2018

Assessing Research Productivity from an Institutional Effectiveness Perspective: How Universities Influence Faculty Research Productivity

Michael M. Rawls
Virginia Commonwealth University

Assessing Research Productivity from an Institutional Effectiveness Perspective: How Universities Influence Faculty Research Productivity

A dissertation submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy at Virginia Commonwealth University

by

Michael M. Rawls
M.A. in History, Virginia Commonwealth University, 2003
B.S. in Business Administration, Virginia Commonwealth University, 1994

Director: Dr. Maike Philipsen
Professor, School of Education

Virginia Commonwealth University
Richmond, Virginia

Acknowledgements

I have to acknowledge my amazement at the broad swath of people across the entirety of Virginia Commonwealth University who were willing to take a selfless, substantive, and earnest interest in guiding me through the process of this dissertation. That begins with Dr. William C. Bosher, an infectiously positive and supportive man who graciously agreed to serve as my advisor, although he was already chairing more dissertation committees than any faculty member of whom I was aware. Upon his unfortunate passing, as I was trying to figure a way forward, Dr. Maike Philipsen agreed to take his place. I was surprised that a faculty member from another school, whom I have never met, would consider investing herself so fully in a student unknown to her, but she deserves more credit than I can offer in guiding me – when not pushing me – through to completion. The other members of my committee also all deserve special praise. As a graduate student who knew just enough about statistics and regression analysis to be dangerous, Dr. I-Shian Suen helped keep me on point. Dr. Ann Nichols-Casebolt not only offered the perspective of an accomplished research administrator, but was kind enough to agree to serve just as she was retiring. My discussions with Dr. Henry Brownstein provided me with insight into the practical and applied world of qualitative research that was very helpful in preparing me to conduct a type of research that I had theretofore only read about. Finally, when we realized that my committee drew so broadly across the campus that an additional Widler School faculty member was needed, Dr. Sarah Jane Brubaker, my program director, selflessly stepped in at the last minute. The dedication that I received from each of you has made a permanent impression.

I also received enormous help from VCU people beyond my committee. I owe a debt of gratitude to Dr. John Leonard, of the School of Engineering, for introducing me to a mountain of engineering data, along with the insights of someone who has invested no small amount of time in contemplating faculty productivity in the field of engineering. I also owe special thanks to Julie Arendt of VCU Libraries – although she reminds me that I have already thanked her – for determining an ingenious approach to collecting a vital portion of my data when my plan had fallen flat. I would also like to thank Jeanne Hammer, John Duke, and John Ulmschneider – my supervisors at VCU Libraries – for all allowing me the flexibility and space to juggle this dissertation with my work schedule. I really have no idea how other full time employees complete a dissertation, but I know I would not have completed this one without your support and understanding.

Most importantly, I would like to thank my family. My wife, Jennifer, never stopped encouraging me, never had a doubt in me, and offered me patience during times of stress. My mother, father, and sister – educators all – are responsible for instilling in me a lifelong interest in learning. Finally, I would like to thank all the other members of my family and my friends who have helped cheer me on with their encouragement and positive thoughts.

Table of Contents

Abstract

ASSESSING RESEARCH PRODUCTIVITY FROM AN INSTITUTIONAL EFFECTIVENESS PERSTPECTIVE: HOW UNIVERSITIES INFLUENCE FACULTY RESEARCH PRODUCTIVITY

Michael M. Rawls, Public Policy and Administration, Ph.D.

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Virginia Commonwealth University

Virginia Commonwealth University, 2018

Director: Dr. Maike Philipsen
Professor, School of Education

Faculty research productivity studies typically focus on the scholarly performance of the individual researcher, although environmental and organizational conditions that are largely outside of the researcher's control can significantly impact that performance. From an institutional effectiveness perspective, it is imperative for the higher education administrators and leaders who share the responsibility of managing and supporting their university's research enterprise to understand how the institutional environment itself impacts the productivity of its research community. In this sequential mixed methods study, a quantitative framework was tested for assessing institutional effectiveness in research administration based on the assertion that this concept can be measured indirectly, at the departmental level, based on the calculation of a program's residual scholarly output. This is the difference between the actual amount of

scholarly output a program produces compared to the predicted amount of scholarly output that its resources suggest it is capable of producing. The assumption is that the institution's effectiveness in supporting research is largely reflected by the extent to which a program over- or under-produces scholarship based on its level of resources. The residual scholarly output was calculated for each Ph.D.-granting biomedical engineering program in doctoral universities with a Carnegie classification of "highest research activity" for the period of 2014 through 2016. A sampling of those programs that achieved among the highest and lowest residual productivity levels then became the subject of a qualitative inquiry where researchers and administrators were interviewed with two goals in mind. The more ostensive goal was to reveal what factors, characteristics, resources, and conditions distinguish under- and over-producing programs for the purpose of informing best and worst practices in research administration. Equally important, the second goal was to determine if the quantitative framework was actually successful in distinguishing institutional effectiveness in supporting research. The study concludes that the quantitative framework proved to be a successful method for detecting institutional effectiveness in supporting research, and that the primary distinguishing characteristic between high and low-functioning environments was how well programs were able to reduce the general administrative burdens that researchers face, particularly in grant management and the operation of research laboratories.

CHAPTER I. INTRODUCTION

This study explores how institutional policies, practices, climate, resources and other organizational characteristics impact the scholarly research productivity of a university's faculty. Its goal is to measure institutional performance in research administration, rather than the performance of faculty researchers, and to understand what drives the success of those institutions that are among the most effective at positively influencing productivity. While the primary responsibility to produce scholarship rests with individual faculty members, the organizational environment in which they labor to produce new knowledge is a responsibility that is shared by multiple layers of leadership. The adequacy of that environment is shaped by the effectiveness of each institution's approach toward supporting its research community at all organizational levels. Focusing on the discipline of biomedical engineering (BME), this study is designed to identify those institutions that are the most effective at producing scholarly research in this field, given their resources, and to then explore those organizational factors and conditions that differentiate such institutions from their less effective peers.

To accomplish this task, the study tested a new approach designed to detect institutional effectiveness in supporting research. The approach involves calculating the residual scholarly output, a concept described later, for each BME program based on the theoretical assumption that those programs with the highest residual scholarly output would reside in institutions that were among the most effective at supporting the research efforts of their scholars, while those with the lowest would reside in institutions that were hampering researcher productivity. While the more conspicuous goal of this study is to inform practice in research administration, an equally important goal was to determine if the proposed approach actually worked. In the event that the approach did not function as intended, and if residual scholarly output proved not to be a useful means of gauging institutional effectiveness, then the findings of the study would be essentially null. On the other hand, if the residual approach did work, not only would the findings be potentially beneficial in informing practice, but it would represent a new methodological approach to assessing institutional research productivity. Ultimately, the findings of the study did suggest that the residual approach was successful, as is argued in chapter 5. As such, the implications of the method itself is equally as significant a contribution, if not more so, than the findings regarding institutional effectiveness in support of research.

Rationale for the Study

Academics are evaluated in terms of their accomplishments in teaching, scholarship, and service - the three criteria codified in promotion and tenure guidelines across academia which illustrate the expectations for those in the profession (Fairweather, 2002). What Massy and Wilger (1995) found in interviewing faculty members and administrators at numerous universities, however, was that research and publications are viewed as what matters most when it comes to institutional rewards. This uneven valuation is not a secret, nor is it entirely without good reason. Advancing research and generating new knowledge is a fundamental role of research universities. Accordingly, institutions attach great value to the research and scholarly productivity of their faculty. This pursuit of scholarship is the source of tremendous public benefit by forging discoveries in science and medicine, exploring socially relevant issues, enriching our understanding of the arts and humanities, creating technology transfer and impacting economic development.

Societal goals aside, however, the quantity and quality of scholarship that a research university produces also has a beneficial impact on its reputational status. The more accomplished a university's research and publishing profile, the better positioned it will be to secure new grant funding, to attract and retain quality faculty, and to appeal to more prospective students (Bland & Ruffin, 1992). It is for this reason that the responsibility to provide the policies, practices and resources necessary to foster the type of environment capable of sustaining high levels of scholarly production rests squarely with university leadership. As Bland and Ruffin (1992) contend, "personal characteristics are essential but insufficient by themselves" and that "to be productive, researchers, it seems, must have certain personal characteristics and in addition must work in environments conducive to research" (p. 386). This means going beyond resting upon the scholarship requirements of the promotion and tenure guidelines as a means of motivating individual productivity and towards optimally furnishing the support, services, and workplace environment that helps ensure that "researchers are focused on research, not the ancillary things that surround it" (Research Information Network, 2010).

At its core, this study is concerned with institutional effectiveness. It is essential for those academic leaders who share the responsibility of planning, developing, and sustaining the various aspects of their organization's research mission to be aware of those factors that define productive research environments. A university's research enterprise is large, complex, and continuously evolving. It includes investment in research infrastructure, management of research space, support for locating and securing grants, grant management and accounting, institutional review, the negotiation and allocation of facilities and administrative cost recovery funds, investment in library materials, data management and curation, assistance in technology transfer, the management of intellectual property rights, and the oversight of research integrity. Some of these supports and resources are provided centrally, others by individual academic units, and many are provided by a combination of both. How well these components are strategically assembled, coordinated, prioritized, and managed are bound to either aid or hinder the productivity of researchers. Likewise, the organizational culture and climate in which the researcher is placed can also combine to create an atmosphere that is conducive to research productivity or to create one that confounds it (Bland et. al., 2005). To the degree that research productivity is influenced by a sum of various institutional dimensions, it warrants investigation as to which of those factors are most associated with productive research environments. This is true from both the perspective of providing good stewardship of the public funds directed toward biomedical engineering research and in running a cogent research enterprise that achieves a reputation for high-level performance.

Rationale for Biomedical Engineering Focus

There are several contributing factors that make biomedical engineering an appealing discipline of focus for the purposes of this study. As a STEM-H field, research in BME has a propensity to be dependent on a greater number of institutional factors which are beyond the control of the individual researcher than some other disciplines, such as those in the humanities or social sciences. In addition to the usual challenges faced by researchers across disciplines, BME requires adequate access to lab space and equipment prior to funding; it is fundamentally dependent on the labor of graduate students and; by its own nature, BME requires substantial interdisciplinary collaboration. BME also features a significant commercialization component that presents an additional layer of complexity. The philosophical approach that a university or engineering school takes toward commercialization of research, management of intellectual property rights and the facilitation of technology transfer can add or relieve the additional pressures BME faculty members face in the course of conducting research.

Another reason for selecting BME is one of practicality. As the methodology section will address, this study employs a research design where individual departments will serve as the unit of analysis. It seems more feasible to attempt to understand the workings, climate, and issues that define a single department rather than to attempt to understand the vast sum of moving parts that define entire schools or other large organizational units. So, while other STEM-H programs may also face similar conditions and challenges as those listed above, a particular discipline had to be selected and BME appears particularly suitable.

Use of Key Terms

Several terms are used throughout this study that have similar meanings. In order to add clarity, some of the key terms and how they interrelate are summarized in this section. These are not complete definitions, but rather a description of how these terms are used for the purposes of this study specifically. Each of the terms is more fully addressed elsewhere in the study. *Research productivity*, or just *productivity*, refers to the amount and quality of scholarly output collectively produced by a BME program. *Productive efficiency* refers to a program's *productivity* relative to the level of resources, or inputs, that the program has dedicated toward producing scholarly output. *Residual scholarly output*, or just *residual output,* is the means by which *productive efficiency* is detected and measured. *Institutional effectiveness* is how well an institution – or subunit therein – goes about achieving its goals and mission.

The Institutional Effectiveness Perspective and Research Questions

This study conceives of institutional effectiveness as the ability of a university, or any unit therein, to meet the objectives of its mission, including its research mission (O'Meara, 2005; Volkwein, 2010; Head, 2011). Approaching faculty research productivity from an institutional effectiveness perspective is somewhat distinct from traditional research productivity studies or inquiries. The goal is not to understand what drives individual researchers or what personal qualities make them more or less productive. Nor is it a question of which BME programs are producing the greatest volume of high-impact scholarship, bringing in the most grant funding, or how certain programs rank amongst peer groupings – all of which can be answered using readily available data. Instead, this perspective is most concerned with how effective the institution is at facilitating research. How well are the resources functioning that have been put in place to foster productivity? What processes need improving? What resources are missing? Where would additional investment provide most impact? Overall, is the level research productivity of the institution, or unit therein, reasonable given the resources that are being invested?

