Measuring Faculty Productivity: Towards a Unified Methodology at the Departmental Level

by

Patricia Bartholomew

A dissertation submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Auburn, Alabama
December 12, 2015

Keywords: higher education departmental productivity, faculty productivity, productivity management, methodology, university mission

Approved by

Josè R. Llanes, Chair, Professor of Educational Foundations, Leadership and Technology
David C. DiRamio, Associate Professor of Educational Foundations, Leadership and Technology
Maria M. Witte, Professor of Educational Foundations, Leadership and Technology
Iryna Y. Johnson, Associate Director of Assessment

Abstract

A disproportionate number of departmental and faculty productivity indices in higher education rely solely on research. Productivity in other areas of faculty workload areas, like teaching and institutional and community service are either measured separately or ignored all together. This dissertation proposes a unified methodology to calculate higher education departmental productivity inclusive of the three primary areas of faculty responsibility: teaching, research, and service. This information better informs higher education administrators and department heads and helps to focus decision making in resource allocation by providing adequate benchmarking against peer and aspirant institutions while accounting for institutional mission.

Acknowledgements

I would like to express my deepest gratitude to my mentor, **Dr. José Llanes**, without whom this dissertation simply would not exist. The road to this point simply would not have been navigable without his expert guidance, wealth of knowledge, and generous spirit, not only in good times, but, crucially, when obstacles seemed insurmountable.

I would also like to thank the members of my dissertation committee, **Dr. David DiRamio, Dr. Maria Witte**, and **Dr. Iryna Johnson** for their insightful direction and invaluable feedback in the shaping this research. Additional special thanks to **Dr. John Jahera** for his time and efforts on my behalf in serving as University Reader.

The Department of Institutional Research at **Auburn University** deserves special recognition for the incredible work they do and the support they provided me, especially in helping me understand the (sometimes conflicting) challenges that university administrators face. Special recognition is reserved for **Dr. Iryna Johnson** and **Dr. Drew Clark** for their generosity in sharing their time and resources with me. Further I would like to express my appreciation to the department for the provision of my graduate research assistantship, was invaluable in allowing me to pursue my studies and crucial in shaping my research.

David Ramsey at **Academic Analytics** also merits special acknowledgement for his encouragement to think outside the box and his time and efforts on my behalf.

Altamese Stroud-Hill helped me immensely in complying with the Byzantine formatting requirements critical for dissertation success. Thank you for your professionalism, courtesy, and speed of response. I am indebted to you.

Lastly, but certainly not least, special thanks go to my children**, Maximilian and Lily Salow**, for many things, but especially the understanding beyond their years in practicing patience with their mother when this dissertation conflicted with playtime or when the quality of meals suffered. This dissertation is therefore dedicated to them. They earned it.

Table of Contents

Abstract ii

Acknowledgments iii

List of Tables v

List of Figures vi

List of Abbreviations vii

CHAPTER 1. INTRODUCTION 1

Historical Background 3

Statement of the Problem 8

Why a New Index is Important 9

Purpose of the Study 11

Research Questions 12

Significance of the Study 13

Limitations of the Study 14

Assumptions of the Study 15

Definition of Terms 15

Organization of the Study 16

CHAPTER 2. LITERATURE REVIEW 18

CHAPTER 3. METHODOLOGY 29

Purpose 31

CHAPTER 4. APPLICATION OF UNIFIED INDEX AND RESULTS 38

CHAPTER 5. SUMMARY AND DISCUSSION OF FINDINGS 45

References 50

Appendix A: Sample Delaware Study of Instructional Cost Data Collection Form 61

Appendix B: Methodological Matrix for U.S. News & World Report Rankings 63

Appendix C: Comprehensive Matrix of H-Index and its Variants 74

Appendix D: JCAR Definitions of Faculty Activities 79

Appendix E: Auburn University Institutional Review Board Application and Approval 81

List of Tables

Table 1 Example Statistics for Delaware Study: Individual Institution vs. National Averages 23
Table 2 Direct Instructional Cost per Student Credit Hour Taught: Delaware Study Benchmarks for Research Universities 25
Table 3 Workload Allocation Weights 33
Table 4 Data and Data Sources 36
Table 5 Comparison of Productivity Scores: University One: Teaching Mission 44
Table 6 Comparison of Productivity Scores: University Two: Service Mission 44

List of Figures

Figure 1 Frequency of "Faculty Productivity" in Published Books, 1900–2008 6

Figure 2 Departmental Comparison: 100% Research Metric 39

Figure 3 Departmental Comparison: Teaching, 50% and Research, 50% 40

Figure 4 Departmental Comparison: Research (60%), Teaching (20%), Service (20%)........ 41

Figure 5 Composite Z-Scores – Varying Weights by University Mission.............................. 42

Figure 6 Composite Z-Score for University 2 with Variable Service Weight 43

CHAPTER 1. INTRODUCTION

"The call for effectiveness … will be perceived by many in the university as the best current definition of evil," declared Clark Kerr, first Chancellor of the University of California, Berkeley. He then noted that higher education institutions cannot "escape the imperative of more effective use of resources" (Cohen & Kisker, 2010, p. 497). Groccia and Miller (2005) quote Tierney, "The word [productivity] itself is unsettling to many in the academic community, akin a dirty word or taboo subject" (p. xix). Departmental and faculty productivity are not immune. A full measure of derision for a concept lies in the number of jokes. A current one being, "there is a new P-word that has replaced Parking as the central campus evil. Its name is productivity" (Townsend & Rosser, 2007, p. 7). The attempt to measure faculty productivity at the departmental level and make improvements, i.e., assessment-for departmental management purposes is held in even more derision. "I have never met a faculty member who was excited about assessment, although rumor has it they exist" (Wehlburg, 2008, p. 39). The derision has reached such levels that the stages through which an institution cycles when dealing with assessment have been likened to the psychological steps, pioneered by Elizabeth Kübler-Ross, through which one cycles when dealing with death and dying. This observation elicits many knowing chuckles from faculty and administrators:

1. Denial – "It's a fad. If I ignore it, it will go away."
2. Acceptance – "Okay, I guess we have to do it."

3. Resistance – “I feel threatened. My department feels threatened. My campus feels threatened. Can I subvert it by not participating in the process or in some other way?”
4. Understanding – “Maybe we can learn something useful. Can we use what we’ve already been doing?”
5. Campaign – “We have a plan. Maybe it’s not perfect, but let’s get moving!”
6. Collaboration – “We have a plan with long range objectives that are clearly defined and based on our experience with assessment. We believe it works.” (Wehlburg, 2008, p. 40)

University faculty see their duty as tri-fold: teaching, conducting research, and service to the institutional and community (Middaugh 2001). The measure of success lies in the quality aspect of that work, which is notoriously difficult to objectively measure. So, it is no surprise that in a higher education setting, productivity has a negative connotation. Further, tackling the subject of faculty productivity is akin to the touch of the third rail, a potential career-ending action in academic administration where successful political management of the collaborative decision making process is key to advancement. No one likes to talk about it, much less base a research agenda on such a politically charged topic. So it is no wonder that there is a dearth of academic literature on the subject (Groccia & Miller, 1998). Yet state (and now the federal) legislatures continue to demand measurability of institutional productivity, transparency on expenditures and the translation of those expenditures into tangible outcomes. This pressure is particularly acute in times of economic downturn. At these times, it is critical that politicians show their constituencies that taxpayer money has been and is being well spent. Due to the lack of quality metrics to easily, succinctly, and clearly communicate productivity at the departmental and institutional level, education makes an easy target.

Historical Background

The large-scale shift away from the concept of faculty duties as a sole teaching activity can be traced back to eighteenth century Prussia. During this period, public confidence in universities critically eroded. This widespread public questioning of universities' educational relevance for society required a reactionary evolution in pedagogy for the university system to survive. *Wissenschaftsideologie* was developed by a group of German intellectuals of the time (i.e., Wilhelm von Humboldt, Johann Gottlieb Fichte) to improve educational productivity of its students and by extension, the reputation of the universities. Thus, interest in gauging faculty productivity was born (Turner, 1973, p. 247–253; Watson, 2011, p. 226–237).

Wissenschaftsideologie's main innovation was the introduction of the research imperative to higher education as a way to ensure faculty remain up-to-date on current evolutions in their respective disciplines and be required to contribute to further evolution. This new ideology viewed teaching and research as an inseparable "unity" (Turner, 1973, p. 247–248). Although the idea was revolutionary at the time, it was largely successful in reforming the Prussian educational system and was instrumental in propelling Prussian universities into leading intellectual centers in 1800s. The requisite research imperative for top colleges and universities still reverberates today in the structure and incentives of not only the modern U.S. higher education system, but worldwide.

When the scientific method began to flourish in the United States (perhaps in response to Darwin's *Origin of the Species* in 1859), the more intellectually curious U.S. students began to travel to Germany to study these new approaches to scholarship, bringing back with them an interest in the German model to education and scholarship (Heydinger & Simsek, 1992, p. 16). Influential names in US higher education at the time also had traveled to Germany and brought back the same persuasive ideas that underpinned *Wissenschaftsideologie*. George Ticknor (Dartmouth), Daniel Gilman (John's Hopkins), Andrew White (Cornell), Charles Eliot (Harvard), Theodore Woolsey (Yale), and G. Stanley Hall (Clark) are all examples of Americans influenced by German universities who were instrumental in converting US colleges to universities along the German model in order to improve education in the United States (Cohen & Kisker, 2010, p. 111–112). This interest pressured the faculty of clergy-dominated model of U.S. instruction of the time and its associated concept that teaching (Christianity) as the sole measure of faculty productivity.

The agricultural and industrial revolution further cemented the idea that research was critical to faculty duties and added the additional aspect of practical application of theoretical concepts, i.e. service and outreach activities. With a dawning comprehension of the importance of the need to compete internationally in food production as key to national strength and independence, the U.S. Congress enacted the Morrill Act of 1862. The Act created land grant institutions of higher education in the United States whose missions were now not only tied to teaching, but also the practical application of those teaching concepts and a service component to the local community to teach those applied concepts to the local farming community. This expanded the idea of what faculty should be producing from a life devoted to teaching Christian dogma (as in the clergy model) to a combined teaching and research institution (as under the Prussian model) began to include a service component designed to bolster community economies (Cohen & Kisker, 2010, p. 111–112; Heydinger & Simsek, 1992, p. 16; Shin, Toutkoutshian, & Teichler, 2011, p. 123–125).

Interest in defining and quantifying faculty productivity in the US has been around at least since the early 1900s in the United States. The topic appeared in 1910 with the Carnegie Foundation's initial report on medical schools, commonly referred to as the Flexner report. This report is widely recognized as the first US (and perhaps worldwide) attempt "to measure the efficiency and productivity of educational institutions in a manner similar to that of industrial factories" (Massy, Sullivan, & Mackie, 2013, p. 9). The excesses of the Mass Higher Education Era (1945–1970) and the G.I. Bill shifted the definition of productivity to "passing more students through to graduation" (Cohen & Kisker, 2010, p. 306). Then the economic problems of the 1970s began pre-occupying public and political attention in the United States when interest in faculty productivity exploded.

Figure 1 shows the number of times the terms "faculty productivity", "departmental productivity", and "academic productivity" shows up in print as a percentage of all two-word (bigram) phrases in Google's online book collection that are written in English and published in the United States. Google's attempt to digitally scan all known books and more information on its book collection (and the ensuing copyright controversy) and cooperation with libraries can be found on the Google Books website "about" tab.

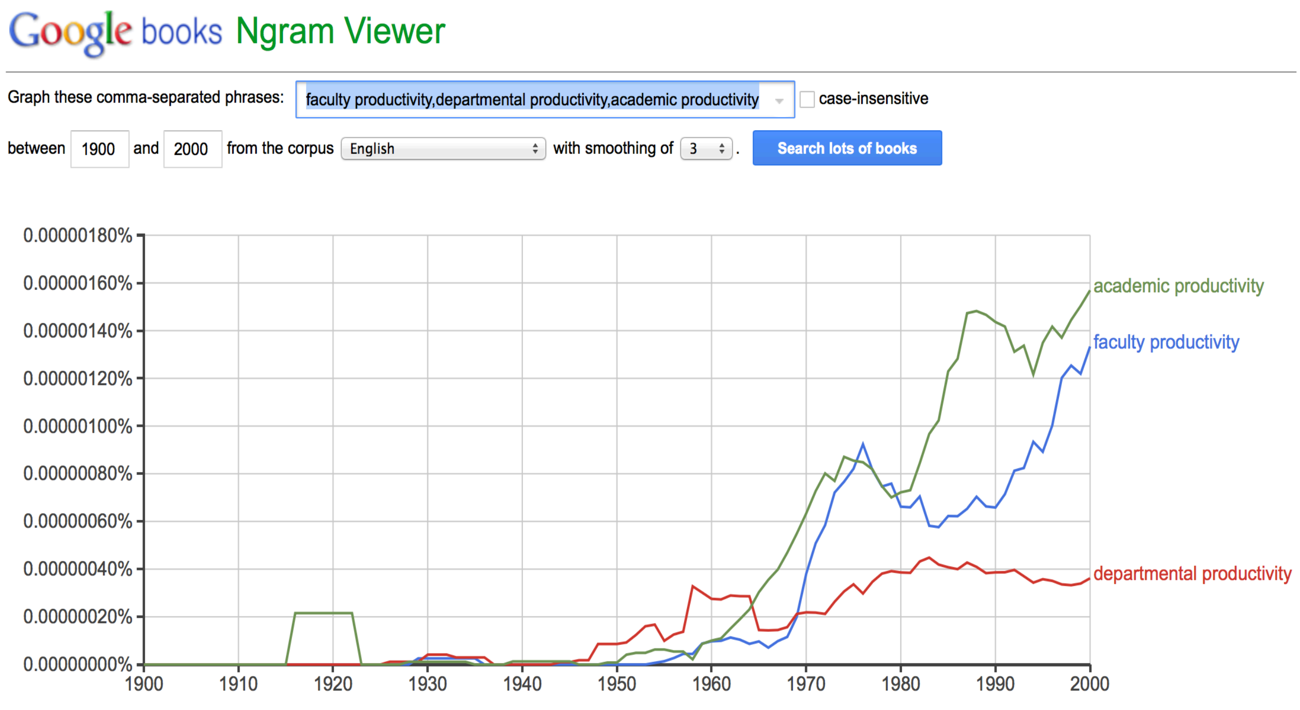

Figure 1. Frequency of "Faculty Productivity" in Published Books, 1900–2008

The number of times "faculty productivity" appears increases dramatically in the 1970s. Indeed the graph shows interest in faculty productivity to fluctuate in apparent positive correlation with economic hardship. The appearance of the phrase "faculty productivity" show an increase in books in 1930s and then largely disappear—only to reappear in the 1970s, when the dual occurrence of inflation and stagnant growth (stagflation) renewed interest in accountability.

Due to the "third rail" type concerns about tackling faculty productivity inside institutions of higher education, outside consultants were brought in to salve public and political demands for accountability. Businesses picked up where faculty feared to tread. Application of business principles to state institutions of higher education in the United States became a favorite revenue source for consultants during this time, spawning an increase of interest into methods of quantifying academic productivity by outside consulting firms (Cohen & Krisker, 2010, p. 495-498). Interest in the topic reached a peak in 1996 and began to fall off again in 2000 as can been seen in Figure 1. Unfortunately the data ends in 2008, although past experience and anecdotal evidence suggest that renewed interest in educational accountability and faculty productivity would have increased as a result of the 2008 crisis. Indeed the initiative led by the Obama administration in 2011 lends credence to this pattern. See White House document "Bringing Transparency to College Costs."

