dauphinee, T. L. and Vecchio, A. D. (1995) ‘‘The development and validation of the survey of antitudes toward statistics’’, *Educational and Psychological Measurement*, vol. 55(5), p. 868-875.

Schmitt, D. P., A. Realo, M. Voracek, and J. Allik. (2008) ‘’Why can't a man be more like a woman? Sex differences in Big Five personality traits across 55 cultures’’, *Journal of Personality and Social Psychology,* Vol. 94(1) .168. https://psycnet.apa.org/doi/10.1037/a0014651

Schram, C. M. (1996) ''A meta-analysis of gender differences in applied statistics achievement'', *Journal of Educational and Behavioral Statistics*, Vol. 21(1), p. 55-70.

Sesé, A., Jiménez, R., Montaño, J. J. and Palmer, A. (2015) ''Can attitudes toward statistics and statistics anxiety explain students' performance'', *Revista de Psicodidáctica,* 20(2), p. 285–304 [online] https://doi.org/10.1387/RevPsicodidact.13080

Shi, J., Wu, C. and Qian, X. (2020) 'The Effects of Multiple Factors on Elderly Pedestrians' Speed Perception and Stopping Distance Estimation of Approaching Vehicles', *Sustainability*, vol. 12(13), 5308. https://doi.org/10.3390/su12135308

Stanisavljevic, D., Trajkovic, G., Marinkovic, J., Bukumiric, Z., Cirkovic, A. and Milic, N. (2014) ''Assessing attitudes towards statistics among medical students: psychometric properties of the Serbian version of the Survey of Attitudes Towards Statistics (SATS)'', *PLOS One*, vol. 9(11),. 112567. https://doi.org/10.1371/journal.pone.0112567

Stoet, G. and Geary, D. C. (2018) ''The gender-equality paradox in science, technology, engineering, and mathematics education'', *Psychological Science*, vol. 29(4), p. 581-593 https://doi.org/10.1177/0956797617741719

Trapmann, S., Hell, B., Hirn, J. O. W., & Schuler, H. (2007). ,'Meta-analysis of the relationship between the Big Five and academic success at university''. *Zeitschrift für Psychologie/Journal of Psychology*, vol. 215(2), p. 132-151. https://doi/org/10.1027/0044-3409.215.2.132

Tempelaar, D. and Nijhuis, J. (2007) ''Commonalities in attitudes and beliefs toward different academic subjects'', in *The Challenges of Educating People to Lead in a Challenging World,* Springer, Dordrecht, p. 225-249.

Utvær, H. T. (2019) ''Elevenes motivasjon for matematikk i barneskolen [Pupils' motivation for mathematics in primary school]'' (Bachelors thesis)'. NTNU, 2019 (In Norwegian).

VanEs, C., and Weaver, M.M. (2018) ''Race, sex, and their influences on introductory statistics education'', *Journal of Statistics Education*, vol. 26(1), p. 48–54. https://doi.org/10.1080/10691898.2018.1434426

Weisberg, Y. J., DeYoung, C. G. and Hirsh, J. B. (2011) 'Gender differences in personality across the ten aspects of the Big Five', *Frontiers in Psychology*, 2,178https://doi.org/10.3389/fpsyg.2011.00178

Worthington, A. and Higgs, H. (2004) ''Factors explaining the choice of an economics major: the role of student characteristics, personality and perceptions of the profession'', *International Journal of Social Economics*. vol. 31(5/6), p. 593-613.

Zell, E. and Lesick, T. L. (2021) 'Big five personality traits and performance: A quantitative synthesis of 50+ meta-analyses', *Journal of Personality.* https://doi.org/10.1111/jopy.12683

BIOGRAPHY

Leiv Opstad is professor in economics at NTNU (Norwegian University of Science and Technology) Business School in Trondheim, Norway. Previous work is five years at the Norwegian Ministry of Finance. The author has long experience in teaching at different levels and has written textbooks in macroeconomics and management within the public sector. His research is mainly connected to marked analysis and educational issues.

Business Education & Accreditation
Vol. 14, No. 1, 2022, pp. 17-31
ISSN: 1944-5903 (print)
ISSN: 2157-0809 (online)

IBFR
www.theIBFR.COM

A LEXICAL ANALYSIS OF MISSION STATEMENTS FROM AACSB ACCREDITED BUSINESS SCHOOLS

James Suleiman, University of Southern Maine

ABSTRACT

College of Business mission statements can be a means to differentiation or an exercise in conformity. This article uses n-gram analysis to show that there are some lexical patterns distinctive to specific types of institutions and then employs Latent Dirichlet Analytics, a specific form of unsupervised topic modeling, to examine mission statement characteristics by a variety of institutional characteristics for institutions accredited by the Association to Advance Collegiate Schools of Business. There are certain words that were more common to specific types of institutions based on characteristics including region, Carnegie classification, initial accreditation year, and institutional control. A variety of topic models are examined but due to potential conformity in mission writing information and process sharing, there wasn't sufficient variety in mission to differentiate adequate models based on the set of institutional characteristics used. Suggestions for further research are discussed.

JEL: M0, M1

KEYWORDS: AACSB Accreditation, LDA, Topic Modeling, Mission Statement

INTRODUCTION

A well-written mission statement describes why an organization exists, differentiates it from others of its type, guides future actions, and shows us what image the organization wishes to project. Forming the mission is an essential component of strategic planning and is often accompanied by the definition of a vision (Pearce, 1982). Companies create, implement, and revise mission statements because they are a cornerstone of strategic planning, and there is an expectation that the process will generate benefits.

Vision and mission statements should causally influence employee decision making and help achieve organizational goals. Mission statements have an internal focus, and the primary audience is leadership, team, and stakeholders. They have been shown to be important in achieving results (Taiwo et al., 2016). Employees need to know the mission exists and what it means. Employee awareness and buy-in are critical to a successful mission, and they shouldn't believe that the mission is solely owned by senior management (Darbi, 2012).

The rationale for developing a mission statement is essential. The first step in formulating a mission statement should be asking why you want to create a mission statement. A mission statement misaligned with organizational structure is of little value. Rationales for developing mission statements may lead to improved performance. Some of these are consistent across industry, while others may be firm-specific (Bart & Tabone, 1998). While there are no set standards for mission statements, it is simple to search the management literature for guidelines. In providing guidelines for developing mission, (Powers, 2012) suggests:

"A mission statement should simply identify the broad customer need(s) that an organization is going to satisfy. It indicates the organization's fundamental reason for existence. Using this guideline, examples of good mission statements are: Wal-Mart: To help people save money so they can live better; Harvard: To educate leaders who make a difference in the world; United Methodist Church: To make disciples of Jesus Christ for the transformation of the world."

The mission preparation process is more important than the actual mission statement. It needs to encompass a wide variety of stakeholders and have top management's commitment (Mullane, 2002). Much of the early management research examined mission statement formulation in for-profit, US-based firms. More current research has shown that mission statements are not static enduring proverbs. Environments shift, organizations merge, innovation disrupts entire business models, and organizations must continuously revise the mission. There are similarities and differences across countries (King et al., 2012) and comparing for-profits to non-profits (Bart & Tabone, 1998).

Periodically, researchers and executives question the alleged necessity of creating mission statements. Simply put, they ask "do mission statements matter?". While many authors extol the virtues of the mission statement, (Piercy & Morgan, 1994) argue that there is a lack of empirical evidence showing they improve firm performance even though they are de rigueur. Perhaps a more scathing critique of mission statements can be found in (Goett, 1997):

"Every last one of them (mission statements), extols Mom, apple pie, quality, and teamwork. Every last one of them is written in excruciatingly formal prose. And every last one of them, when reduced to essentials, simply states the obvious. What's really sad is that most of the newer mission statements are the products of the labors of some very smart executives...So a lot of firms packed their most senior people off to expensive retreats to prepare this vital document...And so they worked very hard and then came home from the very expensive retreat with a brief document suitable for calligraphy...And the document...got tacked up on the wall and promptly forgotten...The fact is, mission statements are rarely useful."

This paper is the first to attempt to categorize mission statements computationally. Human beings are wired to categorize data, however imperfectly the data might map to a specific category. Algorithms, on the other hand, make potentially far less subjective categorical assignments. This paper also provides a re-examination of mission statements as the AACSB accreditation process has placed more emphasis on linking mission to activities.

The remainder of this paper is organized as follows: The next section examines the literature in mission statement formation with a specific focus on changes in guidance and focus provided by AACSB over time. The data and methodology are then described followed by results then conclusions.

LITERATURE REVIEW

The formation of mission statements in higher education lagged the corporate trend by at least a decade. Now, nearly all universities have mission statements -- higher education accreditation agencies require them. While university missions are similar across Carnegie classifications, their emphasis tends to differ by institutional control. Private universities place more emphasis on liberal arts and diversity while public institutions emphasize serving the local area (Morphew & Hartley, 2006). The most common element in public university mission statements is providing services for a qualified workforce's education. Research is typically emphasized in the vision statement (Özdem, 2011). Mission statements from religious universities are more likely to explicitly express ethical values and moral character traits than those from secular universities. Graduates from religious universities are more likely to exhibit those attributes (Davis et al., 2007).

Possibly due to public pressure and growing diversity in board representation, some researchers have been more critical of these mission statements as being “amazingly vague, vapid, evasive, or rhetorical, lacking specificity or clear purpose…full of honorable verbiage signifying nothing” (Newsom & Hayes, 1991).

In the early 1990s, colleges and schools of business started formulating mission statements. The Association to Advance Collegiate Schools of Business (AACSB), an accrediting agency, moved to "mission-linked" standards in 1991, which were fully implemented in 1994 (Jantzen, 2000). The modification of standards was an attempt to increase the focus on strategic management and move away from prior ratio-based standards. AACSB added even more strategic emphasis to the 2003 and 2013 standards. There were additional revisions to the standards throughout, but 1991, 2003, and 2013 were the years that AACSB adopted a new set of standards. They had a phase-in period of three years for all but 2013.

The 1991 standards added a section titled "MISSION AND OBJECTIVES" that explained the expectations for missions and mission statements -- mainly resource allocation decisions (e.g., faculty priorities, educational objectives, intellectual and service priorities) be consonant with the school's mission. The 2003 standards are well known for implementing assurance of learning standards (Miles et al., 2004), but there was also significant expansion with the introduction of "Standard 1 -- MISSION STATEMENT." The mission statement became even more central to the accreditation process and added a strategic management plan requirement. Finally, the 2013 standards expanded Standard 1 to "MISSION, IMPACT and INNOVATION." Missions are now a statement or set of statements that describe the school, including the mission statement, vision statement, and statement of values. Standard 2 was also modified to suggest that intellectual contributions impact theory, practice, or teaching, which is consistent with the school's mission. There was, however, decreased prescriptive language regarding the mission statement. This summary of changes from the 1991 standards, through the 2003 standards, and the 2013 standards is far from comprehensive. The emphasis is that since the introduction of mission-linked standards, AACSB has consistently strengthened the linkage of the mission to all of a school's activities.

While the linkage between mission and activities have strengthened, there is decreased prescriptive language regarding the mission statement (Jantzen, 2000). The term "mission statement" is mentioned 88 times in the 2003 standards, three times in the 2013 standards, and only once in the 2020 standards. The focus, clearly, is moving from mission formation to mission implementation. (Palmer & Short, 2008) conducted a comprehensive review of AACSB accredited Schools and Colleges of Business in the United States. Using a typology for analyzing the content of mission statements by manual encoding (Pearce & David, 1987), they categorized mission statement components, found variance between the statements for U.S. colleges of business, and linked them to performance-related attributes of schools. Since then, several analytic tools have made the analysis of text more robust, allowing us to further examine the differences and similarities in mission statement lexical patterns while attempting to link them to performance and resource-related attributes of their respective colleges.

DATA AND METHODOLOGY

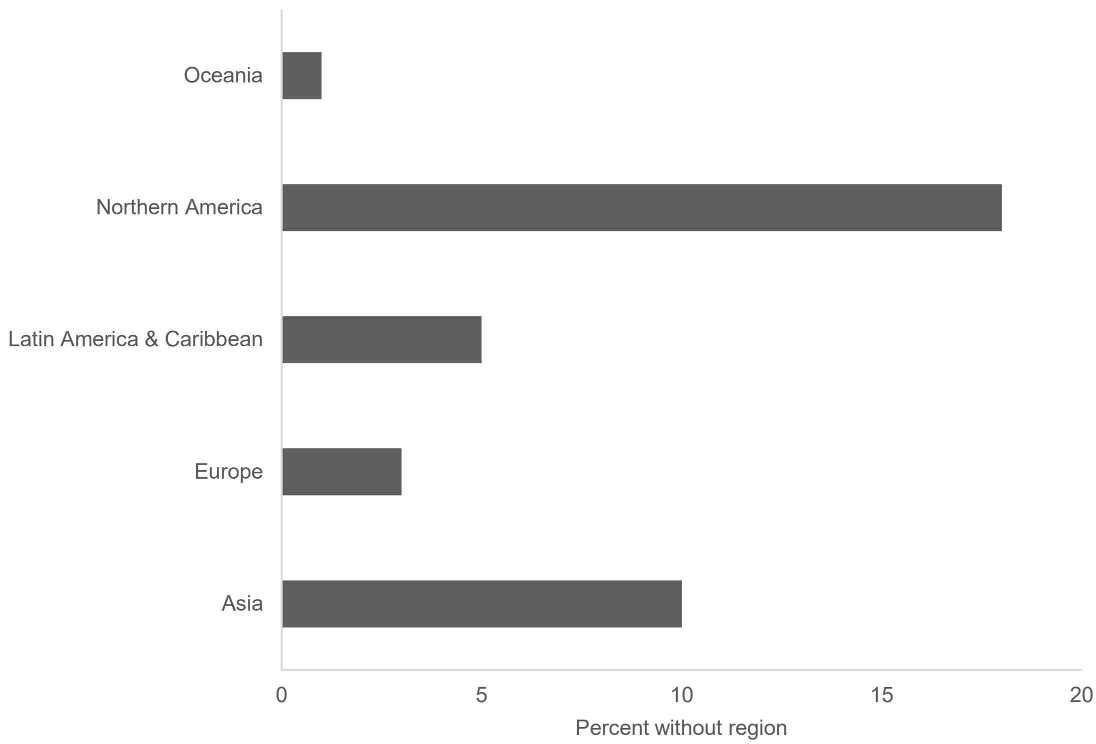
Mission statements were collected from the web over the 2018 and 2019 academic years for AACSB accredited institutions. First, I examine patterns in mission statements by region of the country. Second, I see if recently accredited schools use more common phrases in their mission statements and whether they employ language that is peculiar to the AACSB standards. For example, if you Google the term "intellectual contribution," the first five hits will be AACSB accredited business schools. Finally, I use Latent Dirichlet Analysis (LDA) to create clusters of U.S. schools based on mission phrases and see if there are any correlations between these clusters and resource-based characteristics of the institutions.

For each AACSB-accredited institution, an attempt was made to locate a mission statement by navigating the business college, school, or department website. The statement was often found in a section labeled with a strategic title (e.g., mission/vision/values). If this was unsuccessful, a Google search specific to that site was conducted. If also unsuccessful, a non-site specific Google search was conducted along with examining the institution's graduate and undergraduate catalogs or bulletins. Mission statements were found for 752 out of 789 accredited institutions. This 95.3% sample of school mission statements is consistent with the 95.1% seen in (Palmer and Short 2008), although the prior study examined only accredited schools in the United States and used the AACSB website as the sole source of mission statements during a time when AACSB published school mission statements.

Many business schools have a section on their website dedicated to strategy, often listing a vision, mission, and core values. Sometimes those missions specifically delineate a mission statement while others are less clear. If a school has a multi-paragraph mission without a clear indication of mission statement definition vs. a discussion of how a school might accomplish its mission, a subjective determination was made to delineate the mission statement. Statements used to trace these mission components from the actual statement include self-referencing words like "in order to accomplish this mission."

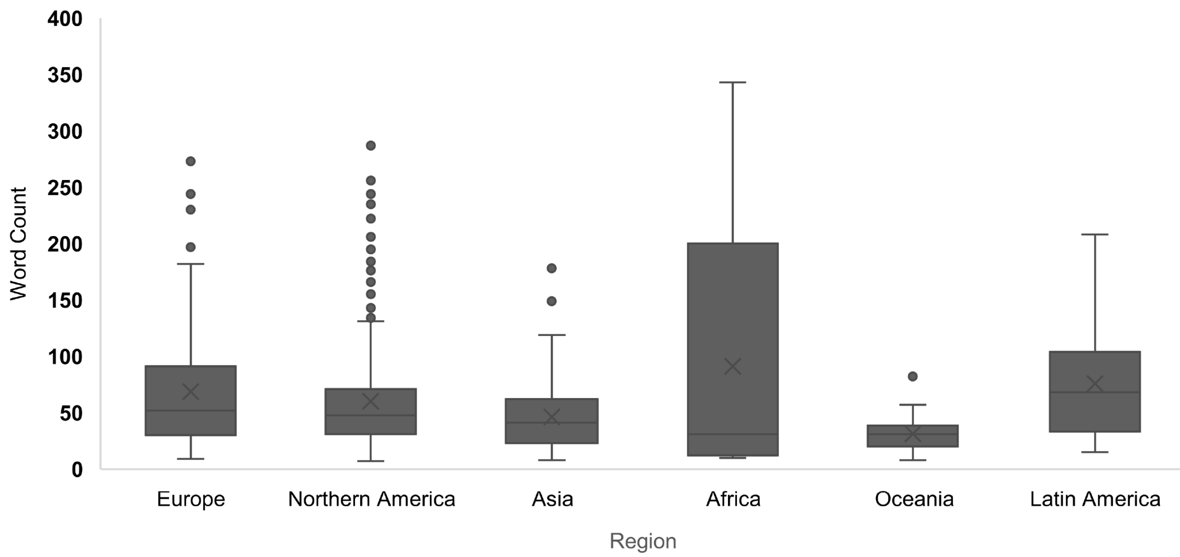
Upon further examination, of the 37 schools where a mission statement could not be found, Figure 1 shows a higher proportion in Latin America and a somewhat higher proportion in Asia. I suspect that language may be a contributing factor as the search for mission statements was done in English, and there is additional complexity in managing multilingual websites.

Figure 1: No Mission Statement Found by Region



We can observe lexical patterns in international mission statements, including the distribution of word counts, common and unique n-grams, and institutional clusters. Still, once we attempt to link them to performance and resources, we must limit the dataset to U.S. based business schools for comparative purposes. Since the dataset is nearly the entire population of AACSB schools, for proportional comparisons there is no need for statistical tests since we are examining differences in population groups. Figure 2 shows the distribution of word counts by region. The North America region has the most accredited institutions and, subsequently, the widest distribution of word counts although we see a majority of all mission statements have fewer than one hundred words.

Figure 2: Mission Word Counts by Region



An n-gram is a contiguous sequence of words from a given corpus of text – in this case, mission statements. The Latin numerical prefix refers to the size of the n-gram, so a unigram is a single word, a bigram is two words in sequence followed by trigrams, tetragrams, etc. Stop words are words that are so common that they provide minimal value to any lexical analysis and are filtered out before any processing. Typically, these are standard articles and prepositions like "a," "the," "and," etc. For this analysis, common stop words were filtered out of the corpus along with a stop word custom lexicon containing the words mission, business, school, college, university.

The analysis ended with trigrams since groupings of tetragrams only applied to a small number of business schools (i.e., five or less). Table 1 shows the top 10 uni, bi, and tri-grams by count. Terms from the top 10 unigrams were further refined by creating a custom dictionary that combined multiple words that impart similar meanings. For example, “education” captures the words “educate,” “education,” “educating,” etc. In some cases, the combinations went beyond word stems (e.g., global and international). We can see that terms incorporating teaching, research, and, to a lesser extent, service are fairly popular since they are also prevalent in the AACSB standards.

Table 2 shows the top ten unigrams by region. We can see that most of these terms, except for “teaching,” are relatively common across regions. Students and community tend to be more prevalent in North American mission statements, management in African mission statements, and research in Oceanic and European mission statements. Unigrams that are somewhat common to regions (i.e., is in greater than 10% of a region’s mission statements) but rarely observed in other region’s mission statements include: “sustainable” in Europe and Latin America; “Christian” in Africa; “cultivate” in Asia, “service” in North America and Latin America.

Table 1: Top 10 N-grams in Business School Mission Statements

Unigram	Count	Bigram	Count	Trigram	Count
students	512	teaching research	52	teaching research service	21
research	390	undergraduate graduate	52	excellence teaching research	13
education	310	prepare students	52	diverse student body	13
global	306	global environment	44	undergraduate graduate education	13
knowledge	291	quality education	40	provide quality education	12
community	281	learning environment	39	undergraduate graduate programs	12
leaders	271	economic development	39	undergraduate graduate students	11
learning	246	experiential learning	38	students successful careers	10
management	211	theory practice	37	students diverse backgrounds	10
teaching	180	research service	31	diverse student population	9

Table 1 shows the count of unigrams, bigrams, and trigrams in business school mission statements.

Table 2: Proportion of Mission Statements Containing Common Unigrams by Region

Word	Africa	Asia	Europe	Latin America	Northern America	Oceania
community	20%	21%	16%	18%	42%	27%
education	60%	38%	44%	29%	53%	50%
global	40%	46%	52%	35%	46%	41%
knowledge	20%	42%	43%	35%	28%	14%
leaders	60%	52%	28%	53%	42%	27%
learning	20%	16%	22%	12%	36%	14%
management	60%	39%	42%	24%	16%	14%
research	40%	40%	56%	35%	35%	68%
students	20%	27%	35%	24%	59%	14%
teaching	*0%*	9%	27%	24%	23%	23%

Table 2 shows the common unigrams found in mission statements by the designated AACSB regions.

For the remainder of the unigram analyses, we will restrict the data to U.S.-based colleges for consistent demographic comparisons. Table 3 shows us that institutional control differences are somewhat more minor, with private schools emphasizing leaders more while public schools place more emphasis on research. Public colleges were much more likely to use the words “economic” and “region” than their private counterparts.

Table 3: Proportion of U.S. Mission Statements Containing Common Unigrams by Institutional Control

Term	Private	Public
community	0.36	0.46
education	0.54	0.53
global	0.46	0.45
knowledge	0.24	0.29
leaders	0.6	0.34
learning	0.36	0.36
management	0.21	0.13
research	0.18	0.41
students	0.52	0.65

Table 3 shows the common unigrams found in mission statements by institutional control for US accredited institutions.

There are ten distinct Carnegie Classifications given to universities. Table 4 shows counts for US colleges along with the abbreviations for classifications. There are very few (39) undergraduate-only accredited institutions.

Table 4: Count of U.S. Mission Statements Containing Common Unigrams by Carnegie Classification

Code	Carnegie Classification	Count
BCAS	Baccalaureate Colleges--Arts & Sciences	17
BCDF	Baccalaureate Colleges--Diverse Fields	20
BAC	Baccalaureate/Associate's Colleges	2
DRU	Doctoral/Research Universities	48
MCL	Master's Colleges and Universities (larger programs)	173
MCM	Master's Colleges and Universities (medium programs)	61
MCS	Master's Colleges and Universities (smaller programs)	18
RUH	Research Universities (high research activity)	93
RUVH	Research Universities (very high research activity)	88
SBM	Schools of business and management	4

Table 4 shows the abbreviated code for each Carnegie classification along with a count of accredited US institutions in that category. This table can also be used as a reference for tables 5 and 9.

Table 5 shows common unigrams by most of the Carnegie Classifications listed in Table 4. Baccalaureate/Associates colleges and Schools of business and management are not included as there are so few of them. We can see that research universities are less likely to have “students” in their mission statements and only somewhat more likely to have “research” in their mission statements. More prevalent words include “alumni” for research universities; “prepares” for non-baccalaureate institutions; “diversity” for smaller master’s programs; “accessible,” “communication,” “curriculum,” “professions,” and “department” for Arts & Sciences; and “pedagogical” for diverse field baccalaureate colleges.