While programs are commonly ranked in terms of research expenditures or publication levels, such rankings are unable to indicate which programs are producing the most high-impact scholarship *given their resources*, thus exceeding expectations. We do not know which programs are making the most of what they have or, more precisely, which programs are producing more high quality and high visibility scholarship than their resources suggest they out to be able to produce. Conversely, programs that may appear to be productive by traditional measures may actually be low-achieving when their scholarly output is weighed against the level of resources that they have dedicated to supporting research activity. The idea that scholarly output – as the most fundamental measure of research productivity – should be weighed against the resources or inputs that go into supporting the research process can be more accurately described as productive efficiency rather than research productivity. The concept of productive efficiency, which forms the basis of this dissertation, begins to describe the institutional effectiveness perspective because it is primarily concerned with identifying the extent to which programs are making effective use of their resources. By establishing the productive efficiency of each BME program, both high- and low-productive efficiency programs can be identified. Once identified, the characteristics and environmental conditions that distinguish the most efficient programs from their least efficient peers is the topic this dissertation will address, using the following research questions:

R1: What factors of faculty research productivity are most strongly correlated with scholarly output as measured by weighted article count per tenured or tenure track faculty at PhD-granting biomedical engineering programs in Carnegie-classified “highest research activity” doctoral universities?

R2: Which Ph.D.-granting biomedical engineering programs in Carnegie-classified “highest research activity” doctoral universities exhibit the highest and the lowest levels of productive efficiency in creating scholarly output?

R3: Which institutional factors most influence the research efficiency of faculty research productivity efforts in biomedical engineering departments in US doctoral institutions?

R4: How do these institutional factors influence the research efficiency of faculty research productivity efforts in biomedical engineering departments in US doctoral institutions?

Conceptual Framework

The reasoning behind the line of questioning above is based on the assumption that some portion of the residual success attained by programs that outperform their measurable resources is attributable to effective administrative agency – whether it is in the form of cultivating the right workplace environment, putting forth the right policies (or discarding the wrong ones), investing in the best supports and resources to maximize productivity, placing proper emphasis on professional development, or any other number of actions that can positively shape the environment in which the researcher pursues his or her endeavors. This approach borrows from the economic concept of total factor productivity (Hulten, 2007) where the residual output between predicted and actual production levels is assumed be driven by factors not accounted for in the model, such as managerial competence or the incorporation of technology among other things. In this sense, the central idea of the study the *residual scholarly output* is an indirect measure of institutional effectiveness in facilitating scholarly productivity. According to this line of thinking, if the collective level of scholarly output produced by a program's faculty aligns with the level of resources the program puts into the research process – that is to say it is producing at the expected capacity – then the assumption, from an institutional effectiveness perspective, would be that the program is operating at an acceptable level. Moreover, this assumption would hold true regardless of where the program resides on the continuum of programs when ranked by overall research output. However, if a program is not meeting its expected output, then there is an institutional effectiveness problem to be found in that program's operations. And finally, as mentioned above, if a program is significantly out-producing its expected capacity, then it can be judged as highly effective and must be doing something better than its peers which warrants investigation – again, regardless of where it may reside in the overall ranking of programs by scholarly output or research funding. In this sense, a central assertion of this study is that the basis for measuring research productivity from an institutional effectiveness perspective should center on measuring a program's *productive efficiency*, or the relation of a program's scholarly output to the level of inputs invested into its research process. This concept of research efficiency is the basis by which this study argues that institutional effectiveness in research productivity can be empirically distinguished from the cumulative success of individual researchers as measured by sheer volume.

By focusing on the maximization of research output in relation to inputs, rather than the maximization of output alone, this approach aligns well with the tenets of institutional effectiveness in general. The review of the literature surrounding that topic demonstrates that the concepts of efficiency and continuous improvement are of great practical importance in measuring, assessing and, ultimately judging programmatic success. As an example, when an individual attains a position of leadership within any organization, it presumably comes with the minimum expectation they will do what is necessary to maintain the effectiveness of that part of the organization that falls within their purview. More realistically, the expectation is that they would improve upon what they inherited. This is true of academic leaders, whether the position in question is a department chair, associate dean, dean, vice president, or so on. Upon assuming a leadership role, each person is tasked with improving upon the situation and resources they have acquired so that their department, school, or university will be better positioned for the future. Whether the organizational unit in question is cash-strapped or has an embarrassment of riches, administratively speaking, is relative. The measure of success for leaders is whether or not they improved upon the circumstances they inherited. It is for this reason that the phrase "given their resources" is employed frequently in this paper. It summarizes the goal of neutralizing the size effect that grant funding, departmental expenditures, number of faculty researchers, and other resources have on the overall output of faculty research in order to get at who is doing the best at maximizing productivity of the resources they have.

To illustrate this position, consider the spectrum of BME programs in Carnegie-classified research institutions as ranked by research expenditures. The top ten highest spending programs recorded an average of $31.2M in research expenditures in 2014, roughly six times above the national average for BME programs (NSF, 2014). Logically, these top-spending programs should be expected to achieve a level of scholarly research productivity on a scale somewhat proportional to their funding. Being so well-positioned to excel, it should be considered a failure if one of these programs were to produce only an average amount of scholarly output in terms of quantity, quality and visibility. Alternatively, if a middling program, in terms of funding, were to generate scholarship on a level approaching the top spending programs, it should be considered a significant success in terms of effectiveness. Of course grant funding will not be the only program characteristic explored in this analysis, but this hypothetical middling program can serve as one example of the sort of program that this study will attempt to quantitatively identify and then qualitatively investigate in contrast with under-performing programs serving as antithetical examples.

A viable counter argument to this research efficiency approach is the view that the most important measure of institutional effectiveness in research administration is how much grant funding a department can consistently attract. The suggestion is that grant funding is the key ingredient and, if a program is continuously successful in attracting grants, then all else should follow. Further, it could be suggested that to devise a quantitative model that may neutralize total grant funding to some degree by instead focusing on the amount of research output relative to grant funding (and other inputs), is to potentially miss the point. This is a reasonable position and a study that focuses on institutional effectiveness as it relates specifically to the acquisition of extramural funding would certainly have its place. Such a study, however, does not contemplate how two programs with equal funding might have unequal productivity.

To that end, this dissertation serves as a study in contrast. It relies on the use of a statistical model capable of predicting the approximate level of scholarly output each PhD-granting BME program should be able to produce given their inputs into the research process. Rather than focusing on the characteristics of highly productive BME programs or even productivity in BME programs in general, this study qualitatively compared BME departments with the greatest levels productive efficiency to those that have among the lowest. This contrast was intended to make evident the key differences between both environments and to understand how opposing traits can influence research productivity. This approach of focusing on direct opposites is unique in the study of faculty scholarly research productivity as is the proposed approach for defining and identifying high- and low-productive efficiency programs.

Brief Overview of the Literature

The literature review explores the body of scholarship pertaining to faculty research productivity, concepts of productivity that can be borrowed from the field of economics in support of the study's objectives, and the concept of institutional effectiveness in higher education.

Faculty research productivity literature can be generally divided between those studies that approach the question from the context of the individual researcher and those that explore it from an institutional perspective (Abouchedid & Abdelnour, 2015). Both strains of inquiry are typically concerned with identifying the factors associated with higher or lower levels of research productivity, at their respective units of analysis, in order to explain observed differences in productivity between like entities (Helsi & Lee, 2011). While this dissertation examines program-level research productivity exclusively, such studies are much less common than studies pertaining to individual faculty members. As such, both types of studies will be reviewed in an effort to understand the extent to which they can inform this proposed dissertation.

Studies pertaining to individual productivity inform this dissertation in several ways. First, many researchers in this area have applied economic concepts to their studies, including the recognition that productivity should not be conceived of as a raw count of outputs, but as a ratio inputs to outputs (Levin & Stephan, 1991; Bieber & Blackburn, 1993; Masey & Wilger, 1995; Eagen & Garvey, 2015). Similarly, Levin and Stephan (1991) delved more deeply into economic concepts by applying aspects production theory to faculty research productivity – an approach that this study also employsf. A second implication of the literature on individual research productivity is the use of journal articles produced by each faculty member over a finite time period as the most common measure of productivity (Blackburn & Bieber; Jones & Preusz, 1993; Steurly & Maranto, 1994; Bellas & Toutkoushian, 1999; Hu & Gill, 2000, Hesli & Lee, 2011). Journal articles are far from the only form of scholarly expression, but they are the most frequent and common form and, as such, they can be expected to be distributed more evenly across a period of a few years than could books, book chapters, patents, or other forms of scholarship. They are also more accessible and easily quantified than other forms of scholarship.

Studies that explore productivity at the program or institutional level are not only rare, but among the few that were located, most were concerned only with ranking programs for prestige purposes rather than understanding the impact that organizational influences exert on productivity. As with many of the individual studies, however, some institutional studies also draw on economic theory, echoing the concept that productivity should be conceived of as a ratio of output to inputs rather than output alone (Dundar & Lewis, 1998; Toutkoushian & Porter, 2005; Miller et al., 2013). The primary manifestation of this concept can be found in programmatic studies that recognize that the effects of department-size should be offset by using a measure of scholarly output per faculty member rather than scholarly output per department (Dundar & Lewis, 1998; Toutkoushian & Porter, 2005).

Dunbar and Lewis (1998) come closest to the objectives of the project proposed here. Their study of the relationship between research productivity in STEM-H programs and institutional factors shares the same stated goal of informing institutional effectiveness efforts. In studying 1,834 programs at 90 Carnegie classified Research 1 universities, they found a variety of characteristics that correlated with higher productivity, including number of program graduate students, library expenditures, program size (faculty count), percentage of faculty with a recent publication, and percentage of faculty with a grant. As with Levin and Stephan (1991) in their individual productivity study, this study also regressed scholarly output for each department against an array of department inputs and characteristics, placing both studies into a small group within the productivity literature that applies the concept of the production function to the research environment. To better understand the concept of the production function and its applicability to the research productivity at the institutional level, it is necessary to turn directly to the economics literature.

In production theory, the various factors of production are combined into a production function that expresses the relationship between the quantity of inputs and the quantity of outputs

(Robinson, 1955; Hulten, 2001; Fioretti, 2007). One approach is to regress actual output measures against actual input measures to derive parameter estimates that each of the factors of production have in relation to output (Hulten, 2001). This approach is relevant to studying research productivity at the departmental or institutional level more so than to individual faculty productivity, because organizational entities are more analogous to the concept of a firm, which is the unit of analysis commonly used in production theory (Fioretti, 2007). Furthermore, production theory offers a means of identifying when a particular firm is exhibiting greater production efficiency than its industry counterparts. This is possible due to a residual effect that is produced when a firm's output levels exceed the amount that is expected given the level of capital and labor inputs the firm employs in production. As Hulten (2007) summarizes the concept of this residual further:

> "…the residual captures change in the amount of output that can be produced by a given quantity of inputs. Intuitively, ***it measures a shift in the production function*** [caused by] technical innovations, organizational and institutional changes, shifts in societal attitudes, fluctuations in demand, changes in factor shares, omitted variables, and measurement errors" (p. 40).

This residual is at the heart of a concept known as total factor productivity (TFP). According to Black (2012), "TFP reflects how efficiently the inputs and the given technology are utilized…[by identifying] that portion of the output not explained by the quantity of the inputs into production." Of course, as Hulten points out above, that residual may also include components of measurement error or other influences not related to efficiency. Nevertheless, the effect that superior organization, management, and incorporation of technology have on production efficiency are also contained within this residual, even though the parameters of their effects cannot be precisely defined (Hulten, 2007). Identifying the residual is what makes it conceivable to quantitatively identify, albeit indirectly, the effects that administrative, environmental, and other immeasurable characteristics can exert on scholarly production.

Measuring and assessing how institutional efforts relate to institutional outcomes is at the core of institutional effectiveness and, as such, is what ties the concept of the residual above to the concept of institutional effectiveness. While the term *institutional effectiveness* is most frequently associated with the assessment of student learning outcomes, most authors describe it in broader terms, recognizing that it has assumed a variety of meanings while lacking a consistently agreed upon definition (Bauer, 2001; Head, 2011). In particular, Head (2011) views institutional effectiveness as an umbrella term that encompasses "a broad-scaled, institution-wide process consisting of specific components, including the evaluation of all academic programs, administrative units, and support services; the assessment of student learning outcomes; and data-driven support by the institutional research arm of a college or university" (p.10).