What few modern *academic* attempts to measure faculty productivity there are, have focused primarily on research productivity as the sole proxy for overall productivity (Baird, 1986; Creswell, 1985) at the expense of other important faculty activities like teaching and service. (It must be noted that classroom hours and student contact hours have been also widely used as a "crude index of gross productivity according to Cohen & Kisker, 2010, p. 358).

More recently Jorge Hirsch's H-index (2005) improved on the methodology by introducing a quality dimension of research. While this improvement is useful, it does not address the fact that most of the solutions and research into quantitative measures of faculty productivity at the worst ignore the teaching and service aspects completely, at the least calculate individual indices for two of the three aspects, much less all three aspects, leaving the administrator without guidance on how to evaluate both indices in tandem. (See Cohen and Kisker, 2010, p. 359–360, for an enlightening explanation as to why this is so.)

Time and time again surveys of faculty workload allocation show that faculty at institutions with differing missions allocate their time differently (Jordan, 1994; Blackburn, Bieber, Lawrence, & Trautvetter, 1991; Serpe, Large, Brown-Large, Newton, Kilpatrick, &

Mason, 2002). Not one quantitative method to date has attempted to combine measurement of all three faculty responsibilities into a properly weighted, single, easily understandable index.

Statement of the Problem

Lack of academic research into comparable productivity indices is the main problem. Cohen and Kisker (2010, p. 359–360) neatly summarize the reason for lack by identifying the stumbling blocks to the construction of comparable productivity indices:

> Productivity indices simply conflict with too many variables. Within the same department some instructors work harder at classroom teaching, whereas others are highly entrepreneurial, bringing in grants for research or student support. One professor may turn out a half-dozen short papers a year, whereas another takes five years to complete a book. Some spend many hours counseling students, whereas others find the activity not worth the time they could better devote to preparing a distance education program. Add to all that the differences among academic departments or disciplines in the same institution, and the differences among institutions in the same higher education system and it is easy to conclude that neither the faculty as a whole nor any enlightened administrators would accept an index of productivity applied indiscriminately across the board.

Yet, accountability is essential. To make rational decisions on resource allocation, one must have good data and reliable metrics on which to base decisions. Education is no different in this regard. Measurement of faculty productivity is important. The unique nature of education presents challenges to quantifying and comparing faculty productivity, and these challenges will be discussed in detail in the coming chapters. However, those challenges do not negate the fact that having good metrics on which to base decisions is critical to the successful allocation of resources and management of any institution of higher education. Legislators and the public deserve transparent accountability for the taxes they pay to support education. Groccia and Miller (1998) make the valid point that "…for some institutions, survival may depend on developing a reasoned approach to increasing productivity" (p. 2).

The recent emphasis on continuous assessment methods and assessment's rise in importance in the departmental and institutional management process, reflects a passing of the burden of proof for transparency and accountability from the institution at the macro level to the program level, placing the burden of proof on individual departments. A unified index would be a welcome tool to add to the departmental toolbox as evidence of performance measurement and attempts at continuous improvement. Groccia and Miller (2005) and many others argue that the "ongoing assessment of productivity and quality gains is essential for making rational and productive decisions about program continuation" (p. xx). Failure of academics to contribute to this discussion out of fear, derision or other motivation results in a missed opportunity to shape the discussion around which faculty themselves will be ultimately judged.

Why a New Index is Important

Individual indices exist mostly to measure faculty and departmental research productivity. (See Appendix C for a list of over 40 H-index variants and an explanation of their uses). Measurement of teaching productivity has traditionally been limited to classroom evaluations or classroom hours taught (Hattie & Marsh, 2002; Middaugh, 2001). Not only does a measurement of service hour productivity not exist, even data on the subject is difficult to obtain. The problem is not one of lack of productivity measures at the departmental level, but one of too many and lack of a way to intellectually combine three measures into a single quantifiable index that is easily comparable across universities and departments.

Absence of an index unifying productivity into a single measure across these three dimensions results in a less information and methodological bias as administrators try to weigh the three dimensions in some sort of internal, subjective mental calculus, obscured from outside and objective reasoning; which creates problems for administrators and faculty alike (Heydinger & Simsek, 1992; Mezrich & Nagy, 2007). Decision-making and resource allocation are impaired due to varying missions and the resultant varying workload allocations of diverse institutional types making decisions based on a skewed methodology: one that primarily values research, when the institution itself may not value research as its primary function.

Over-emphasis on a research metric as a sole measure of faculty productivity despite evidence that faculty workload and goals differ by institutional mission "may unfairly disadvantage individuals in disciplines that do not follow the traditional models of publication and extramural funding as primary measures of faculty work" (Webber, 2011, p. 118).

Across all types of institutions faculty respondents most often defined their own productivity by refereed publications and research grants (Massy & Wilger, 1995, p. 15–16). Winston (1994) attributes this finding to a national labor market for faculty members, which reinforces "the use [of] similar research-oriented criteria in hiring and rewarding existing faculty" (p. 43) across institutional type (Fairweather, 2002, p. 32).

There is some evidence that point to the fact that research productivity is not a proxy for productivity in other areas of a faculty member's workload allocation. Hattie and Marsh (2002) and Feldman (1987), for example, found either a weak correlation or none at all between research and teaching productivity (as measured by student evaluations).

The most comprehensive attempt to measure teaching productivity at institutions of higher education to date has been the Delaware Study. "There was a real need ... for comparative benchmarking data that would enable inter-institutional comparisons, at the academic department level, of variables that measured faculty instructional productivity, costs, and externally sponsored faculty activity" (Middaugh, 2001, p. 91). A further explanation of the Delaware Study is found in Chapter 2.

Few metrics to measure instructional productivity and time spent on service and outreach exist; however, despite this fact, no attempt has been made to combine the three dimensions of faculty duties into a single productivity metric. This not only makes comparison of across disciplines in a single university and between universities with differing missions difficult but also puts individuals whose primary focus is not research at a comparative disadvantage. It has already been shown that faculty at different institutions weight their workload allocations accordingly (Jordan, 1994; Blackburn, Bieber, Lawrence, & Trautvetter, 1991; Serpe, Large, Brown-Large, Newton, Kilpatrick, & Mason 2002).

Purpose of the Study

A comparison statistic based on institutional mission and discipline is critical. "Time and time again the context of discipline surfaces as an important variable" (Webber, 2011, p. 109). The two main objectives of this study are to present a well-defined concept of departmental productivity in higher education that encompasses the three dimensions (teaching, research, and service) of departmental workload, and to recommend empirically valid and operationally practical guidelines for weighting, measuring, and calculating it.

In this regard, it will be argued that creating a unified index of departmental productivity will be needed to improve on existing measurements. That index will ideally be able to be in a form that allows comparison of departments not only across a single campus, but also across multiple campuses in the United States.

Research Questions

The overarching research question is: What is an appropriate and tractable method for combining the existing measures of faculty departmental productivity into one metric that provides fair comparability across disciplines and colleges? In order to approach the answer, the topic will be investigated as a series of answers to the following sub-questions:

a) What is productivity and how can the concept of productivity be applied to higher education?

b) What limitations and complexities are confronted when attempting to do so? Why is the measurement of productivity important to education policy?

c) How can the measurement of productivity be improved? Under this last question, the following sub-questions will be addressed:

    1) Which facets of faculty workload will be measured?

    2) Which measures of productivity should be used to measure the different facets of faculty workload?

    3) Devising a methodology for a fair combination of these facets (i.e., what faculty workload weights are fair given differing missions across higher educational institutions.)

    4) Calculating a practical application of the methodology to a sample of five departments in at least two universities of differing missions in the United States.

Significance of the Study

A typical problem that higher education administrators have over and above not having a methodology to combine diverse productivity metrics is not being able to compare productivity of departments across their own campus or with outside universities (Cohen & Kisker, 2010, 359–360). "Time and time again the context of discipline surfaces as an important variable. Several scholars, including Adams and Roberts (1993), Baird (1986), and Becher (1994), discuss the importance of discussing productivity by discipline" (Webber, 2011, p. 109). Citation levels are typically lower in the English and liberal arts disciplines than in say biology or other scientific fields, social science lower than life and physical scientists (Webber, 2011, p. 111). Therefore, straightforward comparisons using productivity measures such as H-indices introduce unfair bias. A further common problem is one of comparison with outside institutions complicated by variant missions and department size. Several surveys of faculty workload allocation show that faculty at institutions with differing missions allocate their time accordingly resulting in variant workload allocation (Blackburn, Bieber, Lawrence, & Trautvetter, 1991; Jordan, 1994; Serpe, Large, Brown-Large, Newton, Kilpatrick, & Mason, 2002). Selecting peer institutions in some way moderates this problem, but does not solve it completely. Therefore this study would contribute to tools available to higher education administrators to better understand productivity in their own institutions and how they compare to others. Further utility would be introduced by knowing what the impact of hiring decisions would have on the department's overall current productivity when making decisions regarding resource allocation by being able to use the equation to measure the impact of prospective faculty on the current departmental index.

Limitations of the Study

Due to tractability concerns, this study will limit its calculations to the departmental level and not calculate down to the micro-level of individual faculty members.

No one measure would be able to adequately adjust for all the peculiarities of comparing faculty productivity. Decisions about what to include and exclude, quality concerns, and the like will ultimately affect outcomes. Limiting the index to include the most common three metrics of faculty workload (research, teaching, and service) by definition excludes other facets of productivity like professional development and administrative work (although some administrative work is logged as university service). The method could be easily extended to include those facets if and when the data are included in the faculty work allocation statistics and available databases. Confusion between self-reported service and extension and outreach activities and double reporting may also be limiting factors (although all attempts have been made to exclude instances of double reporting).

Quality is a naggingly subjective, but a critically important, dimension to measure. While some research productivity indices now include a quality aspect (Hirsch's H-index), assessing teaching and service quality remains problematic. It is an issue that continues to challenge researchers.

Due to the difficulties in measuring learning outcomes for students, undergraduate classroom hours as measured by the Delaware Study will be used as a proxy for teaching productivity. Weights will vary by university mission only; despite the fact that certain research (Middaugh, 2001) has suggested workload allocation also varies by discipline. This limitation only affects the application of the unified index, as the application of the index itself only posits weights as fixed. The weights actually selected are only limited by user's access to data that points to accurate weighting for each individual's situation. It is assumed that the department head will know how she or he would like the faculty in the department to be spending their time.

Data is by department. No attempt has been made to tease out the number of levels of faculty (i.e., assistant professor, full professor, tenured, tenured track, full time, part time, temporary, teaching assistants, etc.).

Assumptions of the Study

A critical assumption is that each university values every department equally. However, if this is not the case, the bias could be easily remediated by assigning qualitative weights reflecting the relative importance of individual departments. This requires close consultation with university decision makers and assumes there will be general agreement on this topic among administrators.

A further assumption of the study is that input teaching hours are a good proxy for learning outcomes. Given the complications in measuring student learning outcomes, it is difficult to know how much of an assumption this is and how much variability there is across departments that is not reflected in the data. This assumption also holds true for service hours, i.e., that more service hours reflect higher productivity. Although intuitive, this does carry the implicit assumption that the input hours are all uniformly productive, which, of course, is somewhat tautological.

Definition of Terms

Productivity and efficiency are often confused. Faculty Productivity as defined by Meyer (1998, p. 45–46) distinguishes productivity from workload and time allocation: “Workload ... captures how their [the faculty] time is spent, while productivity is a measure of what is produced with that time” (Fairweather, 2002, p. 31). However, “efficiency is the barometer of the ‘how’ of operations. It links strategy, effective delivery, and quality. Operations are at maximum efficiency when they achieve strategic outcomes, at an agreed level of quality, for the lowest expenditure of whatever is defined as the scarce resource. The scarce resource is often money, but it could be faculty effort or elapsed time or classroom and laboratory space or any other limiting resource" (Hubbell, 2007, p. 6). Therefore, efficiency is commonly a measure of the ratio of dollar value of inputs to output over time. Productivity measures output without a dollar value attached. This dissertation will focus on productivity, and although the model can be easily modified to account for salaries, this research will restrict itself to productivity and leave efficiency for future research.

Organization of the Study

This dissertation comprises five chapters. Chapter One has provided an introduction to the dissertation by introducing and defining the problem at hand. In that vein, Chapter One has placed the problem under investigation in historical context and described the questions to be addressed by the study. Further, Chapter One has expressly outlined the limitations, assumptions, and the structure of the study.

Chapter Two provides in-depth contextualization of the thesis by presenting a review of relevant academic literature to date that has tackled the problem of faculty productivity as defined in Chapter One and addressed by this study.

Chapter Three presents the framework, methodology and procedures used to design a unified index for departmental faculty productivity. The chapter discusses the data sources and the manipulations to the data that is required by the process.

Chapter Four presents a practical application of the new model to five similar departments in two universities with different missions. The chapter discusses the results of that application and implications of that research.

Chapter Five offers a summary and discussion of the researcher's findings, implications for practice, and recommendations for future research. Chapter Five concludes the research.

CHAPTER 2. LITERATURE REVIEW

Ideas of faculty productivity have been around since at least the 1860s when the US adopted the Prussian-style education system (Turner, 1973; Watson, 2011). Wilhelm von Humboldt is credited as identifying research as key to excellence in the professoriate and re-orienting faculty work toward the research imperative. "Only he is effective as a teacher who he himself is actively productive in science" (Turner, 1973, p. 7). That focus on research productivity as a proxy for total productivity still reverberates in US higher education today.

Much of the modern research on faculty productivity focuses on the causes or determinants of productivity (Baird, 1986; Creswell, 1986; Porter, 1981; Roades, 2001). In these works, the proposed models use research productivity as their measure of choice for total productivity (in line with the aforementioned historical view of this phenomenon). Leonard Baird, one of the first to quantitatively analyze faculty productivity, utilizes six measures of productivity: 1) articles and book chapters over entire career; 2) articles, book chapters, and book reviews over entire career; 3) books and monographs over entire career; 4–6) same as 1–3 but only over the last three years to correlate with academic reputation. His main findings were that outcomes vary by discipline. He was the first to suggest discipline-specific approach with respect to measurement and evaluation of productivity. In line with many subsequent publications (Blackburn, 1974; Braxton, 1996; Feldman, 1987; Hattie & Marsh, 1987; Porter, 2001), Baird also found that "research productivity (was) unrelated to the quality of teaching" (p. 14). It must be noted here though that many publications show research and teaching quality are positively related—just as the aforementioned researchers found negative or zero correlation (Allen, 1995; Bressler, 1968; Stallings & Singhal, 1969).

Creswell (1986) identified the following, albeit largely general, factors as determinants of research productivity: faculty work environment, discipline, and institution attitude toward research. He was the forerunner in proposing weighting objective and subjective measurements in an effort to address the problems associated with aspects of productivity. Porter (2001) has used both grant dollars and publications as a proxy for overall productivity. Using a multi-level model (in order to avoid results that are biased by aggregated data), Porter finds that number of undergraduate courses taught is significantly, negatively correlated with both publications and grant dollars while the ordinary least squares (OLS) formulation found an insignificant relationship. This type of multi-level modeling approach to the data might explain the inconsistent outcomes found in the literature previously mentioned. The mathematical paradox known as Simpson's Paradox (or Yule-Simpson Effect) reveals the appearance of a trend in the analysis of disaggregated groups of data, but the trend disappears or shows the opposite relationship when these groups are aggregated. The aggregation of data could simply give the opposite results to that which is inherent in the data, which would explain the conflicting results. This result is avoided in this dissertation, as there is no aggregation in the denominator portion of any statistic. When averages are summed, they are done so in a weighted average manner, rather than adding the numerators and denominators and then taking that aggregate average.

Consistent with Hirsch's (2005) and Baird's (1986) findings, Porter (2001) finds that discipline is also a significant indicator of publications using the random coefficients model (RCM).