Table 5: Proportion of U.S. Mission Statements Containing Common Unigrams by Carnegie Classification

Term	BCAS	BCDF	DRU	MCL	MCM	MCS	RUH	RUVH
community	0.53	0.15	0.42	0.49	0.41	0.44	0.45	0.36
education	0.71	0.6	0.54	0.56	0.67	0.33	0.47	0.43
global	0.24	0.35	0.44	0.52	0.44	0.56	0.53	0.33
knowledge	0.29	0.1	0.19	0.25	0.21	0.11	0.41	0.36
leaders	0.35	0.35	0.38	0.41	0.41	0.39	0.37	0.58
learning	0.47	0.45	0.35	0.45	0.38	0.44	0.31	0.18
management	0.18	0.05	0.15	0.12	0.1	0.06	0.17	0.27
research	0.24	0.35	0.46	0.32	0.23	0.22	0.45	0.34
students	0.82	0.7	0.58	0.69	0.72	0.67	0.53	0.39
teaching	0.35	0.15	0.31	0.24	0.25	0.33	0.24	0.12

Table 5 shows common unigrams by Carnegie classification code (shown in Table 4). BCAS = Baccalaureate Colleges - Arts & Sciences, BCDF = Baccalaureate Colleges - Diverse Fields, DRU = Doctoral/Research Universities, MCL = Master's Colleges and Universities (larger programs), MCM = Master’s Colleges and Universities (medium programs), MCS = Master’s Colleges and Universities (smaller programs), RUH = Research Universities (high research activity), RUVH = Research Universities (very high research activity).

The AACSB accredited its first school outside of North America (ESSEC) in 1997 and started expanding internationally in earnest in 1998. Table 6 shows us common unigrams for pre and post expansion accredited US universities. We can see substantial differences in the proportions for the more common mission statement terms. Other terms that are more prevalent in pre-1998 schools include: academic; create; practice; world.

Table 6: Proportion of U.S. Mission Statements Containing Common Unigrams by Initial Accreditation Year

Term	Post 1998	Pre 1998
community	0.18	0.43
education	0.25	0.5
global	0.2	0.45
knowledge	0.1	0.29
leaders	0.19	0.42
learning	0.17	0.33
management	0.06	0.15
research	0.13	0.36
students	0.3	0.54
teaching	0.11	0.22

Table 6 shows common unigrams by initial accreditation year.

Analyzing bigrams is a bit more complicated since there are far fewer common bigrams. Table 7 shows us common bigrams by region. The only bigrams that are seen in more than 10% of a region’s mission statements are: “economic development” in Africa and “teaching research” in Oceania. Since stopwords and punctuation are removed, bigrams like "teaching research" may appear as "teaching, research," "teaching and research," etc.

Table 7: Proportion of Mission Statements Containing Common Bigrams by Region

Term	Africa	Asia	Europe	Latin America & Caribbean	Northern America	Oceania
economic development	0.2	0.04	0.03	0	0.05	0
global environment	0	0.05	0.04	0	0.06	0
learning environment	0	0.06	0.04	0	0.05	0
prepare students	0	0.02	0.04	0	0.08	0
quality education	0	0.06	0.03	0	0.07	0.05
socially responsible	0	0.05	0.07	0	0.04	0
teaching research	0	0.03	0.07	0.06	0.07	0.14
theory practice	0	0.06	0.02	0	0.05	0
undergraduate graduate	0	0.03	0.02	0.06	0.08	0
experiential learning	0	0	0	0	0.07	0

Table 7 shows the common bigrams found in mission statements by the designated AACSB regions.

For the remainder of our bigram analysis, we will, once again, restrict our data to the US. Table 8 shows common bigrams by institutional control. We see that "socially responsible" is much more common in private institutions, while "undergraduate, graduate" is more common in public schools.

Table 8: Proportion of U.S. Mission Statements Containing Common Bigrams by Institutional Control

Term	Private	Public
experiential learning	0.04	0.09
global environment	0.05	0.07
learning environment	0.05	0.06
prepare students	0.06	0.09
quality education	0.04	0.09
socially responsible	0.09	0.01
teaching research	0.04	0.08
theory practice	0.06	0.05
undergraduate graduate	0.04	0.11
economic development	0	0.08

Table 8 shows the common bigrams found in mission statements by institutional control for US accredited institutions.

Table 9 shows us bigrams by Carnegie classification. "Experiential learning" is more common in non-research universities, while "undergraduate, graduate" is nonexistent at very high research universities.

Table 9: Proportion of U.S. Mission Statements Containing Common Bigrams by Carnegie Classification

Term	BCAS	BCDF	DRU	MCL	MCM	MCS	RUH	RUVH
experiential learning	0.12	0.05	0.06	0.09	0.08	0.17	0.05	0.03
learning environment	0.12	0	0.15	0.05	0.07	0.17	0.03	0.01
prepare students	0.06	0.05	0.1	0.1	0.03	0.11	0.1	0.05
quality education	0.06	0	0.08	0.08	0.11	0	0.08	0.06
undergraduate graduate	0.12	0.15	0.1	0.11	0.11	0.11	0.09	0
economic development	0	0.05	0.06	0.06	0.03	0.06	0.08	0.06
teaching research	0	0.1	0.08	0.06	0.07	0.11	0.1	0.05
global environment	0	0	0.04	0.09	0.05	0.17	0.09	0.03
socially responsible	0	0	0.06	0.06	0.03	0.06	0.01	0.01
theory practice	0	0	0.04	0.06	0.08	0.17	0.06	0.01

Table 9 shows common bigrams by Carnegie classification code (shown in Table 4). BCAS = Baccalaureate Colleges - Arts & Sciences, BCDF = Baccalaureate Colleges - Diverse Fields, DRU = Doctoral/Research Universities, MCL = Master's Colleges and Universities (larger programs), MCM = Master's Colleges and Universities (medium programs), MCS = Master's Colleges and Universities (smaller programs), RUH = Research Universities (high research activity), RUVH = Research Universities (very high research activity)

Table 10 shows us common bigrams for pre and post 1998 initially accredited US universities. Bigram proportions are all somewhat low when categorizing U.S. schools by initial accreditation year. We do see all of the top bigrams more prevalent in schools accredited before 1998.

Table 10: Proportion of U.S. Mission Statements Containing Common Bigrams by Initial Accreditation Year

Term	Post 1998	Pre 1998
economic development	0.02	0.07
experiential learning	0.04	0.06
global environment	0.03	0.06
learning environment	0.02	0.06
prepare students	0.04	0.07
quality education	0.03	0.08
socially responsible	0.01	0.04
teaching research	0.02	0.08
theory practice	0.02	0.06
undergraduate graduate	0.05	0.07

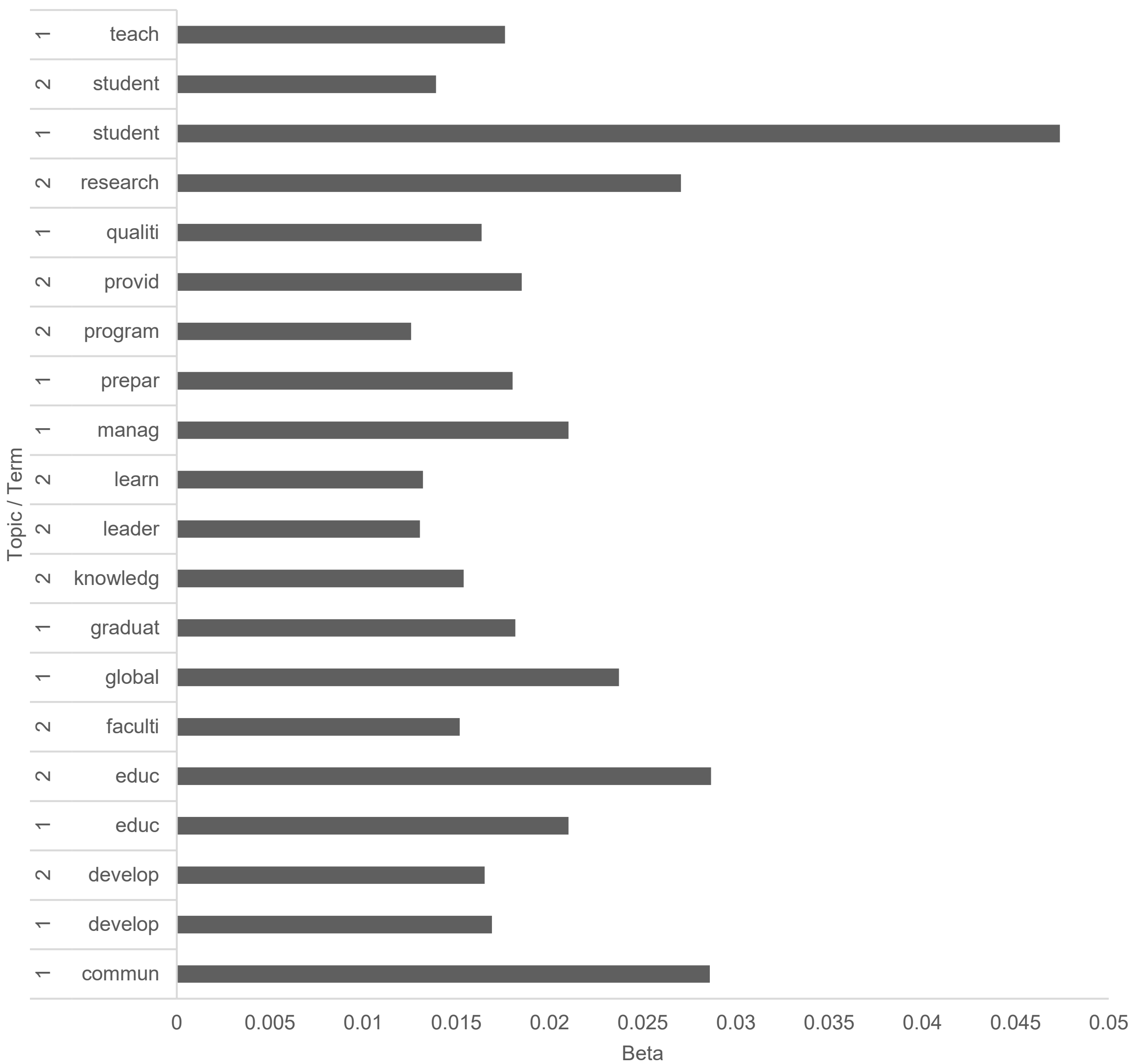
Table 10 shows common bigrams by initial accreditation year.

RESULTS

Latent Dirichlet allocation (LDA) is a popular machine learning algorithm used in topic modeling. In this case, we consider every mission a potential combination of topics, and every topic a combination of words. Topics are defined by a probability distribution of words and the same word can be used in multiple topics. A document, or in this case, a mission statement can be defined by a probability distribution of topics. For a particular topic, the per-word probability is called β("beta"), and the per-document topic probability is called γ ("gamma"). The number of topics that LDA attempts to fit must be defined in advance. I will cover a simple two-topic model and then discuss n-topic models.

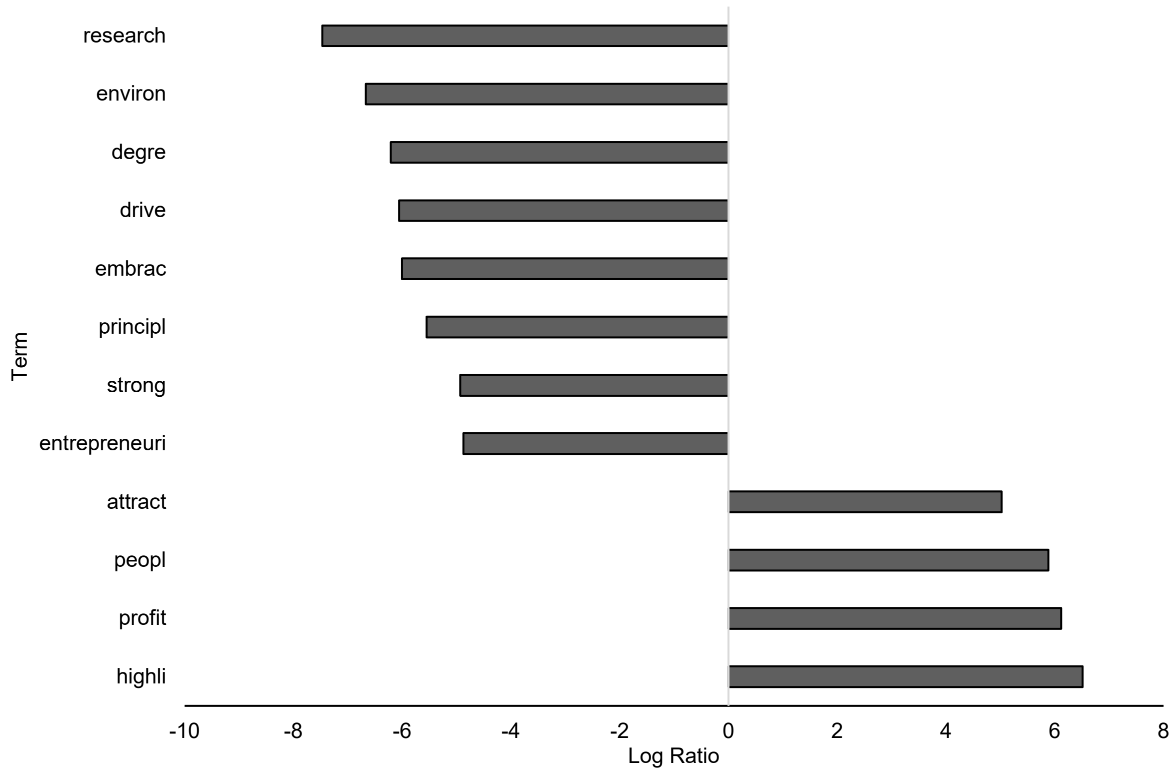
Figure 3 shows the ten largest β values for a two-topic model. These are the most common or densest words, or word stems, for each of the two topics. In LDA, we often see words common to multiple topics. For example, the word "student" is prevalent in both topics but more so in topic one ($\beta_1 = 0.047$, $\beta_2 = 0.014$). The word "teach" appears in the top ten words for topic 1 ($\beta_1 = 0.018$). While "teach" is not in the top ten list for topic two, it does exist at a significantly lower density ($\beta_2 = 0.001$)

Figure 3: Top Ten Betas – Two Topic Model



It is often helpful to look at the largest differences in beta between topics. Figure 4 shows us a beta spread chart, which contains the log ratio $log_2\left(\frac{\beta_2}{\beta_1}\right)$. Reporting these differences as log ratios is valuable as it makes the difference symmetrical. A log ratio of 1 means β_2 is twice as large as β_1 while a log ratio of -1 means β_1 is twice as large as β_2 (Silge & Robinson, 2017). In this two-topic model, one might look at the words comprising the topic and say that topic 1, which places more emphasis on words like student and teach might be more applicable to institutions that have a teaching focus or private institutions while the more prevalent words in topic 2, including research and "knowledg" might be more applicable to research institutions.

Figure 4: Top Beta Spreads – Two Topic Model



Examining the words and their corresponding probabilities allows us to attribute descriptive names to topics and potentially test for correlations between topics and other variables. If we were performing LDA on a book with multiple chapters, we would call the book a "corpus" and each chapter a "document" and examine the probability that a topic belongs to a chapter by discussing the per-document topic probability, or γ. In this study, our corpus is a collection of mission statements and our documents can conceivably be any attribute that groups institutions, including those examined earlier (i.e., region, institutional control, Carnegie classification, and initial accreditation year). γ that are similar to the inverse of the number of categories of the comparison variable are indicative of a poor mapping. In two topic model, a γ near 0.5 would be considered a poor mapping.

To see if our two-topic model potentially mapped to institution control, we could examine a simple boxplot of the γ values for each topic by institutional control, as shown in Figure 5. We conclude that our two topics aren't indicative of institutional control. In this case, the number of topics, two, matches the number of categories in the institution control variable (i.e., public/private). This need not be the case and, there is no requirement that the number of topics match the number of variables.

Figure 5: Gamma Values by Institutional Control – Two Topic Model

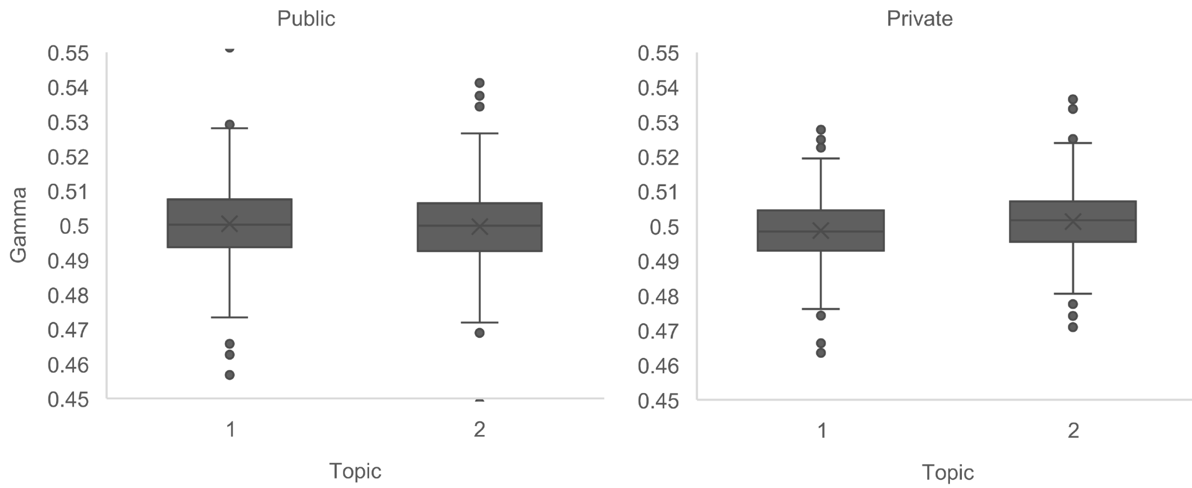
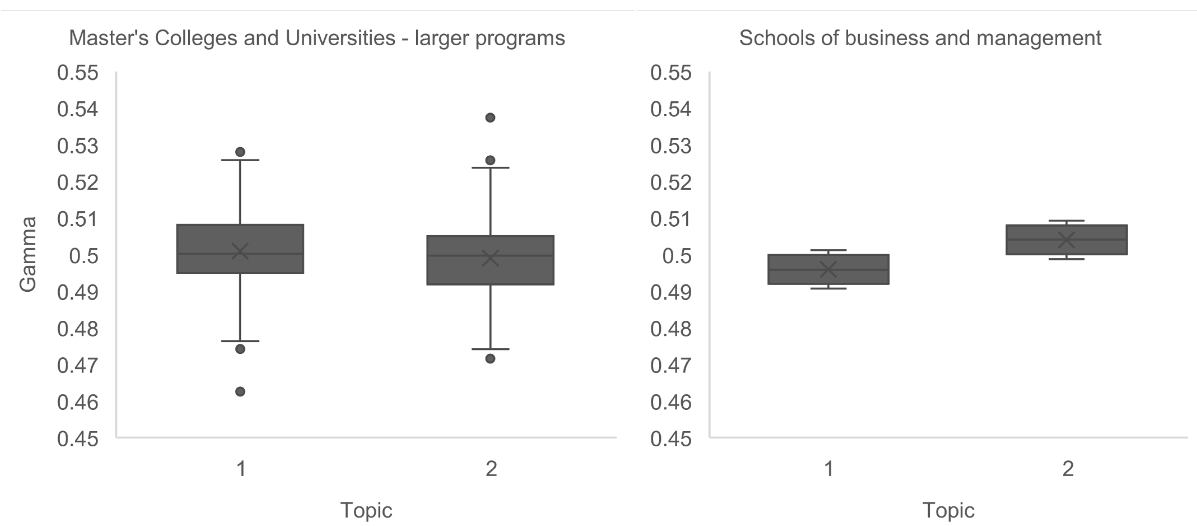


Figure 6 shows our two-topic model by select Carnegie Classification. While this mapping is slightly better for less represented schools, it is still poor. We only see slight differentiation in categories with very few observations. There are 173 AACSB accredited "Master's Colleges and Universities – larger programs," while there are only four with the Carnegie Classification "Schools of business and management."

Figure 6: Gamma Values by Select Carnegie Classification – Two Topic Model



CONCLUDING COMMENTS

This paper is the first to attempt to categorize mission statements computationally using n-gram analysis and Latent Dirichlet Allocation (LDA) on the corpus of known mission statements for AACSB accredited institutions. While we have shown certain words and phrases are prevalent in specific types of institutions, after iterating through topic models ranging from two to ten topics two the institutional variables region, institutional control, Carnegie classification, and initial accreditation year, there were no cases where topics mapped well to the variables. This leads me to conclude that while there are some distinctive lexical patterns in mission statements, topic modeling – specifically LDA, does not allow us to group topics to variables in a way that differentiates institutional characteristics.

Prior studies that have had success grouping mission statements by institutional categories had processes that were not automated and perhaps somewhat subjective. The only subjective component to LDA, and other unsupervised learning models, is choosing the number of topics. There are no human-driven judgment calls that are common in more subjective forms of classification.

The process of writing mission, vision, and values statements at AACSB accredited schools is informed by other accredited schools, peer institutions, and AACSB-run seminars and conferences. Some level of conformity in mission statements is only natural. Evidence exists that some institutions felt compelled to emphasize research even if they were teaching-focused (Stepanovich et al., 2014). One can assume that as an institution emphasizes something, even if due to external pressures, it is more likely to eventually be incorporated as a part of its mission. In a more pessimistic light, perhaps the Newsome and Hayes quote from earlier in this paper is applicable, and these mission statements might be "amazingly vague, vapid, evasive, or rhetorical, lacking specificity or clear purpose…full of honorable verbiage signifying nothing." Institutions and other interested parties can use this information to develop processes that mitigate the pull towards conformity in mission statements if they so desire. Accrediting agencies can also be of service here in developing methods that lead towards more diversity in mission statements.

Finally, a third possible explanation, and call for further research, is that LDA and other forms of topic modeling can yield good mappings between topics and institutional variables, just not for the variables examined. A limitation to this study is that characteristics like religious affiliation, third-party institutional rankings, specific location-based characteristics, and others were not examined and it is entirely possible that LDA or other topic modeling algorithms may provide better models using other institutional characteristics.

REFERENCES

Bart, C. K., & Tabone, J. C. (1998). Mission Statement Rationales and Organizational Alignment in the Not-for-Profit Health Care Sector. *Health Care Management Review*, *23*(4), 54–69.

Darbi, W. P. K. (2012). *Of mission and vision statements and their potential impact on employee behaviour and attitudes: The case of a public but profit-oriented tertiary institution*. *3*(14), 15.

Davis, J. H., Ruhe, J. A., Lee, M., & Rajadhyaksha, U. (2007). Mission Possible: Do School Mission Statements Work? *Journal of Business Ethics*, *70*(1), 99–110.

Goett, P. (1997). Mission impossible. *The Journal of Business Strategy*, *18*(1), 2. http://dx.doi.org.wv-o-ursus-proxy02.ursus.maine.edu/10.1108/eb039820

Jantzen, R. H. (2000). AACSB Mission-Linked Standards: Effects on the Accreditation Process. *Journal of Education for Business*, *75*(6), 343–347. https://doi.org/10.1080/08832320009599038

King, D. L., Case, C. J., & Premo, K. M. (2012). An International Mission Statement Comparsion: United States, France, Germany, Japan, and China. *Academy of Strategic Management Journal*, *11*(2), 93–119.

Miles, M. P., Hazeldine, M. F., & Munilla, L. S. (2004). The 2003 AACSB Accreditation Standards and Implications for Business Faculty: A Short Note. *Journal of Education for Business*, *80*(1), 29–34. https://doi.org/10.3200/JOEB.80.1.29-34

Morphew, C. C., & Hartley, M. (2006). Mission Statements: A Thematic Analysis of Rhetoric across Institutional Type. *The Journal of Higher Education*, *77*(3), 456–471. https://doi.org/10.1080/00221546.2006.11778934

Mullane, J. V. (2002). The mission statement is a strategic tool: When used properly. *Management Decision*, *40*(5), 448–455. https://doi.org/10.1108/00251740210430461

Newsom, W., & Hayes, C. R. (1991). Are Mission Statements Worthwhile? *Planning for Higher Education*, *19*(2), 28–30.