While most authors view the scope of institutional effectiveness to include the assessment of any activity or program of importance to a university, no studies were found that apply this principle exclusively to the university's capacity to support the research productivity of their faculty. There are studies, however, that draw on methodologies similar to the one proposed for the quantitative portion of this dissertation in evaluating other aspects of the university. Most notably, this includes a study by Horn and Lee (2016) regarding the use of residuals between expected and actual graduation rates as a means for testing institutional effectiveness in that area. The impetus for their work stemmed from their concern that "despite the widespread use of graduation rates in accountability systems, it is doubtful that the relevant dimensions of institutional effectiveness are being adequately assessed" (p.470). By this, the authors mean that although a college or university can control some of the factors that influence academic performance and persistence, other factors – such as the socioeconomic profile of the institution's students – were also major determinants. They concluded that "raw graduation rates may thus better reflect advantages and circumstances than strengths or policy and practice" (p.470). In some ways the aim of this study could be viewed as an inversion of the aim of this dissertation, because the authors seek to demonstrate how factors outside of an institution's control can influence important outcome. In a more importance sense, however, this dissertation views research productivity in precisely the same way that Horn and Lee (2007) view "raw" graduation rates – which is that you cannot measure the institutional effectiveness of either concept without considering the inputs.

Additionally, the literature includes several studies that partially or fully draw on the concept of the production function and apply it to institutional outcomes such as graduation rates, retention rates, student engagement, etc. Again, none of the studies dealt with research productivity, but Powell, Gilleland, and Pearson (2012) used data from the U.S. Department of Eduction's Intergrated Postsecondary Education Data System to examine how institutional characteristics and categories of institutional expenditures compared to both effectiveness (retention and graduation rates) and efficiency (faculty workload data). They found effectiveness increased as spending went up, but efficiency – the amount of institutional outputs per unit of faculty effort – declined. The primary concern of the authors was that most measures of success demanded by and reported to stakeholders are associated with productivity only. The policy implication raised by the results of the study suggests that by failing to account for efficiency while focusing only on productivity could contribute cost increases in higher education. As with Horn and Lee's residual, Powell et al. (2012) indicates that institutional effectiveness assessments must weigh university outcomes against institutional inputs to be meaningful.

Finally, the literature suggests that leadership and workplace climate, particularly at the department level, can have a significant impact on a program's research productivity. In his attempt to understand how best to evaluate academic departments as units, Wergin (2003) came to conclude that department climate was more important to improving departmental quality than any means of assessing quality. Regardless of what aspect of a department's performance is being evaluated, he found that those departments that collectively develop a clear vision of what they want to accomplish were the best at assuring quality. He suggests that this requires a department chair that can communicate vision clearly and who understands that establishing and meeting goals must be a process that is openly negotiated with the faculty. In this sense, he suggests that a chair who can cultivate a meaningful engagement between a program's faculty and it goals is central to establishing high-performing departments. Contrastingly, he argues that goals and evaluative frameworks imposed from the university level without the process of negotiation tend only to produce a "compliance mentality."

Chapter 2 shows how the literature surrounding these somewhat disparate subject areas of faculty research productivity, economics production theory, and institutional effectiveness can be synthesized in such a way as to help identify how institutions influence the productivity of their faculty in both positive and negative ways.

Methodology

This study employed a sequential mixed methods approach to identify those BME programs that exhibit among the highest and lowest levels of productive efficiency in producing scholarly research, given their resources, and then to explore those organizational factors and conditions that most differentiate these institutions from one another. The quantitative portion of the dissertation was designed to create a cross-sectional statistical model capable of ranking

BME programs by their residual scholarly output, which is the extent to which a program's actual research output exceeds or lags behind the model's predicted research output based on a program's inputs into the research process. The purpose was to provide a basis for identifying programs that qualify as having either particularly high or low levels of productive efficiency which then could become the focus of the qualitative component of the study, where the goal was to determine the defining characteristics and themes that surround each type of program. Based on the literature, the driving assumption was that some portion of the margin by which the productivity of these programs exceeded or fell below more typical productivity levels would be attributable to institutional and environmental characteristics that are not directly measurable in such a quantitative model. A secondary purpose of the quantitative component of the study was to understand how each of the model's independent variables, which represent the various measurable inputs of a program's research productivity, correlated with each program's scholarly output, in order to better understand the magnitude to which the different components are associated with increases in productivity.

Quantitative Research Design

This quantitative component of this study employed a non-experimental, exploratory cross-sectional research design. The goal was to develop a regression model capable of measuring the productive efficiency of each BME program at Carnegie-classified "highest research activity" doctoral universities in the United States. This was accomplished by allowing the amount of scholarly output produced by each program to serve as the dependent variable of the model while a variety of programmatic and institutional characteristics representing inputs presumed to influence that scholarly output served as independent variables. The resulting regression equation was then used to calculate predicted levels of scholarly output given a BME program's actual inputs and constraints. Those programs with that produced the largest residual scholarly output levels – where actual output most exceeded predicted output – were regarded as having the greatest productive efficiency, while those whose actual output fell farthest below their predicted output were regarded as having the lowest productive efficiency. As a comparison of under- and over-producing programs was the goal of the qualitative component of the study, this statistical model was intended to serve as the basis for identifying suitable programs. Because the design is exploratory, rather than experimental in nature, there is no quantitative hypotheses to be tested. However, it is important to note that manner in which the different factors of scholarly production correlate with scholarly output in the quantitative model is also of particular interest and was used to inform and shape the nature of the qualitative inquiry.

Once each program's productive efficiency was established, maximum variation sampling was used to identify which institutions to include in the qualitative inquiry. This strategy was not intended to produce a sample that was representative of an entire population, but rather to produce a sample that could best support theory building. By studying programs that contrasted sharply with one another, it was hoped that more could be learned than from studying high-productive efficiency programs alone.

Qualitative Research Design

Identifying residual levels of production that exceed the amount which can be explained by the quantity of inputs is an established means for identifying production efficiency in a firm (Hulten, 2007). Yet, measuring the residual does nothing to decipher the conditions that cause productive efficiency in a firm or, in the case of this study, a university department's production of scholarship. A purely quantitative study is unable to answer the "why" or "how" types of questions (Patton, 2002; Given, 2016; Edmonds & Kennedy, 2017). Rather, the quantitative methods described in this paper were useful for empirically discerning, from a substantial amount of data, those institutions and BME programs that were making the best use of their resources as well as those who were under-producing. This winnowing down of BME programs allowed for qualitative inquiry focused on a manageable number of contrasting programs for the purpose of identifying the defining characteristics of both types of environments.

Therefore, the study includes a basic qualitative research component featuring semi-structured one-on-one interviews with BME researchers and research administrators with BME affiliations to explore these questions. As Merriam (1998, p. 11) explains, basic qualitative research “seek[s] to discover and understand a phenomenon, a process, or the perspectives and world views of the people involved.” Themes and patterns found in the analysis of the researcher interviewer transcripts were used develop conclusions about what factors differentiated low from high productive efficiency programs.

Initially the study proposed to use a more involved grounded theory design, but circumstances dictated that this approach had to be abandoned for practical reasons. Had the situation allowed, this approach would likely have produced fuller results and, as such, its applicability to this study should remain in the discussion. Both basic qualitative research and grounded theory are inductive processes that involve developing theories from data collection and analysis, rather than analyzing data for the purposes of testing pre-conceived theories (Glaser & Strauss, 1967; Merriam, 1998; Patton, 2002; Charmaz, 2004; Silverman, 2016; Edmonds & Kennedy, 2017; Hesse-Biber, 2017). Either application would have been suitable to this dissertation because the study began with no firm theoretical assumptions about the conditions that define high or low productive efficiency environments. However, grounded theory involves the additional step of developing theory during the midst of the data collection stage – not after it

– for the purpose of informing subsequent data collection (Charmaz, 2015; Edmonds & Kennedy, 2017). It is meant to be an iterative process that encourages the emergence and exploration of unanticipated themes. As is discussed further in chapter 3, however, due to difficulties encountered during interview recruitment process, which subsequently lead to challenges of timing, the data collection and analysis in this study had to be truncated. As a result, collection and analysis did not end up being an iterative process, but rather one where analysis occurred after the interview process had concluded. Had circumstances allowed for the grounded theory approach, the topics identified in these interviews would have informed subsequent interviews as key themes and issues began to emerge, allowing for, perhaps, a fuller discussion (Patton, 2002; Charmaz 2015; Givens, 2016; Silverman, 2016).

Nonetheless, the analysis centered on identifying common themes among institutions with high productivity efficiency as well as those themes that are common to institutions with low productive efficiency. The specific aim was to ask researchers in both environments to describe the quality of support services and resources provided to aid their research as well as the nature of their work place culture. By asking a consistent set of questions to both groups, the assumption was that a disparate sets of themes would emerge that would demonstrate which resources, support services, and/or organizational culture factors most distinguish a well-run research program from one that needs improvement. Because the institutions studied in the qualitative inquiry already been identified in quantitative terms as exhibiting either exceptionally strong or weak productive efficiency, it was assumed that the characteristics that distinguished them from one another are the institutional characteristics that matter most.

Pilot Interviews

Prior to beginning work in earnest on this dissertation, pilot interviews were conducted with two senior researchers in Virginia Commonwealth University's department of biomedical engineering. The interviews were conducted separately using the interview guide found in appendix C. The goal of these interviews was to help develop an understanding of the general nature of BME scholarly research and how external institutional factors might impact research productivity. Each researcher underscored that research productivity in their field was highly dependent on the quality and quantity of graduate assistants available to researchers. They both also perceived an improvement in the quality of services offered by VCU related to grant management, procurement, and other centrally-provided research supports over the years, which they each believed had a positive impact on their productivity. They both suggested that their department's culture values discovery over commercialization of research, although one researcher indicated a belief that departments that do emphasize patents and other commercialized research create pressure on researchers that negatively impacts productivity. Both pilot participants also indicated that newer NIH policies that favor interdisciplinary work between medical researchers conducting clinical trials and BME researchers has led to significant increases in NIH funding directed toward BME, and hence increase productivity.

Additionally, at the proposal defense of the dissertation, the suggestion was made that the economic concepts at the core of the quantitative methodology should be reviewed by an economist. This was to insure that the concepts – particularly that of total factor productivity – were being properly represented and applied in a reasonable fashion. Dr. Robert Lacy, who teaches economics for both the VCU School of Business and the Wilder School of Government and Public Affairs, agreed to both review the quantitative methodology section of the proposal and to meet with the researcher to discuss the dissertation's proposed economic approach. In the meeting he indicated that the concepts of the production function and TFP were both being used and described in a manner consistent with the economic theory. Particularly, he agreed that the calculation of the residual production should capture a portion of output not explained by the model, and that efficiency and technology are generally assumed to be the most common drivers of residual production.

Definition of Terms

- Cobb-Douglas Function (Black, Hashimzade, and Myles (2012): A functional form, named after its originators, that is widely used in both theoretical economics and applied economics as both a production function and a utility function. Denote aggregate output by Y, the input of capital by K, and the input of labour by L. The Cobb–Douglas production function is then given by $Y = AK^{\alpha}L^{\beta}$ where A, α, and β are positive constants. If $\alpha + \beta = 1$ this function has constant returns to scale: if K and L are each multiplied by any positive constant λ then Y will also be multiplied by λ. The Cobb–Douglas production function has also been applied at the level of the individual firm.
- Faculty Scholarly Research Productivity: The amount scholarly publications produced by a faculty researcher or research unit over a period of time.
- Institutional Effectiveness (Welsh and Metcalf, 2003): "[R]efers to initiatives that are oriented toward the measurement and realization of an institution's progress toward fulfilling its mission or the fulfilment of expectations for student learning as measured by outcomes"
- Institutional Effectiveness (Head, 2011): "Institutional effectiveness is a broad-scaled, institution-wide process consisting of specific components, including the evaluation of all academic programs, administrative units, and support services; the assessment of student learning outcomes; and data-driven support by the institutional research arm of a college or university."
- Production Function (Encyclopedia Britannica, 1998): "Production Function, in economics, equation that expresses the relationship between the quantities of productive factors (such as

labour and capital) used and the amount of product obtained. It states the amount of product that can be obtained from every combination of factors, assuming that the most efficient available methods of production are used. The production function can thus answer a variety of questions. It can, for example, measure the marginal productivity of a particular factor of production (*i.e.,* the change in output from one additional unit of that factor). It can also be used to determine the cheapest combination of productive factors that can be used to produce a given output."

- Research Efficiency: The relation of an academic unit's scholarly output to the level of inputs invested into its research process.
- Scholarly Output: While this term could denote several different forms of scholarly expression, for the purpose of the dissertation, scholarly output is the total number of journal articles produced by a BME department over a given period of time, as weighted by journal impact factor and normalized by the BME's departmental faculty count.Total Factor Productivity (Black, Hashimzade, and Myles (2012): The portion of output not explained by the quantity of inputs into production. TFP reflects how efficiently the inputs and the given technology are utilized. For example, if the production function is Cobb–Douglas with output, Y, related to capital, K, and labour, L, by $Y = AK^{\alpha} L^{\beta}$, then A is the value of TFP. The rate of growth of TFP is measured by the Solow residual, under the assumption of competition in factor markets.