Rhoades (2001) explores both teaching and research as aspects of productivity. While not defining methods to measure productivity, he opts for listing a series of managerial recommendations designed to increase productivity independent of defining an initial level of productivity. An example recommendation is "RPMP #2. Joint Production Principle: Productivity measures should promote and factor in efficiencies in teaching and research obtained through joint production" (p. 622).

In 2005, Jorge Hirsch revolutionized the area of quantifying research productivity by devising the H-index. In his seminal paper, "An Index to Quantify an Individual's Scientific Research Output", Hirsch cracks the formerly intractable nut of quality vs. quantity in a faculty member's research. He was the first to incorporate citation data into a research productivity matrix, rather than just relying on publication counts. His formula, "A scientist has index h if h of his/her Np papers have at least h citations each, and the other (Np − h) papers have no more than h citations each" (1) balances the importance of research by how often others used it with how many publications a faculty member has. After Hirsch published his index, a great number of alternative indices emerged. Alonso, et al. (2009) documented the most influential of variants included the "g-index", "R-index", and m-index among others. In addition to utilizing an alphabet soup approach to naming h-indices variants, an adjectival approach was also popular in names like: normalized h-index, tapered h-index, and citation weighted h-index to name a few. All of the variants attempt to re-dress shortcomings in the original h-index (i.e., timeliness, journal impact factors, removes self-citation, etc.). There are so many of these alternatives that it has its own subgenre of literature called "bibliometrics." See Appendix C for a comprehensive matrix of the h-index and its variants.

In 2012 Google Scholar began publishing real time h-indices for individuals. More information can be found at <http://googlescholar.blogspot.com>.

Alonso, et al. (2009) reiterate the point conceded by Hirsch that, "h-index cannot directly be used to compare research workers of different areas, mainly due to lack of normalization for reference practices and traditions in the different fields of science" (p. 284). Hirsch found h-indices in biological sciences to be higher, on average than physics (p. 5), which further reiterates the need for discipline specific indexing. The reoccurring finding of incomparability across academic discipline has proven a stumbling block to using index numbers to compare even research productivity across departments of differing disciplines.

Private companies have also gotten into the game of measuring faculty research productivity, but again only limits itself to research as a proxy for all types of productivity. Academic Analytics releases a "Faculty Scholarly Productivity Index" on a subscription basis. The product ranks both individuals and departments on important research variables and then compares the department with the national averages in a radar plot to give department heads and administrators a quick visual on where departments might be weak or strong in comparison with the nation or other peer or aspirant institutions. Further information on Academic Analytics can be found at <academicanalytics.com>.

Surprisingly little research has been done to include faculty activity other than research in the productivity calculus. The groundbreaking researcher on teaching productivity has been Michael Middaugh (2001). He tackled teaching productivity from the perspective of generated student credit hours. His work, known colloquially as the Delaware Study, is considered the gold standard on teaching productivity and is used widely as a benchmarking tool for institutional researchers and higher education administrators. Middaugh's work builds on the earlier work of Joint Commission for Accountability Reporting (JCAR), one of whose major contributions to the field was to specifically define the three aspects of faculty workload (p. 37–38) and define that the sum should total 100%. Middaugh's contribution introduces the managerial aspect of "cost per instructional hour" in an attempt to add an efficiency aspect to productivity. In 1999 the national average cost per instructional unit was $225 and the number of undergraduate student hours taught per FTE was 211(Middaugh, 2001). Summary statistics can be found in Table 1.

Table 1

Example Statistics for Delaware Study: Individual Institution vs. National Averages

	1994	1996	1997	1998	1999
Undergraduate Credit Hours Taught per FTE					
U Delaware	214	218	219	176	134
Nation	267	235	247	220	211
Total Student Credit Hours Taught per FTE					
U Delaware	236	235	242	196	158
Nation	288	275	267	241	230
Total Class Sections Taught per FTE					
U Delaware	2.5	2.6	2	3	2.9
Nation	2.1	2.4	2.5	2.6	2.2
Total Student Credit Hours Taught per FTE Faculty (All Categories)					
U Delaware	249.5	208	217	216	209
Nation	237.5	231	212	211	211
Direct Instruction Expenditures per Student Credit Hour					
U Delaware	$168	$173	$183	$228	$237
Nation	$149	$175	$177	$201	$225
Separately Budgeted Research and Service Expenditures per FTE					
	FY95	FY97	FY98	FY99	FY00
U Delaware	$9,196	$22,017	$40,868	$68,030	$98,204
Nation	$51,876	$58,318	$62,040	$78,854	$77,135

Source: University of Delaware, Office of Institutional Research

One of the strengths of the Delaware study is the ability not only to compare individual institutions to national data, but also to calculate statistics with peer and/or aspirant institutions and/or departments. The most surprising outcome of the study was finding that direct instruction costs vary more by discipline than Carnegie classification. “While Carnegie institutional classification could be expected to account for some of the variance, its explanatory power does not approach that of the disciplinary mix within the institutional curriculum. When Carnegie institutional classification is taken into account, the relative variance due to disciplinary mix ranges from 81.0 to 88.0 percent in the three cycles under examination" (Middaugh, 2003, p. 18–20). This is in line with similar findings for research productivity. See Table 2 for a comparison of direct instructional costs across disciplinary lines.

Table 2

Direct Instructional Cost per Student Credit Hour Taught: Delaware Study Benchmarks for Research Universities

	2000	2001		2002		2003		2001 to 2003 3 Yr Wgtd Average	% Increase From 2000
			%		%		%		
	Cost/SCH	Cost/SCH	Increase	Cost/SCH	Increase	Cost/SCH	Increase	Cost/SCH	
Communications	164	164	0.0	157	-4.3	161	2.5	160	-2.3
Computer Science	203	204	0.5	242	18.6	293	21.1	261	28.7
Engineering	415	417	0.5	438	5.0	453	3.4	442	6.5
Foreign Languages	169	171	1.2	175	2.3	176	0.6	175	3.5
English	138	140	1.4	133	-5.0	141	6.0	138	0.1
Philosophy	134	132	-1.5	132	0.0	128	-3.0	130	-3.0
Chemistry	255	264	3.5	263	-0.4	245	-6.8	254	-0.3
Economics	145	154	6.2	144	-6.5	144	0.0	146	0.5
Geography	155	164	5.8	169	3.0	165	-2.4	166	7.2
History	142	149	4.9	141	-5.4	148	5.0	146	2.7
Political Science	168	164	-2.4	160	-2.4	161	0.6	161	-4.1
Sociology	130	124	-4.6	123	-0.8	120	-2.4	122	-6.4
Business Administration	175	199	13.7	208	4.5	209	0.5	207	18.3
Financial Management	184	187	1.6	201	7.5	215	7.0	206	11.8
Avg for 24 Disciplines	**107**	**110**	**1.3**	**112**	**0.7**	**115**	**1.3**	**113**	**5.3**

Source: Michael Middaugh University of Delaware, Office of Institutional, Research Presentation, www.udel.edu/IR/.../ContainingCosts.ppt

Bardes (1998) proposed a weighted average method of evaluating teaching productivity at medical schools according to a relative value scale in teaching. Weights are increased by difficulty of courses taught (new courses have a higher weight than old, problem based learning courses have a higher weight than ordinary courses).

Crosta and Packman (2005) modeled productivity based on number of doctoral students supervised at Cornell and found departmental prestige and longevity were significant predictors of productivity. As found in research productivity, they also found unequal distribution of productivity among discipline lines, with faculty members in the Humanities being the most productive when judged by committee chair-ship. Weighting differing level classes properly remains controversial, as there are as many ideas on proper weighting as there are faculty members. It is generally acknowledged that upper division classes should carry more weight; however, how much more weight is a contentious issue. The same controversy exists also for class size and type of class.

In addition to complications of measuring teaching productivity, the same issues of tempering quality with quantity exist as in research productivity. As yet, no breakthrough on measurement in this area has been achieved.

Although there is not much research into teaching productivity, even less exists on service. The exception is Blackburn (1975, 1991) who divides service into three components: public (non-academic), professional (peer-review), and campus (committee). Blackburn attempts to find significant determinants of service hours like professorial rank (full professor was significant), self-reported self-efficacy, and career age. Consistent with some of the literature, finding time spent on teaching reduces research productivity, more time spent on service was also found to decrease research productivity.

Townsend and Rosser (2007) studied workload allocation and its effects on productivity. They limited themselves to research and teaching productivity as proxies for total productivity, ignoring the service component for tractability's sake. Their key finding is that institutional type was a key determinant of faculty workload allocation (faculty at teaching colleges spend more time teaching for example).

Looking into faculty salary research (as perhaps its own proxy of faculty productivity), Fairweather (2005) found "the relationship between hours spent in the classroom per week and pay was substantially more negative in 1998–99 than 1992–93" even in teaching-oriented institutions (p. 412). Toutkoushian (2007) found more demographic reasons ("white, married, male") for higher pay while controlling for publications and discipline, which were also significant.

Tangentially, Dundar and Lewis (1995) use cost data to predict economies of scale and scope (diversity of disciplines) in cost and productivity in teaching and research structures. The productivity variables modeled include annual student hours per department (separated by undergraduate, master, and PhD) and number of publications by department. Their results confirmed the conventional wisdom that research universities have the higher cost structure, and advanced education is the most costly from a teaching perspective (except for social sciences where masters students cost more to educate than doctoral students).

A review of the literature would not be complete without a review of the sub-genre of literature that argues the futility of measuring faculty productivity. Blackburn (1974) wrote one of the earliest and most persuasive arguments against applying productivity indices to faculty. He posits several arguments, but the main thrust is that free time and work time are indistinguishable in the life of an academic. Academia is a way of life for faculty and thus measuring hours becomes a fruitless exercise in splitting hairs about such topics as to whether thinking about one's work while walking about campus counts as scholarship. He uses four arguments against research productivity in particular: 1) discipline differences in publishing rates, 2) gender differences favor males, 3) age of earning the PhD, and 4) type of institution.

Massy and Zemsky (1994) wrote an influential article on the negative side effects of focusing solely on productivity and efficiency. They coined the popular catch-phrase "academic rachet" to mean that emphasis on research and teaching productivity causes faculty to move "away from their traditional teaching and student advising responsibilities to focus on research and scholarly activity directed at meeting their own needs rather than those of the institution" (Middaugh, 1992 p. 61) and the associated increasing costs. This of course is the fault of the metric by which faculty are measured. A unified index that reflected the value of traditional teaching and student advising as well as other faculty activities would discourage the distortion of faculty emphasis solely toward publication.

Yet despite all this work on academic productivity, no one methodology exists to combine the three main faculty activities of teaching, research, and service in a comprehensive index. For an index to be useful to university administrators, it needs to address comparability across differing disciplines and university missions as well as encompass the three main activities of faculty. This dissertation proposes to model such an index.

The following chapters will develop a theoretical model to address these major issues, populate it with data and calculate example indices for a range of disciplines. Chapters 4 and 5 will discuss the advantages and limitations of the model as well as how universities can adopt it to aid in decision making processes campus-wide.

CHAPTER 3. METHODOLOGY

As Paul Krugman, Nobel Prize economist, famously commented, "Productivity isn't everything, but in the long run it is almost everything." But what is faculty productivity? Any discussion of productivity needs to begin with defining what the term actually means. The definition, then, will drive the metrics by which it can be measured.

Productivity has been conceptualized in many ways. In education, productivity at the institutional level has been established primarily in the public's mind by ranking criteria made popular by publications like the *US News and World Report*. The original factors in the *US News and World Report* were primarily input variables many of which remain in the ranking system today. Current variables now include: Student selectivity (input) variables i.e., acceptance rates, matriculation rates, high school class rank, and ACT/SAT scores. Output variables include graduation rates, and time to completion. Interim variables: faculty salaries, percent full time faculty, retention rates (after first year), class size, expenditures per student, percent of faculty with a terminal degree, and peer assessments. A current matrix of weights and criteria used by US News & World Report can be found in Appendix B. International universities have not escaped the ranking mania. There are many ranking institutions including US News & World Report, the Shanghai Index, International Champion League of Institutions, and QS World University rankings.

Individual departments at the undergraduate level have largely escaped this type of public scrutiny; however, graduate departments and programs are ranked by *U.S. News and World Report*, utilizes input/output variables, which differ from program to program.

Some have attempted to measure learning as output variables. The two largest attempts at this type of assessment of learning are the Collegiate Learning Assessment (CLA) and the National Survey of Student Engagement (NSSE). The Collegiate Learning Assessment has critics who cite problematic scoring and validity as two of many issues facing the instrument's dubious use. See Possin (2013) for a general discussion of these concerns. National Survey of Student Engagement (NSSE) critics cite self-evaluation bias as key to comparability among other concerns. See Campbell and Cabrera (2011) for a good overview of the methodological psychometrics associated with the survey.

Groccia and Miller (1998) divide the topic of academic productivity into approach-based methodologies as an effective way in systemizing conceptualization of the myriad of approaches to increasing productivity in US higher education. By 2005, six strategic areas were identified and explored by Groccia and Miller (organization, assessment, faculty development, technology, curriculum, and classroom strategies). For an excellent review of the six strategic areas and the history of the evolution of thought on faculty productivity see Groccia and Miller (1998, 2005).

This dissertation will follow Groccia and Miller's systemization and restrict thinking on faculty productivity to the strategic area of assessment for departmental managerial purposes. In addition, essential to the assessment process is "a required method for gathering information to show accountability to certain stakeholders, most notably the accreditation process" (Wehlburg, 2008, p. 1). In this vein, the methodology in this thesis will suggest the departmental management purposes (as opposed to other assessment purposes i.e., learning, teaching, etc.) to prove accountability to interested stakeholders.

Purpose

The purpose of this study is to devise a method to calculate a single index number by which departmental productivity can be measured in three areas of faculty responsibilities. The four purposes of this chapter are to (1) describe the procedure devised for index design, (2) explain the component variables, (3), describe the data that populate the model (4) provide an explanation of the statistical procedures used to transform the data.

Method of Index Design and Component Variables

In order to calculate a unified index, measures of the traditional three faculty activities need to be identified. In line with Isaac and Middaugh's work (2005), those activities are defined as: teaching, research, and service. (See Appendix D for JCAR's definition of those activities.) To combine three measures of different units into a unified index, two transformations need to occur: 1) the measures will have to be normalized into measures that can be easily added together, and 2) their components will need to be weighted. The weights can be determined by any method, by the workload allocation set by the department heads for each professor ideally would be the best proxy for individual faculty time allocation. Such a measure would be readily available to college administration. A company called Digital Measures collects self-reported data on workload allocation by individual professor and department and makes it available to institutions under their Activity Insight module. Sometimes data entry is tasked to the individual faculty member, which may cause problems with the overall departmental weights if each weighs themselves differently; however, this dissertation assumes the department head knows and understands what the workload allocation for the department under investigation and can accurately assign the workload allocation weights. Another criticism of this method might be that if teaching is 100% of the workload allocation, then a faculty member (or department) with the same teaching hours but who is also doing research for example would be rated the same as the faculty member not doing research. This is a valid concern; however, the rejoinder to this reasoning is that if a faculty member is spending time on activities not in his workload allocation, then that the department head has determined a priori that that activity has no value to the department and thus should NOT be counted toward productivity because in fact it is a waste of the department's resources in allocating them to activities that the department has deemed in the faculty's workload.