Özdem, G. (2011). An Analysis of the Mission and Vision Statements on the Strategic Plans of Higher Education Institutions. *Educational Sciences: Theory & Practice*, *11*(4), 1887–1894.

Palmer, T. B., & Short, J. C. (2008). Mission statements in US colleges of business: An empirical examination of their content with linkages to configurations and performance. *Academy of Management Learning & Education*, *7*(4), 454–470.

Pearce, J. A. (1982). The Company Mission As a Strategic Tool: What Is a Company Mission? *Sloan Management Review (Pre-1986)*, *23*(3), 15.

Pearce, J. A., & David, F. (1987). Corporate Mission Statements: The Bottom Line. *The Academy of Management Executive (1987-1989)*, *1*(2), 109–115.

Piercy, N. F., & Morgan, N. A. (1994). Mission Analysis: An Operational Approach. *Journal of General Management*, *19*(3), 1–19. https://doi.org/10.1177/030630709401900301

Powers, E. L. (2012). Organizational Mission Statement Guidelines Revisited. *International Journal of Management & Information Systems (Online)*, *16*(4), 281.

Silge, J., & Robinson, D. (2017). *Text Mining with R: A Tidy Approach*. O'Reilly Media, Inc.

Stepanovich, P., Mueller, J., & Benson, D. (2014). AACSB Accreditation and Possible Unintended Consequences: A Deming View. *Journal of Education for Business*, *89*(2), 103–109. https://doi.org/10.1080/08832323.2013.763754

Taiwo, A. A., Lawal, F. A., & Agwu, P. E. (2016). Vision and Mission in Organization: Myth or Heuristic Device? *The International Journal of Business & Management*, *4*(3), 127–134.

BIOGRAPHY

Dr. James Suleiman is an Associate Professor of Business Analytics and Information Systems at University of Southern Maine. He can be contacted at School of Business, Portland, ME 04104.

Business Education & Accreditation
Vol. 14, No. 1, 2022, pp. 33-42
ISSN: 1944-5903 (print)
ISSN: 2157-0809 (online)

www.theIBFR.COM

EXAMINING IF GRADING BIAS EXISTS IN A PROFESSOR OF BUSINESS COURSES: AN INDIVIDUAL STUDY

Gary F. Keller, Eastern Oregon University, LaGrande, OR

ABSTRACT

This research paper investigated if grading bias based on gender, course modality (face-to-face and online) and grade level (undergraduate vs. graduate) existed in classes the researcher taught over a three-year period. During the course of the literature review for this investigation, a meta-analysis suggested that more research was needed on a micro level regarding bias rather than reliance on results from metanalysis reports. A second goal was the extension of two studies conducted by the researcher that examined if relationships existed between the amount of time students spent on course assignments and final grades and if the quantity of time a faculty member spent grading student assignments affected final grades given. In this study grades assigned during the academic years 2018 – 2020 derived from the University's Canvas Learning Management System were analyzed. The population studied consisted of 912 students enrolled in undergraduate and graduate courses taught at a public university located in the Pacific Northwest. The main results from this examination showed statistically significant differences were found in aggregate totals between the grades of males and females and the grades of females enrolled in online and face to face courses in two of the three years of the study.

JEL: I23, J16

KEYWORDS: Discrimination, Gender Stereotypes, Natural Experiment, Biased Grading

INTRODUCTION

The subject of grading bias, raises a myriad of questions as grades have a significant bearing not only on students' self-image and future employment opportunities but also on faculty members' promotions and institutional reputations. Ascertaining the causes of grading biases at a university level has been the subject of numerous, meta-analysis studies (primarily European institutions) often attempting to support/challenge the validity of anonymous grading schemes. However, a gap in the scholarly literature was discovered. Scant attention has been paid to studying individual faculty members grading tendencies rather than on an institutional level leading one researcher (Keyser, 2018) to recommend grading bias research should be conducted by individual faculty members. The chief goal of this exploratory research project was to contribute new research to the subject domain by examining courses taught by this faculty member during the academic years 2018 – 2020 to determine, to what extent if any, grading bias existed. Secondary research goals were: a) to learn if there were statistically significant differences between the final course grades for students enrolled in online and face to face courses; b) to inquire if differences exist between the final course grades of males and females and; c) to find out if disparities exist between the final course grades of undergraduate and graduate students. A review of the scholarly literature regarding grading bias is presented, an explanation of the methodology used in this research project is described, the results of statistical tests conducted to determine if grading bias existed in courses taught by the researcher are shown and recommendations for future grading bias research on the part of individual faculty members, institutions and accrediting agencies are made.

Numerous definitions exist for bias ranging from conscious to unconscious partiality. Research bias is defined as "any tendency which prevents unprejudiced consideration of a question. In research, bias occurs when "systematic error [is] introduced into sampling or testing by selecting or encouraging one outcome or answer over others" (Pannucci & Wilkins, 2011, para. 3). In grading students work, bias may be considered "as a technical term most often refers to a characteristic of tests that present advantage or disadvantage to a particular subgroup (e.g., by gender or ethnicity) (Nitko, 2004; Popham, 2005 as cited in Hardré, 2014, p. 1). In common usage, bias is "a particular tendency, trend, inclination, feeling, or opinion, especially one that is preconceived or unreasoned" (Dictionary.com, 2021). It is assumed that faculty members act with personal and professional integrity. Institutions of higher learning generally do not require faculty members to rigorously explore their grading trends other than grade distribution reports contained in promotion portfolios (which merely exhibit how many As, Bs, Cs, etc. over a period of time are assigned). Adjunct faculty members, who account for an increasing percentage of university faculty, are rarely, if ever, included in grade bias studies. While many grading tools, such as rubrics embedded in learning management systems, are available and helpful, they merely assist how one can justify assignment of grades but do not enable a faculty member to locate if grading biases exist. There appears to be few if any organizational efforts, requirements or positive individual incentives for faculty members to rigorously examine if grading biases exists in their courses. Student grade appeals and complaints, which occur on an ad hoc basis, are the generally accepted process to at least raise the question of grading unfairness. However, the grade appeal is a singular, after the fact process to raise the question of a faculty member's grading predispositions rather than a proactive, ongoing, quantitative improvement method.

Exploring grading bias and its impact on faculty and student performance is the focus of this study. It evolved from an exploration of the topic which indicated the existence of large-scale bias research projects and meta-analysis but a scant number of investigations conducted by individual faculty members. A second motivation was from two previous investigations the author conducted that sought to determine if a relationship existed between the amount of time students and this faculty member spent on the institution's LMS and final course grades. Part of the data analysis included in this study comparing face-to-face students with online students, undergraduate and graduates and male/female contrasts. One variable of this author's previous two studies that was not explored was if biases may have influenced the way this professor allocated student grades. The same subject groupings from the previous two studies were used in this research project to ascertain if grading bias may exist in courses taught by this faculty member over a three-year time period. A review of the scholarly literature encompassing the domains of existing studies regarding grading bias on a university level, the role that gender may play in grading bias and techniques that sought to decrease/eliminate grading bias are discussed followed by a description of the methodology used in this study, the results of the data analysis and recommendations for future research are offered.

LITERATURE REVIEW

Three scholarly spheres were examined in this exploratory research topic. The first domain involved locating studies that examined grading bias on a university level. A consistent theme in the literature was large-scale studies and many were European based. Malouff and Thorsteinsson (2016) conducted a meta-analysis of 23 analyses of 20 experimental research studies on populations ranging from grade school children to university students. The authors remarked that "bias can occur in subjective grading when graders are aware of irrelevant information about the students" (para.1). One important limitation noted by Malouff and Thorsteinsson was meta-analysis studies "did not examine grading of student work by the actual teachers of the students" (2016, para. 25). Malouff, Emmerton and Schutte (2013) explored the *halo effect* on grading. A cross disciplinary group of 159 professors and teaching assistants was tasked to first grade oral presentations and then written presentations of the same students. Results of this research indicated that the faculty/TAs gave significantly higher scores to written work following the better oral presentation than following a poor demonstration displaying a halo effect. Malouff, et al. recommended that keeping students anonymous helps prevent bias in grading.

Another massive study by Hinton and Higson (2017) involved a study of more than 30,000 university students over a 12-year period to primarily ascertain if grading assignments namelessly would reduce performance differences between a variety of student classifications (ex. gender, ethnicity, social status, etc.). While Hinton and Higson concluded that anonymous marking had a negligible effect in reducing performance disparities, they urged future researchers to understand performance differences in terms of the *mechanisms* by which performance differences manifest at the group level…and it is far from easy in practical terms to obtain these kinds of data in large quantities" (para. 60). Bygen (2019) reviewed data from five undergraduate courses at Stockholm University, from 2005 and 2013 to see if examination results were biased based on the *foreignness* of one's name and gender. Results showed that grades were not biased based on these two factors. The above studies tended to be centered on utilizing large data sets to assess if grading bias existed. A study by Keyser (2018) took a different approach. The researcher centered his research on determining if course grades differed between on-campus and distance (online) students taking the same courses over a one-year period; other elements such as gender, age, etc. were also examined. Two significant differences, class standing and GPA, accounted for disparities in grades. Keyser's study was an interesting approach because it examined a single professor seeking to determine if grades may have been influenced by the medium of instruction. Additionally, Keyser recommended that professors throughout the university where he teaches should conduct studies similar to his for comparative evaluation and later expand that research to additional universities.

The second domain reviewed was the role that gender may play in grading bias. Exploring the part gender may play in university admissions, academic and social life is an important issue as American institutions continue the long effort to improve equity and inclusion in university life. Two key factors to enable fairness and inclusiveness to succeed is creating an environment that provides a transparent and reliable admission process but also supports students' educational pathways once they are admitted. Sometimes the best intentions obscure rather than clarify high minded admission goals. One study (Breda & Ly, 2015) revealed that in entrance exams of a French higher education institution, the appraisal favored females in traditionally male-dominated subjects (e.g., math, philosophy) and the reverse in customarily female-dominated areas (e.g., literature, biology). A micro study (n = 12 students) was conducted to investigate if student gender influenced feedback. The context for this study was to test the validity of those who purport that anonymity is the optimal solution to grading bias. The results of the study indicated that gender had little to no effect and thus contested the merits of the anonymity advocates.

Jansson and Tyrefors (2018) affirmed that there are few studies investigating grading bias at the university level. In their study (conducted in Sweden) evidence was found that "TAs correcting exams at the university favor students of their own gender. However, the size of the in-group bias was only approximately 20 % of the total effect. Interestingly, both the in-group bias and the general bias disappear when exams are graded anonymously" (p. 21). Krawczyk (2018) reported the results of a study on grades awarded for bachelor and master theses at a large Polish university. The research's purpose was to detect gender or physical attractiveness bias. Approximately 15,000 students were included in the study. The conclusions noted that some evidence existed to indicate that females received relatively high grades from advisors and no evidence of influence of physical attractiveness. Additionally, Krawczk noted "gender seems to play some role, with male students getting relatively higher grades from referees and females from advisors. This is consistent with the hypothesis of males being perceived as more competent but less likeable, and may thus be a manifestation of a bias… It does suggest that grading of term papers, which are only read by a single evaluator who knows the student personally, may be biased against male students" (p. 158). The third scholarly literature field examined focused on techniques that sought to decrease/eliminate grading bias. The French, Swedish and Polish studies in particular raise a healthy question; do universities regularly examine grade distributions based on gender, composition of the student body, modality (i.e., face-to-face or online), and course level (freshman – senior, undergraduate – graduate)? Other queries derived from the literature review include how many faculty members regularly examine grade distributions to provide insights into their grading patterns and practices and should regional and professional accreditation agencies require grade distribution analysis as part of their endorsement process? Finally, the literature review showed that queries into the grading bias question appeared to be singular events demonstrating the need for university leadership to institute regular, consistent and quantitative research into the issue. The literature review contributed to this research project to show the problem with analyzing aggregate data rather than carefully segmentation. The results from this study indicated that some grading bias based on gender, modality or grade level existed in courses taught by this faculty member from 2018-2020.

METHODOLOGY

The chief goal of this exploratory research project was to determine if grading bias existed in courses taught by this faculty member during academic years 2018 - 2020. Secondary research goals were: a) to determine if there were statistically significant differences between the final course grades for students enrolled in online and face to face courses; b) to what extent did differences exist between the final course grades of males and females; c) to what extent did differences exist between the final course grades of undergraduate and graduate students.

This investigation encompassed 51 College of Business courses (12 face-to-face - 24%, 39 - 76% online) taught by this researcher at a small, liberal arts, rural public university located in the Pacific Northwest from the start of the Winter 2018 term to the completion of the Fall 2020 term). The total population consisted of 912 students (430 males - 47%, 482 females - 53%). Of the 912 students 780 (86%) were undergraduates, 132 (13%) graduates; 236 (26%) participated in face-to-face courses; 676 (74%) were enrolled in online courses. 116 (88%) graduate students were enrolled in online courses and 16 (12%) choose face to face courses. It should be noted that enrollment in online courses dramatically enlarged beginning in the Spring term of 2019 due to the Covid-19 pandemic. The university essentially transferred all courses to an online format; however, in the Fall 2020 term, the institution reopened with a variety of restrictions. The pandemic's impact on this faculty member's teaching load did not have a material effect on the size of the research population.

Course assignments were generally composed of quizzes, discussions, essays and a final "capstone" assignment. Data for this research was derived from the University's Canvas Learning Management System (LMS). Individual and final course grades were recorded in the Canvas LMS and recovered by the researcher. The University utilizes a quarter course scheduling system; each term is composed of 10 weeks of instruction and one week allocated for final exams. A single factor ANOVA test was used to determine differences between subjects in this study. A student t-test was used to explore differences resulting from the ANOVA analysis. A significance value of 0.05 was used to determine whether to accept or reject the null hypothesis in each hypothesis analyzed (the use of ** was used to designate significance at the 0.05 level in tables). To analyze the data, a grade point average number was assigned to each letter grade in the data set. For example, 4.0 represented "A", 3.67 represented "A-", 3.33 represented "B+", 3.0 represented "B" and so on. Finally, the data was organized by gender for each course in order to examine the hypotheses. The main research questions for this exploratory research are listed in Table 1.

Table 1: Research Questions

RQ 1	To what extent if any are there statistically significant differences between the grades of male students and female students enrolled in courses taught by this faculty member between 2018 - 2020?
RQ2	To what extent if any are there statistically significant differences between the grades of male graduate students enrolled in face-to-face courses and male graduate students enrolled in online courses taught by this faculty member between 2018 – 2020?
RQ3	To what extent if any are there statistically significant differences between the grades of female graduate students enrolled in face-to-face courses and female graduate students enrolled in online courses taught by this faculty member between 2018 – 2020?
RQ4	To what extent if any are there statistically significant differences between the grades of male undergraduate face-to-face students and male online undergraduate college students in courses taught by this faculty member between 2018 - 2020?
RQ5	To what extent if any are there statistically significant differences between the grades of female undergraduate face-to-face students and female online undergraduate students in courses taught by this faculty member between 2018 – 2020?
RQ6	To what extent if any are there statistically significant differences between the grades of face-to-face undergraduate students and online undergraduate students in courses taught by this faculty member between 2018 - 2020?

Table 1 shows the main research questions that guided this study.

RESULTS AND DISCUSSION

A single factor ANOVA test was used to determine differences between the subjects in this study. A student t-test was used to explore differences if the ANOVA results were significant. A significance value of 0.05 was used to determine whether to accept or reject the null hypothesis in each hypothesis set.
RQ 1: To what extent if any are there statistically significant differences between the grades of male students and female students enrolled in courses taught by this faculty member between 2018 - 2020? (**N** = 912, 430 males, 482 females)

Hypothesis: H1o: There are no statistically significant differences between the grades of male and female students enrolled in courses taught by this faculty member between 2018 – 2020.

Result: As depicted in Table 2, statistically significant differences were found between the grades of male and female students enrolled in courses taught by this faculty member between 2018 – 2020.

Table 2: Grade Comparison – Male/Females 2018 – 2020

Hypothesis-Year	**T-Critical**	**T-Statistic**	**P-Value**	**Decision**
(Male vs. Female 2018 - 2020)	1.962	2.0467	0.0041	Reject Null

Statistically significant differences were found between the grades of male and female students enrolled in courses taught by this faculty member between 2018 – 2020. The chief reasons for the variances may be due to combining all types of courses (undergraduate/graduate), modalities (face to face and online, disparities between the number of males (430) and females (482) and effect of aggregating dissimilar groups that indicated grading bias on a macro level while the micro-outcomes indicate little if any grading bias.

RQ 2: To what extent if any are there statistically significant differences between the grades of male graduate students enrolled in face-to-face courses and male graduate students enrolled in online courses taught by this faculty member between 2018 – 2020? (N= 69, 11 Face to Face, 58 Online).

Hypothesis H2o: There are no statistically significant differences between the grades of male graduate students enrolled in face-to-face courses and male graduate students enrolled in online courses taught by this faculty member between 2018 – 2020.

Result: Data displayed in Table 3 shows that no statistically significant differences were found between the grades of male graduate students enrolled in face-to-face courses and male graduate students enrolled in online courses taught by this faculty member between 2018-2020.

RQ 3: To what extent if any are there statistically significant differences between the grades of female graduate students enrolled in face-to-face courses and female graduate students enrolled in online courses taught by this faculty member between 2018 – 2020? (*N=6 Face to Face, 57 Online)*.

Hypothesis H3o: There is no statistically significant differences between the grades of female graduate students enrolled in face-to-face courses and female graduate students enrolled in online courses taught by this faculty member between 2018 – 2020.

Result: Data displayed in Table 3 shows that no statistically significant differences were detected between the grades of female graduate students enrolled in face-to-face courses and female graduate students enrolled in online courses taught by this faculty member between 2018 – 2020. Female graduate students enrolled in face-to-face courses earned higher grades than female graduate students enrolled in online courses. A key reason for the disparity was very low enrolled face to face courses compared to online.

Table 3: Grade Comparison – Male/Female Graduate Students 2018 – 2020

	T-Critical	T-Statistic	P-Value	Decision
Panel A: Male Graduate Face to Face/Male Graduate Online 2018 – 2020				
	2.07	0.090	0.9288	Accept Null
Panel B: Female Graduate Face to Face/Female Graduate Online 2019 – 2020				
	1.67	1.76**	0.042	Reject Null

*Note: ** was used to designate significance at the 0.05 level. The results of the analysis in Table 3 shows that no statistically significant differences were found between the grades of male graduate students enrolled in face-to-face courses and male graduate students enrolled in online and no statistically significant differences were detected between the grades of female graduate students enrolled in face-to-face courses and female graduate students enrolled in online courses taught by this faculty member between 2018 – 2020. Female graduate students enrolled in face-to-face courses earned higher grades than female graduate students enrolled in online courses due to very low enrolled face to face courses compared to online.*

RQ 4: To what extent if any are there statistically significant differences between the grades of male undergraduate face-to-face students and male online undergraduate college students in courses taught by this faculty member between 2018 - 2020?

Hypothesis H4o: There are no statistically significant differences between the grades of male undergraduate face-to-face students and the grades of undergraduate male online students in courses taught by this faculty member between 2018 - 2020. (Face to Face N = 118 – Online N = 256)

Result: Table 4 below shows that no statistically significant differences were found between the grades of male undergraduate face-to-face students and the grades of undergraduate male online students in courses taught by this faculty member between 2018 - 2020.

RQ 5: To what extent if any are there statistically significant differences between the grades of female undergraduate face-to-face students and female online undergraduate students in courses taught by this faculty member between 2018 – 2020?

Hypothesis H5o: There are no statistically significant differences between the grades of female undergraduate face to face students and undergraduate female online students in courses taught by this faculty member between 2018 – 2020. (*Face to Face N = 74 – Online N = 346)*

Result: Table 4 below shows that no statistically significant differences were found between the grades of male and female undergraduate regardless of modality in courses taught by this faculty member from 2018 - 2020. However statistically significant differences were found between the grades of females enrolled in undergraduate online and face to face courses taught by this faculty member between 2019 - 2020. The disparity could be attributed to the types of courses (modality) assigned to the faculty member and major decreases in enrollment in face-to-face courses due to the Covid-19 pandemic.

Table 4: Grade Comparisons – Male/Female Undergraduate Students 2018 – 2020

	T-Critical	T-Statistic	P-Value	Decision
Panel A: Undergraduate Male Face-to-Face Students and Undergraduate Online Students				
(Male OL vs. FF 2018)	1.9812	-0.1993	0.8424	Accept Null
(Male OL vs. FF 2019)	1.9803	-1.7756	0.0784	Accept Null
(Male OL vs. FF 2020)	1.9774	0.2540	0.7999	Accept Null
Panel B: Female Face- to-Face Students and Undergraduate Online Students				
(Female OL vs. FF 2018)	1.9757	-1.3627	0.1749	Accept Null
(Female OL vs. FF 2019)	1.9808	-3.0185	0.0031	Reject Null
(Female OL vs. FF 2020)	1.9762	-2.3235	0.0215	Reject Null

Table 4 above shows that no statistically significant differences were found between the grades of male and female undergraduate regardless of modality in courses taught by this faculty member between 2018 - 2020. However statistically significant differences were found between the grades of females enrolled in undergraduate online and face to face courses taught by this faculty member between 2019 - 2020. The disparity could be attributed to the types of courses (modality) assigned to the faculty member and major decreases in enrollment in face-to-face courses due to the Covid-19 pandemic.

RQ 6: To what extent if any are there statistically significant differences between the grades of face-to-face undergraduate students and online undergraduate students in courses taught by this faculty member between 2018 - 2020?

Hypothesis: H6o: There are no statistically significant differences between the grades of face-to-face undergraduate students and online undergraduate students in courses taught by this faculty member between 2018 – 2020.

Result: Table 5 below illustrates that in general no statistically significant differences were found between the grades of undergraduate face-to-face students and the grades of undergraduate face to face online students in courses taught by this faculty member in 2018 and 2020. The primary reason for the disparity between grades between U/G and G online/face to face was attributed to the near termination of face-to-face courses due to the Covid-19 pandemic resulting in a reject the null decision in 2019. It should be noted that while the decision was to accept the null in year 2018, the p-value of 0.12 was trending towards significance as was the p-value of 0.14 in 2020.

Table 5: Grade Comparison Between Face-to-Face Undergraduate Students and Online Undergraduate Students 2018-2020 (N = 806, 285 Face-to-Face, 521 Online)

Hypotheses - Year	T-Critical	T-Statistic	P-Value	Decision
U/G Online v. U/G Face to Face 2018	1.97	1.56	0.12	Accept Null
G Online v. U/G Face to Face 2019	1.97	3.02	0.0009**	Reject Null
U/G Online v. U/G Face to Face 2020	1.97	0.91	0.14	Accept Null

*Note: ** was used to designate significance at the 0.05 level. The data analysis in Table 5 above illustrates that no statistically significant differences were found in the aggregate between the grades of undergraduate face-to-face students and the grades of undergraduate face to face online students in courses taught by this faculty member in 2018 and 2020. The primary reason for the disparity between grades between U/G and G online/face to face was attributed to the near termination of face-to-face courses due to the Covid-19 pandemic resulting in a reject the null decision in 2019. It should be noted that while the decision was to accept the null in year 2018, the p-value of 0.12 was trending towards significance as was the p-value of 0.14 in 2020.*

CONCLUSION

The chief goal of this exploratory research project was to determine if grading bias existed in courses taught by this faculty member during academic years 2018 - 2020. Six research questions were posed to respond to the research questions. A single factor ANOVA test was used to determine differences between subjects in this study. A student t-test was used to explore differences. A significance value of 0.05 was used to determine whether to accept or reject the null hypothesis in each hypothesis. To analyze the data, a grade point average number was assigned to each letter grade in the data set. Finally, the data was organized by gender for each course to examine the hypotheses. Of the six research questions, segmented findings resulted in accepting 7 null hypotheses that no grading biases between subjects existed. There were 5 instances where the null hypothesis was rejected indicating a disparity between the aggregate grades of male and female students and 4 occurrences when the subject populations were analyzed in detail in courses taught by this faculty member between 2019 and 2020. Two contributing reasons for bias could be attributed; a) the types of courses (modalities) assigned to the faculty member and changes in enrollments by modalities caused by the Covid-19 pandemic.