CHAPTER II. LITERATURE REVIEW

Methodology of the Review

For the research productivity portion of the literature review, a cross-database search of the VCU Libraries collections was conducted for the term "faculty research productivity," using the Primo discovery tool in the spring of 2016. The search, which was limited to English language peer-reviewed materials, produced 90 results. These results were compared to the same search in the ERIC database, where it yielded 43 hits that were almost entirely duplicates, which successfully confirmed that the search results had included materials drawn from the most comprehensive education-focused database in addition to the other databases that produced the remaining 47 results. Items were omitted under the following circumstances: duplicated titles (12); results that were not peer-reviewed studies (e.g. abstracts, white papers, editorials, reviews, etc.) (7); results with bad links (4); and results that included the search term, but where faculty research productivity was not the focus of the study (2). The abstracts of remaining 65 titles where reviewed and classified by type as is revealed in the body the review. Of particular interest were studies that dealt with faculty research productivity from an institutional perspective. Additionally, a general search was conducted filtered for monographs that produced 10 results, though none dealt directly with faculty research productivity. The same search was repeated in the spring of 2017 to determine any relevant articles had been published in the previous year.

Surprisingly, the search for “faculty research productivity” produced 537 results. I made an inquiry to Thomas McNulty, head of VCU Libraries’ Enterprise Systems, to determine what might have caused this large difference. He indicated that prior to the fall of 2016 “Primo would only search the full text of articles if there was a relatively small pool or results,” but Primo had subsequently “expanded its search capabilities to be truly ‘full text’.” When the new results were ranked in order of relevance, most of the articles produced in the original search were near the top. Many of the new articles found beyond that range made only incidental use of the search terms, which were indicative of the reality that if the term was not contained in the title or the abstract, it was not likely useful. The new search did uncover some additional articles regarding individual faculty productivity in particular fields, but that topic was already adequately addressed in the literature review. Nonetheless, there were some useful article found in the expanded search, particularly one that dealt with relations between researchers and research administrators.

The following search terms were used in the Primo cross-database discovery tool to conduct the economics and production function literature search within the VCU Libraries collection. The search was limited to English language peer-reviewed materials only: “total factor productivity” (42 hits); “production function” (9,683 hits); “production function” AND “total factor productivity” (15 hits); and “Cobb-Douglass” and “Total Factor Productivity” (9). The primary purpose for reviewing the literature on production functions in general was to look for definitions of the term in order to be certain that it was described correctly in the dissertation. As such, priority was given to economic dictionaries and texts that were aimed primarily at education in economics rather than traditionally scholarly papers that were not intended to explain basic concepts already well-known by those in the field. The “Total Factor Productivity” and “Cobb-Douglass” literature review was conducted more systematically and in greater depth. Most of the search results for “Total Factor Productivity” were economic studies that used or referenced the concept without attempting to explain it to a lay audience. As such, most of the dissertation’s discussion of TFP comes from a particularly useful book chapter authored by Charles R. Hulten (2007) in *New Developments in Productivity Analysis* which is dedicated to describing the history of the concept.

The following search terms were used in the Primo cross-database discovery tool to conduct institutional effectiveness literature search within the VCU Libraries collection. The search was limited to English language peer-reviewed materials only: "institutional effectiveness" AND "higher education" (131 hits); "institutional effectiveness" AND "research productivity” (0); “institutional research” AND “research productivity” (18); and "institutional effectiveness" AND "definition" (10). The results of the "institutional effectiveness" AND "higher education" search were reviewed with the goal of establishing the definition and scope of the term “institutional effectiveness” in the context of higher education and to determine if the literature included direct examples of the concept being applied to the topic research productivity, which it did not. Likewise, while the search “institutional research” AND “research productivity” produced 18 hits, none represented studies that featured research productivity as the topic. Nonetheless, I reviewed the abstracts of the hits produced by each of the search terms to see if the methods or frameworks that the authors were using to apply to other institutional effectiveness topics could be applied or used to inform my research. Of particular note was a conceptual framework for institutional effectiveness offered by Volkwein (2010) and a study by Horn and Lee (2012) that applied the concept of the residual to graduation rates.

Finally, a primo search for "department chair" AND "research productivity" was conducted that yielded 31 hits. Despite limiting the results to peer reviewed materials, the majority of these works were in the form of primers or survival guides based on the personal experiences of department chairs for the purpose of mentoring others entering academic administration. It did, however, include a very useful monograph which dealt explicitly with departmental research productivity, *The Research Productive Department* (Bland, Weber-Main, Lund, & Finstad, 2005). Additionally, Wergin (2005) *Departments that Work: Building and Sustaining Cultures of Excellence in Academic Programs* was sought based directly on a recommendation.

Institutional Effectiveness

The goal of the literature review surrounding institutional effectiveness is to define its scope, explain what how the concept applies to the current study, and to explore the extent to which studies in this area have applied similar methodologies toward similar goals. The term institutional effectiveness has various connotations in the context of higher education, though it is most closely associated with the assessment of student learning and outcomes. The term *institutional effectiveness* began to take hold in the higher education community during the early 1980s (Head, 2011). Rogers (1997) argues that the terms was adopted as an alternative to the term "assessment," which had become contentious because it was viewed by many in higher education as the representation of an encroachment by public demands for educational accountability on the historic autonomy of colleges and universities. At that time both regional accreditors and state stakeholders began to demand greater accountability in demonstrating actual student learning outcomes rather than input measures such as faculty-student ratios or indirect measures such as retention and graduation rates (Rogers, 1997). In that sense, the meaning of institutional effectiveness was initially synonymous with student assessment and viewed as a necessary component of the accreditation process (Head, 2011).

Most authors, however, describe institutional effectiveness in broader terms, recognizing that it has assumed a variety of meanings while lacking a consistently agreed-upon definition (Bauer, 2001; Head, 2011). In studying the literature surrounding institutional effectiveness, Bauer (2001) found that while assessment and institutional effectiveness are often used interchangeably, the meaning applied to each term ranges "from ones narrowly focused on student learning to broader ones encompassing entire systems of higher education" (Head, 2011, p. 9). According to Welsh and Metcalf (2003), "institutional effectiveness increasingly refers to initiatives that are oriented toward the measurement and realization of an institution's progress toward fulfilling its mission *OR* [emphasis added] the fulfilment of expectations for student learning as measured by outcomes" (p.35). McLeod and Atwell (1992) define institutional effectiveness in part as "the condition of achieving the set of goals of an institution and being able to verify the attainment of these goals with specific data which show the degree or quality of their attainment" (p.34). O'Meara (2005) defines it simply as "the ability to meet goals and objectives" (p.485). Volkwein (2010) conceptualizes institutional effectiveness as comprised of dual purposes: "the *inspirational*, which is oriented toward internal improvement, and the *pragmatic*, which is oriented toward external accountability." Head (2011) views institutional effectiveness as an umbrella term that encompasses "a broad-scaled, institution-wide process consisting of specific components, including the evaluation of all academic programs, administrative units, and support services; the assessment of student learning outcomes; and data-driven support by the institutional research arm of a college or university" (p.10). The common theme in these expanded definitions of institutional effectiveness is that it encompasses the evaluation and outcomes measurement associated with any and all activities that a university values.

Head's inclusion of "the evaluation of all academic programs, administrative units, and support services" within his definition of institutional effectiveness can be interpreted to include research productivity, research administration and research support, though he does not specifically mention those concepts. Volkwein (2010), however, does specifically address these topics in his model for assessing institutional effectiveness. He indicates that the enormous costs that go into research necessitate accountability to demonstrate value and that "some universities use productivity and performance to assess various administrative and academic support services" (p.17). Volkwein's model includes four levels of evaluation and quality assurance: the institutional level; the department or program level; the individual faculty level; and the classroom, course and student level. At each level, regardless of the type of activity being assessed, Volkwein believes that institutional effectiveness "generally seeks answers to one or more of these five questions" (p. 15): (1) Is the institution or program meeting its goals?; (2) Is the institution or program improving?; (3) Does the institution or program meet professional standards?; (4) how does the institution or program compare to others?; (5) is the institution or program cost-effective?

Focusing on the department level, and to some degree the institutional level, questions 2, 4, and 5 are particularly relevant to this dissertation. In explaining question 2, regarding whether programs are improving, Volkwein emphasizes that all programs are at different starting points and that programmatic or administrative success is relative to this starting point, rather than in comparison to similar programs at other institutions. At the same time, he asks how a program compares to others in question 4. Here he emphasizes the importance of having an established means for benchmarking for comparative purposes. In the case of this dissertation, the means for comparison is the measurement of productive efficiency in research. The argument is made in subsequent chapters that this measure is a valid means for inter-departmental comparison. In expanding on question 5, Volkwein describes cost-effectiveness as "the harsh productivity question that compares costs with benefits, expenditures and resources with results" (p. 17). He cites both internal and external quality assurance reasons as to why this type of evaluation is necessary, including analyzing the cost of research in comparison to research productivity.

While the institutional effectiveness literature makes clear that the measurement and assessment of research productivity as well as the administrative and support components of a university research enterprise fall within the purview institutional effectiveness, no clear examples of studies or methodologies that do so were discovered. Particularly, no examples were found that included the application of a production function as a means for assessing research productivity nor the use of residuals between expected and actual output as a means of identifying research efficiency. Both techniques, however, have been applied to various student-learning concepts as well as degree completion and graduation rates. A study by Pike, Kuh, McCormick, Ethington, and Smart (2011) compared various institutional expenditures and outcomes. They found a positive correlation between expenditures focused on undergraduate education and National Survey of Student Engagement's measures for academic challenge and student-faculty engagement. Powell, Gilleland, and Pearson (2012) examined how institutional characteristics and categories of institutional expenditures from IPEDS compared to both effectiveness (retention and graduation rates) and efficiency (faculty workload data obtain from the National Study of Postsecondary Faculty). They found a positive correlation between expenditures and effectiveness, but a negative correlation between expenditures and efficiency.

The primary concern of the authors was that most measures of success demanded by and reported to stakeholders are associated with effectiveness only. The policy implication raised by the results of the study suggest that by failing to account for efficiency while focusing only on effectiveness could contribute to cost increases in higher education.

Horn and Lee (2016) tested the reliability and validity of using residuals between expected and actual graduation rates as a means for testing institutional effectiveness. While their study was primarily focused on methodological questions, the impetus for their work stemmed from the concern that "despite the widespread use of graduation rates in accountability systems, it is doubtful that the relevant dimensions of institutional effectiveness are being adequately assessed" (p.470). Their conceptual framework was based on the assumption that although institutional dimensions under the control of a college or university can influence academic performance and persistence, other factors such as the socioeconomic profile of the institution's students were also major determinants. They argued that "raw graduation rates may thus better reflect advantages and circumstances than strengths or policy and practice" (p.470). Horn and Lee's approach has remarkable similarities to this study. Rather than focusing on outcomes only – retention and graduation rates in this instance – the authors suggest that the socioeconomic profile of the students should be factored into the calculation to derive the true measure of institutional effectiveness in the rates that they are able to achieve. Institutions with students from more privileged backgrounds would be expected to achieve higher rates, while there would be different expectations from schools whose enrollment draws from less privileged socioeconomic groups. By accurately accounting for that factor in calculation of such rates, the effectiveness of these different institutions could become more validly comparable.

Organizational Climate and the Role of Department Chairs

As Volkwein (2010) indicates above, the concept of institutional effectiveness is just as applicable at the departmental or programmatic level as it is at the institutional level.

While the body of literature that examines how academic departments function is treated as a separate subject from that of institutional effectiveness, the two areas share many of the same goals in terms of assessing performance and stimulating improvement. The obvious difference is that the institutional effectiveness literature is generally written from a university-wide perspective. Wergin's *Departments that Work: Building and Sustaining Cultures of Excellence in Academic Programs* (2003) focuses squarely on the individual academic unit in a manner that is more concerned with the point of view of the department looking out toward the institution, rather than the institution peering into the department. In particular, he is interested in understanding the characteristics and climatic conditions that are necessary to build a quality department and then how to apply such an understanding toward the effective evaluation and improvement of departmental performance. To be clear, he does not use the term institutional effectiveness, but the purpose of his work places him squarely within the scope of that concept based on the definitions above. Furthermore, his work is primarily concerned with teaching quality and student learning. Nonetheless, by focusing solely on how departments function and behave in regard to achieving quality results, he offers insights that are applicable to the current study in terms of what types of qualities and characteristics distinguish a positively functioning department that is capable of achieving meaningful improvement.