For the purpose of this dissertation, in Chapter 4's initial comparison, the estimated workload allocation will be taken from the 2012 Faculty Workload National Survey. The national average workload from that survey shows Research at 49.8%, Instruction at 20.1%, Service at 20.4%, and Service related to research at 9.7% (p. 18). As service related to research may be conceptualized as helping in research (serving on editorial boards, networking), this percentage is added to research, so the percentage time allocation spent on research for the purposes of this study will be 60%. Instruction (20%) and Service to the community and university (20%) round out the 100 percent of time calculation. Other earlier surveys show differing workload allocations, most notably the National Center for Education Statistic's 2004 National Study of Post-secondary Faculty. However, service components were not surveyed or reported. Service components are surveyed in UCLA's HERI Undergraduate Faculty Survey (as range data). See Table 3 for an estimate of those service weights and a comparison of the time allocation statistics from the various studies. The table shows a large variation in reported workload allocation, which further emphasizes the need for a flexible benchmark system to account for varying workload allocation. Middaugh (2001) reported changing weights on faculty workloads across time, so some of the variation may be accounted for as a function of time.

Table 3

Workload Allocation Weights

Activity	2012 Faculty Workload Study	2004 NCES National Study of Post-secondary Faculty	HERI Undergraduate Faculty Survey 2013–2014	JCAR 1997	NCES Profiles of Faculty in Higher Education 1988
Research	60%	18%	27%	30.2%	16%
Instruction	20%	61%	46%	52.4%	56%
Service (Community)	20%	-	6.7%	17.4%	–
Administration	–	20%	13.4%		13%
Advising & Counseling Students	–	–	8.5%		–
All Other					16%

To normalize the magnitude of measurements, a statistic similar to a Z-score will be calculated. For the purposes of this thesis, it will be termed a P-score (P for productivity). The Z-score (sometimes referred to as a standard score) is a very useful statistic because it enables comparison between two or more scores that are from different distributions. The standard score does this by converting (in other words, standardizing) scores in a normal distribution to z-scores in what becomes a standard normal distribution. The P-Score will be calculated according to the following familiar Z-score formula from statistics according to the following formula:

$$1.\quad P - score_{i,j} = \frac{x_{i,j} - \overline{x}_{i,j}}{\sigma_{i,j}}$$

Where: i = workload allocation area (1= research, 2=teaching, 3 = service)

j= discipline or department in which the faculty member is working

To weight the components, the equation will be modified to include the workload allocation, resulting in equation 2.

$$2. Weighted\ Average\ P - score = \sum_{i=1}^{3} p_i \frac{x_{i,j} - \overline{x}_{i,j}}{\sigma_{i,j}}$$

The research measure (i=1) include: average publications per faculty member, average citations, average grant dollars, average number of grants, average number of books published, average number of conference proceedings, and average number of awards over the active academic lifetime of the faculty. Therefore those departments with longer-tenured faculty may see higher productivity in their index. Restricting the data to the last five years would reduce this bias.

The teaching measure will be undergraduate student credit hours per academic year divided by full time equivalent faculty as measured by the Delaware Study. The service measure will be service hours per academic year. The data will be taken from the most recent UCLA Higher Education Research Institute's (HERI) Faculty Survey.

Data

Data will be taken from a combination of sources: including Academic Analytics, the Delaware Study, HERI's Faculty Survey, the National Study of time allocation. See Table 4 for a complete reference of index components, sources, and dates.

The data for the research portion of the calculation were taken from Academic Analytics (2014) data set. Total journal publications, total citations, total dollar amounts of grants, total number of grants, total awards, total conference proceedings, and total books variables were each divided by the number of faculty to arrive at average numbers for the seven measures by department. Means and standard deviations were then calculated by discipline. The seven Z-scores for research were then averaged (this assumes that each institution and department values the variables equally, i.e., number of publications is just as important as citations. Although it would be straightforward to change that assumption by assigning weights by institutional and/or department value.) Also note this method averaged seven Z-scores into a single number Z-score to avoid the trap of Simpson's Paradox.)

For teaching hours, data was taken from the 2014 Delaware study for undergraduate teaching hours. Due to the difficulties in measuring learning outcomes for students, classroom hours as measured by the Delaware Study will be used as a proxy for teaching productivity. Further only undergraduate teaching hours were used due to the controversy in weighting higher-level courses.

For community service, the Z-statistics were estimated based on the mission statement of the universities. The mean and standard deviations by department were extrapolated from the HERI study (2013–2014). Weights were estimated from range data. The calculations took the number of hours from the most frequent responses, multiplied by the mid-point, then summed and divided by the total. When frequencies were nearly equal, ranges were combined.

Table 4

Data and Data Sources

Index Component	Source	Year
Research	Academic Analytics	2014
Teaching	Delaware Study	2014
Service	Higher Education Research Institute Faculty Survey	2013–2014

Five departments commonly found in most universities were chosen for analysis: English, Biology, Mathematics, Sociology, and Computer Science. They were chosen primarily because they were represented in both data samples at the institution level and range from soft to hard skill, liberal arts, social sciences and hard sciences. Thus they represent the disciplinary differences most commonly cited as troublesome for inter-departmental comparison according to the literature. Two institutions with differing missions are analyzed and then compared. The composite Z-score for each of the five departments at both universities are calculated and compared with varying faculty workload percentages.

A national average statistic would be equal to zero on a Z-score. Any score above zero (positive) reflects a department preforming above national averages. Any score in negative represents a department preforming below the national standard. Additionally, the higher the Z-score, then the more productive the department.

Further comparisons could be made across disciplines for administrative decisions that require resource allocation across not only departments, but also various colleges on a university campus.

Note that the above equations can be modified to include any other information that is deemed important to faculty productivity including but not limited to grant dollars, number of graduate/doctoral students supervised, etc. in accordance with the values of the user. The only limitation is data to calculate the averages and standard deviations of the p-scores.

SPSS version 22 was used to compute standard deviations and means from the data set.

CHAPTER 4. APPLICATION OF UNIFIED INDEX AND RESULTS

Two universities whose data for the three-workload areas as described in Chapter 3 were available were selected to compare. Both universities are classified has RVH (Research-Very High) in the Carnegie classification. One is in an urban area (top quartile of MSA by population) and one is in a rural community. Both are public institutions. University Two has land grant (community service) status. University One was established as a normal (teaching) school.

The initial statistic calculated is one showing a single Z-score for research. This is done with the reasoning that, as mentioned previously, research has been used as a proxy for overall departmental productivity (research weight equals 100%). Under this calculation, the second university excels the nation in English, is below national averages in mathematics, sociology, and computer science. The first university is under national averages in almost all departments; however, remains mostly in a quarter of a standard deviation from the national mean (see Figure 5). None of these measures accurately reflects the actual relative performance of the departments due to the fact that both universities' missions are not solely research (despite their RVH Carnegie designations). The next set of calculations will add in a teaching component in the hopes of moving the indices more in line with an accurate representation of the departments in each institution actual performance.

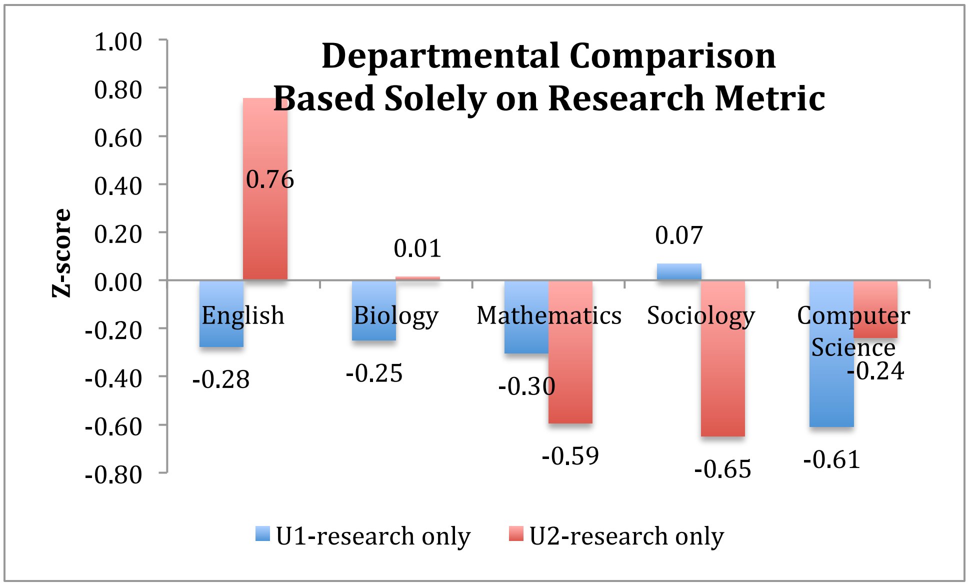

Figure 2. Departmental Comparison: 100% Research Metric

Adding teaching into the productivity equation with equal weight with research improves the model as seen in Figure 2. Now most departments at both universities are at or above national means. Note the dramatic change in the computer science statistics for universities, going from a negative half and quarter standard deviation respectively to a positive Z-score above 1 for both. Reflecting extreme productivity in student credit hour production per full time equivalent faculty.

This addition improves the comparison statistics for both universities and is a better metric as both universities are public and as such teaching is an important part of the mission. In the next figure, the service component will be added to the equation.

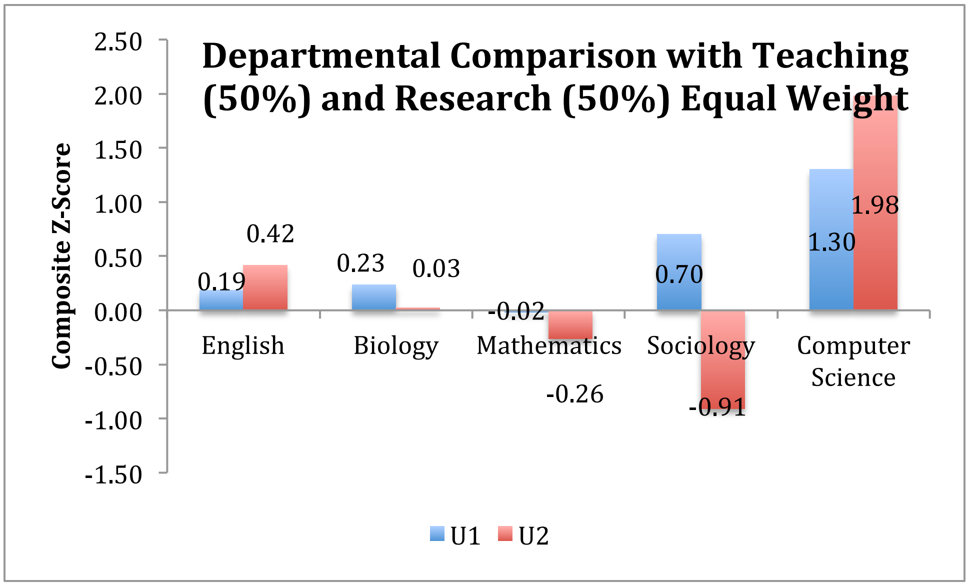

Figure 3. Departmental Comparison: Teaching, 50% and Research, 50%

Figure 3 adds the percentages workload allocation from the 2012 Faculty Workload National Survey. The national average workload from that survey shows research at 49.8%, instruction at 20.1%, service at 20.4%, and service related to research at 9.7% (p. 18). For research related universities, the service related to research will be added to research, to bring the research percentage to 60%, instruction at 20%, and service at 20%. The service Z-scores are estimated based on the universities' missions. As University One is a public school with a normal school heritage, we estimate that Z-score to be -1. The second university, a land grant, whose primary existence is community service, is granted a +1.

Due to the high percentage placed on research (60%), this figure brings the Z-scores back in line with the original 100% research figures in Figure 2.

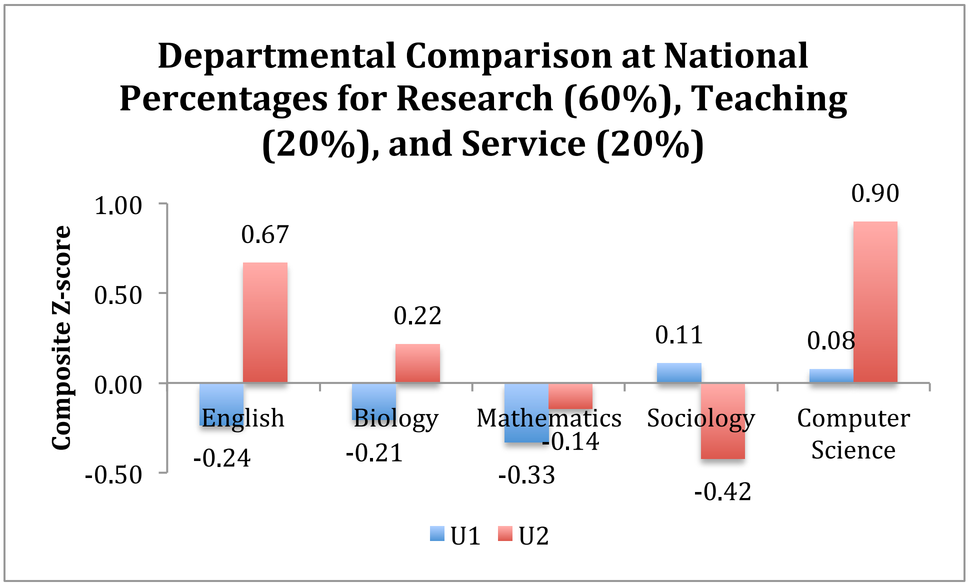

Figure 4. Departmental Comparison: Research (60%), Teaching (20%), Service (20%)

The percentages will next be varied to match each universities' own mission. The case of University One, the teaching portion will be increased at the expense of research and service, while at University Two, teaching and service will be increased at the expense of research. University One's percentages are teaching 60%, research 30%, service 10%; University Two's percentages are set at teaching 50%, research 20%, and service 30%. The resulting composite Z-scores remain near national averages except for computer science, which increases dramatically to over 2 for both universities. Sociology however remains negative for both universities.

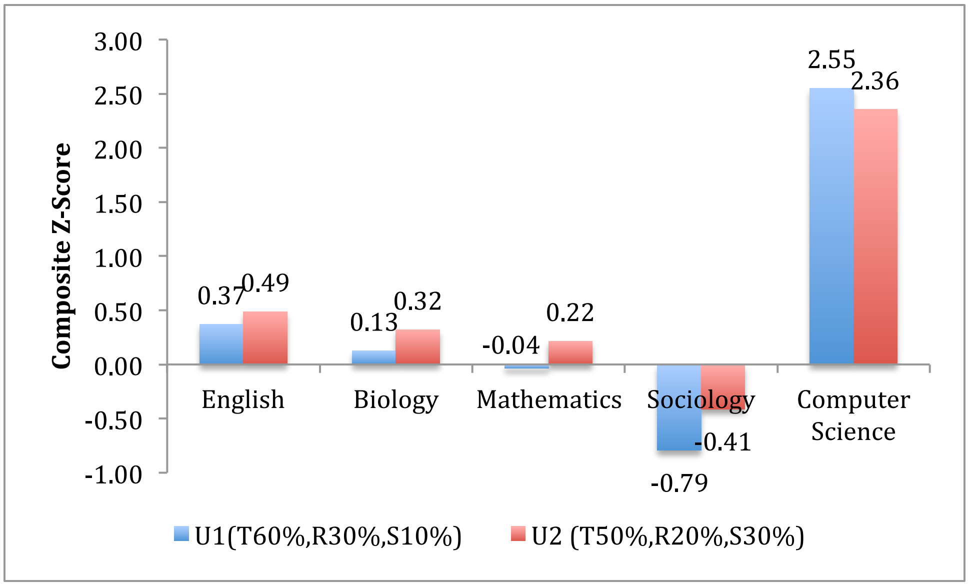

Figure 5. Composite Z-Scores – Varying Weights by University Mission

The model shows widely swinging metrics for university benchmarking depending on the workload allocation weights. This highlights the importance of assigning weights that correctly reflect the university and departmental mission in correct benchmarking and ranking for faculty productivity. Figure 5 illustrates the point by showing how the composite Z-score for University Two would increase as the service weight is increased (at the expense of research). Teaching is held constant at 20%. For those departments in beginning in the negative, a service weight of 0.4 or higher turns the composite Z-score positive for even the stubborn Sociology department.