The implications of this work for educators are; a) further basic study needs to be done by individual faculty members as a matter of post course analysis about grading tendencies recorded in the institution's LMS or other grade records; b) a standard grade reporting methodology needs to be developed to aggregate individual faculty members grading outcomes, ex. by course segmented by genders, face-to-face, grade level, online, etc. for use by individual departments, programs, colleges, schools, etc. to enable higher level evaluation to evaluate grading trends at individual institutions; c) accreditation agencies should consider requiring reporting of grading changes and movements and actions taken by institutions to ensure that students are evaluated by faculty members utilizing fair, effective, translucent and consistent grading practices and; d) the research population (faculty and students) should be enlarged, to capture a multiplicity of populations, courses and disciplines to create a reference point from which further research can be conducted to foster improvements regarding faculty grading practices.

There were three limitations to this study. The first was gender which was presumedly based on interpersonal interactions as participants did not self-report their gender. Gender was assumed based on personal interactions and presentations. Gender for online students was grounded on commonly accepted name conventions. However; if gender identity was not clear from interactions or easily recognized from the individual's name, the participant's data was excluded from the analysis. The second limitation was the size of the subject population. The study involved only courses taught by this researcher during three calendar years and as a result did not reflect a representative sample of the entire student population of the University or College of Business. A third restraint was class size, mix of types of assignments (example an automatically graded quiz, discussions, etc.) and level (undergraduate compared to graduate level) influenced the faculty member's grading expectations. Some courses such as the BA 321 online had more quizzes and fewer discussions than their face-to-face counterparts which did not have any quizzes and more discussions. Four suggestions for future research on this topic are suggested; a) further basic study needs to be done by individual faculty members as a matter of post course analysis about grading tendencies recorded in the institution's learning management system or other grade records; b) a standard grade reporting methodology needs to be developed to aggregate individual faculty members grading outcomes, ex. by course segmented by genders, face-to-face, grade level, online, etc. for use by individual departments, programs, colleges, schools, etc. to enable higher level evaluation to evaluate grading trends at individual institutions; c) accreditation agencies should consider requiring reporting of significant grading changes and actions taken by institutions to ensure that students are evaluated by faculty members utilizing fair, effective, translucent and consistent grading practices and; d) the research population (faculty and students) should be enlarged to capture a wider range of course subjects and modalities.

REFERENCES

Breda, T., & Ly, S. T. (2015) “Professors in Core Science Fields are not Always Biased Against Women: Evidence from France” *American Economic Journal: Applied Economics*, *vol. 7*(4), p. 53–75. Retrieved June 3,2021 from: https://doi.org/10.1257/app.20140022

Bygren, M. (2019) “Biased Grades? Changes in Grading after a Blinding of Examinations Reform,” *Assessment and Evaluation in Higher Education*, vol. *45*(2), p. 292–303. Retrieved May 30,2021 from: https://doi.org/10.1080/02602938.2019.1638885

Dictionary.com (2021) Retrieved June 19,2021 from: https://www.dictionary.com/browse/bias

Hardré, P. L. (2014) “Checked Your Bias Lately? Reasons and Strategies for Rural Teachers to Self-Assess for Grading Bias,” *Rural Educator*, *vol. 35*(2), Retrieved May 15, 2021 from https://doi.org/10.35608/ruraled.v35i2.352

Hinton, D. P., & Higson, H. (2017) “A Large-Scale Examination of the Effectiveness of Anonymous Marking in Reducing Group Performance Differences in Higher Education Assessment,” *PLOS ONE*, *vol. 12*(8), Retrieved July 7, 2021 from https://doi.org/10.1371/journal.pone.0182711

Jansson, J., & Tyrefors, B. (2018) “Gender Grading Bias at Stockholm University: Quasi-Experimental Evidence from an Anonymous Grading Reform,” (SSRN Scholarly Paper ID 3334460). Retrieved June 14, 2021 from Social Science Research Network, https://papers.ssrn.com/abstract=3334460

Keyser, R. S. (2018) “Analyzing Course Grades in a Converged Classroom Environment--A Follow-up Study,” *Journal of Higher Education Theory & Practice*, *vol. 18*(4), p. 102–113. Retrieved August 6, 2021 from http://search.ebscohost.com.access.library.eou.edu/login.aspx?direct=true&db=eft&AN=132449553&site=ehost-live

Krawczyk, M. (2018). Do gender and physical attractiveness affect college grades? Assessment & Evaluation in Higher Education, *43*(1), p. 151–161. Retrieved August 31, 2021 from http://search.ebscohost.com.access.library.eou.edu/login.aspx?direct=true&db=eft&AN=126358822&site=ehost-live

Malouff, J. M., Emmerton, A. J., & Schutte, N. S. (2013) “The Risk of a Halo Bias as a Reason to Keep Students Anonymous During Grading, *Teaching of Psychology*, *vol. 40*(3), p. 233–237.

Malouff, J.M., Thorsteinsson, E.B. (2016) “Bias in Grading: A Meta-Analysis of Experimental Research Findings. Retrieved September 2, 2021 from https://doi-org.access.library.eou.edu/10.1177/0004944116664618

Nitko, A.J. (2004) Educational Assessment of Students. 4th Edition, Pearson Education, Inc., Upper Saddle River, NJ. Retrieved on June 22, 2021 from https://www.scirp.org/(S(351jmbntvnsjt1aadkposzje))/reference/referencespapers.aspx?referenceid=2411143

Pannucci, C. & Wilkins, E. (2011) “Identifying and Avoiding Bias in Research,” Retrieved on August 18, 2021 from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2917255/#:~:text=I.,or%20answer%20over%20others%E2%80%9D%207.

ACKNOWLEDGEMENT

The author acknowledges the support of Dr. Allen Klingenberg and his Statistics 3050 course students (Kyle Czerak, Brian Dean, Siobhan Haggerty, Noah Jensen and Olivia Masso) of Carthage College, Kenosha, Wisconsin in this research project.

BIOGRAPHY

Dr. Keller teaches management courses in the College of Business at Eastern Oregon University. He also teaches for Monarch Business School – Switzerland.

Business Education & Accreditation
Vol. 14, No. 1, 2022, pp. 43-60
ISSN: 1944-5903 (print)
ISSN: 2157-0809 (online)

IBFR
www.theIBFR.COM

ARE STUDENT-MANAGED FUNDS CLOSET INDEXERS?

Lynda S. Livingston, University of Puget Sound and Four Horsemen Investments
Shirley Mazaltov-Ast, University of Puget Sound and Four Horsemen Investments

ABSTRACT

Many business schools offer finance students the opportunity to run student-managed funds, which are meant to give participants experience running real money in real time. The usefulness of such an experience, however, depends on the structure of the fund: is it grounded in economic principles; does it mitigate or exacerbate behavioral investment biases; does it focus on stock picking, or emphasize portfolio management; does it promote true active management, or encourage benchmark-mimicking closet indexing? In this paper, we present a description of our university's fund, highlighting critical industry measures of active management to assess its performance. Some of our main generalizable findings are that funds' benchmarks must be consistent with the actual investment approach employed, that performance metrics must be clearly and explicitly determined in advance, and that the fund's investment policy statement must be reflective of empirical market realities.

JEL: G11, G21

KEYWORDS: Student-Managed Investment Fund, Active Share, Closet Indexing

INTRODUCTION

These days, if a university has a business program, it probably has a student-managed fund (SMF, sometimes called a student managed investment fund, SMIF). These funds offer participants the "ultimate" experiential learning opportunity to run real money in real time (D'Souza and Johnson, 2019), and—as they become ubiquitous—they are playing "an increasingly significant role in business college curriculums" (Phillips, *et al.*, 2020). Charlton, *et al.* (2015) note that SMFs develop not only students' ability to work with money but also with people, as students undertake group leadership roles and responsibilities; these authors assert that the high levels of student involvement encourage knowledge retention and the development of life-long skills. Clinebell and Murphy (2016) show that SMF alumni demonstrate superior communication skills and a deeper understanding of investments even eight years into their careers. Students seem to like these funds, and to gain relevant financial knowledge from running them (Hysmith, 2017). However, do they really gain useful money management experience from buying stocks with a university's money? Even if we stipulate that actively managing an equity portfolio of a few hundred thousand dollars can be a good idea, do the student managers actually manage actively? Or do they succumb to the plague of real-world active managers: closet indexing?

Petajisto (2013) defines closet indexing as "the practice of staying close to the benchmark index while claiming to be an active manager." This obviously undermines the portfolio goals of the client. It is also a "heinous" example of moral hazard (Taylor, 2004), since managers of closet index funds charge as if they were truly active, getting paid for something they do not do. Only the manager's active positions can contribute alpha, but she charges her fees on all of her stocks, even the index-replicators. Student fund managers do not impose this fee burden on their clients, but if their mandate is active management, they should not be trained to hug a benchmark. In this paper, we examine our own university's student managed fund, the PTA fund, for evidence of closet indexing. We use Cremers and Petajisto's (2009) active share measure, as well as tracking error, attribution, and traditional portfolio metrics, to assess how actively our students are actually managing the money entrusted to them. We find that while they are not closet indexers—which is consistent with the *spirit* of the fund—their active management actually violates the *letter* of their investment policy statement, whose quantitative objectives almost compel closet indexing. We offer our example as a case study in the necessity of ensuring consistency across the initial design and ongoing monitoring of a university student-managed fund. The paper proceeds as follows. In the next section, we briefly review the burgeoning literature on student-managed funds, then describe the active share measure and other relevant industry metrics. We then characterize our university's SMF using many of these standard approaches, as well as more unusual techniques such as polynomial goal programming. We discuss these results next, paying particular attention to the relationship between the fund's actual performance and its policy mandate. We conclude with a few suggestions for other SMF participants.

LITERATURE REVIEW

Our paper draws on several strains of research. Since we wish to identify any discrepancies between the promise and reality of student portfolio management, we begin by reviewing some of the prior literature on student-managed funds. We then describe work on relevant industry metrics of active management, including those specifically employed to identify closet indexers.

Student-Managed Funds

Student-managed funds provide students the opportunity to manage real money in real time. Indeed, they "not only offer unique learning opportunities, but also are valuable for resume enhancement, school promotion, alumni networking, and involvement of practitioners with finance programs" (Gradisher, *et al.*, 2016). The number of funds has grown quickly: in his early 1990 sample, Lawrence (1990) identified 22 extant SMFs, but by 2014, Bruce and Greene (2014) had cataloged 338. This growth is a testament to the experiential learning potential afforded by SMFs. However, as D'Souza and Johnson (2019) note, the specific structure of the fund is critical to all aspects of its success. The literature has identified three salient aspects of this structure: where the fund's seed money comes from, how its investment decisions are made, and how it is integrated into its university's standard curriculum (Gradisher, *et al.*, 2016). We will consider each of these for our PTA fund, then consider one final element essential for a successful fund: the investment policy statement. For our PTA portfolio, the initial investment came from a donor, who gave it specifically to establish the fund, which therefore operates like a restricted endowment. Students make the investment decisions. Faculty advisors provide guidance, but do not choose investments; advisors do, however, have veto power. Block and French (1991) argue against such administrative power, asserting that it "undermines the premise for establishing the fund in the first place."

Nonetheless, while we certainly acknowledge advisors' need to avoid controlling and overriding the students' actions and for allowing mistakes as students are learning (see Charlton, *et al.*, 2015), we believe that having a veto is prudent and consistent with the university's fiduciary duty. The veto has never been used. While the donor created the fund as a learning tool, it is run as an extracurricular activity, with no academic credit. This is our most controversial choice. According to Gradisher, *et al.* (2016), it would be better to have fund management explicitly linked to a class (as are 71% of funds in Lawrence's 2008, sample). This would ensure that students are engaging in educational, rather than professional, management, and that professors are acting solely as educators; these are important considerations for complying with the Uniform Prudential Management of Institutional Funds Act (UPMIFA). Links to specific courses would also ensure a baseline level of financial knowledge for the students; for example, Block and French (1991) suggest that fund managers should take corporate finance and investments (and preferably portfolio theory and derivatives as well).

On the other hand, Bergquist, *et al.* (2020) advocate for an extracurricular approach to SMF management. For them, "[i]mplementing a student investment group as a club as opposed to a course can provide an inclusive, non-restrictive, interdisciplinary environment for learning which closely resembles work environments in which most students will enter upon graduation." Since these authors nonetheless believe that some background is required, they train their incoming students using six peer-led "modules." This sort of fund-specific training seems to be becoming more common. It is also often preceded by stringent application requirements. For example, D'Souza and Johnson's (2019) school's fund requires aspiring student managers to submit an application, demonstrate substantial relevant academic and practical experience, and have an interview. At Ammermann, *et al.*'s (2020) university, interested students must compete for SMF management slots through an innovative request-for-proposal process.

Our fund neither requires explicit coursework nor SMF-specific training. Instead, making two presentations to the members, along with approval from the faculty advisor, is sufficient for admission to the investment board. This allows students from non-business disciplines to participate, but it prohibits the assumption of a common baseline of knowledge. Our lack of a rigorous screening mechanism may be a weakness to our fund's structure, since high demands encourage similarly high levels of student engagement and commitment (Tashjian, 2020). Regardless of funding source or curriculum integration, all studies of successful funds agree that an SMF should have a governing document. Indeed, Boughton and Jackson (2020) consider an investment policy statement (IPS) so fundamental to the SMF experience that they conclude that as long as the fund has such a governing policy, there is no other requisite common structural feature for an effective fund. Most SMFs do appear to have a policy statement. In Neely and Cooley's (2004) 61-fund sample, 49 funds had a traditional IPS, three others used their university's endowment's, and only three funds had no policy at all. Our PTA fund is both discussed in our university endowment's IPS and has its own fund-specific version (although the latter devotes a large proportion of its length to descriptions of internal fund affairs and officer descriptions, topics which might be better relegated to bylaws.) Involving students in the writing or editing of their fund's IPS can be extremely educational (Neely and Cooley, 2004; see Horstmeyer, 2020, for a discussion on how to write an IPS for a student-managed fund). Active involvement affords student-managers the opportunity to become familiar with their own fund's constraints and risk objectives—such as the adherence to Catholic values, capitalization and portfolio weight guidelines, and sector ranges that Daugherty and Vang's (2015) students handle in their annual rewrite. It also can help them learn broader lessons; for example, Gradisher, *et al.* (2016) stress that incorporating IPS guidelines like those of the CFA Institute introduces students to the standards of prudence and fiduciary care mandated by UPMIFA. Given the potential benefits, Phillips, *et al.* (2020) suggest that students write the IPS from scratch every year. We do not do this for our fund; while we have all students in our portfolio management class write a *de novo* IPS for the PTA fund as an exercise, writing an actual IPS each year would require far too much administrative staff time to be practical.

Despite the centrality of the IPS to the fund's management—and of return and risk objectives to the IPS—we did not find in the literature any specific references to other schools' relative or absolute return targets. Every fund has some sort of benchmark, but at most the associated IPS sets an objective to "exceed" it (for example, see Horstmeyer, 2020). Many funds just seek to match their benchmarks (see, for example, Ghosh, *et al.*, 2020, and Betker and Doellman, 2020; see also Haddad, *et al.*, 2020, for the objectives set for the Tennessee Valley Authority's program, which is associated with the funds of 25 universities). In contrast, our PTA fund has both a relative return target (100 bp above the Russell 3000) and a tracking error target (200 bp). These targets are an important consideration in our empirical work, as we will discuss below. Next, however, we consider the relevant literature on closet indexing.

<u>Closet Indexing and Industry Metrics</u>

Closet indexers pretend they actively manage their portfolios, but they really hug an index. Taylor (2004) tests a strategy of shorting suspected closet indexers while buying true index funds, which—if the shorts are really indexing—should allow the trader to benefit from the difference in the longs' (low) and shorts' (higher) fees. However, he finds that the strategy's Sharpe ratio is almost the same as the market's, leading him to conclude that "widespread closet-indexing does not exist in the mutual fund industry." In contrast, Petajisto (2013) finds that closet indexing has been increasing since 2007, accounting for about one-third of mutual fund assets at the time of his study. He uses several metrics to identify such behavior, as will we. We begin with the most novel: active share. Cremers and Petajisto (2009) note that a truly active equity fund can be viewed as an index fund core with a long-short portfolio satellite appended to it. Thus, they and Petajisto (2013) investigate the prevalence and ramifications of closet indexing using active share, which they define as the proportion of the fund that differs from its index, given the weighting differences of index stocks. For example, consider a concentrated fund that chooses to plunge into a single stock, out of a universe of three stocks. Compared to an equally weighted index of the universe, the concentrated fund has an active share of $(0.5)*[|(1 - 0.333)| + |(0 - 0.333)| + |(0 - 0.333)|] = 67\%$. (The absolute values abstract from the direction of the difference; the multiplication of the sum by 0.5 keeps the active share of non-shorting portfolios between 0% and 100%.) Such a fund's manager "fishes" only in the top third of his universe. In contrast, a fund that only slightly changed the index weights—say to (.40, .20, .40)—would have an active share of $(0.5)*[|(0.40 - 0.333)| + |(0.20 - 0.333)| + |(0.40 - 0.333)|] = 13\%$. This latter fund would be a closet indexer. In general, Petajisto (2013) categorizes a fund with less than 50% active share as a combination of an actively managed fund and an index fund. Assuming that half of all funds perform better than the index and the other half perform worse, holding less than 50% active share means some of a manager's positions "cannot exist because the manager expects them to outperform the index; they exist only because he wants to reduce his risk relative to the index, even when that means including negative-alpha stocks in the portfolio." 50% is the "theoretical minimum" active share than a pure active manager can have. However, not every fund with low active share is a closet indexer. Instead, Petajisto (2013) categorizes funds into four types—described below—using both active share and a more traditional comparison metric: tracking error. He defines tracking error as the standard deviation of the difference between the fund's return and the index's (not adjusting for beta exposure), since he asserts that the simple difference is the more common performance measure for active managers. He then associates tracking error with exposure to systematic factors (e.g., picking sectors, choosing styles, holding cash), and active share to stock selection—the only two possible ways an active manager can add value.

Given these two dimensions of active management, he creates his fund categories. Active stock pickers will have high active share. If they diversify those bets, they will also have low tracking error, making them "stock pickers"; in contrast, more concentrated stock bets mean more tracking error for the "concentrated" group. "Factor bets" have low active share but high tracking error, which comes from their high exposure to systematic factor risk. Finally, closet indexers have both low active share and low tracking error—they stick close to their benchmarks, hoping that their mediocrity will be inoffensive. In his sample, the factor-bet/high tracking error funds "destroyed value." In contrast, funds with high active share—concentrated funds and diversified stock pickers—outperform. Only stock pickers, however, add value after fees. These active stock pickers can do especially well when the cross-sectional dispersion in index returns is high. This "cross-vol" measure is based on the weighted differences in the individual stocks' returns relative to the index. High idiosyncratic risk may drive high cross-sectional volatility, providing active stock pickers with target-rich environments: their active returns will be high if the firm-specific mispricings resolve themselves over the managers' holding period. Indeed, Petajisto (2013) finds that high cross-sectional volatility is positively related to stock pickers' future returns.

However, Petajisto (2013) does not link cross-sectional volatility with the performance or prevalence of closet indexing. Brown and Davies (2015a, 2015b), on the other hand, put volatility—albeit the volatility of the individual funds, rather than that of the index stocks in general—at the heart of their theoretical model. They conclude that closet indexers strategically increase the volatility of their funds to help themselves masquerade as truly active funds. In so doing, they limit the efficacy of Petajisto's (2013) active share metric as an identifier of closet indexing behavior. In the Brown and Davies (2015a, 2015b) model, highly skilled managers—those with more than a threshold level of skill—will invest the resources necessary to run a truly active fund. Less-skilled managers will become closet indexers. These "charlatans" hope their clients mistake them for true active managers; otherwise, the high fees will stop. If the true active managers' returns are more variable, it becomes harder for investors to determine whether a fund did poorly because its manager failed or because she did not try—to distinguish a true active manager from a closet indexer. Performance becomes a noisier signal of value. This not only encourages more closet indexing, but also induces the charlatans to inject more variability into their own returns, by "signal jamming." They do this by making uninformed changes to their funds' weights of index stocks, just so their portfolios will differ from the index. (Similarly, Cremers and Petajisto, 2009, suggest closet indexers may also "mask" their essentially passive funds with excess turnover.) If the managers do not play this game, they risk outing themselves as closet indexers. Thus, in this model, every charlatan plays: "[n]oise injection and closet indexing are complements; as the moral hazard problem becomes more severe, closet indexers inject more noise, attenuating investors' abilities to identify skilled managers." Since this strategic behavior may make tracking error and active share less effective indicators of true active management, Cremers and Petajisto (2009) suggest that investors should use <u>both</u> measures when choosing active managers, as well as the prior year's returns (to take advantage of empirically demonstrated performance persistence in the most active funds). That is the framework that we will use in this paper. We now turn to characterizing our fund using these metrics, as well as both some more traditional and some more novel ones.

DATA AND METHODOLOGY

In this section, we first provide some background on the structure our school's PTA fund and on the characteristics of our past student managers. We then describe the performance data we will use to characterize our fund. Our fund is run by undergraduate students who study business within a general liberal arts curriculum. Since its inception in late 2016, the fund has had five executive teams, typically consisting of a president, vice president, treasurer, market research chair, outreach chair, and secretary. Executive positions are typically spread across class cohorts, with the presidency held by a junior or senior, and students may hold executive roles multiple years in a row. Approximately 85% of all executive positions have been held by male students. This includes all five presidents, as well as every market research chair and secretary. No women were on the executive team until spring of 2019; half of the female participation we have had has been in the role of outreach chair. As noted in the literature review section, no formal coursework is required for participation in the club that runs the fund, and all trading decisions (subject to an as-yet never used administrative veto) are made by the students. Having described the structure of the fund and the types of students who run it, we turn now to the data we will use to evaluate its performance. Our sample period runs from the date of the fund's inception, 8/29/16, through 4/15/21. Our SMF started investing in individual stocks on 3/10/17, having held its assets only in an index "parking place" for the first six months. We chose our annual analysis periods to conform roughly to the active parts of our university's school years, therefore defining our year-end to be around April 15th (specifically, 4/13/17, 4/13/18, 4/15/19, 4/15/20, and 4/14/21). We ascribe the initial 24 trading-day period from 3/10/17 through 4/13/17 to our "incubation period," focusing our analysis on the last four years (see Cremers and Petajisto, 2009; we are following their very loose interpretation of the incubation period from Evans, 2004). Data for all but two of the portfolio's stocks, as well as for all indexes, comes from Yahoo! Finance. The remaining two portfolio stocks, The Michael's Companies (MIK) and Instructure (INST), went private during our sample period; we obtained date for these stocks from Quandl. For MIK, which was bought out in April of 2021, just after our sample period, we have data for the full period; for INST, however, we have data only through 3/23/20. We therefore measure all covariances with INST using the subperiod 8/29/16-3/32/20.

RESULTS

In this section, we summarize the results of our investigations into the performance of our student-managed fund. The various approaches we use will help us determine whether our fund is being run as a truly active fund. We begin with returns- and holdings-based analyses, turnover, and tracking error, then turn to metrics and techniques that may be more novel for SMFs: active share, attribution, and polynomial goal programming.