A central observation in Wergin's work is that building a departmental climate that is both engaged in and capable of supporting quality improvement is a prerequisite to any attempt to actually improve quality. Successfully building such a climate is dependent on both the right kind of leadership and a collective buy-in by the department faculty. Without these qualities, which he labels the "leadership of engagement" and "the engaged department," efforts to improve departmental quality will only elicit a "compliance mentality" that is doomed to failure. By "leadership of engagement," Wergin means that the department chair must be able to communicate vision clearly and make a case for change the faculty agree with and are willing to adhere to. To achieve this, the faculty must be active participants in defining the vision for the department through negotiation with the chair and other leaders. In other words, departments most decide together what they want to achieve, how they want to achieve it, and how to assess their progress. Goals, quality measurement efforts, and frameworks that are simply imposed from the outside are unlikely to produce results.

Buller (2012) regards the chair's responsibility as acting as an intermediary between the faculty and university stakeholders (e.g. trustees, legislators, funders, etc.) in matters of productivity and quality. In his view, the chair must simultaneously understand that "a college professor is not a machine for producing student credit hours or refereed articles," while at the same time recognizing that "as administrators, we need to be accountable for our use of resources – including human resources – for the public good, and that many accreditation or program review processes require that our departments remain viable" (p. 215). Buller suggests that the chair is responsible for identifying faculty members with low-performing research productivity, and then he offers strategies for improving their performance. These strategies include clearly communicating expectations, assigning peer mentors, removing obstacles where possible, looking for possible sources of pride and motivation, and documenting the progress – or lack thereof – of the under-producing faculty member in order to form a paper trail.

In this sense, Buller seems to advocate that improvement is only for poor performing faculty and, further, that it should take place in the form a one-on-one effort between the chair and the individual faculty member – and not the department as a whole. He seems to paint a department chair as a caretaker who patches problems as they arise, rather than Wergin's view where the chair should be capable of providing authentic transformative leadership. The implication for the present dissertation is that an attempt must be made during the qualitative component of the study to assess the degree to which faculty members and chairs within a department share a collective vision of productivity goals. It may be that large positive residuals in productivity detected in the quantitative analysis could be a manifestation of the positive effects of this sort of departmental climate.

Bland, Weber-Main, Lund, and Finstad (2005) draw many of the same conclusions as Wergin, in their work *The Research Productive Department: Strategies from Departments that Excel* which focuses exclusively on research productivity. From their review of the literature surrounding research productivity the authors developed a framework that characterizes the qualities and traits of a highly productive research organization at three levels: (1) the individual researcher; (2) environmental characteristics; (3) leadership characteristics; with the latter two being of particular importance to the purposes of this dissertation. The authors then used this framework to guide interviews with approximately 40 academic leaders across several disciplines "from the most comprehensive and highly ranked research institutions in the country" (p. *xiv*). Like Wergin, they find that clear research goals must be identified and frequently communicated and that such goal-setting should take place through a process of assertive participative governance. Furthermore, they argue that this requires a leader who "internalize[s] the group's mission and keeps the research emphasis clear to the group" (p.191). Expanding on leadership qualities, they believe that for a department to be especially productive, the chair must successfully play the role of "keeper of the vision" in addition to being an effective manager of resources, a peer model, and advocate for the department. While the authors contend that all of these roles must be performed well, they suggest that the "keeper of the vision" role cannot be overstated because in addition to the importance of developing shared research goals for a department, those goals must be constantly reinforced. One of the study's participants stated that "I still am amazed at how much vision you have to provide as the chair. It just doesn't happen another way. You need to express it and you need to be personally involved" (p. 193).

Bland, Weber-Main, Lund, & Finstad (2005) also dedicate attention specifically to the resources that are needed to support research productivity. They found that local peer networks can be one of the most important resources for building a culture of high achievement, but "a strong sense of *esprit de corps* must be cultivated for it to be effective" (p. 109). The implication for this dissertation is that there should be an attempt to gauge both the extent and the quality of such activities taking place within the subject BME programs. The authors also include student quality as a consideration under the topic of resources, citing studies that suggest that one of the main predictors of faculty research productivity was the number of doctoral students advised to completion in the past five years (p. 110). Finally, the authors found that the issue of access to and quality of research space came up in almost every interview they conducted. They suggest that effective leaders emphasize the need to take space into consideration ahead of hiring, while noting that the issue of space is often more important in making the determination to hire new faculty than the issue of funding. Again, the emphasis that the authors place on this consideration suggest that it must be incorporated into the interview guide as it has the potential to be a major theme for the qualitative component of the dissertation.

In expanding the question of climate and culture beyond the department, Cole (2007) explored the perceptions that faculty researchers hold regarding their interactions and relationships with administrators in university research offices. She collected opinions from 34 faculty researchers from various research universities who had each received at least $1 million in federal grant funding to learn about their interactions with research administrators. The goal was to better understand how to improve the process of research administration. Despite the difference in topic and approach of the departmental-centered studies discussed above, her primary conclusion fits the themed of developing shared goals. Cole's concludes from her findings that "administrators and research faculty should view each other as team members whose objectives are to discover and understand how to achieve common goals" (p. 139). From the faculty perspective, this means primarily that "research administrators should provide more help and be less focused on enforcing rules" (p. 151). Faculty members were particularly interested in help identifying and securing grants as well as improvement in reducing bottlenecks around institutional review, procurement, and accounting processes related to grants. Further, faculty respondents recognized that they needed to improve their side of the relationship by submitting proposals with more adequate lead-time and by showing greater respect and understanding towards research administrators. So, as with the departmental literature, this study suggests that gauging BME faculty's relationship with the central research administration units of their university could be an important component in explaining high or low research productivity. Is the relationship supportive or authoritative? Is it marked by a collaborative spirit? Are operations marked by efficiency or disorganization?

The institutional effectiveness literature informed the current study in a number of substantial ways. First, the broader view of institutional effectiveness suggests that the concept is germane to all university endeavors and at all organizational levels. Second, that properly assessing institutional effectiveness involves measurement that should account for inputs or costs in relation to outcomes, rather than just outcomes alone, implying that efficiency is a value of effectiveness. Third, that it has been argued the calculating the residual between predicted and actual outcomes is a valid means for detecting institutional effectiveness in graduation and retention rates. The literature surrounding academic department informed the study further by emphasizing that a department's climate and how it functions is likely to impact its collective outcomes. The themes that were stressed in the literature – assertive participative governance, authentic goal-setting, transformational leadership, and general engagement – are all concepts that were incorporated into the interview guide to insure that these themes were explored qualitatively.

Overview of Faculty Research Productivity

Broadly defined, faculty productivity is concerned with the three primary duties of the professoriate – teaching, research, and service – and the extent to which these activities are successfully carried out by individual faculty members, departments or entire institutions (Massy & Wilger, 1995; Dundar & Lewis, 1998; Fairweather, 2002; Betsey, 2007; Helsi & Lee, 2011). The precise expectations of faculty members in regard to these three responsibilities vary by institution type, career stage, discipline and other factors. Of these three responsibilities, however, research productivity receives the most attention in the literature and is widely perceived to be among the most important sources of institutional reputation as well as individual rewards such as tenure, promotion and other recognition (Massy & Wildger, 1995; Dundar & Lewis, 1998; Tierney, 1999; Helsi & Lee, 2011). Therefore, as Helsi and Lee suggest, "the justification for studying faculty research productivity is that it affects individual advancement and reputation within academe, as well as departmental and institutional prestige" (p. 393).

Just as it is recognized that individual faculty members and institutions both accrue benefits through generating scholarly output, the literature surrounding faculty research productivity can generally be divided between those studies that approach the question from the context of the individual researcher and those that explore it from an institutional perspective (Abouchedid & Abdelnour, 2015). Both strains of inquiry seek to identify the factors that are associated with higher or lower levels of research productivity, at their respective units of analysis, in order to explain observed differences in productivity between like entities (Helsi & Lee, 2011). This review examines both strains of literature in an effort to demonstrate the extent to which they informed the study. It also explores concepts of productivity from the field of economics that were borrowed in support of the study's objectives.

Implications of Individual Faculty Research Productivity Literature

Because the factors that predict or influence the research productivity of individual faculty members is not the focus of this study, it could be suggested that this strain of the literature provides limited applicability. As the vast preponderance of research productivity studies are focused on the individual researcher, however, the literature in this area does offer certain useful elements in terms of research design, methodology, or concepts that can be applied to the study of research productivity at the programmatic and institutional levels. Such elements include the application of economic concepts to describing and understanding the scholarly research environment, insight into different means of operationalizing the concept of research productivity, and the introduction of institutional characteristics to serve as control variables.

In terms of applying economic concepts, several studies borrow their definition of productivity from economics, which generally recognizes the concept cannot be described by output alone, but rather should be conceived of as a ratio of output to inputs (Levin & Stephan, 1991; Bieber & Blackburn, 1993; Masey & Wilger, 1995; Black, 2009). To some degree, this is reflected in opportunity cost studies that account for time that is dedicated to non-research responsibilities, because such studies are not concerned with raw output, but rather output per unit of time dedicated to research (Jordon, 1989; Carstensen, 1992; Dundar & Lewis, 1998; Bellas & Toutkousian, 1999; Hu & Gill, 2000; Eagen & Garvey, 2015). Other studies go further by adopting the concept of the production function from economics as a quantitative framework for analyzing research productivity. In its most basic form, the production function is a comparison of production output with the production inputs of labor and capital (Hulten, 2007). In this context, Levin & Stephan (1991) combine measures of time spent on research (labor), a variety of variables representing the quality of the researcher (capital) to compare to research output as a true measure of productivity.

Another common theme through the literature is the use of journal articles as the primary means for operationalizing the concept of research productivity. Several studies rely solely on the number of journal articles published by individual faculty members as their measure of research productivity (Bieber & Blackburn, 1993; Steurly & Maranto, 1994; Tien & Blackburn, 1996; Hu & Gill, 2000; Azad & Seyyed, 2007; Long et al., 2009; Hesli & Lee, 2011). A lesser number of studies feature journal articles as one measure in a composite among other forms of scholarly output such as books, book chapters, conference proceedings, etc. (Blackburn & Bentley, 1993; Betsey, 2007; Tien, 2007; Kwiek, 2016). One justification for using journal articles as the sole measure of research productivity is that there are established methods for assigning values to the quality of both articles (citation counts) and the journal in which they are published (impact factor) that are not available for other forms of scholarly publishing. Several studies that relied on journal counts also incorporated some measure of quality into their design (Levin & Stephan, 1991; Steurly & Maranto, 1994; Long et al., 2009).

Finally, the individual faculty productivity literature can help inform the qualitative component of this study by providing an awareness of the main non-institutional factors that influence research productivity. While the goal will be to understand how institutional characteristics and policies affect departmental research productivity in BME, it must be recognized that departmental output is the sum of individual output, the determinants of which may be more influential than institutional dimensions. It should also be noted that some of the considerations raised in the individual productivity literature such as teaching loads, working environment, and motivational factors, can be viewed as institutional factors that are likely to impact departmental productivity as much or more than individual productivity.

Institutional Determinants of Faculty Research Productivity

There are far fewer studies that explore the concept of research productivity at the institutional level where the unit of analysis is that of a program, department, or university rather than an individual researcher. Of such studies, most are concerned only with identifying which departments or universities produce the most scholarship, rather than identifying the organizational determinants that may be behind such productivity. The goal of such studies is to simply establish a ranking of programs for reputational purposes, rather than an understanding of determinants. Examples include an attempt by Grover, Seagars, and Simon (1992) to identify the top 50 management information systems programs from 1982 to 1991. More recently, Barrow, Settlage, and German, (2008) used similar methods to identify and rank the top 30 science education programs in the United States for the decade of the 1990s.

As such, the primary focus of these works is to determine the best method for quantifying scholarly outputs. This centers on questions such as which types of scholarship should be included, how to weight different forms of scholarship, how to count co-authored articles, etc. In a typical example, Barrow, Settlage, and German (2008) analyzed the publication records of eight leading journals in science education to determine which programs had contributed the most to that field in the 1990s. Their analysis included both a raw count of articles attributable to each program as well as a count that was weighted by order of authorship. However, no effort was made to normalize the data on an article per faculty basis, meaning that department size was likely to influence the rankings. This approach is nearly identical to that used by Grover, Seagars, and Simon (1992) in ranking management information systems programs between 1982 and 1991, with the exception that the authors weighted the MIS journals by reputation. Similar studies have focused on quantitative psychology from 1976 to 1982 (Maxwell & Howard, 1986); management from 1980 to 1985 (Stahl, Leap, and Wei, 1988); production operation management from 1980 to 1994 (Malhotra & Kher, 1995) and psychology for the period of 1986 to 2008 (Mahoney et al., 2010).