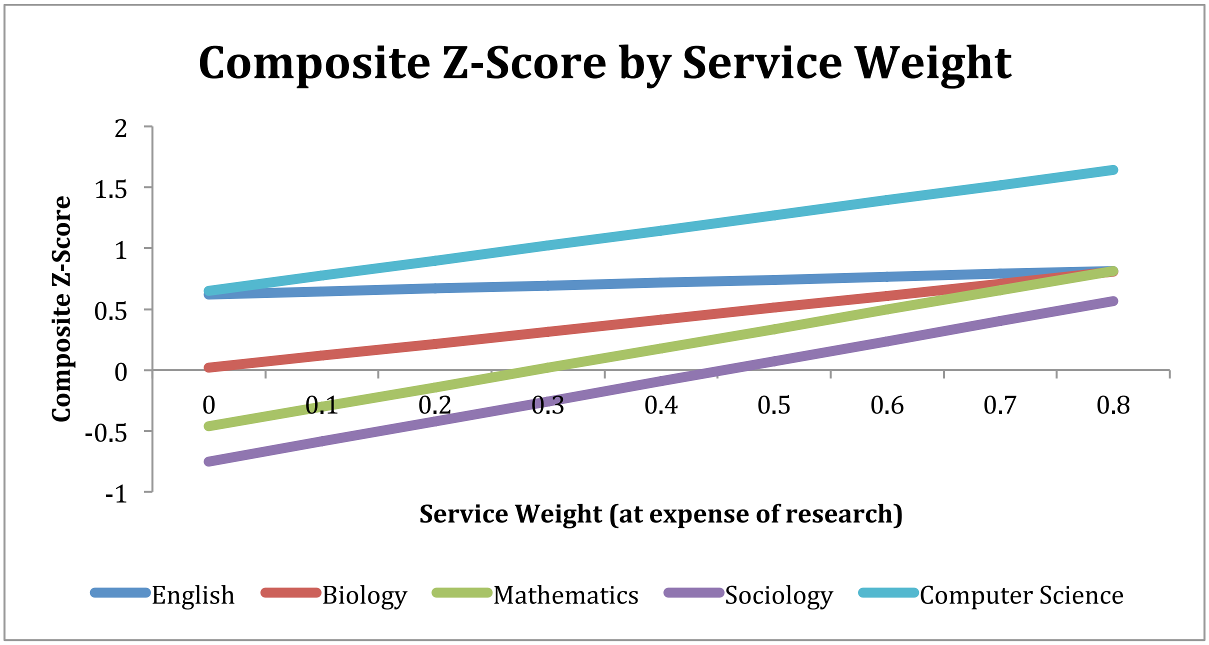

Figure 6. Composite Z-Score for University 2 with Variable Service Weight

Isaacs and Middaugh's (2001) work on faculty workload found differing workload allocation by department, which would argue for making weights variable not just by university mission, but also by idiosyncrasies of departmental discipline.

Tables 5 and 6 show the improving metrics for both universities when the workload allocations are adjusted more in line with university mission. The metrics improved in four out of five of the departments in both universities-showing greater productivity than when only judged based on a sole research metric. This lends weight to the thesis that a sole research metric biases productivity estimates and comparisons for universities with missions inclusive of other goals.

Table 5

Comparison of Productivity Scores: University One: Teaching Mission

Department	100% Research	50% Each Teaching and Research	National Workload Allocation *	Teaching 60%, Research 30%, Service 10%
English	-0.28	0.19	-0.24	0.37
Biology	-0.25	0.23	-0.21	0.13
Mathematics	-0.30	-0.02	-0.33	-0.04
Sociology	-0.65	0.70	0.11	-0.79
Computer Science	-0.61	1.30	0.08	2.55

*Research 60%, Teaching 20%, Service 20%

Table 6

Comparison of Productivity Scores: University Two: Service Mission

Department	100% Research	50% Each Teaching and Research	National Workload Allocation *	Teaching 50%, Research 20%, Service 30%
English	0.76	0.42	0.67	0.49
Biology	0.01	0.03	0.22	0.32
Mathematics	-0.59	-0.26	0.14	0.22

Sociology	-0.65	-0.91	-0.42	-0.41
Computer Science	-0.24	1.98	0.90	2.36

CHAPTER 5. SUMMARY AND DISCUSSION OF FINDINGS

Chapter One provided an introduction to this dissertation by introducing and defining the problem at hand: the absence of a unified methodology for measuring faculty productivity at the departmental level. Further, Chapter One placed the problem under investigation in historical context by tracing the shift in faculty duties from teaching to research to and service in the Nineteenth Century and the evolution of colleges and universities with differing missions in the United States. Then Chapter One has expressly outlined the limitations, assumptions, and the structure of the study. Those limitations include restricting the study to measuring productivity for assessment purposes. Challenges in measuring quality were also discussed.

Chapter Two provided an in-depth contextualization of the thesis by presenting a review of relevant academic literature to date that has tackled the problem of faculty productivity as defined in Chapter One and addressed by this study. Academic literature has focused mainly on defining and measuring research productivity at the expense of the other two components of faculty work: teaching and service as evidenced by the surplus in literature on how to measure academic productivity as a sole function of research and the determinants of productivity with productivity defined solely as a function of research outcomes.

Chapter Three presented the framework, methodology and procedures used to design a unified index for departmental faculty productivity. The chapter discussed the three main data sources, Academic Analytics, the Delaware Study of Instructional Costs, and the Faculty Survey of the Higher Education Research Institute at UCLA. The chapter went on to explain the manipulations to which the data was subjected to achieve a unified index.

Chapter Four presented a practical application of the new model to five similar departments in two universities with different missions. It was shown how the scores of the institutions varied with different weights assigned to their three individual scores in research, teaching, and service. The chapter also showed how each university excelled in their metrics when a weighting commensurate with their institutional mission was applied compared to a sole research criteria, which sent both universities below national averages for all except two departments (Sociology at University One and English at University Two). Adding in a teaching component moves the indices more in line with an accurate representation of the departments in each institution actual performance.

With the inclusion of a teaching metric, six out of ten of the departments at both universities are at or above national means. Computer science statistics for both universities changed dramatically — going from a negative half and quarter standard deviation respectively to a positive Z-score above 1 for both. This reflects the extreme productivity in student credit hour production per full time equivalent faculty at both institutions.

Next, the percentages were allowed to vary by institutional mission. University One's percentages were allocated as teaching 60%, research 30%, service 10%, reflecting the primacy of teaching in its mission; while University Two's percentages were set at teaching 50%, research 20%, and service 30%, reflecting its service mission as a land grant university. The resulting composite Z-scores show both universities slightly above national averages for all departments — except for computer science, which increases dramatically above national averages, and Sociology which slips below national averages into slightly negative territory.

It must be stressed that the metrics improved in four out of five of the departments in both universities when weights were more in line with their missions. The departments showed greater productivity than when only judged based on a sole research metric. This lends weight to the thesis that a sole research metric biases productivity estimates and comparisons for universities with missions inclusive of other goals.

Finally, the impact of varying service weights at the expense of research weights (with teaching weights held constant) were shown. This is to highlight the importance of assigning weights that correctly reflect the university and departmental mission in correct benchmarking and ranking for faculty productivity. Composite Z-scores for University Two increase as the service weight is increased (at the expense of research). Teaching is held constant at 20%. For those departments in beginning in the negative, a service weight of 0.4 or higher turns the composite Z-score positive for even the stubborn Sociology department. By the time service is fully weighted, all departments' productivity indices are above national averages, ranging from 0.5 to 1.64.

Within the context of the thesis the following research questions introduced at the beginning of the dissertation were answered in the following manners.

1. *What is productivity and how can the concept of productivity be applied to higher education?* Productivity was defined as the weighted average of a combination of the three departmental responsibilities of teaching, research, and service in line with Isaac and Middaugh's (2005) work. Appendix D contains the definitions of these activities.

2. *What limitations and complexities are confronted when attempting to do so?* To keep the thesis tractable, the scope was limited to the analysis of productivity for assessment purposes. In this vein, input variables were used for teaching and service and output variables were used

for research in the calculations.

3. *Why is the measurement of productivity important to education policy*? Accountability is important. In order to make rational decisions on resource allocation, one must have good data and reliable metrics on which to base decisions. Education is no different in this regard. Therefore, measurement of faculty productivity is important. The unique nature of education presents challenges to quantifying and comparing faculty productivity, and these challenges have been discussed in the previous chapters. However, those challenges do not negate the fact that having good metrics on which to base decisions is critical to the successful allocation of resources and management of any institution of higher education. Legislators and the public deserve transparent accountability for the taxes paid in support of education.

4. *How can the measurement of productivity be improved*? The following sub-questions were addressed:

> a) *Which facets of faculty workload will be measured?* The three agreed upon aspects of faculty workload were chosen: teaching, research, and service. While there are other activities and responsibilities of faculty, it was determined that these three activities were the most important to faculty workload
>
> b) *Which measures of productivity should be used to measure the different facets of faculty workload?* Tractability determined the selection. The Delaware study data on instructional hours per department were used for the teaching component. An average of the Academic Analytics data for research was used. HERI data was used for service.
>
> c) Devising a methodology for a fair combination of these facets (i.e., *what faculty workload weights are fair given differing missions across higher*

educational institutions?). Differing weights were applied, which showed the impact on the benchmark statistics and hence ranking of the departments. An argument for variable weighting based on institutional mission was made in order for benchmark indices to have more true reflective power on where the universities and their departments rank.

d) Calculating a practical application of the methodology to a sample of five departments in at least two universities in the United States of differing missions.

The two main objectives of this study were to present a well-defined concept of departmental productivity in higher education that encompasses the three dimensions (teaching, research, and service) of departmental workload, and to recommend empirically valid and operationally practical guidelines for weighting, measuring, and calculating it. The unified p-index accomplishes these objectives. However, further research into improving this simple index is more than warranted. Useful research into how to improve the index would include how to account for quality (especially in the teaching metric); inclusion of more output-oriented metrics into the model; inclusion of other facets of productivity, like professional development and administrative work (although some administrative work is logged as university service); and survey work to establish the benchmark weights not only by university or college, but by department. Further, it would make sense to calculate the Z-scores based not only on departmental discipline, but break the population into mission subsets for better and more precise comparison.

REFERENCE LIST

Allen, M. (1995). *Research productivity and positive teaching evaluations: Examining the relationship using meta-analysis*. Portland OR. Retrieved from http://files.eric.ed.gov/fulltext/ED379705.pdf

Alonso, S., Cabrerizo, F. J., Herrera-Viedma, E., & Herrera, F. (2009). h-Index: A review focused in its variants, computation and standardization for different scientific fields. *Journal of Informetrics*, *3*(4), 273–289. doi:10.1016/j.joi.2009.04.001

Baird, L. (1980). *The relationship between ratings of graduate departments and faculty publication rates*. Retrieved from https://www.ets.org/Media/Research/pdf/RR-80-27-Baird.pdf

Baird, L. L. (1986). What characterizes a productive research department? *Research in Higher Education*, *25*(3), 211–225. Retrieved from http://spot.lib.auburn.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ349450&site=ehost-live

Bardes, C., & Hayes, J. (1995). Are the teachers teaching: Measuring the educational activities of clinical faculty. *Academic Medicine : Journal of the Association of American Medical Colleges*, *70*, 349–354.

Bardes, C., Hayes, J., Falcone, D., Hajjar, D., & Alsonso, D. (1998). Measuring teaching: A relative value scale in teaching. *Teaching and Learning in Medicine*, *10*(1), 40–43.

Barnes, M. (2011) “Bringing Transparency to College Costs.” White House Blog . Retrieved from https://www.whitehouse.gov/blog/2011/06/30/bringing-transparency-college-costs

Basken, P. (2012). Cost savings appear elusive in push for faculty productivity. *Chronicle of Higher Education*, *59*(7), A14–A14. Retrieved from http://spot.lib.auburn.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=aph&AN=82316072&site=ehost-live

Bauerlein, M. (2013). How not to measure faculty productivity. *Phi Kappa Phi Forum*, (Spring), 14–17.

Becher, T. (1994). The significance of disciplinary differences. *Studies in Higher Education*, *19*, 151–161. doi:10.1080/03075079412331382007

Blackburn, R. (1974). The meaning of work in academia. *New Directions for Institutional Research*, 75–99.

Blackburn, R., Bieber, J., Lawrence, J., & Trautvetter, L. (1991). Faculty at work: Focus on research, scholarship, and service. *Research in Higher Education*, *32*(4), 385–412.

Bland, C. J., Center, B. A., Bruce A., Finstad, D. A., Risbey, K. R., & Staples, J. (2006). The impact of appointment type on the productivity and commitment of full-time faculty in research and doctoral institutions. *The Journal of Higher Education*, *77*(1), 89–123. doi:10.1353/jhe.2006.0002

Bowen, H. (1978). *Academic compensation: Are faculty and staff in American higher education adequately paid?* New York, NY.

Bresler, J. B. (1968). Teaching effectiveness and government awards. *Science*, *160*, 164–168.

Campbell, C., & Cabrera, A. (2011). How sound is NSSE. *Review of Higher Education*, *35*(1), 77–103. doi:10.1353/rhe.2011.0035

Cohen, A., & Kisker, C. (2010). *The shaping of American higher education* (2nd ed.). San Francisco: Jossey Bass.

Creswell, J. (1986). Concluding thoughts: Observing, promoting, evaluating, and reviewing research performance. In J. Creswell (Ed.), *Measuring faculty research performance* (p. 87–102). San Francisco: Jossey Bass.

Crosta, P. M., & Packman, I. G. (2005). Faculty productivity in supervising doctoral students' dissertations at Cornell University. *Economics of Education Review*, *24*, 55–65. doi:10.1016/j.econedurev.2004.03.011

Dundar, H., & Lewis, D. R. (1995). Departmental productivity in American universities: Economies of scale and scope. *Economics of Education Review*, *14*(2), 119–144. doi:10.1016/0272-7757(95)90393-M

Eagan, K., Aragon, M. C., & Eagan, K. (2014). *Undergraduate teaching faculty: The 2013–2014 HERI faculty survey*. Los Angeles, CA: Higher Education Research Institute, UCLA.

Fairweather, J. S. (2002). The mythologies of faculty productivity. *Journal of Higher Education*, *73*(1), 26–48.

Fairweather, J. S. (2005). Beyond the rhetoric: Trends in the relative value of teaching and research in faculty salaries. *The Journal of Higher Education*. doi:10.1353/jhe.2005.0027

Fairweather, J. S., & Beach, A. L. (2002). Variations in faculty work at research universities: Implications for state and institutional policy. *The Review of Higher Education*, *26*(1), 97–115. doi:10.1353/rhe.2002.0024

Flanagan, S., & Morse, R. (2015). How U.S. News calculated the 2016 best graduate schools rankings. Retrieved August 18, 2015, from http://www.usnews.com/education/best-graduate-schools/articles/how-us-news-calculated-the-rankings?page=2

Flexner, A. (1910). *Medical education in the United States and Canada*. New York, NY.

Gilmore, J. (1992). Evaluating academic productivity and quality. *New Directions for Institutional Research*, (75), 35–47.

Groccia, J., & Miller, J. (Eds.). (1998). *Enhancing productivity: Administrative, instructional, and technological strategies*. San Francisco: Jossey Bass.

Groccia, J., & Miller, J. (Eds.). (2005). *On becoming a productive university*. Bolton, MA: Anker.

Marsh, H.W., & Hattie, J. (2002). The relation between research productivity and teaching effectiveness—Complementary, antagonistic, or independent constructs? *Journal of Higher Education, 73*(5), 603-641.