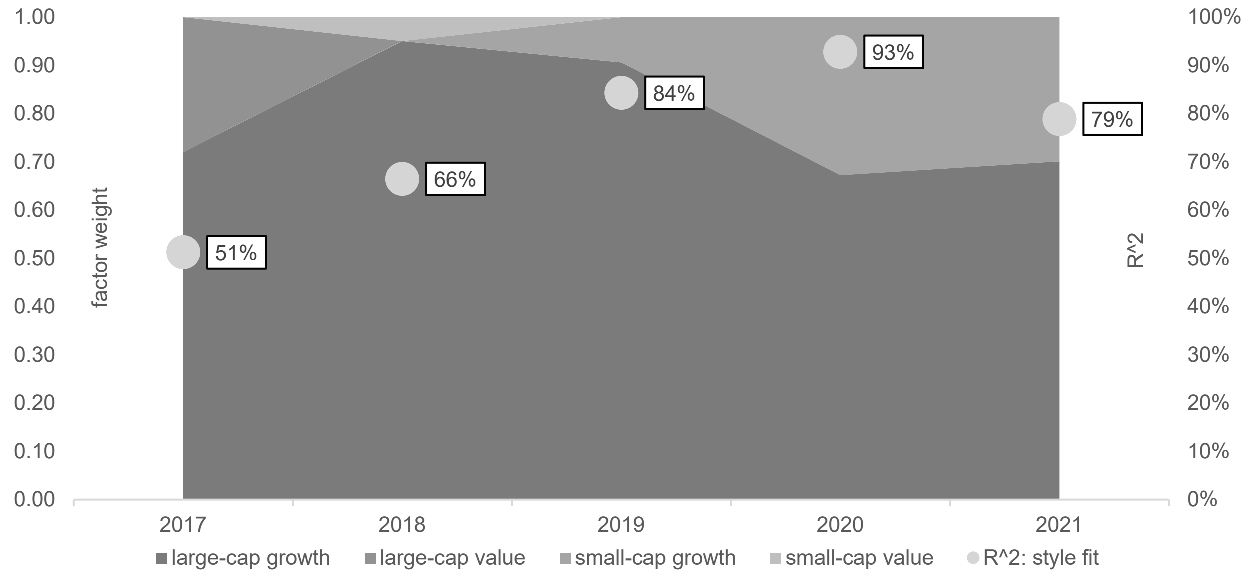
Returns- and Holdings-Based Analysis, Turnover, and Tracking Error

We performed both holdings-based and returns-based analysis, which allow us to identify any style biases in the portfolio (see Gastineau, *et. al.*, 2007; see Taylor, 2004 for a recommendation that holdings-based analysis be used to identify closet indexing). We will first discuss the latter. For each annual subperiod and for the full period, we regressed the returns of the PTA portfolio on four ETFs spanning the U.S. domestic equity market: VTWV (small-cap value), VTWG (small-cap growth), VONV (large-cap value), and VONG (large-cap growth) (see Gastineau, *et al.*, 2007). We constrained the coefficients to be nonnegative and required them to add to 1, so that we can interpret them as factor weights. The resulting rolling style chart is shown in Figure 1. The fund is clearly being run as a growth fund, despite having a broad Russell 3000 mandate. The white dots in the graph identify the constrained regressions' R^2 values, which can be interpreted here as the style fit (the complement, [$1-R^2$], measures the security selection component). These factor tilts—84% since inception—explain the vast majority of the fund's performance, overwhelming any contribution from active stock picking. Taylor (2004) notes that a high R^2 from a comparison of fund and index returns is a necessary indicator of closet indexing, and identifies his "suspect" closet indexers from among funds with R^2 values of at least 97%. However, R^2 is not sufficient, since it may not persist. Thus, the PTA fund may not be a closet indexer, but it is clearly not a balanced fund.

We confirmed the fund's style bias with holdings-based analysis. Using market value, trailing twelve-month earnings per share, year-over-year quarterly revenue growth, trailing price/earnings, price/book, dividend yield, and sector, we characterized the PTA fund's stocks by size and style (see Gastineau, *et al.*, 2007). Large-cap growth stocks always represented a majority of the portfolio. During the "incubation period" from August 2016 through April 2017, the PTA fund held only growth stocks, two-thirds of them large-cap. From April, 2017 to April, 2018 it had its highest weight on value, at one-third, evenly split between small- and large-cap; two years later, it was at its most evenly balanced, with a since-inception minimum of 58.8% in large-cap growth (plus 11.8% small-cap growth, 23.5% large-cap value, and 5.9% small-cap value). Most recently, from April, 2020 through 2021, large-cap growth stocks have hit their since-inception maximum: 75%; only 18.8% of all securities are in value, with 6.3% small-cap value and 12.5% large-cap value. Overall, it is evident that the PTA is heavily overweight in large-cap growth stocks, which is inconsistent with their benchmark of the Russell 3000.

Later in this section, we consider this tilt's impact on the fund's tracking error. First, however, we examine the PTA fund's turnover. Turnover is our initial measure of the PTA fund's active management. In general, we expect less turnover from less active funds (Cremers and Petajisto, 2009); for example, the Center for Research in Security Prices (2021) notes that some "funds that have a passive management style may… show very low or zero turnover." However, there are at least two caveats. First, we do not really expect *no* turnover from a true indexer, since even these funds must accommodate investor flows and adjust to events such as the annual Russell reconstitution. Thus, in Frino, *et al.* (2004), index funds average 6.47% annual turnover, while enhanced indexers average 12.26%; in Cremers and Petajisto (2009), the most passive category of funds has turnover of 18.1% per year.

Figure 1: Rolling Style Chart for PTA Fund



The figure shows the proportional effective exposure of the PTA fund to four mutually exclusive and exhaustive factor tilts. The shaded areas represent the weights of the four market-spanning equity-market indicators representing those factors, which were used as independent variables in a regression of the PTA fund's returns. The labeled white dots are the regressions' R^2 values, which here measure the style fit. The large area accounted for by the large-cap growth indicator, and the high R^2 values, indicate that the fund is being run as a growth fund, rather than as the blend that its benchmark would suggest.

Second, as we noted earlier, higher trading activity might just be a way to make a manager look busy and mask their closet indexing—the tactic Cremers and Petajisto (2009) call "signal-jamming." Unless trading enhances good active bets, it will not create value. To assess the PTA's fund's level of trading activity, we used Frino, *et al.*'s (2004) definition of turnover: the minimum of annual dollar sales or purchases, divided by the daily average portfolio size. We find turnover of 7% and 16%, in 2019 and 2020, respectively; for all other years, the managers made no sales, so that the numerator of the turnover ratio—the minimum of purchases and sales—was zero. Thus, while these low levels of measured turnover land solidly in Cremer and Petajisto's (2009) indexing category (and between Frino, *et al.*'s, 2004, indexing and enhanced indexing strategies), we suspect that the turnover this early in the fund's life is an artifact of its "incubation," rather than a reflection of its active management style. The donor released his funds in tranches, and the members' early priority was on investing those funds, not on optimizing the portfolio as a whole. Members still do not have a codified sell discipline; indeed, two of the five sales events were from stocks being taken private, not from club decisions. Even for years that do have sales, the magnitude of the purchases was at least four times higher (and at least twice as large as the average purchase proportion reported in Frino, *et al.*, 2004), indicating the members' focus on buys. Thus, while turnover may be an important metric to watch going forward, we do not believe it sheds much light on the past behavior of our student-managers. Of course, we could be observing some behavioral biases in the turnover numbers. Members may be reluctant to sell assets bequeathed to them by prior groups (endowment bias). From anecdotal evidence, we suspect that this may be a factor, especially since the fund is still not fully invested. On the other hand, Block and French (1991) suggest that incoming managers may deem their predecessors' choices "unwise," choosing to sell them immediately to get the fund back "on… track." They attribute their own SMF's high turnover to such inclinations (although they do not quantify how "high" their turnover actually is). We plan to explore these sorts of attitudes in a future survey of current and former PTA fund members. However, for now, we turn to the first of two active-management measures used by Cremers and Petajisto (2009): tracking error. For each school year, as well as for the full since-inception period, we calculated tracking error (TE) using both the beta-adjusted and the common simple-difference definitions.

The former uses the volatility of the error term from a regression on the benchmark, for which we use the following:

$$R_{PTA,t} = intercept + beta * (R_{benchmark,t})\ error_t \qquad (1)$$

(The relatively small variability of the Treasury series over this period implies that using the excess return form of this regression would have an immaterial effect on our estimated betas and errors; for example, betas using the two different regression forms for the normal benchmark regression are identical out to the fifth decimal place. We do verify these betas using excess returns in the discussion section of the paper, below.) The simple-difference version (dubbed "common" in both Petajisto, 2013, and Cremers and Petajisto, 2009) is simply the standard deviation of the difference between the fund's return and the benchmark's; this assumes that the beta from equation (1) is one. For our benchmarks, we used the PTA fund's actual Russell 3000 benchmark (proxied by the ETF VTHR), a large-cap growth benchmark (the ETF VONG), and a normal benchmark composed of 70% VONG and 30% VTWG, our small-cap growth ETF. In all cases, the betas from equation (1) were close enough to one so that the two tracking-error measures differed by only a few basis points, so we will focus our discussion on the common measure. We report our results in Table 1.

Table 1: PTA Fund Summary Statistics and Tracking Error Data

	2017	**2018**	**2019**	**2020**	**2021**	**Since Inception**
arithmetic mean	-0.00153	0.00027	0.00080	0.00050	0.00201	0.00084
variance	0.00004	0.00011	0.00016	0.00040	0.00023	0.00022
standard deviation	0.00593	0.01048	0.01284	0.01991	0.01506	0.01484
holding period return (chain-linked)	-43.2%	5.7%	19.9%	7.9%	61.0%	112.1%
annualized standard deviation	9.4%	16.6%	20.3%	31.5%	23.8%	23.5%
RUSSELL 3000						
beta	0.90	1.06	1.19	1.01	0.98	1.03
intercept	-0.0010	-0.0004	0.0003	0.0005	0.0002	0.0001
simple difference tracking error	0.00424	0.00677	0.00692	0.00715	0.00926	0.00753
beta-adjusted tracking error	0.00422	0.00676	0.00671	0.00714	0.00925	0.00752
annualized TE: simple difference	6.71%	10.71%	10.95%	11.30%	14.64%	11.90%
LARGE-CAP GROWTH						
beta	1.19	1.05	1.02	0.97	0.92	0.97
intercept	-0.0011	-0.0006	0.0002	0.0001	0.0002	-0.0001
simple difference tracking error	0.00401	0.00605	0.00512	0.00601	0.00776	0.00626
beta-adjusted tracking error	0.00395	0.00604	0.00511	0.00597	0.00768	0.00625
annualized TE: simple difference	6.34%	9.57%	8.09%	9.50%	12.26%	9.90%
NORMAL						
beta	0.83	1.04	1.02	0.96	0.95	0.97
intercept	-0.0012	-0.0006	0.0003	0.0003	0.0000	0.0000
simple difference tracking error	0.00452	0.00622	0.00522	0.00534	0.00692	0.00594
beta-adjusted tracking error	0.00445	0.00621	0.00522	0.00527	0.00689	0.00625
annualized TE: simple difference	7.14%	9.83%	8.26%	8.44%	10.94%	9.39%

This table provides summary statistics and tracking error (TE) values for the PTA fund. In the top panel, we provide periodic and annualized risk and return metrics for the fund. To put the holding period returns in perspective, we note that the 2021 values for the Russell 3000, the large-cap growth ETF, and the custom benchmark were 56.1%, 55.6%, and 63%, respectively; since-inception values were 88.8%, 137.4%, and 96.2%. The remaining three panels provide tracking error data for the three benchmarks, using both the simple-difference method and the beta-adjusted method. In all cases, for years after the "incubation" year of 2017, these tracking error values put the PTA fund in the top 20% of the tracking-error sample in Cremers and Petajisto (2009).

The top panel of Table 1 gives the summary statistics for the PTA fund over all subperiods. The other three panels present inputs and results for the tracking error measurements, using our three benchmarks. Across all five years and all benchmarks, the maximum TE is 14.64%, and the minimum is 6.34%. Ignoring the "incubation" year of 2017, the minimum rises to 8.09%. These ranges place the PTA fund beyond Petajisto's (2013) characterization of closet indexers, which in his sample of 1,124 funds have a mean TE of 3.5% with a standard deviation of 0.9%. If we create a range of (mean ± 1 standard deviation), using Petajisto's (2013) summary statistics and his classifications of fund types (stock pickers, concentrated, factor bets, moderately active, and closet indexers), we find that 93.3% of the PTA's 15 index-year combinations fall within the range defined for factor bets (characterized as high TE/low active share), while none do for the closet indexer range (low on both dimensions). (The others are stock pickers: 53.3%; concentrated: 13.3% and moderately active: 20% [all in incubation year]). This result is consistent with the strong loading of the PTA fund's returns on the two growth ETFs, and strongly suggests that the fund is not a closet indexer.

Having considered the more standard measures of portfolio performance, we turn now to the more novel.

<u>Active Share, Attribution, and Polynomial Goal Programming</u>

The second dimension that Cremers and Petajisto (2009) use to categorize funds—along with tracking error—is active share. As noted earlier, active share is half the absolute difference in asset weights between the fund and its benchmark. (See also Frino, *et al.*, 2004, for a similar measure they call the "absolute deviation.") For long-only funds like the PTA fund, this measure will fall between 0% and 100%, with higher values indicating a higher level of active management. Since the PTA fund has between three and sixteen stocks during our analysis period, its active share versus the benchmark Russell 3000 is obviously extremely high: over 90% for each year (as were 21% of Cremer and Petajisto's, 2009, sample). This would place the PTA fund firmly in the "concentrated" category of Cremers and Petajisto's (2009) taxonomy. The difference between concentrated and factor bets categories is the high active share. But is that active share measure really as meaningful a contributor to the PTA fund's results as is the tracking error? To explore further the implications of the fund's active share, we perform the following simple attribution:

$$\left(R_f - R_b\right) = \sum_{i=1}^{n}(R_i - bogey) * \left(wt_{i,f} - wt_{i,MV}\right) + \sum_{i=1}^{n}(R_i - bogey) * (wt_{i,MV} - wt_{i,b}), \quad (2)$$

where R_f is the return for the PTA fund for the period; R_b is the benchmark return; the wt_i terms are the weights of index stock *i* in the *f*und, *MV*-weighted, and *b*enchmark portfolios; the "bogey" is simply the return for our Russell 3000 benchmark, VTHR; and *n* is the number of stocks in that benchmark. (Measuring the stocks' returns relative to the bogey allows easy identification of over- and underperformers.) Since the PTA fund is a long-only domestic equity fund benchmarked to a market capitalization-weighted index of the broad market, the only way it can generate outperformance is to choose outperforming stocks (and omit underperformers) and/or to overweight the "good" stocks. Our simple attribution therefore has two parts. *Security selection* is measured using the difference between the performance of the stocks the fund chooses and that of the full set of stocks in the benchmark. The weights of the chosen stocks in this term are "naïve": they are simply the weights that would be used in a market cap-weighted index using only that subset of stocks. (Thus, if the fund chose only two stocks, both of which had the same market cap, then each of their $wt_{i,MV}$ terms would be 0.50, and the weight of all other benchmark stocks would be zero.) This is the second term in equation (2). Having used the naïve weights to measure security selection, we then use the fund's deviation from those weights as the measure of the *intensity* of the student managers' commitment to each of their selections. Thus, the first term in the attribution is based on the difference between a selected security's actual fund weight and its naïve weight. The "actual" weights are the average daily weight over the period, so there is asynchrony between them and the year-end values of the benchmark weights.

The PTA fund beat its benchmark in the last three of the five periods (four years, plus the 24-day 2016 "incubation" period). In all but the last year, stock selection was a positive contributor (contributing from 1.4% of outperformance in the incubation period to 27.5% in 2020), while intensity was negative (contributing from -3.5% in both the incubation period and 2019, to -11.2% in 2020). This pattern occurred in both good and bad market years. In 2021, when intensity determined 103% of the 193 bp excess return, the portfolio benefitted tremendously from its positions in Square (SQ), Axon Enterprise (AXON), and Etsy (ETSY): none of these stocks makes up more than 0.3% of the Russell 3000, but the PTA fund weighted them at 5%, 15%, and 3% (approximately 2, 8, and 5 times their naïve values, respectively). Our holdings-based analysis categorized these three stocks as growth stocks. Their impact on the portfolio may therefore have been more an artifact of a good period for growth than a reflection of superior active insights about these specific stocks. To continue our examination of the fund's stock weighting insights (intensity), we compared the actual weights chosen for the PTA fund at the end of our sample period (called $wt_{i,f}$ in the attribution above) to the optimal weights for that set of stocks. To determine these optimal weights, we used Excel's MMULT and Solver functions to identify the unconstrained and constrained (nonnegative) weight sets, respectively, that would give the lowest variance to a portfolio that had the same mean return as the PTA fund's 2021 portfolio. We used the sixteen included stocks' returns for our full data set, 8/30/16-4/14/21, to create the covariance matrix.

The differences were stark: Solver assigned a zero weight to fully half of the PTA fund's holdings; MMULT would short-sell six of them. The MMULT and Solver solutions share the same top five stocks, but the PTA fund has only two of those in its top five. MMULT's second most heavily weighted stock, Wal-Mart at 37%, is thirteenth for PTA (at 4%). (The Solver weight is 16%, the fourth highest.) In contrast, the PTA fund's top stock, Axon, is weighted almost twice as heavily as it is in either of the optimized portfolios (15% v. 8%). The PTA fund does better with its smaller holdings: four of its lowest-weighted five stocks are in Solver's "omit" category; three of these have negative weights in MMUL's unconstrained optimum. And all portfolios agree in putting ATT last in the rankings. To quantify these differences, we compared the PTA fund's weights to those of the optimized MMULT and Solver portfolios using the Spearman rank correlation coefficient. Higher values for this metric imply better correspondence between the PTA fund's choices and the optimized portfolios'. PTA and MMULT's result was 34%; PTA and Solver's—both of which prohibited short sales—was 41%. Although the PTA fund's weighting scheme differs markedly from the optima determined by MMULT and Solver, it may nonetheless confer benefits not considered in traditional mean-variance analysis. We therefore used polynomial goal programming (PGP) to incorporate skewness and kurtosis into the optimization objective function, so that we could compare the fund's weights to those optimized over four moments (see Lai, *et al.*, 2006, for an overview of the procedure, and Livingston, 2019, for an application to an SMF). To implement this approach, we minimized the following objective function:

$$Z = \left|\frac{d_1}{R*}\right|^{\lambda_1} + \left|\frac{d_2}{V*}\right|^{\lambda_2} + \left|\frac{d_3}{S*}\right|^{\lambda_3} + \left|\frac{d_4}{K*}\right|^{\lambda_4} \tag{3}$$

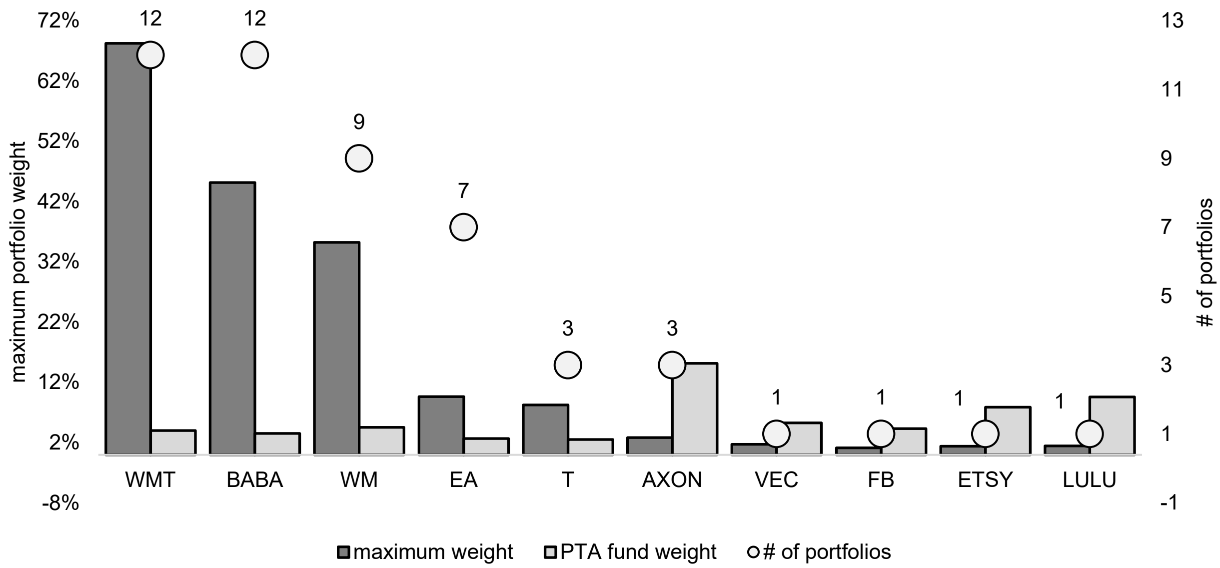
Here, R*, V*, S*, and K* are the "aspired" levels for mean, variance, skewness, and kurtosis, respectively; thus, V* and K* are the minimum possible values of <u>v</u>ariance and <u>k</u>urtosis, given our set of stocks, and S* and R* are the maximum possible <u>s</u>kewness and <u>m</u>ean return. (All of these aspired values are subject to our non-negativity constraint. As we will see, unless we also limit maximum weights, the max-return portfolio will plunge into the highest-mean asset, and R* will equal its mean.) Once the market opportunities are characterized, we add representations of the investor's preferences. These are the λ values: λ_1 reflects the investor's preference for the first moment, mean; λ_2 reflects variance; λ_3, skew; and λ_4, kurtosis. Preferences guiding portfolio creation are therefore described as (λ_1, λ_2, λ_3, λ_4) sets; for example, (1100) is the traditional mean-variance portfolio, while (1110) adds skewness as a equivalently valued criterion and (1111) adds both skewness and kurtosis. An investor with a stronger preference for a particular moment will tend to see more attractive values for that moment in her optimized portfolio (Davies *et al.*, 2004). We used the twelve λ_i preference sets used by Kemalbay, *et al.* (2011), including (1100), the mean-variance portfolio. (See caption of Figure 2 for full list.)

The PGP optimization proceeds by choosing the *d* terms to minimize the objective function. These *d* terms measure the distance between the portfolio's moment and the optimal level of that moment. Thus, d_1 represents the amount by which the portfolio's mean falls below M*: d_1 = (M* - portfolio mean), and d_3 measures the difference in skew: d_3 = (S* - portfolio skew). The d_2 and d_4 values measure portfolio distances from V* and K*, respectively, defined by subtracting the aspired values (which will be smaller) from the portfolio values. (Thus, *d* values are defined to be positive.) The absolute value functions correct for possible negative values of M* and S*. (For our optimizations, M* and S* are positive, so we can ignore the absolute values; we therefore optimize using Excel Solver's GRG Nonlinear/Multistart engine rather than its Evolutionary engine. See Winston, 2016.) Figure 2 summarizes our results. Five of the PTA fund's sixteen stocks were never used in any of the PGP optimized portfolios (Home Depot [HD], Microsoft [MSFT], Intuitive Surgical [ISRG], BlackRock [BLK], and TPI Composites Inc

[TPIC]). Another, Square (SQ), was used only once: in the mean-optimized "aspired" portfolio (given the constraint against nonnegative weights, this portfolio plunges into the highest-mean asset). Four stocks (Vectrus [VEC], Facebook [FB], Etsy [ETSY], and Lululemon [LULU]) have trivial impacts, appearing in only the mean-variance optimized portfolio, and all at less than a 2% weight. (ETSY does have an 8% weight in the aspired minimum-kurtosis portfolio, however.) AXON shows up in three portfolios emphasizing low variance, but again, its weights are negligible (less than 3%). AT&T (T) enters the same portfolios as AXON, but generally has higher weights; its strongest presence is in the aspired minimum-variance portfolio, where its weight is just over 16% (this weight is cut in half in the mean-variance optimized portfolio). Electronic Arts (EA) also is included in the variance-focused portfolios, but is most important in the kurtosis-focused portfolios: it makes up 32% of the aspired low-kurtosis benchmark, as well as almost 10% in the 1311 and the 1113 portfolios.