While these studies provide some useful insight into methods of aggregating and attributing values to research output at the departmental level, they are unconcerned with the institutional factors that may be driving the variations in output they observe. As Levin and Stephan (1991) note, however, “research output is not affected only by the attributes of scientists but also by attributes of the employing institutions” (p.118). There are a small number of studies that approach the question of research productivity from an organizational perspective that are interested in more than just rankings programs within an academic discipline or subfield.

Toutkoushian and Porter (2005) explore the relationship between student quality and research output as the driving factors of institutional prestige by collecting a wealth of data on 203 universities and 143 liberal arts colleges in the United States. They begin with the assumption that institutions of higher learning seek to maximize reputation in a fashion similar to that of private firms that seek to maximize profits. The authors argue that reputation is a function of the per capita research output of the faculty, average quality of the students (SAT scores), as well as some influential institutional characteristics (e.g. age of institution, public or private status, geographic location and size). This relationship is represented as:

$$Rep = (R_F, Q_S, IC_{Rep})$$

where Rep is institutional reputation, R_F is the per capita research output, Q_S is student quality, and IC_{Rep} is institutional characteristics presumed to influence either average faculty or student quality (p. 607). While the primary goal of the study is not to understand how research productivity is influenced by institutional factors, a component of their research design does share that objective with this study. In order to build their model to the point where they could test that research productivity, student quality and institutional characteristics are determinants of reputation, they first had to quantify each of these three proposed factors of reputation separately. This lead them to develop a measurement of per capita research productivity that is conceptually represented as:

$$R_F = (Q_F, Q_S, IC_F)$$

where R_F is per capita research output, Q_F is the average quality of the faculty researchers, Q_S is the average quality of students, and IC_F are institutional characteristics that could affect the per-capita research output of an institution. The authors measured “the average human capital of the faculty as measured by age, educational attainment, and gender composition of the faculty…” (p.607). The authors offer no explanation for their inclusion of student quality as a predictor of research output beyond positing that it is. Likewise, there are no other examples in the literature where gender is considered as a surrogate measure for human capital. The only other statement the authors make about gender in the article is in their literature review, where they state that “differences in research output by gender is also a common finding, even after controlling for field and other differences between men and women (see, for example, Creamer 1999)” (p. 609). So it is odd that Toutkoushian and Porter would so haphazardly suggest that human capital diminishes as the proportion of women in the composition of the faculty increases, despite the fact that there is disagreement in the literature over the relationship between gender and productivity and that several studies offer explanations as to the factors that might influence the gender disparities where it has been observed (Sax et al., 2002; Hesli & Lee, 2011; Aiston & Jung 2015). As baffling and causally prejudiced as their logic appears, their work is important to this dissertation in the sense that this represents one of two studies that fully applies the concept of the production function to research productivity. Their approach is based on the premise that institutional-level research productivity can be predicted by a combination of labor and capital. Labor is represented by their choice to use per-capita research productivity (as opposed a total) “measured as the log of the ratio of institutional publications to the number of full time faculty” (p.608). By dividing the number of faculty into the number of publications, the authors effectively moved the labor unit variable (number of faculty) from the independent variable side of the equation to the dependent variable side. Capital is represented by their measures of the average quality of faculty members and institutional characteristics that are presumed to influence per-capita productivity. They found that their model explained over 70% of the variation of per capita research productivity from one university to the next, but only 41% of the variation at liberal arts colleges. Regarding research productivity in their overall findings, the authors concluded “that institutions with better reputations have on average more publications per faculty member; a one-point increase in reputation score is associated with an 80% increase in publications per faculty member” (p. 613). While this study did apply the concept of the production toward analyzing a collective institutional outcome, this study differs from this dissertation in that it was only concerned with how institutional characteristics impacted reputational status. There was not effort to learn how institutional effectiveness impacted research productivity.

Closer to the goals of the proposed study, Dundar and Lewis (1998) focus specifically on “the effect of program and organizational factors as powerful attributes for enhancing productivity...in four broad fields of the biological sciences, engineering, the physical sciences and mathematics, and the social and behavioral sciences” (p.606) drawn from Carnegie-classified Research I universities in the United States. In their effort to “uncover those policy and institutional factors that not only associate with but also facilitate enhanced productivity” (p.607) the authors developed a model that featured the ratio of publications to faculty members over a four year span as the dependent variable which was then regressed against a variety of independent variables representing different institutional characteristics. These independent variables included program size, percentage of faculty with a publication, percentage of faculty members with grant funding, library expenditures, ratio of graduate student to faculty, percentage of graduate students holding assistantships, and a dummy variable for institutional control as either a public or private university.

The study included 1,834 programs in total, including 380 in engineering. Research output was measured by the journal articles attributed to each institution in the *Science Citation Index* and the *Social Science Citation Index*, from 1988 through 1992. The authors found that larger programs, as measured by the number of faculty, had higher per capita output, though the effect diminishes as size increases. The authors found supporting arguments in the literature for this result that suggested a combination of factors such as better facilitation of collaboration in terms of cooperation and joint research projects, increased likelihood of attracting higher-quality researchers, and greater accumulation of resources in support of research. The authors suggested that the policy implication for their department size findings is that an increase in faculty size by itself can increase individual productivity across the board. Not surprisingly, the results also showed that extramural grant funding was highly related to increased productivity in each of the fields studied. In terms of institutional support of critical resources, the only measure available to the authors was library expenditures. They found departmental research productivity to be significantly related to library expenditures except in social sciences and engineering. The authors speculated that increased library expenditures could potentially suggest “that institutions with more resources provide better resources in other infrastructure ways…[that] should contribute to increasing their research productivity” (p.624). Finally, when the four disciplines were pooled together, the results indicated that both the number of graduate students and increased graduate student to faculty ratios were each significantly and positively related to departmental research productivity. When these variables were analyzed by discipline, however, these results held only for the physical sciences and engineering. The authors argue that this indicates that “fields such as engineering and physical sciences do in fact use graduate students more effectively in both their teaching and research activities” (p.625). The policy implication, beyond that of enrollment, is that graduate assistants are an important factor of production of departmental scholarly output in these fields.

Implications of Institutional Determinants of Faculty Research Productivity

As was the case with the individual determinants of research productivity, the brief literature surrounding institutional determinants has conceptual and methodological implications for the study proposed here. Toutkoushian and Porter (2005) and Dundar and Lewis (1998) both employ models that are deliberately borrowed from economics and based on production theory where scholarly output is regressed against the factors of scholarly production. Both studies recognize that an accurate measure of departmental productivity relies on a per capita measure of scholarly output rather than total output in order to satisfy the economic concept of productivity. Both studies also acknowledge institutional characteristics as important determinants of productivity. Only Dundar and Lewis, however, employ such characteristics in an attempt to understand how institutional conditions can drive productivity.

What these studies do not attempt, however, is to apply their models to run predicted values for the amount of scholarly output that an institution should be capable of producing given their inputs. This can be accomplished by using the coefficients produced by regressing actual scholarly output against actual institutional inputs – as both studies did – and then computing a predicted output value based on the actual institutional inputs. Comparing the predicted output to the actual output for each department provides a basis for determining which institutions are achieving the highest levels of faculty research productivity given their resources. Although Dundar and Lewis's model demonstrates the correlations and parameter effects that different institutional variables have on research productivity, it does not indicate how effectively each institution makes use of these inputs – understanding, of course, this was not their objective. While such an approach is not represented in the faculty productivity literature, the institutional effectiveness literature review did find a single study that took this approach toward analyzing institutional effectiveness in retention and graduation rates. To determine how such concepts can be applied towards measuring institutional effectiveness in research productivity, it is necessary to turn the economics literature for a fuller explanation of the concept of the production function as well as total factor productivity and its focus on how residual productivity can serve to measure effectiveness and efficiency.

Economic Concepts of Productivity

In production theory, the various factors of production are combined into a production function that expresses the relationship between the quantity of inputs and the quantity of outputs produced (Robinson, 1955; Hulten, 2001; Fioretti, 2007). Capital and labor are the most primary components of the production function (Hulten, 2007). At the macroeconomic level, for example, it is the national aggregate of these two factors that make up the producer's side of the gross domestic product (Betancourt, 1986; Hulten, 2001). At the level of the firm (or that of a university department) Fioretti (2007) uses the first expression below to generalize the production function, where *y* represents output and *K* and *L* represent the categorization of all factors of production as either capital or labor (p.708). In the second expression he shows the production function where the factors of production are enumerated as individual inputs, as represented by *X*.

$$y = f(K, L)$$

$$y = f(X_1, X_2, X_3, \dots, X_n)$$

The Cobb-Douglas production function model is ubiquitous in production theory literature as a prevalent approach to estimating output based on the factors of production (Robinson, 1955; Basu & Fernald, 2007; Hulten, 2007; Biddle, 2011; Rasmussen, 2013). First introduced in the 1920s, Biddle (2011) suggests that the researchers' goal was to develop an "understanding of the relationship between the level of output and the quantities of inputs employed in production…[by] the empirical procedure of regressing a measure of outputs on a measure of inputs [to evaluate] what it could reveal about the parameters of firm-specific production function" (p.235). In other words, Cobb and Douglas assembled historic production data – both actual inputs and actual outputs – for various industries into a regression model to determine the relationship between a set of inputs and the quantity produced that, in turn, could be used to analyze the output of a specific firm in the context of a specific industry at a specific time. This exercise can indicate how efficient a firm is at converting its inputs into output, but it cannot identify what is driving efficiency. According to Hulten (2007), identifying the change agents that cause observed increases in production efficiency is an area of sharp disagreement among economists, but it is also central to the discussion of how university policies and practices could be adding to an institution's scholarly capacity.

Total Factor Productivity and Residual Production

Different schools of economic thought hold different views as to what constitutes the primary drivers of increased efficiency and how those increases translate into productivity growth over time. Hulten (2007) describes this divergence of thought as a dichotomy between those explanations that are technology-based and those that attribute increases in efficiency to the effects of capital formation, or the process of adding to or developing the productive capacity of existing capital used in production (Black et al., 2012). Technology-based explanations of efficiency center on how breakthroughs in technical innovation or know-how might improve production, which includes the organization of production –the managerial governance of how the production process is facilitated (Hulten, 2007). Capital formation, on the other hand, could lead to increased efficiency by developing the skills of the labor force without increasing the actual level of labor input. Regardless of whatever effects capital formation and/or technology improvements may respectively have on productive efficiency, a confounding issue in economic discourse is that neither's contribution can be directly measured (Hulten, 2007). Instead, they produce a residual effect (along with anything else that might be influencing efficiency) whereby output levels exceed their expected amounts based on the level of capital and labor inputs. Hulten (2007) summarizes the concept of this residual further:

> "…the residual captures change in the amount of output that can be produced by a given quantity of inputs. Intuitively, ***it measures a shift in the production function*** [caused by] technical innovations, organizational and institutional changes, shifts in societal attitudes, fluctuations in demand, changes in factor shares, omitted variables, and measurement errors" (p. 40).

This residual is at the heart of a concept known as total factor productivity (TFP). According to Black (2012), "TFP reflects how efficiently the inputs and the given technology are utilized…[by identifying] that portion of the output not explained by the quantity of the inputs into production." Of course, as Hulten points out above, that residual may also include components of measurement error or other influences not related to efficiency. Nevertheless, the effect that superior organization, management, and incorporation of technology have on production efficiency are also contained within this residual, even though the parameters of their effects cannot be precisely defined.

Identifying this production residual is what makes it conceivable to explore the effect that institutional characteristics, policies and practices can have on scholarly production. The residual can be identified by using the Cobb-Douglas production function to assign a predicted value of the scholarly output to each BME program and then comparing the predicted output to actual output. Those programs whose actual output most greatly exceeds their predicted output would be most worthy of qualitative investigation. This is based on the TFP assumption that some part of the observed production residual is likely attributable to production efficiency, which in turn is attributable to the operation of a research enterprise that is better organized, coordinated, resourced, and able to maintain a thriving research and work environment than that of the average university. On the other hand, those programs exhibiting a negative residual, or lower than anticipated productivity, could reasonably be expected to contain inefficiencies or obstacles in their research operations that inhibit them from reaching their potential, making those programs of equal interest for qualitative investigation.