Marsh, H. W. (1996). The relationship between research and teaching—a meta- analysis. *Review of Educational Research*, *66*, 507-542.

Henry, B., & Allen, L. (2008). Faculty workload and productivity: The next generation's plight—and opportunity. *The NEA 2009 Almanac of Higher Education*, 55–64. Retrieved from http://www.nea.org/assets/img/PubAlmanac/ALM_09_05.pdf

Heydinger, R. & Simsek, H. (1992). An agenda for reshaping faculty productivity. *Education Commission of the States*, *November* (8).

Hilton, C., Fisher, W. Jr., Lopez, A., & Sanders, C. (1997). A relative-value-based system for calculating faculty productivity in teaching, research, administration, and patient care. Academic Medicine, (September), 787–791.

Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(46), 16569–16572. doi:10.1073/pnas.0507655102

Howard, R., & Stephen, K. (1974, April 30). The effective use of resources: Financial and human. In *Speeches presented at the Association of Governing Boards of Universities and Colleges National Conference on Trusteeship* (p. 33). New Orleans, LA.

Hubbell, L. L. (2007). Quality, efficiency, and accountability: Definitions and Applications. *New Directions for Institutional Research*, (140), 5–14. doi:10.1002/he

Isaacs, H. K., & Middaugh, M. F. (2001). *Out-of-classroom faculty activity study: A discussion of the results, utilization, and lessons learned*. San Diego, CA. Retrieved from http://www.udel.edu/IR/focs/2004 Forum Contributed Paper.pdf

Isaacs, H. K., & Middaugh, M. F. (2005). *Out-of-classroom faculty activity study: A look at two data collection cycles and a verification study*. San Diego, CA. Retrieved from http://www.udel.edu/IR/focs/2005 Forum Contributed Paper.pdf

Jackson, J. (2004). The story is not in the numbers: Academic socialization and diversifying the faculty. *NWSA Journal*, *16*(1), 172–185. doi:10.1353/nwsa.2004.0035

Jordan, S. (1994). What we have learned about faculty workload. *New Directions for Institutional Research*, *83*, 15–23. Retrieved from http://highered.colorado.gov/Publications/General/StrategicPlanning/Meetings/Resources/Mission/Mission_100217_Faculty_Workload.pdf

Klein, S., Liu, O. L., & Sconing, J. (2009). *Test validity study (TVS) report: Education*. Retrieved from http://cae.org/images/uploads/pdf/13_Test_Validity_Study_Report.pdf

Lazaridis, T. (2010). Ranking university departments using the mean h-index. *Scientometrics*, (82), 211–216. doi: 10.1007/s11192-009-0048-4

Mamiseishvili, K., & Rosser, V. J. (2009). International and citizen faculty in the United States: An examination of their productivity at research universities. *Research in Higher Education*, *51*(1), 88–107. doi:10.1007/s11162-009-9145-8

March, E. (2011). *National study of instructional costs and productivity results from the 2009–10 Delaware study*. Retrieved from http://www.albany.edu/UDel_Tables_for_BAG4_3_2_11.pdf

Massy, W. F., Sullivan, T. A., & Mackie, C. (Eds.). (2013). *Improving measurement of productivity in higher education. The National Academies Press* (Vol. 45, p. 15–23). Washington, DC: The National Research Council. doi:10.1080/00091383.2013.749140

Massy, W., & Wilger, A. (1995). Improving productivity. *Change*, *10*(27.4). Retrieved from http://spot.lib.auburn.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=aph&AN=9508180353&site=ehost-live

Massy, W., & Zemsky, R. (1994). Faculty discretionary time: Departments and the academic ratchet. *Journal of Higher Education*, *65*(1), 1–22.

McLendon, M., Hearn, J. C., & Deaton, R. (2006). Called to account: Analyzing the origins and spread of state performance-accountability policies for higher education. *Educational Evaluation and Policy Analysis*, *28*(1), 1–24. doi:10.3102/01623737028001001

Mezrich, R., & Nagy, P. G. (2007). The academic RVU: A system for measuring academic productivity. *Journal of the American College of Radiology*, *4*, 471–478. doi:10.1016/j.jacr.2007.02.009

Middaugh, M. (2001). *Understanding faculty productivity*. San Francisco: Jossey Bass.

Middaugh, M. F. (1992a). Examining academic and administrative productivity measures. *New Directions for Institutional Research*, (75), 61–75.

Middaugh, M. F. (1992b). Examining academic and administrative productivity measures. *New Directions for Institutional Research*, *19*(753), 61–.

Middaugh, M. F., Graham, R., Shahid, A., & Carroll, C. D. (2003). A study of higher education instructional expenditures: The Delaware study of instructional costs. *Higher Education*.

Morse, R. (2015). Best Colleges Ranking Criteria and Weights. Retrieved August 18, 2015, from http://www.usnews.com/education/best-colleges/articles/2014/09/08/best-colleges-ranking-criteria-and-weights

Mortimer, K. P., "On Accountability Reporting of State Colleges, Universities, and Community Colleges" (1997). *JCAR faculty assignment reporting*. American Association of State Colleges and Universities. Retrieved from https://books.google.com/books?id=7H_kGwAACAAJ

Musick, M. A. (2011). An analysis of faculty instructional and grant-based productivity at the University of Texas at Austin. (November).

Noble, J. H., Cryns, A. G., & Laury, B. S. (1992). Faculty productivity and costs. *Evaluation Review*, *16*(3), 288–314.

Porter, S. R., & Umbach, P. D. (2001). Analyzing faculty workload data using multilevel modeling. *Research in Higher Education*, *42*, 171–196. doi:10.1023/A:1026573503271

Possin, Kevin (2013). A Fatal Flaw in the Collegiate Learning Assessment. Assessment Update. Retrieved from: http://www.assessmentupdate.com/sample-articles/a-fatal-flaw-in-the-collegiate-learning-assessment-test.aspx.

Preskill, H., & Torres, R. (1999). *Evaluative inquiry for learning in organizations*. London: Sage Publications.

Rhoades, G. (2001). Managing productivity in an academic institution: Rethinking the whom, which, what, and whose of productivity. *Research in Higher Education*, *42*, 619–632. doi:10.1023/A:1011006511651

Richard, B. (1992). An agenda for reshaping faculty productivity: State policy and college learning. *Education Commission of the States*, *Nov*(8), 38.

Rockwell, S. (2014). 2012 Faculty workload survey (April). Retrieved from http://sites.nationalacademies.org/cs/groups/pgasite/documents/webpage/pga_087667.pdf

Russel, S., Fairweather, J., & Zimbler, L. (1988). *Profiles of faculty in higher education institutions*. Washington, D.C. Retrieved from http://files.eric.ed.gov/fulltext/ED336058.pdf

Serpe, R., Large, M., Brown-Large, L., Newton, R., Kilpatrick, K., & Mason, R. (2002). *CSU faculty workload report*. Retrieved from https://www.calstate.edu/acadres/docs/csu_facwrkldrpt.pdf

Shin, C. J., Toutkoushian, R. K., & Teichler, U. (Eds.). (2011). *University rankings*. Springer. doi:10.1007/978-94-007-1116-7

Stallings, W. M., & Singhal, S. (1970). Some observations on the relationships between research productivity and student evaluations of courses and teaching. *American Sociologist*, *5*, 141–143.

Tab, E. D., Cataldi, E. F., Zimbler, L., & Whitehurst, G. J. (2005). *2004 national study of postsecondary faculty (NSOPF: 04) Report on faculty and instructional staff in fall 2003*. *National Center for Education Statistics*. Retrieved from http://nces.ed.gov/pubs2006/2006176.pdf

Toutkoushian, R. K., Bellas, M. L., & Moore, J. V. (2007). The interaction effects of gender, race, and marital status on faculty salaries. *The Journal of Higher Education*, *78*(5), 572–601. doi:10.1353/jhe.2007.0031

Townsend, B. K., & Rosser, V. J. (2007). Workload issues and measures of faculty productivity. *Thought and Action*, 7–20. Retrieved from http://www.nea.org/assets/img/PubThoughtAndAction/TAA_07_02.pdf

Trigwell, K. (2011). Measuring teaching performance. In C. J. Shin, R. Toutkoushian, & U. Teichler (Eds.), *University rankings* (p. 165–181). London, New York: Springer Verlag.

Turner, S. R. (1973). *The Prussian universities and the research imperative, 1806 to 1848*. Princeton.

Von Humbolt, W. (1809). Über die innere und äußere Organisation der höheren wissenschaftlichen Anstalten in Berlin. *Wilherlm von Humbolt - Werke IV - Schriften Zur Politik Und Zum Bildungswesen*, 255–266.

Wasley, P. (2007). Faculty-productivity index offers surprises. *Chronicle of Higher Education*.

Watson, P. (2011). *The German genius: Europe's third renaissance, the second scientific revolution, and the twentieth century*. New York: Harper Perennial.

Webber, K. (2011). Measuring faculty productivity. In J. C. Shin, R.K. Toutkoushian, & U. Teichler (Eds.), *University rankings* (p. 105–121). New York: Springer Verlag. doi:10.11007/978-94-007-1116-7_6

Wehlburg, C. (2008). *Promoting integrated and transformative assessment*. San Francisco: Jossey Bass.

Appendix A

Sample Delaware Study of Instructional Cost Data Collection Form

Data Collection Form

2002-03 Delaware Study of Instructional Cost and Productivity

Institution: ____________________

Department/Discipline: ____________________

Associated CIP Identifier: ______

Please indicate the average number of degrees awarded in this discipline at each degree level over the three year period from 1998-99 through 2000-01.

Bachelor's: ______
Master's : ______
Doctorate: ______
Professional: ______

Place an 'X' in the box below if this discipline is non-degree granting.

______

Place an "X" in the box below that describes your academic calendar:

______ Semester
______ Quarter

A. INSTRUCTIONAL COURSELOAD: FALL SEMESTER, 2001

Please complete the following matrix, displaying student credit hours and organized class sections taught, by type of faculty, and by level of instruction. Be sure to consult definitions before proceeding. Do not input data in shaded cells except for those mentioned in the important note below that pertains to (G) and (J).

Faculty				Student Credit Hours								Organized Class Sections				
Classification	FTE Faculty (A) Total	(B) Separately Budgeted	(C) Instructional	(D) Lower Div. OC*[1]	(E) Upper Div. OC*[1]	(F) Undergrad Indv. Instruct.	(G) Total Undergrad SCH	(H) Grad. OC*[1]	(I) Graduate Indv. Instruct.	(J) Total Graduate SCH	(K) Total Student Credit Hours	(L) Lab/Disc/Rec. Sections	Other Section Types (Lecture, Seminar, etc.) (M) Lower Div.	(N) Upper Div.	(O) Graduate	(P) Total
Regular faculty: -Tenured/Tenure Eligible																
- Other Regular Faculty																
Supplemental Faculty		NA														
Teaching Assistants: - Credit Bearing Courses		NA														
- Non-Credit Bearing Activity		NA		NA	NA	NA	NA	NA	NA	NA	NA					
TOTAL																

B. COST DATA: ACADEMIC AND FISCAL YEAR 2001-02

1. In the boxes below, enter the total number of student credit hours that were generated during Academic Year 2001-02 during terms that were supported by the department's instructional budget. (NOTE: Semester calendar institutions will typically report fall and spring student credit hours; quarter calendar institutions will usually report fall, winter, and spring student credit[2] hours.)

______ A. Undergraduate
______ B. Graduate

2. In the boxes below, enter total direct expenditures for instruction in FY 2001-02

______ A. Salaries — Are the benefits included in the number reported for salaries (Y/N)? ______

______ B. Benefits — If the dollar value is not available, what percent of salary do benefits constitute at your institution? ______

______ C. Other than personnel expenditures.

______ D. Total

3. In the box below, enter total direct expenditures for separately budgeted research activity in FY 2001-02

______

4. In the box below, enter total direct expenditures for separately budgeted public service activity in FY 2001-02

[1]OC = organized classes

[2]Summer semesters and quarters are not generally supported by the department's instructional budget.

Appendix B

Methodological Matrix for U.S. News & World Report Rankings

Ranking Indicator	Indicator Weight		Subfactor	Subfactor Weight	
	National Universities and National Liberal Arts Colleges	Regional Universities and Regional Colleges		National Universities and National Liberal Arts Colleges	Regional Universities and Regional Colleges
Undergraduate academic reputation	22.50%	22.50%	Peer assessment survey	66.70%	100%
			High school counselors' ratings	33.30%	0%
Student selectivity for the fall 2013 entering class	12.50%	12.50%	Acceptance rate	10%	10%
			High school class standing in top 10%	25%	0%
			High school class standing in top 25%	0%	25%
			Critical Reading and Math portions of the SAT and composite ACT scores	65%	65%
(con't)			Percent faculty that is	5%	5%

			full time		

Methodological Matrix for U.S. News & World Report Rankings (Con't)

Ranking Indicator	**Indicator Weight**		**Subfactor**	**Subfactor Weight**	
	National Universities	Regional Universities		National Universities	Regional Universities
Faculty resources for 2013-2014 academic year	20%	20%	Faculty compensation	35%	35%
			Percent faculty with terminal degree in their field	15%	15%
			Student-faculty ratio	5%	5%
			Class size, 1-19 students	30%	30%
			Class size, 50+ students	10%	10%
Graduation and retention rates	22.50%	22.50%	Average graduation rate	80%	80%
			Average freshman retention rate	20%	20%
Financial resources	10%	10%	Financial resources per student	100%	100%
Alumni giving	5%	5%	Average alumni giving rate	100%	100%
Graduation rate performance	7.50%	7.50%	Graduation rate performance	100%	100%
Total	100%	100%	—	100%	100%

Definitions of Ranking Criteria

Acceptance rate: The ratio of the number of students admitted to the number of applicants for fall 2013 admission. The acceptance rate is equal to the total number of students admitted divided by the total number of applicants.

Both the applications and acceptances counted only first-time, first-year students. A lower acceptance rate – indicating a school is harder to get into (rejects more students) – scores higher in the ranking model, and a higher acceptance rate – indicating a school is easier to get into (rejects fewer students) – scores lower in the ranking model.

Average alumni giving rate: The average percentage of undergraduate alumni of record who donated money to the college or university. Alumni of record are former full- or part-time students who received an undergraduate degree and for whom the college or university has a current address. Graduates who earned only a graduate degree are excluded.

Undergraduate alumni donors are alumni with undergraduate degrees from an institution who made one or more gifts for either current operations or capital expenses during the specified academic year.

The alumni giving rate is calculated by dividing the number of alumni donors during a given academic year by the number of alumni of record for that same year. The two most recent years of alumni giving rates that were available were averaged and used in the 2015 rankings. The two separately calculated alumni giving rates that were averaged (added together and divided by two) were for giving in the 2011–2012 and 2012–2013 academic years.

The percentage of alumni giving serves as a proxy for how satisfied students are with the school. A higher average alumni giving rate scores better in the ranking model than a lower average alumni giving rate.

Average freshman retention rate: The percentage of first-year freshmen who returned to the same college or university the following fall. Average freshman retention rate indicates the average proportion of the first-year classes entering from fall 2009 through fall 2012 who returned the following fall.

If a school submits fewer than four years of freshman retention rate data, then the average is based on the number of years that are submitted by the school to U.S. News. A higher average freshman retention rate scores better in the ranking model than a lower average retention rate.

Average graduation rate: The percentage of entering freshmen who graduated within a six-year period or less, averaged over the classes entering from fall 2004 through fall 2007. (Note: This excludes students who transferred into the school and then graduated.)

If a school submits fewer than four years of graduation rate data, then the average is based on the number of years that are submitted. A higher average graduation rate scores better in the ranking model than a lower graduation rate.