The most important stocks in the PGP-optimized portfolios are Walmart (WMT), Alibaba (BABA), and Waste Management (WM), in that order. All appear in the aspired minimum-variance portfolio. BABA also makes up 60% of the minimum-kurtosis portfolio; WMT is 100% of the maximum-skew portfolio. In the optimized portfolios, WM primarily contributes to those that are variance-focused (for example, it is weighted at 35% in the mean-variance portfolio 1100, and at 20% in 1110); otherwise, though, its only major appearance is its 21% weight in 3110. WMT and BABA, in contrast, each show up in 12 of the optimized portfolios; in fact, eight of these (all of which have a high-emphasis "3" in any moment position) are essentially two-asset portfolios of about 40% BABA and 60% WMT. However, as is clear from Figure 2, the PTA portfolio makes little use of the potential benefits of WMT and BABA. We therefore cannot conclude that the student-manager's intensity choices have made meaningful contributions to the portfolio's higher-moment characteristics. Thus, the polynomial goal programming results support the conclusion from the attribution: the PTA fund does not demonstrate particular insight in weighting *its own* stocks. Instead, it has derived value from having chosen growth stocks during a period when large-growth funds "trounced" and "smashed" large-value funds (Lynch, 2021). We are thus inclined to attribute the PTA fund's performance to its style tilt, and place it firmly in the "factor bets" category.

Figure 2: Summary of Polynomial Goal Programming Optimal Weights



This figure gives highlights from the PGP optimizations. The PTA fund's stocks are Walmart (WMT), Alibaba (BABA), Waste Management (WM), Electronic Arts (EA), ATT&T (T), Axon (AXON), Vectrus (VEC), Facebook (now Meta; FB), Etsy (ETSY), and Lululemon (LULU). There were 12 preference sets considered: (1100), (1110), (1111), (3110), (3121), (3123), (3131), (1311), (1313), (2331), (1113), and (1232). The dots, which are plotted on the right-hand axis, identify the number of those sets in which the given stock appears. Thus, Walmart and Alibaba are in all 12; each of Vectrus, Facebook, Etsy, and Lululemon is in only one. The dark grey bars (plotted on the left-hand axis) represent the maximum weight assigned to the stock by the PGP technique across these preference sets, while the lighter bars are the actual PTA fund weights. Given the disparity between the PGP weights and the actual weights, we can see that the PTA fund is not taking advantage of any possible skewness or kurtosis benefits from its chosen stocks. (The PTA fund does have a maximum weight of 10%, but nonetheless does not weight its stocks optimally relatively to each other.)

DISCUSSION

Given the tracking error, active share, and attribution results, it seems clear that the PTA fund has *not* been acting as a closet indexer. This conclusion would be edifying for the donor who endowed the fund: he wanted to provide students an opportunity to run real money in an actively managed fund. However, the fund's investment policy statement seems inconsistent both with his wish and with the fund managers' actual behavior. In this section, we consider further evidence that both the IPS's risk/reward objectives and its broad-based benchmark, the Russell 3000, are inconsistent with the intent and practice of the fund. As shown above, the PTA fund has demonstrated a "persistent and prominent" bias toward growth investing. Since its benchmark is supposed to incorporate its "salient investment features," including its "significant exposures to particular sources of systematic risk," the fund should be using a growth benchmark (Bailey, *et al.*, 2007). To underscore the conclusion that the current benchmark is inappropriate, we ran regressions of the PTA fund's excess returns on the excess returns of three possible benchmarks: the actual Russell 3000 benchmark (VTHR); the large-cap growth Russell benchmark (VONG); and the 70/30 custom benchmark which we created based on the rolling style results from our returns-based analysis (we call this benchmark "VMIX"). We used the 20-year constant maturity Treasury series as the risk-free rate, and ran the regressions over the since-inception period (3/13/17-4/14/21).

First, we note that the betas for all three of these excess-return regressions were the same as the raw-return regressions shown in Table 1. Of more interest here are the annualized alphas and the implied information ratios (IRs). The alpha for the custom VMIX benchmark was -0.9% per year, with an IR of -0.35, indicating both a good fit for the benchmark and index-like performance by the fund. In contrast, using the fund's actual Russell 3000 benchmark, the alpha was 4.6% with an IR of 2.25 (the VONG's values were -2.6% and -0.96). Clearly, the assessment of the fund's performance will vary wildly depending on the benchmark used, and the poorly matched Russell 3000 makes it difficult to "differentiate managers...who add value through investment insights from those who do not" (Bailey, *et al.*, 2007). We also see the benefit of the custom VMIX benchmark over the poorly specified VTHR using Bailey, *et al.*'s (2007) active/style breakdown, which describes a portfolio's return as the sum of the market, the style contribution, and the active management contribution. Defining the style contribution as the difference between the portfolio's benchmark and the market (for us, [VMIX – VTHR]) and the active component as the difference between the portfolio and its benchmark [PTA fund – VMIX]), the authors assert that a good benchmark leads to active and style components that are uncorrelated. Our correlation is 0.074 ($R^2 = 0.0055$). The correlation between style and "error" (for us, [PTA fund – VTHR]) should be significantly positive; ours is 0.62 ($R^2 = 0.38$). VMIX therefore functions well as a benchmark for the PTA fund.

In contrast, the Russell 3000 benchmark prescribed in the PTA fund's IPS does not reflect the fund's style. Similarly, the risk/reward objectives reflect neither that style nor the active mandate of the fund in general. The IPS for the PTA fund mandates active, long-only, domestic equity investing. Since as "an active fund, a target return above the agreed upon benchmark is appropriate," the IPS sets an objective of 100 bp above the Russell 3000. As noted earlier, many SMFs do not set excess return targets; for example, the University of Connecticut's fund "does not presume that students will be able to beat the market on a consistent basis," focusing instead "on delivering high-quality, practical education to students" (Ghosh, *et al.*, 2020). Nonetheless, our target is not inconsistent with Cremers and Petajisto's (2009) findings that the most successful active mutual funds—which the authors assert do exhibit active skill—outperform their benchmarks by 1.51% to 2.40%. In particular, they find that high active share/low tracking error active funds in their sample beat their benchmarks by about 134 bp/year.

However, our fund's tracking error is not low. Despite the IPS's specified maximum of 200 bp, we have already seen that its actual since-inception TE against its benchmark is almost 1,200 bp. The prescribed 200-bp maximum places our fund's objective completely within the lowest tracking error bracket in Cremers and Petajisto's (2009) classification (whose highest category is "14%+"). While low tracking error can be associated with active management, the authors assert that this requires "large but diversified" positions away from the index—a requirement neither the conception of (via the IPS) nor the history of the PTA fund supports. In Petajisto (2013), there is no such fund at active share above 50%, the "theoretical minimum a pure active manager could have," and, in Cremers and Petajisto (2009), active share above 60% (the upper cutoff for their "very low" levels of active share) is never associated with tracking error of less than 2%. In fact, Petajisto's (2013) most salient examples of closet indexers have tracking errors of 3.1% and 4.4%. Thus, the tracking error objective for the PTA fund seems wildly inconsistent with its active mandate. Horstmeyer (2020) reports objectives that are more realistic for a fund like ours: the George Mason University SMF is to "exceed" its S&P500 benchmark with a tracking error of less than 750 bp. Thus, both the TE and the alpha targets for this fund are more liberal than ours. They also suggest that our operating history and tracking error record are not unique, and that revising our objectives in light of that history would not put our fund out of the mainstream. In any case, our analysis forces us to conclude that while our PTA fund is not, in fact, a closet indexer, its IPS incongruously apparently wishes it were.

As a final investigation into the appropriateness of our risk/return objectives, we ran two simulations to assess the probability that both the fund's excess return and tracking error targets could be met simultaneously. For the first simulation, we used the actual sixteen stocks held by the fund at the end of our sample period, the fund's Russell 3000 benchmark, and return data for our full sample period (August, 2016 through April, 2021). We used the fund's portfolio weights as of the last day of our sample, and assumed daily rebalancing. For the second simulation, we used the fund's portfolio returns and our VMIX custom benchmark for the period from March, 2017 (when the fund started investing) to April, 2021. In both cases, we used Crystal Ball's batch fit feature to fit distributions and estimate correlations for the return series (see Charnes, 2012), and our outputs were annual "alpha" (PTA fund return – benchmark return), tracking error, and an indicator variable that equaled one when both return and risk targets were met simultaneously. We ran 10,000 trials in each simulation. In *none* of those 20,000 trials did both the tracking error and the incremental return meet the guidelines of the IPS. For the stock-level simulation, the tracking error's distribution was positively skewed, leptokurtic, and had a mean of 22.5%; the incremental return was negatively skewed, leptokurtic, and had a mean of 27.8%. Obviously, these simulated results—especially the returns—were dramatically affected by the extreme experience of 2021 (see Table 1); the mean for the portfolio-level incremental return trials is a very-different -60 bp. Average tracking error for those, though, is still high, at 32.3%. Overall, the simulations underscore what has been clear throughout: the restrictions in the PTA fund's IPS are inconsistent with active management in general and with the fund's management in particular.

CONCLUDING COMMENTS

Student-managed funds are now expected elements of finance education, and while there are various ways to structure these funds, most seem to focus on active management through stock picking. Our university's fund, the PTA fund, has such an active mandate, but it also has a very restrictive tracking error objective. In this paper, we analyze the returns of our fund to determine whether our student managers are acting as true active managers or as effective closet indexers. We find that while they are running an active, highly concentrated fund, they are not—and cannot be—meeting the risk constraints of the fund's investment policy statement, since the IPS effectively mandates closet indexing. To evaluate our fund's performance, we used both traditional professional portfolio assessment techniques, such as returns- and holdings-based analysis, as well as more unconventional methods such as polynomial goal programming. We structured our approach using Cremers and Petajisto's (2009) taxonomy, which is based on tracking error and on their novel metric, active share. We found that—with over 90% active share—the fund is definitely active, in accordance with the initial donor's wishes. However, it is being run as a large-cap growth portfolio, which is inconsistent with its Russell 3000 benchmark. More seriously, its tracking error has been between three and seven times larger than the constraint imposed in the investment policy statement—a constraint that effectively requires the fund to be a closet indexer. Our simulations demonstrate that it is not possible for the fund to simultaneously meet all of the constraints in its governing documents.

Given our results, we offer the following three suggestions to advisors of student-managed funds. First, confirm that the fund's benchmark is consistent with the actual investment approach employed. Second, clearly and explicitly identify the relevant performance metrics in advance. Third, ensure that the fund's investment policy statement reflects empirical market realities. The first two of these are standard procedure for professional funds (see, for example, Maginn, *et al.*, 2007). However, while many prior papers have stressed the need for students to be conversant with their fund's policy statement (see, for example, Phillips, *et al.*, 2020, and Daugherty and Vang, 2015), we know of none that has highlighted the even more basic requirement that the statement be appropriate in the first place. If the purpose of the fund is to have students make active bets—as it is for our fund—then they must have the risk budget to accommodate that mandate.

There are several limitations of our analysis. First, for the last two years, students have been unable to work with each other and with their faculty mentors as envisaged, given that our residential campus has been frequently closed because of the covid pandemic. The resulting disruptions to our academic routine have been profound, and their ramifications for the fund—and for the complete curriculum—are as yet unknowable. We suspect, however, that fewer meetings have meant fewer stock pitches and less turnover than we would otherwise have seen. A more easily appreciated limitation is our (necessary) focus on a young fund. Our sample period starts from the fund's inception, and therefore encompasses its incubation period. Indeed, the fund still does not have a rigorous formal sell discipline, standardized stock-pitch structure, or comprehensive report template. The inconsistencies between the investment policy statement's mandates and the fund's actual investment approach undoubtedly stem partly from the steep initial learning curve for a new fund. Our research is one part of that ongoing development process.

Our next research steps involve characterizing the behavioral biases that are both possible and extant in our fund's management team. For example, are our younger students—who can join the club given our lack of coursework and experience requirements—more averse to active risk, as in Chevalier and Ellison (1999; described in Taylor, 2004)? Are their more experienced counterparts better at picking stocks, as in Daugherty and Vang (2015), or are they more liable to "gravitate toward 'closet indexing,' structuring portfolios with only modest deviations from the market, ensuring both mediocrity and survival" (Swensen, 2009)? Perhaps most importantly, is the make-up of our executive teams—dominated by male students—causing us to accept too much risk and reap too little reward, as in Barber and Odean (2001)? Better understanding who makes our portfolio decisions and how they do it will help us position the fund for improved performance in the future.

REFERENCES

Ammermann, P., P. Gupta, and Y. Ma (2020) "The Learning Experience Continues: Two Decades and Counting for CSULB's SMIF Program," *Managerial Finance*, vol. 46(4), p.513-529.

Bailey, J. V., T. M. Richards, and D. E. Tierney (2007) "Evaluating Portfolio Performance," in *Managing Investment Portfolios: A Dynamic Process*, 3rd ed., by John L. Maginn, Donald L. Tuttle, Dennis W. Mcleavey, and Jerald E. Pinto, John Wiley and Sons, Hoboken, N.J.

Barber, B. M. and T. Odean (2001) "Boys Will Be Boys: Gender, Overconfidence, and Common Stock Investment," *Quarterly Journal of Economics*, February, p. 261-292. Volume 116, Number 1

Bergquist, B. D., D. L. Keig, and T. J. Wilkinson (2020) "Student Investment Groups: Enhancing the Liberal Arts Experience," *Managerial Finance*, vol. 46(4), p. 489-498.

Betker, B. and T. W. Doellman (2020) "The SIM Program at Saint Louis University: Structure, Portfolio Performance, and Use of LinkedIn to Maintain an Alumni Network," *Managerial Finance*, vol. 46(5), p. 624-635.

Block, S. B. and D. W. French (1991) "The Student-Managed Investment Fund: A Special Opportunity in Learning," *Financial Practice and Education*, Spring, p. 35-40.

Boughton, C. M. and K. L. Jackson (2020) "Variation in the Structure and Administration of Student-Managed Investment Funds," *Managerial Finance*, vol. 46(4), p. 481-488.

Brown, D. C. and S. W. Davies (2015a) "The Rise of Closet Indexing," working paper, March 12. Available at: http://www.eief.it/files/2015/03/brown.pdf. Accessed 5/22/21.

Brown, D. C. and S. W. Davies (2015b) "Moral Hazard in Active Asset Management: A Negative Consequence of Index Investing," working paper, September 24. Available at SSRN Website: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2489113. Accessed 5/24/21.

Bruce, B. and Greene, J. (2014), Trading and Money Management in a Student-Managed Portfolio, Elsevier Monographs, Amsterdam.

Center for Research in Security Prices (CRSP) (2021) "Knowledge Base" from CRSP website: http://www.crsp.org/products/documentation/knowledge-base-1; accessed 7-1-21.

Charlton, W., Earl, J., & Stevens, J. (2015). Expanding Management in Student Managed Investment Funds. *Journal of Financial Education*, 41(2), p. 1-23.

Charnes, J. (2012) *Financial Modeling with Oracle Crystal Ball and Excel*, Wiley Finance, Hoboken, N.J.

Clinebell, J., & Murphy, J. (2016). Student Managed Investment Funds: The Perspectives of Alumni. *Journal of Financial Education*, 42(3-4), 304-312.

Cremers, K. J. M. and A. Petajisto (2009) "How Active is Your Fund Manager? A New Measure that Predicts Performance," working paper, March 31. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=891719. Accessed 6/8/21.

Davies, R. J., H. M. Kat, and Sa Lu (2004) "Fund of Hedge Funds Portfolio Selection: A Multiple-Objective Approach," Alternative Investment Research Centre, City University, London, working paper #0018, July 14.

D'Souza, F. P. and M. A. Johnson (2019) "Structuring an Endowment-Allocated Student Managed Fund Course," *Business Education and Accreditation*, vol. 11(1), p. 51-60.

Daugherty, M. S. and D. O. Vang (2015) "Using Performance Data to Evaluate Student Learning in a Student Managed Investment Fund," *Journal of Higher Education Theory and Practice*, vol. 15(2), p. 85-95.

Evans, R. B. (2004). "Does alpha really matter? Evidence from mutual fund incubation, termination and manager change." working paper.

Frino, A., D. R. Gallagher, and T. N. Oetomo (2004) "The Index Tracking Strategies of Passive and Enhanced Index Equity Funds," working paper, March 4. Available at: https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.200.5841&rep=rep1&type=pdf; accessed 7/2/21.

Gastineau, G. L., A. R. Olma, and R. G. Zielinski (2007) "Equity Portfolio Management," in *Managing Investment Portfolios: A Dynamic Process*, 3rd ed., by John L. Maginn, Donald L. Tuttle, Dennis W. McLeavey, and Jerald E. Pinto, John Wiley and Sons, Hoboken, N.J.

Ghosh, C., P. Gilson, and M. Rakatomavo (2020) "Student Managed Fund (SMF) at the University of Connecticut," *Managerial Finance*, vol. 46(4), p. 548-564.

Gradisher, S., D. R. Kahl, J. M. Clinebell, and J. L. Stevens (2016) "Fiduciary and Legal Considerations for Student-Managed Investment Funds," *Journal of Education for Business*, vol. 91(2), p. 83–89.

Haddad, M. M., A. L. Redman, and N. S. Gullett (2020) "IVA Investment Challenge: An Experiential Learning Program," *Managerial Finance*, vol. 46(5), p. 610-623.

Horstmeyer, D. (2020) "Formation of a Student Managed Investment Fund: Risk Management and Oversight," *Managerial Finance*, vol. 46(5), p. 685-691.

Hysmith, R. T. (2017) "Impact of Student-Managed Investment Fund Participation on Financial Knowledge: Financial Satisfaction and Financial Behavior," unpublished dissertation, March. Available at: https://www.proquest.com/openview/b1446890b0e441ee4df67909423f6ad7/1?pq-origsite=gscholar&cbl=18750&diss=y; accessed 7/8/21.

Kemalbay, G., C. M. Ozkut, and C. Franko (2011) "Portfolio Selection with Higher Moments: A Polynomial Goal Programming Approach to ISE-30 Index," Istanbul University Econometrics and Statistics e-Journal.

Lai, K. K., L. Yu, and S. Wang (2006) "Mean-Variance-Skewness-Kurtosis-based Portfolio Optimization," Proceedings of the First International Multi-Symposiums on Computer and Computational Sciences.

Lawrence, E.C. (1990), "Learning Portfolio Management by Experience: University Student Investment Funds," *The Financial Review*, vol. 25(1), p. 165-173.

Lawrence, E. C. (2008) Student-managed Investment Funds: An International Perspective," *Journal of Applied Finance*, 18 (2), p. 67-83.

Livingston, L. S. (2019) "Skewness, Cryptocurrency, and Peer-To-Peer Loans: An Asset Allocation Exercise for a Unique Student-Managed Fund," *Business Education and Accreditation*, vol. 11(1), p. 29-50.

Lynch, K. (2021) "Value vs. Growth: Widest Performance Gap on Record," Morningstar Commentary, 1/11/21. Available at: https://www.morningstar.com/articles/1017342/value-vs-growth-widest-performance-gap-on-record; accessed 7/29/21.

Maginn, J. L., D. L. Tuttle, D. W. McLeavey, and J. E. Pinto (2007) *Managing Investment Portfolios: A Dynamic Process*, 3rd. ed., John Wiley and Sons, Hoboken, New Jersey.

Neely, W. P. and P. L. Cooley (2004) "A Survey of Student Managed Funds," *Advances in Financial Education*, vol.2 (1), p. 1-9.

Petajisto, A. (2013) "Active Share and Mutual Fund Performance," *Financial Analysts Journal*, vol. 69(4), p. 73-93.

Phillips, M., J. Volker, and S. Cockrell (2020) "A Student-Managed Investment Fund Course Design Structure," *Managerial Finance*, vol. 46(4), p. 473-480.

Swensen, D. F. (2009) *Pioneering Portfolio Management: An Unconventional Approach to Institutional Investment*, Free Press, New York.

Tashjian, E. (2020) "The Essence of Investing: Experiential Education with a Student-Run portfolio," *Managerial Finance*, vol. 46(4), p. 530-547.

Taylor, J. (2004) "A Note on Closet Indexing," *Journal of Economics and Business*, vol. 56(6), p. 431-441.

Winston, W. L. (2016) *Microsoft Excel 2016: Data Analysis and Business Modeling*. Microsoft Press, Redmond, WA.

ACKNOWLEDGMENTS

Mazaltov-Ast and Livingston would like to acknowledge the support of the Agricola fund at the University of Puget Sound.

BIOGRAPHY

Lynda S. Livingston is a professor of finance at the University of Puget Sound. She is a founding member of Four Horsemen Investments.

Shirley Mazaltov-Ast is a senior pursuing a Bachelors of Science at the University of Puget Sound studying business leadership and economics with a minor in math. She is a current member of Four Horsemen Investments.

Business Education & Accreditation
Vol. 14, No. 1, 2022, pp. 61-72
ISSN: 1944-5903 (print)
ISSN: 2157-0809 (online)

www.theIBFR.COM

A COMPARISON OF FOUR AND SIX-YEAR GRADUATION RATES AT COLLEGES AND UNIVERSITIES

Jeff Anstine, North Central College

ABSTRACT

Tuition at colleges and universities has increased enormously over the past few decades. Thus, more than ever before, it is very important for students to not just graduate, but graduate in the expected four years to minimize tuition payments and lessen student debt. When institutions of higher education recruit students they state that they will be the 'Class of 202' assuming that they graduate in four years. However, these same schools report graduation rates at 150 percent of the expected completion time, which is in six years. Thus, there is a disconnect between the implied, promised graduation rate and the actual graduation rate provided to the public. This research examines factors that differ on the impact of the reported six-year graduation rates and four-year ones.*

JEL: A20, A22, Z18

KEYWORDS: Graduation Rates, Colleges, Universities, Higher Education

INTRODUCTION

Over the past two decades there has been a significant amount of research on higher education focusing on dozens of areas. The most recent lines of inquiry have explored the inequities faced by historically marginalized groups and developing programs that provide more access to these groups. Related, since a significant number of students feel there is a lack of support, there is a growing interest in addressing Diversity, Equity and Inclusion issues. In addition, there is growing concern about the mental health of college students, particularly depression and anxiety, impacting students, which has only been exacerbated by the pandemic.

However, perhaps the largest concern for most colleges and universities, their administrators and students is the large increases in tuition over the past twenty years. Students, and their parents are increasingly questioning the value of a college degree. In order to minimize the amount of money spent on tuition it is imperative that students graduate in the shortest amount of time, which ideally is the standard, expected, in four years.

There are other obvious additional benefits to completing a degree in four, instead of five or six years. It can greatly reduce a student's debt, which is a huge and growing problem for millions of individuals. The additional year or two can be used for other activities. For some it may be a 'gap year' to travel or to do volunteer work. For others, it means that graduate or law school can begin earlier. And for the majority of individuals it allows them to start their careers sooner, thus giving them more time to advance in their chosen areas, add more human capital and increase wages.

TIAA Institute published a recent report addressing the need for colleges and universities to do a better job of providing improved outcomes for students. Creighton, et. al. (2021) in a study of the New American

Colleges and Universities (NACU) look at the cost of delivering a degree. They look at efficiency based factors such as the cost of attending, expenses, total enrollment, demographics and financial aid. Applying metrics used in other businesses they determine the Return on Investment (ROI) based on Net Tuition revenue minus Instructional Expenses divided by Institutional Expenses. The authors' find that schools that use data to make decisions and focus on cost containment perform better than other schools. In addition to better serving students, being more efficient means that these institutions are better prepared for the future. My paper provides another metric that shows how well colleges and universities are serving their students. This research looks at what variables are more likely to increase graduation rates in four, not six, years. In addition to helping families prepare for college, it is also useful for administrators and policy makers.

The next part of this paper provides a literature review on graduation rates of colleges and universities, mostly focusing on articles using regression. Next, I discuss the data used in the analysis, why this research is unique and summarize the variables. Then I explain, the econometrics used and results. I finish the paper with how my findings add to the literature, the limitations to this research and provide suggestions for future areas of study.

LITERATURE REVIEW

There is a large and growing volume of literature on graduation rates at institutions of higher education in the United States. Early papers focused on how specific programs can impact the graduation rate at a particular school. Over the past decade there has been much about the use of High Impact Practices (HIPS) on the importance of mentoring, advising, new student orientation, tutoring, internships and other actively intensive programs. Some individuals, particularly those from disadvantaged back grounds are not only less likely to attend college but also are also less likely to finish. So there are studies looking at programs improving outcomes for first-generation students and minorities.