Conclusion

Although faculty research productivity, production theory, and institutional effectiveness might be somewhat disparate topics, each element of this dissertation can be found in the literature of these subject areas. A handful of authors have applied the concept of the production function to faculty research productivity. While perhaps recognizing that the generation of scholarship is not an assembly line operation, they also recognize that it is an activity comprised of inputs and intended outputs, and that the latter needs to be measured against the former to gain an adequate understanding of research productivity as a phenomenon. And while some of these authors have gone so far as to regress scholarly output against inputs in the style of the Cobb-Douglass production function, none appear to have applied the concept of total factor productivity and its residual as a means of identifying research efficiency. The institutional effectiveness literature, however, offers examples of the residual concept as applied to graduation and retention rates, but that body of literature generally does not offer much regarding institutional effectiveness as applied to research productivity. This lack of emphasis on research productivity could be caused by a combination of factors such as greater public and stakeholder demand for demonstrating student learning outcomes, research productivity not falling within the purview of most institutional research offices, or the presumption that the responsibility for research productivity is a concern of schools, departments and, perhaps mostly, of individual researchers. Yet, most of the institutional effectiveness literature defines that concept as applying to any major activity in which a college or university is engaged – including administrative and support units. Finally, the literature related to academic departments and the role of the chair identifies several important environmental and culture considerations that can inform the qualitative component of the study in terms of conditions and themes to be on the lookout for during the interview process.

CHAPTER III. METHODOLGY

This study employed a sequential mixed methods approach to understand how institutional factors influence faculty research productivity in PhD-granting biomedical engineering programs at Carnegie-classified “R1 Doctoral Universities: Highest Research Activity” in the United States. The quantitative portion of the study was designed to create a statistical model capable of ranking BME programs by their productive efficiency, which is the extent to which a program’s actual research output exceeds or lags behind the model’s predicted research output based on a program’s inputs into the research process. The purpose was to provide a basis for identifying programs that qualify as having either particularly high or low levels of productive efficiency in producing scholarly output, which then became the focus of the qualitative component of the study. The purpose of the qualitative component of the study was to identify the defining characteristics and themes that surround each type of program. Based on the literature, the driving assumption was that some portion of the margin by which the productivity of these programs exceeded or fell below more typical productivity levels would be attributable to institutional and environmental characteristics that are not directly measurable in such a quantitative model. A secondary purpose of the quantitative portion of the study was to understand how each of the model’s independent variables, which represent the various measurable inputs of a program’s research productivity, each correlated with each program’s scholarly output, in order to better understand magnitude to which the different components are associated with increases in productivity.

Quantitative Research Design

The quantitative section of this study employed a cross-sectional, non-experimental research design in order to explore the relationship between the dependent variable – the scholarly output of BME programs – and the independent variables – the factors of scholarly production that presumably support or influence this output. This was accomplished using existing secondary data from a combination of the Web of Science citation index (WoS), the American Society for Engineering Education (ASEE), the National Science Foundation (NSF), and the Association of College and Research Libraries Trends and Statistics Survey (ACRL).

Quantitative Research Question

R1: What factors of faculty research productivity are most strongly correlated with scholarly output as measured by weighted article count per tenured or tenure track faculty at PhD-granting biomedical engineering programs in Carnegie-classified "highest research activity" doctoral universities?

R2: Which Ph.D.-granting biomedical engineering programs in Carnegie-classified "highest research activity" doctoral universities exhibit the highest and the lowest levels of productive efficiency in creating scholarly output?

Research Population

The quantitative component encompasses each PhD-granting BME program in the United States that resides within universities rated as R1 according the 2015 Carnegie Classification of Institutions of Higher Education. This decision was based on the assumption that non-doctoral granting programs are unlikely to have the same level of research activity or share a research mission, therefore, would not provide a suitable basis for comparison. These programs were identified by downloading the publicly available 2015 excel data file from the Carnegie Classification website and then filtering the data to include "highest research activity" institutions only. The resulting list of institutions was then copied into a spreadsheet containing a list of BME programs that reported awarding doctoral degrees to ASEE. Both lists included institutional IPEDS numbers which served as basis for identifying institutions that were on both lists. This process identified 79 matches that were both "highest research activity" universities and had PhD-granting BME programs to serve as the research population.

Because the population was relatively small and the necessary secondary data was presumably available for each of the programs, there was no reason to draw a sample from the overall population. Once data gathering began, however, it became clear that some programs lacked the necessary data, which lead to some exclusions from the study. Of the 79 programs, three institutions where missing multiple pieces of ASEE data that were crucial for modelling purposes. An additional three institutions did not report sufficient enrollment data for the time period covered by the study. These institutions were excluded because enrollment is considered a key conceptual component. Seven other institutions were excluded for not reporting faculty counts to ASEE for two or more years of the three-year period, which made it impossible to calculate a per-faculty article count used as the dependent variable. Finally, another four programs were excluded because the WoS search for their dependent variable data produced questionable results that were extremely low, suggesting that the search strategy was not be working correctly for those institutions, perhaps because of how the authors' addresses were listed and stored in the Web of Science citation indexes. This left a total of 62 program with complete data to be included in the study.

Dependent Variable

The concept of research productivity was operationalized as the number of journal articles produced by each BME program during the period of 2014 to 2016, with each article being inflated by the impact factor of the journal in which it was published. Serving as the model's dependent variable, this measure will be hereafter referred to as the "weighted article count." The choice to use journal articles to represent research productivity was made because it is the most frequent form of scholarly expression in BME. In fact, this is true of many disciplines, which makes this choice consistent with the majority of research productivity studies in the literature, as noted in Chapter 2. Furthermore, other forms of scholarly output such as monographs or book chapters are less common in engineering than in disciplines such as the humanities or social sciences. Additionally, although the ability to successfully secure grant funding is typically considered an important component of research productivity, the conceptual framework of this dissertation views grant funding as an input into the research process – albeit a vital input – rather than a research outcome. As such, grant funding was included as one of the model's independent variables, where it was treated as an influencer of research productivity rather than an outcome of productivity. The choice of the publication time frame of 2014 through 2016 was intended to ensure that the study focuses only on recent publications that will represent the immediate past of the programs being studied. To reach further back in time would increase the risk that the productivity being measured quantitatively may no longer accurately reflect the current productivity of the programs that will be studied qualitatively in the present. The choice to employ a span of years rather than a single year's journal count was based on the assumption that a program may experience productivity peaks or valleys in a given year – more so than over the course of a three-year period.

Journal Impact Factor

The article-weighting was accomplished by using the five year journal impact factor (IF) of the journal of publication for each article to serve as a proxy measure for the quality of the scholarship being produced. A journal's five year IF is calculated by dividing the total articles published in a journal into the number of total citations those articles received over the five-year period that directly proceeds the year in question (Garfield, 2005). For example, an IF of 2.5 calculated in 2016 would indicate that articles published in the journal in question from 2011 to 2015 received an average of 2.5 citations each. In this example, articles published in that journal in 2016 would be multiplied by 2.5, thus valuing that article two and one-half times greater than an article published in a journal with an IF of 1.0, for the purposes of this study. Conversely, an article published in a lower profile journal with an IF of .5 would be devalued to count only half as much as the article published in a journal with an IF of 1.0.

If a program's researchers are continuously publishing in the top journals, then their work presumably has a higher profile and reaches more professionals in their field. Reaching a wider audience makes their research output potentially more influential than the same number of articles published in lesser known journals. An alternative approach would be to use each article's citation count as a more direct measure of quality. While this approach has appeal, its drawback is that it could inflate or deflate a department's journal count depending on how their publications are distributed across the three-year period being reviewed. A department that was more productive in 2016 might appear less productive than a department that had the exact same output in 2014 because those articles would have more time to collect citations. When the measure is tied to the journal instead of the article, however, articles published in 2015 would be treated on a more equal footing with those published in 2014. Another potential drawback to using citations over impact factor is the possibility that an article could be heavily cited for negative reasons. If an article becomes discredited or otherwise notorious, its citation count could shoot up dramatically, but that count would certainly not be an indication of positive productivity.

Each program's IF-weighted journal count was normalized by dividing it by the number of tenure and tenure-track faculty members the department employs to derive a per capita count of journal articles per program. This was to provide a common scale in the dependent variable data that will make programs of differing sizes more comparable to one another. The faculty count data was provided by the American Society for Engineering Education (ASEE) from its annual survey of schools of engineering in the United States. The actual figure that used was the average tenure and tenure-track faculty count over the three year period. The end result will be a dependent variable value for each BME department, expressed as follows:

$$DV = \frac{\Sigma(article * IF)}{Average\ faculty\ count}$$

Dependent variable data collection and assembly

The original design for dependent variable data collection was to use the WoS online citation index to conduct an article search in for each BME program. The search would have to relied on filters to restrict the results by institution, date range, peer-reviewed articles only, and the WoS-assigned subject category of "engineering, biomedical." The process was followed for each program in the research population. Once the data was exported and explored, however, it became evident that this strategy had not produced usable results. Using a sampling of the journal article records extracted from WoS, a cross reference was made between author's names and the online faculty directory of the departments for which they presumably worked in order to verify accuracy. In most cases, no match was found between the article's author and the supposed department's faculty listings. To research the problem, a review was made of faculty publication records that were linked to some of the online departmental faculty directories to determine the types of journals in which these researchers actually published. Additionally, web searches were made to identify some of the authors in the journal articles records that had been extracted from WoS, but who were not members of their institution's BME department, in order to learn where these faculty members were actually employed. The conclusion was that BME faculty members tend to publish in a wide variety of medical and science journals and that medical and clinical researchers often published in the journals that WoS categorized as "engineering, biomedical." As such, the original data gathering approach proved to be a victim of the interdisciplinary nature of the BME field.

Julie Arendt, the Science and Engineering Research Librarian for Virginia Commonwealth University Libraries, was consulted to help determine a viable path forward. She noted that all author records in WoS include an address featuring a department name line that is recorded using a standardized abbreviation. The standardized abbreviation for BME departments is "Dept Biomed Engn." Ms. Arendt's discovery lead to a new search query based on filtering by each institution individually and limiting the results to those articles that included "Dept Biomed Engn" in the author's address. As with the original strategy, the search was also limited to the time period of 2014 to 2016 and to peer reviewed journals only. The search produced results that were highly accurate when the author's names were cross referenced against departmental faculty directories. The newly devised search was conducted for each of the 79 BME programs (dependent variable data gathering took place before it was clear that some programs would have to be omitted to do lack of data). The complete article records were exported from WoS as text files, which were then individually imported into a single MS Excel file – each article as an individual record occupying a single row. The name of each institution and its US Department of Education's IPEDS identification number were added to each article record to facilitate aggregation. This data was then used to calculated an unweighted article count for each program.

Web of Science's InCites Journal Citation Reports was then used to assign a journal impact factor rating to each article in the dependent variable dataset based on the journal title in which the article was published and the year of publication. A report listing every journal title in the science citation index and each title's total cites, journal impact factor, and eigenfactor for each year from 2014 through 2016 was exported into an excel spreadsheet. Meanwhile, by applying a pivot table to the dependent variable dataset, a complete list of all the journal titles present in the dataset was also produced for each year. These separate sets of lists were then copied and combined by year in a single excel file where conditional formatting was applied to cause any title present in both the dependent variable dataset and in the WoS InCite export lists to be highlighted with green fill. The journal titles from WoS InCite that did not have a match in the dependent variable dataset were then deleted. This left three tables of matching journal titles with annual journal impact factors that could then serve as a value lookup tables to apply the correct impact factor weighting to each of the 19,971 rows of articles contained in the dependent variable dataset. Because each row represents a single article, the JIF figure itself represents the impact-factor weighting without the need for any additional calculation. Once the process was complete, a pivot table was used to aggregate the weighted article count for each program. This figure was then divided by each program's tenure and tenure-track faculty count to derive the final dependent variable for each program.

Independent Variables

Following the concept of the production function, the independent variables are comprised of measures that can be designated as either factors of, or constraints to, the production of scholarly output. Each of these variables are individually described below including justifications, data sources, data collection and assembly, and discussion of potential limitations.

Research Funding

Research expenditure data were collected for each program from the years 2013 through 2015. The totals for 2013 and 2014 were inflated into 2015 dollars based on the consumer price index in order to standardize each year's expenditures. Each year's expenditures were then summed into a single total that will serve as the research expenditures variable. The data was gathered from the Higher Education Research and Development Survey (HERD) which is conducted annually by the National Science Foundation (NSF) to collect research expenditure data annually from all US universities and colleges. It is important to note that the HERD survey collects research expenditures from all funding sources (federal, state, local, business, institutional, and other), not just funding provided by the federal government or the NSF itself. The survey results reported by NSF contain engineering research expenditures by institution and subfield in table 52, which includes a "bioengineering and biomedical engineering" subfield category that provided the data for this variable. A potential limitation is that these figures are self-reported and programs have an incentive to inflate their data reporting because research expenditures figure prominently into the ranking of engineering programs.