Class size, 1–19 students: The percentage of undergraduate classes, excluding class subsections, with fewer than 20 students enrolled during fall 2013. A larger percentage of small classes scores higher in the ranking model than a lower percentage of small classes. In other words, the more small classes, the better.

Class size, 50–plus students: The percentage of undergraduate classes, excluding class subsections, with 50 students or more enrolled during fall 2013. A smaller percentage of large classes scores higher in the ranking model than a larger percentage of large classes. In other words, the fewer large classes, the better.

Expenditures per student: Financial resources are measured by the average spending per full-time-equivalent student on instruction, research, public service, academic support, student services and institutional support during the 2012 and 2013 fiscal years.

The number of full-time-equivalent students is equal to the number of full-time students plus one-third of the number of part-time students. (Note: This includes both undergraduate and graduate students.)

We first scaled the public service and research values by the percentage of full-time-equivalent undergraduate students attending the school. Next, we added in total instruction, academic support, student services, institutional support and operations and maintenance (for public institutions only) and then divided by the number of full-time-equivalent students.

After calculating this value, we applied a logarithmic transformation to the spending per full-time-equivalent student, prior to standardizing the value. This calculation process was done for all schools.

If a school submits fewer than two years of expenditures per student, then the average is based on the one year that is submitted.

Higher average expenditures per student score better in the ranking model than lower average expenditures per student. In other words, financial resources do matter in terms of being able to provide students with a high-quality college experience.

Faculty compensation: The average faculty pay and benefits are adjusted for regional differences in cost of living. This includes full-time assistant, associate and full professors. The values are taken for the 2012–2013 and 2013–2014 academic years and then averaged.

If a school submits fewer than two years of faculty salary data, then only one year is used. The regional differences in cost of living are taken from indexes from Runzheimer International.

Higher average faculty salaries after adjusting for regional cost of living score better in the ranking model than lower average faculty salaries.

Faculty with a Ph.D. or terminal degree: The percentage of full-time faculty members with a doctorate or the highest degree possible in their field or specialty during the 2013-2014 academic year. Schools with a larger proportion of full-time faculty with the terminal degree in their field score better in the ranking model than schools with a lower proportion.

Graduation rate performance: The difference between the actual six-year graduation rate for students entering in fall 2007 and the predicted graduation rate. The predicted graduation rate is based upon characteristics of the entering class, as well as characteristics of the institution.

This indicator of added value shows the effect of the college's programs and policies on the graduation rate of students after controlling for spending and student characteristics such as test scores and the proportion receiving Pell Grants.

If the actual graduation rate is higher than the predicted rate, the college is enhancing achievement or is overperforming. If its actual graduation rate is lower than the predicted rate, then it's underperforming.

A school with a higher ratio of its actual graduation rate compared with its U.S. News predicted graduation rate (actual graduation rate divided by predicted rate) scores better in the ranking model than a school with a lower ratio of its actual graduation rate compared with its U.S. News predicted graduation rate.

In the 2015 edition of the Best Colleges rankings, as in the 2014 edition, this ranking indicator is included for all schools in all the ranking categories. Graduation rate performance has been used in the National Universities and National Liberal Arts Colleges ranking categories since 1997.

High school class standing: The proportion of students enrolled for the academic year beginning in fall 2013 who graduated in the top 10 percent (for National Universities and National Liberal Arts Colleges) or 25 percent (Regional Universities and Regional Colleges) of their high school class.

A higher proportion of students from either the top 10 percent or top 25 percent of their high school class scores better in the ranking model than lower proportions from either the top 10 percent or top 25 percent.

High school counselor rating score: For the fifth consecutive year, U.S. News counted guidance counselor opinions in ranking the National Universities and National Liberal Arts Colleges. These ratings by public and private independent school counselors are used as a separate indicator of academic reputation for these two categories, in addition to ratings by college admissions deans, provosts and presidents.

Scores for each school were totaled and divided by the number of counselors who rated that school.

This year, for the second year in a row, the two most recent years of survey results, from spring 2013 and spring 2014, were averaged to compute the high school counselor reputation score that is used in the rankings. This was done to increase the number of ratings each school received and to reduce the year-to-year volatility in the average counselor score.

The academic peer assessment score, described below, continues to be based only on the most recent year's results. The Regional Colleges and Regional Universities rankings use only the academic peer assessment, with no high school counselor ratings component.

The counselors' one-year response rate was 9 percent for the spring 2014 surveys. A higher average high school counselor reputation score does better in the ranking model than a lower high school counselor reputation score.

[See more on the high school counselor scores in the 2015 rankings.]

Peer assessment: This is used to measure how the school is regarded by administrators at peer institutions. A school's peer assessment score is determined by surveying the presidents, provosts and deans of admissions, or equivalent positions, at institutions in the school's category.

Each individual was asked to rate peer schools' undergraduate academic programs on a scale from 1 (marginal) to 5 (distinguished). Those individuals who did not know enough about a school to evaluate it fairly were asked to mark "don't know."

A school's score is the average score of all the respondents who rated it. Responses of "don't know" counted neither for nor against a school.

The survey for the 2015 edition of the Best Colleges rankings was conducted in spring 2014, and 42 percent of those surveyed responded. This response rate is unchanged from the 2014 edition.

A higher average peer assessment score does better in the ranking model than a lower peer assessment score. The academic peer assessment rating is used in the National Universities, National Liberal Arts Colleges, Regional Universities and Regional Colleges rankings.

Proportion of full-time faculty: This is the proportion of the 2013–2014 full-time-equivalent faculty that is full time. The number of full-time-equivalent faculty is equal to the number of full-time faculty plus one-third of the number of part-time faculty.

(Note: We do not include faculty in preclinical and clinical medicine; administrative officers with titles such as dean of students, librarian, registrar or coach, even though they may devote part of their time to classroom instruction and may have faculty status; undergraduate or graduate students who are teaching assistants or teaching fellows; faculty on leave without pay; or replacement faculty for those faculty members on sabbatical leave.)

To calculate this percentage, the total full-time faculty is divided by the full-time-equivalent faculty (full-time-equivalent faculty is full-time faculty plus one-third of part-time faculty). A higher proportion of faculty who are full time scores better in the ranking model than a lower proportion.

SAT/ACT scores: Average test scores on both the Critical Reading and Math portions of the SAT and composite ACT of all enrolled first-time, first-year students entering in fall 2013 are combined for the ranking model.

Before being used as a ranking indicator, the scores from both tests are converted to the percentile of the national distribution corresponding to that school's scores on the Critical Reading and Math portions of the SAT and the composite ACT. The SAT Writing section was not used in the ranking model.

For the sixth consecutive year, in order to better represent the entire entering class, we used a calculation that combines the values of both the Critical Reading and Math portions of the SAT and the composite ACT of all fall-entering students based on the percentage of the fall entering class that submitted each test.

A higher average entering class test score on the Critical Reading and Math portions of the SAT and composite ACT does better in the ranking model than a lower average SAT and ACT test score.

Student-faculty ratio: This is the ratio of full-time-equivalent students to full-time-equivalent faculty during the fall of 2013, as reported by the school.

(Note: This excludes faculty and students of law, medical, business and other stand-alone graduate or professional programs in which faculty teach virtually only graduate-level students. Faculty numbers also exclude graduate or undergraduate students who are teaching assistants.)

A lower student-faculty ratio (fewer students per each faculty member) scores higher in the ranking model than a higher ratio (more students per each faculty member).

Appendix C

Comprehensive Matrix of H-Index and its Variants

Index Name	**Brief Description and/or link to Research**
h-index(Hirsch JE (2005)	An index to quantify an individual's scientific research output. Proceedings of the National Academy of Sciences 102:16569-16572, doi: 10.1073/pnas.0507655102
g-index: (Egghe L (2006)	Theory and practise of the g-index. Scientometrics 69(1):131-152, doi: 10.1007/s11192-006-0144-7 . PDF Icon) Holding that "a measure which should indicate the overall quality of a scientist ... should deal with the performance of the top article
a-index: (Jin BH, Liang LM, Rousseau R, Egghe L (2007)	The R- and AR-indices: Complementing the h-index. Chinese Science Bulletin 52(6):855-863, doi: 10.1007/s11434-007-0145-9
h(2)-index: (Kosmulski M (2006)	A new Hirsch-type index saves time and works equally well as the original h-index. ISSI Newsletter 2(3):4-6, .
hg-index: (Alonso S, Cabrerizo FJ, Herrera-Viedma E, Herrera F (2010)	hg-index: A new index to characterize the scientific output of researchers based on the h- and g- indices. Scientometrics 82(2):391-400 doi:10.1007/s11192-009-0047-5
q2-index: (Cabrerizo FJ, Alonso S, Herrera-Viedma E, Herrera F (2009)	q2-Index: Quantitative and Qualitative Evaluation Based on the Number and Impact of Papers in the Hirsch Core. Journal of Informetrics 4(1):23-28, doi:10.1016/j.joi.2009.06.005
r-index: (Jin BH, Liang LM, Rousseau R, Egghe L (2007) *The R- and AR-indices: Complementing the h-index*	Chinese Science Bulletin 52(6):855-863, doi: 10.1007/s11434-007-0145-9
ar-index: (Jin B (2007) *T*	*The AR-index: complementing the h-index.* ISSI Newsletter 3(1):6.
m quotient: (Hirsch JE (2005)	An index to quantify an individual's scientific research output. Proceedings of the National Academy of Sciences 102:16569-16572, doi: 10.1073/pnas.0507655102

Index Name	Brief Description and/or link to Research
Trend h-index: (Sidiropoulos A, Katsaros D, Manolopoulos Y (2007)	eneralized Hirsch h-index for disclosing latent facts in citation networks. Scientometrics 72(2):253-280, doi: 10.1007/s11192-007-1722-z
Dynamic h-Type index: (Rousseau R, Ye FY (2008)	A proposal for a dynamic h-type index. Journal of the American Society for Information Science and Technology 59(11):1853-1855, doi: 10.1002/asi.20890
k-index:(Ye FY, Rousseau R (2010)	Probing the h-core: an investigation of the tail-core ratio for rank distributions. Scientometrics. In press,doi:10.1007/s11192-009-0099-6
Seniority-independent Hirsch-type index: (Kosmulski M (2009)	New seniority-independent Hirsch-type index . Journal of Informetrics 3(4):341-347, doi:10.1016/j.joi.2009.05.003
Specific-impact s-index: (De Visscher A	An Index to Measure a Scientist's Specific Impact. Journal of the American Society for Information Science and Technology 61 (2) (2010) 319-328. doi:10.1002/asi.21240)
f-index: (Franceschini F., Maisano D	Analysis of the Hirsch index's operational properties. European Journal of Operational Research 203 (2) (2010) 494-504. doi:10.1016/j.ejor.2009.08.001)
Impact vitality indicator: (Rons N., Amez L.	Impact vitality: an indicator based on citing publications in search of excellent scientists. Research Evaluation 18 (3) (2009) 233-241. doi:10.3152/095820209X470563)
m-index: (Bornmann L, Mutz R, Daniel HD (2008) Are there better indices for evaluation purposes than the h index? A comparison of nine different variants of the h index using data from	Are there better indices for evaluation purposes than the h index? A comparison of nine different variants of the h index using data from biomedicine. Journal of the American Society for Information Science and Technology 59(5):830-837, doi: 10.1002/asi.20806.
hw-index: (Egghe L, Rousseau R (2008)	n h-index weighted by citation impact. Information Processing and Management 44(2):770-780, doi: 10.1016/j.ipm.2007.05.003.
hm-index: (Schreiber M (2008)	o share the fame in a fair way, hm for multi-authored manuscripts. New Journal of Physics 10(040201):1-9, doi: 10.1088/1367-2630/10/4/040201.
Normalized h-index: (Sidiropoulos A, Katsaros D, Manolopoulos Y (2007	Generalized Hirsch h-index for disclosing latent facts in citation networks. Scientometrics 72(2):253-280, doi: 10.1007/s11192-007-1722-z.

Index Name	Brief Description and/or link to Research
Tapered h-index: (Anderson TR, Hankin KSH, Killworth PD (2008	Beyond the Durfee square: Enhancing the h-index to score total publication output. Scientometrics 76(3):577-588, doi: 10.1007/s11192-007-2071-2.
hrat-index: (Ruane F, Tol RSJ (2008) Rational (successive) h-indices: An application to economics in the Republic of Ireland. Scientometrics 75(2):395-405, doi: 10.1007/s11192-	Rational (successive) h-indices: An application to economics in the Republic of Ireland. Scientometrics 75(2):395-405, doi: 10.1007/s11192-007-1869-7.
v-index: (Riikonen P, Vihinen M (2008)	National research contributions: A case study on Finnish biomedical research. Scientometrics. 77(2):207-222, doi: 10.1007/s11192-007-1962-y.
e-index: (Zhang CT (2009)	The e-Index, Complementing the h-Index for Excess Citations. PLoS ONE. 4(5):e5429, doi:10.1371/journal.pone.0005429
Multidimensional h-index: (Garcia-Perez MA (2009	A multidimensional extension to Hirsch's h-index. Scientometrics 81(3):779-785, doi:10.1007/s11192-009-2290-1
f-index: (Katsaros D, Akritidis L, Bozanis P (2009) The f Index: Quantifying the Impact of Coterminal Citations on Scientists' Ranking. Journal of the American Society for Information Science an	The f Index: Quantifying the Impact of Coterminal Citations on Scientists' Ranking. Journal of the American Society for Information Science and Technology 60(5):1051-1056, doi:10.1002/asi.2104
π-index: (Vinkler P (2009)	The π-index: a new indicator for assessing scientific impact. Journal of Information Science 35(5):602-612, doi:10.1177/0165551509103601 .
RC- and CC- indices: (Abbasi A., Altmann J., Hwang J.	Evaluating scholars based on their academic collaboration activities: two indices, the RC-index and the CC-index, for quantifying collaboration activities of researchers and scientific communities. Scientometrics 83 (1) (2010) 1-13. doi:10.1007/s11192-009-0139-2)
ch-index: (Ajiferuke I., Wolfram D	Citer analysis as a measure of research impact: library and information science as a case study. Scientometrics 83 (3) (2010) 623-638. doi:10.1007/s11192-009-0127-6
Citation speed s-index: (Bornmann L., Daniel H.D	he citation speed index: A useful bibliometric indicator to add to the h index. Journal of Informetrics 4 (3) (2010) 444-446. doi:10.1016/j.joi.2010.03.007)

Index Name	Brief Description and/or link to Research
h2-lower, h2-center and h2-upper: (Bornmann L., Mutz R., Daniel H.D.	The h index research output measurement: Two approaches to enhance its accuracy. Journal of Informetrics 4 (3) (2010) 407-414. doi:10.1016/j.joi.2010.03.005)
Environment Hj-indices (Dorta-Gonzalez P., Dorta-Gonzalez M.I.	ibliometric indicator based on the h-index. Revista Española de Documentación Científica 33 (2) (2010) 225-245. doi:10.3989/redc.2010.2.733)
$\bar{h}$-index (Hirsch J.E.	An index to quantify an individual's scientific research output that takes into account the effect of multiple coauthorship. Scientometrics 85 (3) (2010) 741-754. doi:10.1007/s11192-010-0193-9)
Role based h-maj-index (Hu X.J., Rousseau R., Chen J	n those fields where multiple authorship is the rule, the h-index should be supplemented by role-based h-indices. Journal of Information Science 36 (1) (2010) 73-85. doi:10.1177/0165551509348133
2nd generation citations h-index (Kosmulski M	Hirsch-type approach to the 2nd generation citations. Journal of Informetrics 4 (3) (2010) 257-264. doi:10.1016/j.joi.2010.01.003)
n-index (Namazi M.R., Fallahzadeh M.K.	n-index: A novel and easily-calculable parameter for comparison of researchers working in different scientific fields. Indian Journal of Dermatology Venereology & Leprology 76 (3) (2010) 229-230. doi:10.4103/0378-6323.62960)
p-index (Prathap G.	The 100 most prolific economists using the p-index. Scientometrics 84 (1) (2010) 167-172. doi:10.1007/s11192-009-0068
Mock hm-index (Prathap G	Is there a place for a mock h-index?. Scientometrics 84 (1) (2010) 153-165. doi:10.1007/s11192-009-0066-2
w-index (Wu Q	The w-Index: A Measure to Assess Scientific Impact by Focusing on Widely Cited Papers. Journal of the American Society for Information Science and Technology 61 (3) (2010) 609-614. doi:10.1002/asi.21276)
b-index (Brown R.J.C.	A simple method for excluding self-citation from the h-index: the b-index. Online Information Review 33 (6) (2009) 1129-1136. doi:10.1108/14684520911011043)
Generalized h-index (Glanzel W., Schubert	Hirsch-type characteristics of the tail of distributions. The generalised h-index. Journal of Informetrics 4 (1) (2009) 118-123. doi:10.1016/j.joi.2009.10.002)

Index Name	Brief Description and/or link to Research
w-index (Wohlin C.	A new index for the citation curve of researchers. Scientometrics 81 (2)(2009) 521-533. doi:10.1007/s11192-008-2155-z)

Source: http://sci2s.ugr.es/hindex#Early%20indices%20based%20on%20the%20h-index

Appendix D

JCAR Definitions of Faculty Activities

Teaching—includes the direct delivery of instruction, as well as those activities supporting the teaching-learning process. Examples of direct delivery of instruction are lectures, seminars, directed study, laboratory session, clinical or student teaching supervision, and placement/placement supervision. Activities directly supporting teaching include class preparation, evaluation of student work, curriculum development, supervision of graduate student research, including thesis or dissertation, academic and career advising, faculty training, and mentoring. Professional development geared to increasing faculty effectiveness in the foregoing activities would also be included

Research or scholarship—includes an array of activities such as conducting experimental or scholarly research, developing creative works, preparing or reviewing articles or books, preparing and reviewing proposals for external funding, performing or exhibiting works in the fine and applied arts, and attending professional meetings or conferences essential to remaining current in one's field.