The majority of studies analyzing graduation rates rely on case studies, anecdotal evidence, qualitative analysis or look at one or a small number of schools. Regression is a valuable econometric technique used for predicting particular outcomes holding other variables constant. Over the past ten years, there has been more research using regression with data sets examining graduation rates from a macro level.

Some researchers are skeptical about the validity of results using regression. Horn and Lee (2016) show that if the model is identified appropriately it does give correct results. Pike and Graunke (2015) control for time-invariant characteristics, time-varying institutional and time-varying cohort characteristics. Results from this research finds that differences in average student ACT and SAT scores mainly determine retention rate differences.

Anstine (2013) examined almost 1,400 institutions of higher education. Results showed that the percentage of faculty that is full-time positively impact graduation rates. He also examined the relative importance of some HIPs on graduation rates. Holding constant dozens of variables such as standardized test scores, teaching centers and learning groups do not increase graduation rates overall for colleges and universities. But if institutions are examined separately using interaction terms these variables do improve outcomes for comprehensive universities but not for research universities or liberal arts colleges.

Millea et. al. (2018) also use regression to look at retention and graduation rates at one school from 1998 to 2004. As with all research on the topic, they find that academic variables such as higher high school GPA and ACT scores significantly impact graduation rates. In addition, they conclude that if their school had smaller class sizes and provided more financial aid, this would also improve outcomes. This makes sense, however, most schools do not have the resources available to implement this suggestion.

Over thirty states tie some funding of public colleges to their performance, usually measured by six year graduation rates. In Crisp, et. al. (2018) regression analysis they determine the variables of most importance for a non-selective college. As with other studies they find that the percentage of students who are full time, socio-economic variables and institutional revenue all increase graduation rates. In addition, they find that religious affiliation also has an impact.

Hester and Ishitani (2018) believe that efficiency can be measured by determining what variables are statistically significant in predicting six year graduation rates. Results show that executive to staff ratios, faculty teaching load and class size all improve retention and graduation. Importantly, where money is spent matters. Support for instruction improves outcomes, but money spent on research, public service and student services does not.

Hajrasouliha and Ewing (2016) look at retention and graduation through the lens of a college campuses design. As with research using regression, they find that campus living, the percentage of students living on campus in particular increases outcomes. In addition, they look at how land is used, how spread out the campus is, how connected it is, its configuration and greenness. They also find that 'greener' campuses and those in a more urban area also contribute to student satisfaction.

All of the studies discussed in this literature review (and as far I as I know all other research in this area) has only examined graduation rates over a six year period of time. This research adds to the body of knowledge by also looking at the factors that determine graduation rates in four years.

DATA AND METHODOLOGY

This research builds on a previous paper by Anstine and Seidman (2017). In that paper, the author's looked at the relative importance of social and financial variables in determining graduation rates. Variables such as standardized test scores, class size and demographic characteristics all known to be important were held constant. Results showed that while the total number of male and female sports was statistically significant it was financial characteristics that had the largest impact on graduation rates. Schools with higher percentages of students receiving Pell Grants and those with need based financial aid had significantly lower graduation rates than other colleges and universities.

The data used in this paper was obtained from five different sources, then organized in Excel and transferred to SPSS for the econometrics. With the help of a student researcher I started with data from the U.S. News and World Report, since this source had the majority of the variables needed for the analysis. There were some missing variables and observations so we then went directly to the U.S. Department of Education to fix this. The Integrated Post-Secondary Education Data System (IPEDS) has information from four-year institutions on graduation rates, student faculty ratio and dozens of other variables. It should be noted that the U.S. News and World Report also gets its data from IPEDS, thus the data is correct and consistent. All of the data was collected in the summer of 2016.

Information about faculty salaries is from The Chronicle of Higher Education. From 2012 to the present there has been information on almost 4,000 colleges and universities published annually available for purchase. The data includes information on salary by rank, Assistant Professor, Associate Professor and Full Professor. Since there is a large degree of multi-collinearity between the three types, I have only included Associate Professors salary in the data. The link to the document is in the references.

The Carnegie Foundation classification system provided data on the type of institution each school is: National University, Regional University or Liberal Arts College. The Carnegie Foundation for the Advancement of Teaching and the American Council on Education (ACE) have collected this data since 1973 and update it approximately every four or five years to ensure its accuracy.

Table 1: Description of Structural, Selectivity and Demographic Variables

Variable	Description of Variables	Data Source	Numb Obs	Min	Max	Mean	Number
GradRate4years	Four year graduation rate at each school.	US News (from IPEDS)	277	0.02	0.90	0.44	
GradRate6years	Six year graduation rate at each school.	US News	277	0.07	0.95	0.59	
STRUCTURAL							
Regional	If the school is a regional university (yes=1)	Carnegie Foundation	277	0	1	0.56	155
LibArt	If the school is a Liberal Arts College (yes=1)	Carnegie	277	0	1	0.23	64
National	If the school is a national university (yes=1)	Carnegie	277	0	1	0.21	55
Private	If the school is Private (yes=1)	Author	277	0	1	0.67	185
Urban	If the school is in an urban location (yes=1)	Author	277	0	1	0.46	127
Suburban	If the school is suburban (yes=1)	Author	277	0	1	0.25	66
Rural	If the school is in a rural location (yes=1)	Author	277	0	1	0.29	80
IA	If the school is in Iowa (yes=1)	Author	277	0	1	0.09	22
IL	If the school is in Illinois (yes=1)	Author	277	0	1	0.16	45
IN	If the school is in Indiana (yes=1)	Author	277	0	1	0.12	33
KY	If the schools is in Kentucky (yes=1)	Author	277	0	1	0.09	22
MI	If the school is in Michigan (yes=1)	Author	277	0	1	0.11	31
MN	If the schools is in Minnesota (yes=1)	Author	277	0	1	0.08	23
MO	If the school is in Missouri (yes=1)	Author	277	0	1	0.11	31
OH	If the school is in Ohio (yes=1)	Author	277	0	1	0.14	39
WI	If the school is in Wisconsin (yes=1)	Author	277	0	1	0.10	28
SELECTIVITY							
Retention	The percentage of first-time, full-time undergraduate students who returned to school for their second year.	USNews-IPEDS	277	0.46	0.99	0.764	
PerClsU20	Percentage of classes with fewer than 20 students	USNews-IPEDS	277	0.234	0.94	0.570	
StudFac	Student faculty ratio	USNews-IPEDS	277	6	26	14.01	
AccpRate	Percent of students accepted out of those who applied	USNews-IPEDS	277	0.07	1	0.688	
Ave ACT	The average ACT scores of the entering students.	USNews-IPEDS	277	16	33	23	
Fresh10	Percentage of students who were in the top 10 percent of their high school class	USNews-IPEDS	277	0.02	0.98	0.225	
DEMOGRAPHIC							
PerFemale	Percentage of students who are female	USNews-IPEDS	277	0	1	0.557	
OutState	Percent of students from another state.	USNews-IPEDS	277	0	0.93	0.274	
IntNatl	Percentage of students from another Country	USNews-IPEDS	277	0	0.30	0.04	
Black	Percentage of students who reported Black	USNews-IPEDS	277	0	0.83	0.081	
Asian	Percentage of students who reported Asian	USNews-IPEDS	277	0	0.23	0.028	
Hispanic	Percentage of students who are Hispanic	USNews-IPEDS	277	0	0.44	0.058	
White	Percentage of students who reported White	USNews-IPEDS	277	0.25	0.97	0.727	
Other	Percentage of students who reported as Native American, Pacific Islander, Multiracial, or did not report	USNews-IPEDS	277	0	0.3	0.067	

Table 1 provides an overview of the data used in the analysis. The first and second columns in this table lists and defines the structural, selectivity and demographic variables. Structural variables are those Column one and two list the variables. Structural variables do not change. Selectivity are quality indicators. Demographic provides information on the types of students. Column three lists the source of the data, followed by the number of observations, minimum and maximum values, the mean and, if applicable the number in the category. .

The data is from schools in the Midwestern U.S. See the appendix for more details on this. Thus, community colleges, universities that specialize in only upper level transfer students and graduate schools, and all those dedicated to Cooking, Art and Business, in addition to for-profit schools are not included. This paper focuses on the top four-year institutions of higher learning in the Midwest United States. Thus, it is only non-profit, public and private colleges and universities in the study.

There are likely differences between colleges and universities in states due to economic and political characteristics. The location of the state it is in was calculated directly by the author. Geography with respect to density may also be important so I looked up if the schools are located in urban, suburban or rural areas. There are a total of two hundred and seventy seven schools in the data set.

Table 1 and Table 2 define each variable, gives it source, the minimum and maximum values, then the mean and the number in the category if it is qualitative. The first two variables provided are the graduation rates of each school in four years and six years. It is interesting to note that the difference between the two averages is an astonishingly high fifteen percentage points, whereas the minimum and maximum only differ by five percentage points.

Table 2: Description of Faculty, Student Body and Financial Variables

Variable	Description of Variables	Data Source	Numb Obs	Min	Max	Mean
FACULTY						
PerFTFac	Percent of faculty that is full time	USNews-IPEDS	277	0.249	1	0.790
AasocProf	Average Associate Professor Salary	Chronicle of Higher Education	277	36549	117600	66293
STUDENT BODY						
Students	Number of undergraduate students	USNews-IPEDS	277	543	44201	6057
PerLiveOn	Percentage of students who live in campus housing	USNews-IPEDS	277	0	1	0.502
StudOrg	Number of student organizations per capita	USNews-IPEDS	277	4	21	15.41
PerFrat	Percent of male students in a Fraternity	USNews-IPEDS	277	0	0.77	0.088
PerSor	Percent of female students in a Sorority	USNews-IPEDS	277	0	0.67	0.095
MAatlPerCap	Male Athletes per capita	USNews_IPEDS	277	0	0.44	0.11
FAthlPercCap	Female Athletes per capita	USNews_IPEDS	277	0	0.3	0.07
PerStudFT	Percentage of undergraduates who attend full-time	USNews-IPEDS	277	0.007	1	0.834
FINANCIAL						
PerHaveNBA	Percent determined to have financial need	USNews-IPEDS	277	0.37	1	0.715
Pellgrant	Percentage of undergraduates receiving a Pell Grant	USNews-IPEDS	277	0.062	0.926	0.342
PerBorrow	Percent of graduating students who have borrowed	USNews-IPEDS	277	0.08	0.96	0.722
PerCapEndow	End-of-year endowment value per full-time equivalent student	USNews-IPEDS	277	507	950232	45944
ALUMGvRt	Percentage of alumni who give to the school	USNews-IPEDS	277	0.01	0.51	0.129

Table 2 continues providing an overview of the data. Column one and two list the variables. Faculty gives percentage full time and salary. Student body variables look at the relative importance of student groups and other items that connect students to their college. Financial shows how well off schools are and the number of lower income students. Column three gives the source of the data, next is the number of observations, then minimum and maximum values, the mean.

To simplify the large number of variables I have put them into categories. Structural are variables that do not change, such as the type of institution, public or private and location, urban, suburban or rural. Selectivity variables include the retention rate, student faculty ratio and percentage of students in the top ten percent of their high school class.

Demographic variables include the percentage of students who are: female, out of state and international and reported ethnic background. Faculty variables show the percentage of faculty that is full time and the average associate professors salary. Student body variables provide information on social factors of the schools, such as the percentage of students in Greek life, percentage that live on campus and per capita number of athletes. Financial variables include the percentage of students receiving Pell Grants and the school's per capita endowment.

Since this is the first paper comparing four and six-year graduation rates it is worthwhile to examine some simple statistics. The (statistically significant) correlation coefficient between the four and six year graduation rates is 0.9. This makes perfect sense, we would expect it to be high, but not perfect. If the correlation was one there would be no difference in the variables that impact four and six year graduation rates thus showing the importance of this study.

I then calculated the difference between each schools six year and four year graduation rates to see what schools had the smallest difference between the two and those that had the largest to see if any patterns exist. Not surprisingly, of the top thirty-five with the numbers closest together almost all were true Liberal Arts colleges, with Center College, Hanover and Macalester all just two percentage points different between the two years. These schools do not offer degrees in Engineering and other areas that typically take over four years to complete. In addition, they tend to bring in students from more well to do families, from good high schools that prepare them specifically for college. They also have large endowments that enable them to subsidize other high achieving students. The University of Norte Dame and the University of Chicago, both with a five percentage point's difference were the only two schools in the top twenty-five that are not Liberal Arts colleges.

There are seventeen schools with a difference of thirty percentage points or more, between their four and six year graduation rates and another twenty with a difference of twenty-five to twenty-nine. The largest was Kettering University in Michigan which had a difference of fifty-two percent. (I double checked data). The majority of these institutions are comprehensive universities, such as the University of Wisconsin schools and those in Michigan such as Grand Valley State and Western Michigan University. Again, this is not surprising since these schools tend to be less selective, admitting a higher percentage of middle and lower class students, many of whom have to work to pay for school thus taking more time to complete.

RESULTS

Regressions below examine the impact of different variables on colleges and universities four-year graduation rate (equation one) and six-year graduation rate (equation two). Due to multi collinearity not all variables that are listed in Tables 1 and 2 were included. None-the-less there are more variables for this analysis than in the majority of studies using regression. A total of forty-five explanatory variables were included. Both regressions had identical independent variables where I included all control variables such as standardized test scores, demographic characteristics, student body information and socio-economic data. I will first discuss the similarities of the regressions, then the differences.

$$4\, Year\ Graduation\ rate i = \beta 0 + \beta 1\, Structural + \beta 2\, Selectiviy + \beta 3\, Demographic + \beta 4\, Faculty + \beta 5\, Student\ Body + \beta 6\, Financial + \varepsilon i \quad (1)$$

$$6\, Year\ Graduation\ rate i = \beta 0 + \beta 1\, Structural + \beta 2\, Selectiviy + \beta 3\, Demographic + \beta 4\, Faculty + \beta 5\, Student\ Body + \beta 6\, Financial + \varepsilon i \quad (2)$$

One method of determining how good econometric results are is to compare them to comparable existing research. In both of the regressions with six year and four-year graduation rates as the dependent variable, all of the main control variables all have the expected sign and are statistically significant. It is well known that schools with higher standardized test scores (ACT and/or SAT) have higher graduation rates than colleges and universities with lower scores which is the case in both regressions. In addition, socio-economic variables such as the percent of students receiving need based aid and Pell Grants negatively influence graduation rates. This is also consistent with all studies using these variables looking at graduation rates with regression (see the articles in the literature review for some examples).

Table 3: Regression with Graduation Rate in Four Years as Dependent Variable

Independent Vars	Coefficients	Std. Error	T-statistics
Intercept	-0.559	0.169	-3.301
LibArts [a]	0.036	0.018	2.017**
National	0.008	0.017	0.467
Private (yes=1)	0.102	0.024	4.243***
Urban [b]	-0.011	0.013	-0.858
Suburban	-0.015	0.016	-0.933
IA [c]	-0.019	0.026	-0.731
IN	0.030	0.026	1.156
KY	-0.115	0.026	-4.403***
MI	-0.084	0.022	-3.830***
MN	-0.001	0.026	-0.020
MO	-0.067	0.025	-2.721***
OH	-0.044	0.021	-2.036**
WI	-0.072	0.024	-3.025***
PerclsU20	-0.037	0.059	-0.618
StudFac	0.000	0.003	0.129
AccptRate	0.038	0.039	0.963
AVEACT	0.028	0.004	6.275***
PerFemale	0.331	0.063	5.257***
OutState	0.007	0.036	0.198
INatl	-0.090	0.161	-0.561
Black [d]	-0.093	0.085	-1.094
Asian	-0.189	0.218	-0.866
Hispanic	-0.224	0.124	-1.809**
Other	-0.127	0.129	-0.982
PerFTFac	-0.007	0.062	-0.113
AssocProfSal	0.0000024	0.000	3.131***
PerStudsFT	0.001	0.000	1.358
PerLVonCamp	0.085	0.035	2.405***
PerHaveNeed	-0.175	0.082	-2.136**
PerCapOrgs	1.944	1.352	1.438
PerFrat	0.073	0.104	0.709
PerSor	-0.012	0.096	-0.125
MAtlPerCap	0.182	0.130	1.404
FAthPerCap	0.004	0.205	0.018
PellGrant	-0.002	0.001	-2.412***
PerBorrow	0.032	0.061	0.522
PerCapEnd	0.000000059	0.000	0.730
AlumGiveRt	0.180	0.104	1.732

* Number of observations: 276; R-Squared: 0.867; Adjusted R-Squared: 0.846; F-Statistic: 40.8

*This table shows the regression with the Four-Year Graduation Rate as the dependent variable. The excluded category for the type of school is Regional (a). Rural is the omitted classification for location (b). Illinois is the excluded category for states (c). The excluded category for race is white (d). I follow the standard format for levels of statistical significance for one-tailed tests: * the 10% level, ** the 5 percent level, and *** the 1% level.*

There are other independent variables in my regressions that also have the anticipated sign and are statistically significant. We would expect that institutions that pay their faculty more have workers who are more committed to their school and be more involved with students. This is the case, where professors' salaries positively impact graduation rates. In addition, the dummy variable coefficient of if the school is private the graduation rate is higher and statistically significant. It is interesting that the difference is about ten percent higher with the dependent variable as four-year graduation rate and only six percent higher with the dependent variable as six-year graduation rate.

It has been documented that some demographic characteristics are important in determining graduation rates. Both of my regressions confirm this, with colleges that have a higher percentage of female students having higher graduation rates. In addition, schools that are more social or connected have improved outcomes over those that are less connected. The variable, percentage of students living on campus also positively impacts graduation rates.

Though not perfect, R-squared and Adjusted R-squared provide a good measure of how much all of the independent variables explain the schools graduation rates. The R-squared for both regressions is very high, 0.867 (four year) and 0.872 (six year) providing more evidence of the impact of the large number of variables in explaining graduation rates. The corresponding Adjusted R-squared are 0.846 and 0.852 showing the included variables are relevant.

While this data has a large number of variables, there are certainly others that would be useful such as information on unemployment rates and other economic factors that may influence students dropping out of college. While not perfect, the information on states provides a proxy for some of this. Compared to Illinois, Kentucky, Missouri and Ohio all have negative and statistically significant coefficients in both the four year and six year regressions.

It is possible that this is picking up some of the economic, social and demographic differences between states. For example, Illinois has a large number of corporate headquarters, has a significant number of technology jobs and a relatively diversified economy compared to Kentucky, Missouri and Ohio, thus possible making it more attractive to stay in school and graduate. In addition, Illinois is a solid Blue state, consistently voting Democratic compared to the other three states that are Red or Purple possibly showing how education, voting and politics may be related to if students are more or less, likely to stay in school and graduate.

The coefficients that are different between the regressions with the dependent variable as four year and six years graduation rates are if the school is Liberal Arts (positive and statistically significant for four year, not statistically significant for six) and Michigan and Wisconsin (negative and statistically significant for four year, not statistically significant with six.) This is not surprising given that the visual inspection of each school individually showed this to be the case. Controlling for the type of university and other variables, it may be the case that students in these more 'Blue Collar' worker states will graduate but take longer than the four years to do so.

Many First-generation students tend to be pragmatic and realistic when pursuing a degree in higher education. Quite a few seek degrees in Nursing, Management, Accounting and other 'practical' areas. Whereas students attending Liberal Arts colleges get degrees in Philosophy and other disciplines to learn for the sake of learning or are preparing for Law School or another graduate degree. Results imply that students from more privileged backgrounds can focus on their classes more than others from less affluent backgrounds.

It is also well documented that certain minorities, including Hispanic and Black students graduate at lower levels than other groups. The Hispanic coefficient is negative in both regressions, but its level of statistical significance is at the five percent level with the dependent variable as four-year graduation rate, but only at the ten percent level with the dependent variable as six-year graduation rate. The Black coefficient is not statistically significant in Table 3 but is in Table 4, as is the other category.

Table 4: Regression with Graduation Rate in Six Years as Dependent Variable

Independent Vars	Coefficients	Std. Error	T-statistics
Intercept	-0.268	0.69	-1.972
LibArts [a]	0.004	0.009	0.251
National	0.008	0.019	0.570
Private (yes=1)	0.065	0.185	3.349***
Urban [b]	-0.002	-0.006	-0.174
Suburban	-0.006	-0.015	-0.451
IA [c]	0.002	0.003	0.080
IN	0.025	0.050	1.232
KY	-0.088	-0.150	-4.175***
MI	-0.019	-0.037	-1.092
MN	-0.016	-0.026	-0.753
MO	-0.042	-0.081	-2.148**
OH	-0.036	-0.076	-2.076**
WI	0.004	0.008	0.222
PerclsU20	-0.059	-0.056	-1.248
StudFac	0.000	-0.007	-.151
AccptRate	0.015	0.015	0.474
AVEACT	0.031	0.560	8.777***
PerFemale	0.181	0.118	3.589***
OutState	-0.060	-0.081	-2.069***
INatl	0.085	0.021	0.662
Black [d]	-0.195	-0.095	-2.863
Asian	-0.172	-0.036	-0.980
Hispanic	-0.146	-0.050	-1.474*
Other	-0.239	-0.065	-2.309**
PerFTFac	0.017	0.014	0.734
AssocProfSal	0.00000178	0.000	2.898***
PerStudsFT	0.000	0.032	0.958
PerLVonCamp	0.061	0.099	2.164**
PerHaveNeed	-0.183	-0.141	-2.786***
PerCapOrgs	1.590	0.062	1.466*
PerFrat	0.068	0.083	0.820
PerSor	-0.009	-0.007	-0.116
MAtlPerCap	0.034	0.020	0.326
FAthPerCap	-0.058	-0.021	-0.353
PellGrant	-0.001	-0.080	-2.023**
PerBorrow	0.054	0.042	1.103
PerCapEnd	0.000000049	0.000	0.029
AlumGiveRt	0.033	0.017	0.392
Number of observations: 276; R-Squared: 0.872; Adjusted R-Squared: 0.852; F-Statistic: 42.7			

*This table shows the regression with the Six-Year Graduation Rate as the dependent variable. The excluded category for the type of school is Regional (a). Rural is the omitted classification for location (b). Illinois is the excluded category for states (c). The excluded category for race is white (d). I follow the standard format for levels of statistical significance for one tailed tests: * the 10% level, ** the 5 percent level, and *** the 1% level.*

While almost all of the signs on the coefficients and statistical significance of the explanatory variables in both regressions make perfect sense, there are two exceptions. We would expect that alumni who had a good experience at their college would be more likely to donate to it. The coefficient for alumni giving is positive and statistically significant in the four-year dependent variable regression, as expected, but is not in the regression with the six-year dependent variable regression.

Another variable that we would expect to be consistent in the two regressions is the percentage of students that are out of state. This coefficient is negative and statistically significant in Table 4. Perhaps students who go to college in another state are more likely to get homesick and do not have the family support so leave school at higher rates. Then we would expect this to also be the case for the Table 3 regression, but it is not.

There are other variables that were unavailable that I would have liked to include in the regressions, particularly the percentage of students in each school that are First-Generation. It is possible that the regressions do pick up on some of this. For example the Hispanic coefficient is negative and statistically significant for the four year graduation rate, but not for the six year. Schools with a higher percentage of Hispanic students, who are also likely to be First-Generation and have family obligations may graduate but in more years.

CONCLUSION

With fewer students attending colleges and universities every year, bureaucratic administrations growing and tuition rising it is more important than ever to pay close attention factors influencing costs and the value of higher education. The goal of this paper was to compare factors impacting schools reported six year graduation rate with the implied one of four years for prospective students. Results showed that there are some variables that impact four, but not six year graduation rates.

While this research is an important first step in comparing four and six year graduation rates, there are some significant limitations of this study. First is that it would be good to include more variables that likely contribute such as the different majors that are offered at each school, the percentage of First-generation Students and other differences between each institution.

While regression is a very important tool in isolating the impact of one variable on graduation rates, holding other variables constant, additional research should look at a more micro level. For example, it was shown that schools in Michigan and Wisconsin have lower four year graduation rates even holding dozens of variables constant. It would be helpful to determine exactly what some of the factors are and see if there are any policies that might help address the issues. Nonetheless, this is a good first step looking at why the reported six year graduation rate may not accurately mirror the anticipated four year one.