Service—draws on the professional or academic expertise of a faculty member and includes work within the campus community and outside the campus. Departmental and campus service includes work on various committees (for example, governance, recruitment) and department administration. Community or public service includes consulting, giving speeches, and working in organizations or on committees related to a faculty member's academic field.

An illustration of varying workloads by department (and indeed individual faculty) as surveyed by JCAR can be found below (as reported in Middaugh 2001, p. 41).

TABLE 2.1. DEPARTMENTAL REPRESENTATIONS OF FACULTY ACTIVITY USING JCAR PROTOCOLS.

A	B	C	D	E	F	G	H	I	J	K
				Percent Time Allocated to			*Service Months Generated in*			
Department or Program	Faculty Name	Contract Months	Contract Percent Time	Teaching	Research	Service	Teaching	Research	Service	Total Service Months
Biology	Smith	9	100%	50.0%	37.5%	12.5%	4.50	3.38	1.13	9.00
	Jones	9	100%	50.0%	25.0%	25.0%	4.50	2.25	2.25	9.00
	Brown	9	100%	62.5%	25.0%	12.5%	5.63	2.25	1.13	9.00
	White	9	100%	62.5%	25.0%	12.5%	5.63	2.25	1.13	9.00
	Black	9	60%	100.0%	0.0%	0.0%	5.40	0.00	0.00	5.40
	Johnson	9	100%	0.0%	100.0%	0.0%	0.00	9.00	0.00	9.00
	Adams	9	95%	70.0%	10.0%	20.0%	5.99	0.86	1.71	8.55
	Biology Subtotal						**31.64**	**19.98**	**7.34**	**58.95**
Sociology	Wilson	9	100%	50.0%	25.0%	24.0%	4.50	2.25	2.25	9.00
	McCabe	9	100%	62.5%	25.0%	12.5%	5.63	2.25	1.13	9.00
	Davis	12	100%	40.0%	20.0%	40.0%	4.80	2.40	4.80	12.00
	Sociology Subtotal						**14.93**	**6.90**	**8.18**	**30.00**

Source: JCAR Faculty Assignment Reporting, 1997, p. 10.

Appendix E

Auburn University Institutional Review Board Application and Approval

AUBURN UNIVERSITY INSTITUTIONAL REVIEW BOARD for RESEARCH INVOLVING HUMAN SUBJECTS

REQUEST FOR EXEMPT CATEGORY RESEARCH

For information or help completing this form, contact: **THE OFFICE OF RESEARCH COMPLIANCE,** 115 Ramsay Hall
Phone: 334-844-5966 **e-mail:** IRBAdmin@auburn.edu **Web Address:** http://www.auburn.edu/research/vpr/ohs/index.htm

Revised 2/1/2014 **Submit completed form to IRBsubmit@auburn.edu or 115 Ramsay Hall, Auburn University 36849.**

Form must be populated using Adobe Acrobat / Pro 9 or greater standalone program (do not fill out in browser). Hand written forms will not be accepted.

Project activities may not begin until you have received approval from the Auburn University IRB.

1. PROJECT PERSONNEL & TRAINING

PRINCIPAL INVESTIGATOR (PI):

Name Patricia Bartholomew **Title** Ms. **Dept./School** EFLT-Education

Address 3145 Coastal Highway #1138 **AU Email** pzb0009@auburn.edu

Phone 334 329 7378 **Dept. Head** Sheri Downer

FACULTY ADVISOR (if applicable):

Name Jose Llanes **Title** Dr. **Dept./School** EFLT/Education

Address 4080 Haley Center, Auburn, AL

Phone (334) 844-3074 **AU Email** jrl0001@auburn.edu

KEY PERSONNEL: **List Key Personnel (other than PI and FA).** Additional personnel may be listed in an attachment.

Name	Title	Institution	Responsibilities
none			

KEY PERSONNEL TRAINING: **Have all Key Personnel completed CITI Human Research Training (including elective modules related to this research) within the last 3 years?** ☑ **YES** ☐ **NO**

TRAINING CERTIFICATES: **Please attach CITI completion certificates for all Key Personnel.**

2. PROJECT INFORMATION

Title: Measuring Faculty Productivity: Toward a Unified Methodology

Source of Funding: ☑ **Investigator** ☐ **Internal** ☐ **External**

List External Agency & Grant Number: none

List any contractors, sub-contractors, or other entities associate with this project.
none

List any other IRBs associated with this project (including those involved with reviewing, deferring, or determinations).
none

FOR ORC OFFICE USE ONLY

DATE RECEIVED IN ORC: ________ by ________	**APPROVAL #** ________
DATE OF IRB REVIEW: ________ by ________	**APPROVAL CATEGORY:** ________
DATE OF ORC REVIEW: ________ by ________	**INTERVAL FOR CONTINUING REVIEW:** ________
DATE OF APPROVAL: ________ by ________	

COMMENTS:

3. **PROJECT SUMMARY**

a. **Does the research involve any special populations?**

- ☐ YES ☑ NO **Minors (under age 19)**
- ☐ YES ☑ NO **Pregnant women, fetuses, or any products of conception**
- ☐ YES ☑ NO **Prisoners or Wards**
- ☐ YES ☑ NO **Individuals with compromised autonomy and/or decisional capacity**

b. **Does the research pose more than minimal risk to participants?** ☐ YES ☑ NO

Minimal risk means that the probability and magnitude of harm or discomfort anticipated in the research are not greater in and of themselves than those ordinarily encountered in daily life or during the performance of routine physical or psychological examinations or tests. 42 CFR 46.102(i)

c. **Does the study involve any of the following?**

- ☐ YES ☑ NO **Procedures subject to FDA Regulation** Ex. Drugs, biological products, medical devices, etc.
- ☐ YES ☑ NO **Use of school records of identifiable students or information from instructors about specific students**
- ☐ YES ☑ NO **Protected health or medical information when there is a direct or indirect link that could identify the participant**
- ☐ YES ☑ NO **Collection of sensitive aspects of the participant's own behavior, such as illegal conduct, drug use, sexual behavior or use of alcohol**
- ☐ YES ☑ NO **Deception of participants**

If you checked "YES" to any response in Question #3 STOP. It is likely that your study does not meet the "EXEMPT" requirements. Please complete a PROTOCOL FORM for Expedited or Full Board Review.
You may contact IRB Administration for more information. (Phone: 334-844-5966 or Email: IRBAdmin@auburn.edu)

4. **PROJECT DESCRIPTION**

a. **Subject Population** (Describe, include age, special population characteristics, etc.)

Higher Education Faculty members

b. **Describe, step by step, all procedures and methods that will be used to consent participants.**

☑ **N/A (Existing data will be used)**

2 of 3

c. **Brief summary of project.** (Include the research question(s) and a brief description of the methodology, including recruitment and how data will be collected and protected.)

Attempts to measure faculty productivity have focused primarily on research productivity as the sole proxy for overall productivity. Other measures have been developed to calculate productivity for alternative tasks in a typical instructor' s workload allocation. These tasks typically include teaching and service. However these methodologies remain as stand-alone measures of multiple facets of a faculty member' s productivity. This work attempts to consolidate those measures of faculty and departmental productivity into a single indicator comparable across a single campus as well as among several campuses.

The overarching research question is: What is an appropriate method for combining the existing measures of faculty productivity into one metric that provides fair comparability across disciplines and colleges?
In order to approach the answer, the topic will be investigated as a series of answers to the following sub questions:
1) Which facets of faculty workload will be measured?
2) Which measures of productivity should be used to measure the different facets of faculty workload?
3) Devising a methodology for a fair combination of these facets
4) Calculating a practical application of the methodology to at least two faculty members or departments in the United States.

Data will be gathered from existing public sources: The faculty's own CVs that publically available on the respective university's website; individual salary data from Open Alabama and similar sources; average salary data from the census or from national faculty surveys (AAUP); faculty workload allocations averages from the Delaware study or Faculty Workload Survey. All data is already existing and publically available.

d. **Waivers. Check any waivers that apply and describe how the project meets the criteria for the waiver.**

- [x] **Waiver of Consent (Including existing de-identified data)**
- [ ] **Waiver of Documentation of Consent (Use of Information Letter)**
- [ ] **Waiver of Parental Permission (for college students)**

e. **Attachments.** Please attach Informed Consents, Information Letters, data collection instrument(s), advertisements/recruiting materials, or permission letters/site authorizations as appropriate.

Signature of Investigator V. Burtt Date 8/29/14

Signature of Faculty Advisor [signature] Date 9/3/14

Signature of Department Head Sherida Downer Date 9/3/14

3 of 3

CITI Collaborative Institutional Training Initiative

IRB # 2 Social and Behavioral Emphasis - Non-AU Personnel (Blue) Curriculum Completion Report
Printed on 5/24/2013

Learner: Patricia Bartholomew (username: pcbartholomew)
Institution: Auburn University
Contact Information Department: Higher Education
Phone: 9042179311
Email: pzb0009@auburn.edu

IRB # 2 Social and Behavioral Emphasis - Non-AU Personnel (Blue): This course is appropriate for students doing class projects that qualify as "No More Than Minimal Risk" human subjects research.

Stage 1. Basic Course Passed on 05/24/13 (Ref # 10416280)

Required Modules	Date Completed	Score
Belmont Report and CITI Course Introduction	05/24/13	3/3 (100%)
Informed Consent - SBR	05/24/13	5/5 (100%)
Privacy and Confidentiality - SBR	05/24/13	3/5 (60%)
Elective Modules	**Date Completed**	**Score**
Research and HIPAA Privacy Protections	05/24/13	4/5 (80%)

For this Completion Report to be valid, the learner listed above must be affiliated with a CITI participating institution. Falsified information and unauthorized use of the CITI course site is unethical, and may be considered scientific misconduct by your institution.

Paul Braunschweiger Ph.D.
Professor, University of Miami
Director Office of Research Education
CITI Course Coordinator

Return

COLLABORATIVE INSTITUTIONAL TRAINING INITIATIVE (CITI)

IRB # 2 SOCIAL AND BEHAVIORAL EMPHASIS - AU PERSONNEL (BLUE) - BASIC/REFRESHER CURRICULUM COMPLETION REPORT

Printed on 09/03/2014

LEARNER	Patricia Bartholomew (ID: 3537999)
DEPARTMENT	Higher Education
PHONE	9042179311
EMAIL	pzb0009@auburn.edu
INSTITUTION	Auburn University
EXPIRATION DATE	09/02/2017

IRB # 2 SOCIAL AND BEHAVIORAL EMPHASIS - AU PERSONNEL (BLUE) - BASIC/REFRESHER : No direct contact with human subjects.

COURSE/STAGE:	Basic Course/1
PASSED ON:	09/03/2014
REFERENCE ID:	13895821

REQUIRED MODULES	**DATE COMPLETED**	**SCORE**
The Regulations - SBE	09/03/14	5/5 (100%)
Assessing Risk - SBE	09/03/14	5/5 (100%)
Informed Consent - SBE	05/24/13	5/5 (100%)
Privacy and Confidentiality - SBE	05/24/13	3/5 (60%)
Students in Research	09/03/14	8/10 (80%)
Unanticipated Problems and Reporting Requirements in Social and Behavioral Research	09/03/14	3/3 (100%)
Belmont Report and CITI Course Introduction	05/24/13	3/3 (100%)
ELECTIVE MODULES	**DATE COMPLETED**	**SCORE**
Cultural Competence in Research	09/03/14	5/5 (100%)
Research and HIPAA Privacy Protections	05/24/13	4/5 (80%)

For this Completion Report to be valid, the learner listed above must be affiliated with a CITI Program participating institution or be a paid Independent Learner. Falsified information and unauthorized use of the CITI Program course site is unethical, and may be considered research misconduct by your institution.

Paul Braunschweiger Ph.D.
Professor, University of Miami
Director Office of Research Education
CITI Program Course Coordinator

IRB Approval Email

Send Options... HTML

To...

Cc...

Subject: FW: Approval, Exempt Protocol #14-323 EX 1409

Attached: Investigators Responsibil~1.docx (16 KB) [Open as Web Page]

Tahoma 10

From: IRB Administration
Sent: Wednesday, October 08, 2014 3:26 PM
To: Patricia Bartholomew
Cc: Jose Llanes; Sheri Downer
Subject: Approval, Exempt Protocol #14-323 EX 1409

Use IRBsubmit@auburn.edu for protocol-related submissions and IRBadmin@auburn.edu for questions and information.
The IRB only accepts forms posted at https://cws.auburn.edu/vpr/compliance/humansubjects/?Forms and submitted electronically.

Dear Ms. Bartholomew,

Your protocol entitled " Measuring Faculty Productivity: Toward a Unified Methodology " has received approval as "Exempt" under federal regulation 45 CFR 46.101(b)(4).

Official notice:
This e-mail serves as official notice that your protocol has been approved. A formal approval letter will not be sent unless you notify us that you need one. By accepting this approval, you also accept your responsibilities associated with this approval. Details of your responsibilities are attached. Please print and retain.

Consent documents:
Since you do not have to wait to for the return of any consent documents, please conduct your study at your convenience.

Expiration – Approval for three year period:
*****Note that the policy for Exempt approvals is a *three year approval*.** Therefore, your protocol **will expire on September 28, 2017 .** Put that date on your calendar now. About three weeks before that time you will need to submit a renewal request.

When you have completed all research activities, have no plans to collect additional data and have destroyed all identifiable information as approved by the IRB, please notify this office via e-mail. A final report is no longer required.

If you have any questions, please let us know.

Best wishes for success with your research!
Susan