APPENDIX

This paper looks at institutions of higher education in the United States. These are colleges and universities that provide education after high school (post-secondary institutions). Officially a college is different from a university, often a college being part of a university. While these are not the same, I follow conventional custom and use the terms interchangeably.

This paper focuses on states in the Midwest. There are different definitions of the Midwestern United States based on geography. Some of these use locations such as states lying above the 37th parallel between the Appalachian and Rocky Mountains. Others use description of the land such as the Great Plains. The states in this analysis does not follow a specific definition but is a loose combination of various definitions.

REFERENCES

Anstine, Jeff. (2013) "Graduation Rates at U.S. Colleges and Universities: A Large Data Set Analysis," *Business Education & Accreditation* vol. 5(2) p. 55-64

Anstine, Jeff and Megan Seidman. (2017) "Graduation Rates at U.S. Colleges and Universities In The Midwest," *Business Education & Accreditation* vol. 9(1) p. 43-54

The Carnegie Foundation, https://carnegieclassifications.acenet.edu/

The Chronicle of Higher Education, https://store.chronicle.com/collections/che-data/products/faculty-salaries?cid=CS-CHE-EM-FSAL-2022-July27
e03&mkt_tok=otmxluvlqs0ymtgaaagf33eirvjfvca_b6expsj-y31e49lhlljeqzup3xglasbexboa8g2pdc61l0jc9lxtc4_svcacnuuwf-ygdcmckbtxbigzbeyhqcivvpsuxst0_w

Creighton, Sean, LaMendola, Nick and Kechichian, Avo (2021) "Benchmarks for efficiency and sustainability: A research study on the cost of delivery" TIAA Research Institute *Dialogue* issue 179

Crisp, Gloria, Doran, Erin and Salis-Reyes, Nicole (2018) "Predicting Graduation Rates at 4-year Broad Access Institutions Using a Bayesian Modeling Approach" *Research In Higher Education* vol. 59 p. 133-155

Hajrasouliha, Amir and Ewing, Reed (2016) "Campus Does Matter: The Relationship of Student Retention and Degree Attainment to Campus Design" Planning for Higher Education Journal vol. 4(3) p. 30-45.

Hester, Brian and Ishitani, Terry (2018) "Institutional Expenditures and State Economic Factors Influencing 2012-2014 Public University Graduation Rates" *Planning for Higher Education Journal* vol. 4, p. 41-47

Horn, Aaron, and Giljae Lee. (2016) "The Reliability and Validity of Using Regression Residuals to Measure Institutional Effectiveness in Promoting Degree Completion," *Research In Higher Education* vol. 57(4) p. 469-496

Millea, Meghan, Wills, R., Elder, A. and Molina, D. (2018) "What Matters in College Student Success? Determinants of College Retention and Graduation Rates," *Education* vol. 138(4) p. 309-322.

National Center for Education Statistics, The Integrated Post-Secondary Education Data System (IPEDS), 2016, retrieved from http://nces.ed.gov/ipeds/

National Center for Education Statistics, The Integrated Post-Secondary Education Data System (IPEDS), 2021, retrieved from http://nces.ed.gov/ipeds/

Pike, Gary, and Steven Graunke. (2015) "Examining the Effects of Institutional and Cohort Characteristics on Retention Rates," *Research In Higher Education* vol. 56(2) p. 146-165

U.S. News & World Report, retrieved from http://colleges.usnews.rankingsandreviews.com/best-colleges

ACKNOWLEDGEMENTS

The author would like to thank the two anonymous reviewers and the editor for helpful comments that improved this paper.

BIOGRAPHY

Jeff Anstine is a Professor of Management. He teaches a wide range of classes from Business Statistics to Environmental Economics. His research examines topics in higher education and on environmental issues. He can be reached at North Central College in Naperville IL, 60540.

REVIEWERS

The IBFR would like to thank the following members of the academic community and industry for their much appreciated contribution as reviewers.

Hisham Abdelbaki, University of Mansoura- Egypt
Isaac Oluwajoba Abereijo, Obafemi Awolowo University
Naser Abughazaleh, Gulf University for Science and Technology
Nsiah Acheampong, University of Phoenix
Iyabo Adeoye, National Horticultural Research Instittute, Ibadan, Nigeria.
Michael Adusei, Kwame Nkrumah University of Science and Technology
Paolo Agnese, LUISS Guido Carli University
Haydeé Aguilar, Universidad Autónoma de Aguascalientes
Mohd Ajlouni, Yarmouk University
Sylvester Akinbuli, University of Lagos
Anthony Akinlo, Obafemi Awolowo University
Efiong Akwaowo, Ashford University
Yousuf Al-Busaidi, Sultan Qaboos University
Khaled Aljaaidi, Universiti Utara Malaysia
Alawiya Allui, Prince Sultan University
Hussein Al-Tamimi, University of Sharjah
Paulo Alves, CMVM, ISCAL and Lusofona University
Ghazi Al-weshah, Albalqa Applied University
María Antonieta Andrade Vallejo, Instituto Politécnico Nacional
Jeff Anstine, North Central College
Olga Lucía Anzola Morales, Universidad Externado de Colombia
Antonio Arbelo Alvarez, Universidad de la Laguna
Hector Luis Avila Baray, Instituto Tecnologico De Cd. Cuauhtemoc
Graciela Ayala Jiménez, Universidad Autónoma de Querétaro
Fabiola Baltar, Universidad Nacional de Mar del Plata
Samar Baqer, Kuwait University College of Business Administration
Nagib Bayoud, Tripoli University
Ahmet Bayraktar, Rutgers University
Daniel Boylan, Widener University
James E. Briley, Northeastern State University
Kyle Brink, Western Michigan University
Giovanni Bronzetti, University of Calabria
Ana Cecilia Bustamante Valenzuela, Universidad Autónoma de Baja California
Raymond Cairo, University of Surrey
Albanelis Campos Coa, Universidad de Oriente
Carlos Alberto Cano Plata, Universidad de Bogotá Jorge Tadeo Lozano
Alberto Cardenas, Instituto Tecnológico de Cd. Juarez
Edyamira Cardozo, UNIVERSIDAD NACIONAL EXPERIMENTAL DE GUAYANA
Sheila Nora Katia Carrillo Incháustegui, Universidad Peruana Cayetano Heredia
Emma Casas Medina, Centro de Estudios Superiores del Estado de Sonora
Benjamin Castillo Osorio, Universidad del Sinú-Sede Monteria
María Antonia Cervilla de Olivieri, Universidad Simón Bolívar
Priyashni Chand, University of the South Pacific
Surya Chelikani, Quinnipiac University
Yahn-shir Chen, National Yunlin University of Science and Techology, Taiwan
Bea Chiang, The College of New Jersey
Shih Yung Chou, University of the Incarnate Word
Te-kuang Chou, Southern Taiwan University
Monica Clavel San Emeterio, University of La Rioja
Caryn Coatney, University of Southern Queensland
Iyanna College of Business Administration,
Michael Conyette, Okanagan College
Cipriano Domigo Coronado García, Universidad Autónoma de Baja California
Semei Leopoldo Coronado Ramírez, Universidad de Guadalajara
Esther Eduviges Corral Quintero, Universidad Autónoma de Baja California
M. T. Coto, Intl. Academic Sevices
Dorie Cruz Ramirez, Universidad Autonoma del Estado de Hidalgo
Tomás J. Cuevas-Contreras, Universidad Autónoma de Ciudad Juárez
Edna Isabel De La Garza Martinez, Universidad Autónoma De Coahuila
Hilario De Latorre Perez, Universidad Autonoma De Baja California
Javier de León Ledesma, Universidad de Las Palmas de Gran Canaria - Campus Universitario de Tafira
Huang Department of Accounting, Economics & Finance,
Rajni Devi, The University of the South Pacific
Hilario Díaz Guzmán, Universidad Popular Autónoma del Estado de Puebla
Cesar Amador Díaz Pelayo, Universidad de Guadalajara, Centro Universitario Costa Sur
E. M. Ekanayake, Bethune-Cookman University
Mahmoud Elgamal, College of Business Administration - Kuwait University
Avilés Elizabeth, CICESE
Prince Ellis, Argosy University
Esther Enriquez, Instituto Tecnologico de Ciudad Juarez
Ernesto Escobedo, Business Offices of Dr. Escobedo
Zaifeng Fan, University of Wisconsin whitewater
Perrine Ferauge, University of Mons
Olga Ferraro, University of Calabria
Ernesto Geovani Figueroa González, Universidad Juárez del Estado de Durango
Carlos Fong Reynoso, Universidad de Guadalajara
Ana Karen Fraire, Universidad De Guadalajara
Carmen Galve-górriz, Universidad de Zaragoza
Teresa García López, Instituto De Investigaciones Y Estudios Superiores De Las Ciencias Administrativas
Blanca Rosa Garcia Rivera, Universidad Autónoma De Baja California
Helbert Eli Gazca Santos, Instituto Tecnológico De Mérida
Peter Geczy, AIST
Lucia Gibilaro, University of Bergamo
Denisse Gómez Bañuelos, CESUES
Carlos Alberto González Camargo, Universidad Jorge Tadeo Lozano
Hector Alfonso Gonzalez Guerra, Universidad Autonoma de Coahuila
María Brenda González Herrera, Universidad Juárez del Estado de Durango
Ana Ma. Guillén Jiménez, Universidad Autónoma de Baja California
Hongtao Guo, Salem State University
Zi-yi Guo, Wells Fargo Bank, N.A.

REVIEWERS

The IBFR would like to thank the following members of the academic community and industry for their much appreciated contribution as reviewers.

Araceli Gutierrez, Universidad Autonoma De Aguascalientes
Danyelle Guyatt, University of Bath
Glen Hansen, Utica College
Peter Harris, New York Institute of Technology
Shahriar Hasan, Thompson Rivers University
Zulkifli Hasan, Islamic University College of Malaysia
Fariza Hashim, Prince Sultan University
Peng He, Investment Technology Group
Niall Hegarty, St. Johns University
Andreina Hernandez, Universidad Central de Venezuela
Arturo Hernández, Universidad Tecnológica Centroamericana
Jorge Hernandez Palomino, Universidad Autónoma de Mexico
Alejandro Hernández Trasobares, Universidad de Zaragoza
Azucena Leticia Herrera Aguado, Universidad Tecnológica de Puebla
Claudia Soledad Herrera Oliva, Universidad Autónoma de Baja California
Paulin Houanye, University of International Business and Education, School of Law
Daniel Hsiao, University of Minnesota Duluth
Xiaochu Hu, School of Public Policy, George Mason University
Qianyun Huang, City University of New York - Queens College
Jui-Ying Hung, Chatoyang University of Technology
Fazeena Hussain, University of the South Pacific
Luis Enrique Ibarra Morales, Universidad Estatal de Sonora
Alma Delia Inda, Universidad Autonoma Del Estado De Baja California
Stoyu Ivanov, San Jose State University
Shilpa Iyanna, Abu Dhabi University
Mercedes Jalbert, The Institute for Business and Finance Research
Terrance Jalbert, University of Hawaii
Gaspar Alonso Jiménez Rentería, Instituto Tecnológico de Chihuahua
Lourdes Jordán Sales, Universidad de Las Palmas de Gran Canaria
Gricelda Juarez-Luis, Instituto Politécnico Nacional
Tejendra N. Kalia, Worcester State College
Gary Keller, Eastern Oregon University
Ann Kelley, Providence college
Ann Galligan Kelley, Providence College
Ifraz Khan, University of the South Pacific
Katherine Kinkela, Iona College
Halil Kiymaz, Rollins College
Susan Kowalewski, DYouville College
Bohumil Král, University of Economics-Prague
Jan Kruger, Unisa School for Business Leadership
Christopher B. Kummer, Webster University-Vienna
András Kun, University of Debrecen
Mei-Mei Kuo, JinWen University of Science & Technology
Miguel Ángel Latorre Guillem, Universidad Católica de Valencia "San Vicente Mártir"
John Ledgerwood, Embry-Riddle Aeronautical University
Yen-Hsien Lee, Chung Yuan Christian University
Santiago León Ch., Universidad Marítima del Caribe
Victor Lewis, National University
Luis Alberto Limón Valencia, Instituto Tecnológico Superior de Cajeme
Shulin Lin, Hsiuping University of Science and Technology
Yingchou Lin, Missouri University of Science and Technology
Lynda Livingston, University of Puget Sound
Graciela López Méndez, Universidad de Guadalajara-Jalisco
Virginia Guadalupe López Torres, Universidad Autónoma de Baja California
Melissa Lotter, Tshwane University of Technology
Ma. Cruz Lozano Ramírez, Universidad Autónoma De Baja California
Xin (Robert) Luo, Virginia State University
Andy Lynch, Southern New Hampshire University
Angel Machorro Rodríguez, Instituto Tecnológico de Orizaba
Cruz Elda Macias Teran, Universidad Autonoma de Baja California
Eduardo Macias-negrete, Instituto Tecnologico de Ciudad Juarez
Firuza Madrakhimova, University of North America
Aracely Madrid, ITESM, Campus Chihuahua
Deneb Magaña Medina, Universidad Juárez Autónoma de Tabasco
Abeer Mahrous, Cairo university
Tshepiso Makara, University of Botswana
Ioannis Makedos, University of Macedonia
Carlos Manosalvas, Universidad Estatal Amazónica
Gladys Yaneth Mariño Becerra, Universidad Pedagogica y Tecnológica de Colombia
Gladys Marquez-Navarro, Saint Louis University
Omaira Cecilia Martínez Moreno, Universidad Autónoma de Baja California-México
Jesús Apolinar Martínez Puebla, Universidad Autónoma De Tamaulipas
Jesus Carlos Martinez Ruiz, Universidad Autonoma De Chihuahua
Mario Jordi Maura, University of Puerto Rico
Francisco Jose May Hernandez, Universidad del Caribe
Aurora Irma Maynez Guaderrama, Universidad Autonoma de Ciudad Juarez
Romilda Mazzotta, University of Calabria
Mary Beth Mccabe, National University
Linda Margarita Medina Herrera, Tecnológico de Monterrey. Campus Ciudad de México
Marco Mele, Unint University
Alaitz Mendizabal Zubeldia, Universidad del País Vasco
Fidel Antonio Mendoza Shaw, Universidad Estatal De Sonora
Gulser Meric, Rowan University
Miwa Merz, San Jose State University
Avi Messica, Holon Institute of Technology
Cameron Montgomery, Delta State University
Oscar Montiel, Universidad Autonoma de Ciudad Juarez
Oscar Javier Montiel Mendez, Universidad Autonoma de Ciudad Juarez
Juan Nicolás Montoya Monsalve, Universidad Nacional de Colombia-Manizales
Cleamon Moorer, Madonna University
Sandip Mukherji, Howard University
Jennifer Mul Encalada, Universidad Autónoma De Yucatán
Gloria Muñoz Del Real, Universidad Autonoma de Baja California
Alberto Elías Muñoz Santiago, Fundación Universidad del Norte
Tony Mutsue, Iowa Wesleyan College
Ertambang Nahartyo, UGM
Linda Naimi, Purdue University
Arezoo Nakhaei, Massey University

REVIEWERS

The IBFR would like to thank the following members of the academic community and industry for their much appreciated contribution as reviewers.

Cheedradevi Narayanasamy, National University of Malaysia
Erwin Eduardo Navarrete Andrade, Universidad Central de Chile
Gloria Alicia Nieves Bernal, Universidad Autónoma del Estado de Baja California
Bertha Guadalupe Ojeda García, Universidad Estatal de Sonora
Erica Okere, Education Management Corp
Erika Olivas, Universidad Estatal de Sonora
Dennis Olson, Thompson Rivers University
Idowu Emmanuel Olubodun, Obafemi Awolowo University
Godwin Onyeaso, Shorter University
Ramona Orastean, Lucian Blaga University of Sibiu-Romania
Erick Orozco, Universidad Simon Bolivar
Rosa Martha Ortega Martínez, Universidad Juárez del Estado de Durango
Diaeldin Osman, Alabama State University
José Manuel Osorio Atondo, Centro de Estudios Superiores del Estado de Sonora
Carmen Padilla-Lozano, Universidad Catolica de Santiago de Guayaquil
Vera Palea, University of Turin
Julian Pando, University of the Basque Country
Daniel Paredes Zempual, Universidad Estatal de Sonora
Dawn H. Pearcy, Eastern Michigan University
Luz Stella Pemberthy Gallo, Universidad del Cauca
Andres Pereyra Chan, Instituto Tecnologico De Merida
Eloisa Perez, MacEwan University
Iñaki Periáñez, Universidad del Pais Vasco (spain)
Hector Priego Huertas, Universidad De Colima
Pina Puntillo, University of Calabria (Italy)
Rahim Quazi, Prairie View A&M University
Anitha Ramachander, New Horizon College of Engineering
Charles Rambo, University of Nairobi
Eric Amin Ramirez Castillo, Universidad Autónoma Benito Juarez de Oaxaca
Prena Rani, University of the South Pacific
Alma Ruth Rebolledo Mendoza, Universidad De Colima
Kathleen Reddick, College of St. Elizabeth
Oscar Bernardo Reyes Real, Universidad de Colima
Maurizio Rija, University of Calabria.
Carmen Rios, Universidad del Este
Juan Carlos Robledo Fernández, Universidad EAFIT-Medellin/Universidad Tecnologica de Bolivar-Cartagena
Natalia G. Romero Vivar, Universidad Estatal de Sonora
Fabrizio Rossi, University of Cassino and Southern Lazio
Humberto Rosso, Universidad Mayor de San Andres
Matthew T. Royle, Valdosta State University
José Gabriel Ruiz Andrade, Universidad Autónoma de Baja California-México
Antonio Salas, Universidad Autonoma De Chihuahua
Rafiu Oyesola Salawu, Obafemi Awolowo University
Claudia Nora Salcido, Facultad de Economía Contaduría y Administración Universidad Juarez del Estado de Durango
Paul Allen Salisbury, York College, City University of New York
Leire San Jose, University of Basque Country
Juan Manuel San Martín Reyna, Universidad Autónoma de Tamaulipas-México
Francisco Sanches Tomé, Instituto Politécnico da Guarda
Edelmira Sánchez, Universidad Autónoma de Ciudad Juárez
Celsa G. Sánchez, CETYS Universidad
Deycy Janeth Sánchez Preciado, Universidad del Cauca
María Cristina Sánchez Romero, Instituto Tecnológico de Orizaba
María Dolores Sánchez-Fernández, Universidade da Coruña
Luis Eduardo Sandoval Garrido, Universidad Militar de Nueva Granada
I Putu Sugiartha Sanjaya, Atma Jaya Yogyakarta University, Indonesia
Pol Santandreu i Gràcia, Universitat de Barcelona, Santandreu Consultors
Victor Gustavo Sarasqueta, Universidad Argentina de la Empresa UADE
Jaime Andrés Sarmiento Espinel, Universidad Militar de Nueva Granada
Sunando Sengupta, Bowie State University
Jesus Otoniel Sosa Rodriguez, Universidad De Colima
Adriana Patricia Soto Aguilar, Benemerita Universidad Autonoma De Puebla
Smita Mayuresh Sovani, Pune University
Alexandru Stancu, University of Geneva and IATA (International Air Transport Association)
Jonathan Stewart, Abilene Christian University
Jiří Strouhal, University of Economics-Prague
Vichet Sum, University of Maryland -- Eastern Shore
Qian Sun, Kutztown University
Edith Georgina Surdez Pérez, Universidad Juárez Autónoma de Tabasco
Diah Suryaningrum, Universitas Pembangunan Nasional Veteran Jatim
Andree Swanson, Ashford University
James Tanoos, Saint Mary-of-the-Woods College
Jesus María Martín Terán Terán Gastélum, Centro de Estudios Superiores del Estado de Sonora
Ranjini Thaver, Stetson University
Jeannemarie Thorpe, Southern NH University
Maria De La Paz Toldos Romero, Tecnologico De Monterrey, Campus Guadalajara
Alejandro Torres Mussatto, Senado de la Republica; Universidad de Valparaíso
Jorge Torres-Zorrilla, Pontificia Universidad Católica del Perú
William Trainor, East Tennessee State University
Md Hamid Uddin, University of Sharjah
Intiyas Utami, Satya Wacana Christian University
Ozge Uygur, Rowan University
Abraham Vásquez Cruz, Universidad Veracruzana
Rosalva Diamantina Vásquez Mireles, Universidad Autónoma de Coahuila
Angel Wilhelm Vazquez, Universidad Autonoma Del Estado De Morelos
Lorena Vélez García, Universidad Autónoma de Baja California
Alejandro Villafañez Zamudio, Instituto Tecnologico de Matamoros
Hector Rosendo Villanueva Zamora, Universidad Mesoamericana
Oskar Villarreal Larrinaga, Universidad del País Vasco/Euskal Herriko Universitatea
Delimiro Alberto Visbal Cadavid, Universidad del Magdalena
Vijay Vishwakarma, St. Francis Xavier University
Julian Vulliez, University of Phoenix
Janet L. Walsh, Birchtree Global, LLC

REVIEWERS

The IBFR would like to thank the following members of the academic community and industry for their much appreciated contribution as reviewers.

Ya-fang Wang, Providence University
Richard Zhe Wang, Eastern Illinois University
Richard Weaver, National University
Jon Webber, University of Phoenix
Jason West, Griffith University
Wannapa Wichitchanya, Burapha University
Veronda Willis, The University of Texas at San Antonio
Erico Wulf, Universidad de la Serena
Amy Yeo, Tunku Abdul Rahman College
Bingqing Yin, University of Kansas
Paula Zobisch, Forbes School of Business & Technology

HOW TO PUBLISH

Submission Instructions

The Journal welcomes submissions for publication consideration. Complete directions for manuscript submission are available at the Journal website www.theibfr.com/journal-submission/. Papers may be submitted for initial review in any format. However, authors should take special care to address spelling and grammar issues prior to submission. Authors of accepted papers are required to precisely format their document according to the journal guidelines, available at www.theibfr.com/guidelines/.

Authors submitting a manuscript for publication consideration must guarantee that the document contains the original work of the authors, has not been published elsewhere, and is not under publication consideration elsewhere. In addition, submission of a manuscript implies that the author is prepared to pay the publication fee should the manuscript be accepted.

Subscriptions

Individual and library subscriptions to the Journal are available. Please contact us by mail or by email to: admin@theibfr.com for updated information.

Contact Information

Mercedes Jalbert, Managing Editor
The IBFR
P.O. Box 4908
Hilo, HI 96720
editor@theIBFR.com

Website

www.theIBFR.com

PUBLICATION OPPORTUNITIES

REVIEW of BUSINESS & FINANCE STUDIES

Review of Business & Finance Studies

Review of Business & Finance Studies (ISSN: 2150-3338 print and 2156-8081 online) publishes high-quality studies in all areas of business, finance and related fields. Empirical, and theoretical papers as well as case studies are welcome. Cases can be based on real-world or hypothetical situations.

All papers submitted to the Journal are blind reviewed. Visit our website www.theibfr.com/our-journals/ for distribution, listings and ranking information.

BEA Business Education & Accreditation

Business Education and Acreditation (BEA)

Business Education & Accreditation publishes high-quality articles in all areas of business education, curriculum, educational methods, educational administration, advances in educational technology and accreditation. Theoretical, empirical and applied manuscripts are welcome for publication consideration.

All papers submitted to the Journal are blind reviewed. Visit our website www.theibfr.com/our-journals/ for distribution, listings and ranking information.

AT Accounting & Taxation

Accounting and Taxation (AT)

Accounting and Taxation (AT) publishes high-quality articles in all areas of accounting, auditing, taxation and related areas. Theoretical, empirical and applied manuscripts are welcome for publication consideration.

All papers submitted to the Journal are blind reviewed. Visit our website www.theibfr.com/our-journals/ for distribution, listings and ranking information.

REVISTA GLOBAL de NEGOCIOS

Revista Global de Negocios

Revista Global de Negocios (RGN), a Spanish language Journal, publishes high-quality articles in all areas of business. Theoretical, empirical and applied manuscripts are welcome for publication consideration.

AAll papers submitted to the Journal are blind reviewed. Visit our website www.theibfr.com/our-journals/ for distribution, listings and ranking information.