The expectation is that this variable will produce a highly significant, positive correlation to the weighted article count variable, stronger than that of other variables.

Student Enrollment

Student enrollment is not an input to the production of scholarly research, but it belongs in the model because it represents a significant constraint to that productivity. A consistent theme in the literature is that the teaching load faced by a department's faculty is the responsibility most likely to compete for time with their research responsibilities. The greater the teaching load, the more difficult it is for faculty to balance that demand against the time required to effectively conduct research. As such, student enrollment data plays an essential role by approximating each department's teaching load. Engineering schools report enrollment by program and degree level to ASEE annually. Student enrollment was operationalized as the average enrollment for each BME program department over the three-year period of 2013 to 2015.

This variable is expected have a statistically significant negative correlation to the weighted article count variable.

Research and Teaching Support

Student assistants and non-tenure track faculty should presumably increase departmental research productivity, either by directly assisting in the research process itself or by assuming teaching responsibilities that would enable the tenure track faculty to dedicate more time to research. ASEE collects fulltime equivalency data for both non-tenure track teaching and research faculty (counted separately) as well as the number of teaching assistantships, research assistantships and post-doctoral fellowship appointments awarded by each engineering school. Unfortunately, the data is not available at the department or program level. Therefore, both to help remedy this problem and to normalize the data, the value for each category of assistantship and non-tenure track faculty was divided by the total number of tenure and tenure track faculty for the entire school. This resulted in five ratio variables measuring the proportion of tenure track faculty to research assistantships, teaching assistantships, post-doctoral fellowships, non-tenure track research FTE and non-tenure teaching FTE. The obvious limitation is that these are school-wide measures a degree variance may exist from one department to the next may exist.

It was anticipated that these variables would to be prove positively correlated with the weighted article count variable, as each of these variables represents the introduction of additional labor into the research or teaching process. The assumption that the relationships would prove to be statistically significant was much less certain, however, in than it was in respect to research funding or student enrollment levels. Among these variables, expectations were highest for the graduate research assistant variable, as the pilot interviews strongly suggested that these students play an integral role in both conducting research and producing scholarship. Likewise, Dundar and Lewis (1998), in the study that most resembles this one, found that the graduate assistants were more important to the field of engineering than to other broad disciplines.

Program characteristics and demographics

As seen in Chapter 2, the literature is replete with studies that explore whether various program demographics or characteristics are related to research productivity. Of these variables, academic rank fits into the production function concept because the level of professional experience that it represents can serve as measure of human capital. Other variables, such as the gender or ethnic makeup of the program’s faculty fits neither the concept of the production function or the institutional effectiveness perspective underpinnings of the study. In the event that gender or ethnicity would have shown a significant relationship with scholarly output, for both legal and ethical reasons, it would not have been as though that information could or should influence hiring decisions or retention efforts. On the other hand, it was conceivable that the ability (or inability) of a particular work environment to facilitate a diverse workforce could impact the collective productivity of its employees. Therefore, these variables were included because, first, questions of how ethnicity and gender intersect with research productivity are generally common in the broader literature of research productivity and, second, because the question of how workplace climate might influence productivity is important to this study.

All three demographic variables – rank, ethnicity, and gender – were available at the program level from ASEE. Academic rank was operationalized as the percentage of assistant professors to the overall number of tenured or tenure-track program faculty, where both numbers were the average over the three-year period of 2013 to 2015. The expectation was that junior faculty members who have not yet achieved tenure and are just beginning to chart their careers are less likely to be as productive as the tenured faculty members who have had the opportunity to establish more secure sources of funding and to gain more experience in running laboratories and classroom teaching. The expectation was that this variable would be negatively correlated with scholarly output – that is, as the percentage of assistant professors to the whole increases, the weighted article count for the program would decrease. Ethnicity was operationalized at the percentage of the faculty members for each program that are non-Caucasian. Gender was operationalized similarly as the percentage of female faculty members to the overall number of tenure and tenure-track faculty in each program. In the case of both gender and ethnicity, there was no expectation as to how this variable will relate to scholarly output, if at all, as the findings of studies that have explored these issues have been inconsistent and have largely applied to other fields, as demonstrated in chapter 2. In the event that either or both of these variables had produced a negatively correlation with scholarly output, additional questions would have been added to the interview guide in an attempt to identify if workplace conditions may be impacting the success of non-Caucasian or female researchers. As chapter 4 addresses, however, problems were encountered with each of these variables which ultimately excluded their incorporation into the final model.

Additional variables related to program characteristics included whether or not each program's institution has a medical school, the proportion of graduate students to overall enrollment, and the percentage of program research expenditures made up of its school's overall engineering research expenditures. An indicator variable was used for each university based on whether or not that institution includes a school of medicine and/or a hospital. The pilot interviews conducted with VCU biomedical engineering faculty indicated that the presence of a medical school in the same institution as a BME program offers several advantages. This includes a greater potential for establishing interdisciplinary partnerships with clinical researchers and, as suggested in the pilot interview suggested, large funding agencies tend to favor translational research. Finally, the significant research funding that medical schools attract to an institution could create a spillover effect for other disciplines, including the quality of the university's core research facilities or overall support for the office of research.

The percentage that graduate students represent to overall enrollment was introduced based on the assumption that proportionately larger undergraduate populations may represent an increasing teaching responsibilities, while proportionally larger graduate populations may represent an increase in research assistance. The assumption was that percentage of graduate students would be positively correlated with research productivity. ASEE provided enrollment data by program and degree level necessary to calculate this percentage. The variable was operationalized determining the average enrollment for each program from 2013 through 2015 and dividing that number into the average number of masters and doctoral students during the same time period.

Finally, the proportion of BME program research expenditures to its host school's overall research expenditures was intended to indicate whether the proportionality of a program's research enterprise to its school influences productivity. The rationale behind this variable was to introduce some measure of department size. While the dependent variable was standardized on a per-faculty basis for the very purpose of neutralizing the potential imbalance in production totals between larger and smaller departments, there is some indication in the literature that larger departments are generally more productive (Jordan, Meador, & Walters, 1998; Dundar & Lewis, 1998; Jordan, Meador, & Walters, 1999; Toutkoushian & Porter, 2005). What is the BME department's stature in comparison to other engineering programs in the same school? Is it the research powerhouse of the school? Is it a typical department? Or is it relatively undistinguished? By introducing of measure of relative size between the BME program and other engineering departments in the same school, it was hoped that it could provide an approximation of the emphasis that the BME department has within its own school, which, in turn, may help better explain its level of scholarly output.

Library resources

Library expenditure levels have been demonstrated to correlate with measures of research productivity (Dundar and Lewis, 1998; Rawls 2015). The assumption is that most every research project relies to some degree on the use of library resources and that a well-resourced library should aid researchers more effectively and efficiently than a poorly-resourced one. In addition to the potential that library resources have for making a direct impact on research productivity levels, the argument has also been made in the literature that library investment serves as a proxy for the extent to which institutions support their research enterprise. In finding a significant correlation between library expenditures and departmental research productivity, Dundar and Lewis (1998) theorized that such expenditures might be an indicator of an institution's capacity and propensity to investment in its overall research infrastructure in other ways that were less directly measureable – or less widely reported – than library expenditures. Traditionally, volume and titles counts were considered the best metrics for establishing library quality, but in recent years library expenditures have been viewed as a more valid measure in the digital age (Kyrillidou et al, 2012). Total library expenditures was the measure used for this purpose. The Association of College and Research Libraries conducts an annual Trends and Statistics survey that collects this data, which was available through a subscription to the *ACRLMetrics* product from the company Counting Opinions.

Independent variable data handling and assembly

Gathering the data for the independent variables involved combining information from multiple sources and from multiple years. Data from the American Society for Engineering Education's (ASEE) annual survey were assembled into three excel files – one each for aggregating faculty data, enrollment data, and non-tenure and research assistant data. Each spreadsheet included three years' worth of survey results which were combined into a single worksheet that served as a data sources for a pivot table. This allowed the data for each department to be aggregated by year. For the tenure track faculty data, an average faculty count for each year was calculated. This served both as the basis for the per-faculty component of the dependent variable data as well as the base for dividing each department's non-Caucasian and female faculty counts. Likewise, the enrollment data used to produce an average per year for each department. Because the assistantship, non-tenure track faculty, and fellowship data was all reported at the school-level, rather than the department-level, these were assembled somewhat differently. An average of each of these measures per year was divided by the overall school's tenure/tenure-track faculty count to come up with a ratio of each of these measures to full time faculty at the school level. The calculated results from each of the three separate spreadsheets were then combined into a single master data spreadsheet. As each of the components were added to this master data file, both the existing data in the master file and the data being added were sorted by IPEDS number to ensure that the rows – each representing a different institution – were matched up properly aligned. To further confirm that the data was being correctly assembled the IPEDS numbers of the data being added were tested against the IPEDS numbers of existing data on a row-by-row basis to verify each component was belonged to the same institution. This was accomplished by using "if" formulas at the end of each row that would produce word "True" if the IPEDS number from the existing data matched the IPEDS number of the data being added in the same row. If the numbers did not match, the formula would produce the word "False," at which point the problem was identified and corrected. The library expenditures data – consisting of the average expenditures for institution from 2013 through 2015, was collected from the Association of College and Research Libraries (ACRL) annual surveys and were added to the master data file in the same fashion as the ASEE data. A handful of institutions did not report to ACRL, but each of them did report the same data point to the Association of Research Libraries (ARL), so data from both sources are included in the final model. Both organizations categorize expenditures as all expenditures spent from all fund sources excluding fringe benefit payments in US Dollars. Finally, the research expenditure data was collected from the National Science Foundation's (NSF) Higher Education Research and Development (HERD) expenditures survey. Each department's research expenditures for 2013, 2014 and 2015 were added to the master data file in a separate column. Unlike the ASEE and ACRL/ARL data, the HERD survey does not include an IPEDS number, so the data had to be carefully added to ensure that the correct figures were being added to the correct rows by using an eyeball comparison of each institutions name (an "if" formula could not be used because institutions names were listed differently in the HERD survey than in the ASEE data – e.g. University of California - Berkely and U. Calif. - Berkeley). Once added to the master data list, the research expenditures for 2013 and 2014 were inflated to match 2015 dollars based on the CPI inflationary factors.

Data Analysis

Ordinary-least-squares regression analysis was used to develop the statistical model to serve as a production function describing program-level research productivity in the BME field. Each program's weighted article count was regressed against the array of programmatic, school-level, and institutional variables described above, in order to produce a regression equation empirically specifying the relationship between input and output levels in the production of scholarly output. This regression equation was then used to calculate a predicted weighted article count for each department based on their actual inputs levels. This predicted value was then subtracted from each program's actual output to produce a measure, either positive or negative, of that program's residual output, which is viewed as an indirect indicator of research efficiency.

The original intent was to identify the subset programs based on those whose residual scholarly output fell three standard deviations above or below the mean residual output, from which a sample would be drawn for the qualitative component. Once the actual residuals were calculated, however, the dispersion of the residuals was condensed too close to the mean to use three standard deviations as the distinguishing criteria. To solve this problem, percentile rankings were employed whereby those programs ranking above the 90th percentile were deemed high research efficiency programs and those below the 10th percentile were considered low research efficiency programs. The ramifications presented by the compactness of the range of residuals is discussed in Chapter 4, but the primary takeaway is that generation of scholarly output by BME programs is generally efficient across PhD-granting programs at large research universities when considered from a Cobb-Douglas perspective.

Pre-modeling data analysis included the examination of the descriptive statistics of the dependent and independent variables. This analysis consisted of the number of articles produced by the programs, the types of the journals in which they are published, and the inflationary effect that those journal's impact factor had on the calculation of the final dependent variable. It also includes an examination of the means, ranges, and standard deviations of each of the independent variables. In both instances, the descriptive statistics help illustrate the nature and scope of scholarly output and the factors of production that typify BME programs.

Data Modeling

According to Allen (1997), the only necessary assumptions of regression modeling are that the dependent variable must be continuous and that its relationship with any predictor variable(s) must be linear. By meeting these two assumptions, a regression model serves as a descriptive summary of the variable relationships within the dataset itself. It is only when the researcher wishes "to make valid statistical inferences about population parameters from sample statistics" (p.181) that a larger bevy of assumptions must be met. Allen suggests a minimum threshold of four assumptions that should also be met, while others, such as Long (2008) list as many as ten. Because this study encompasses the entire population of PhD-granting BME programs at "highest research activity" universities, there is no concern as to whether or not the results can be inferentially conferred beyond the cases in the study. According to Rawlings, Pantula, and Dickey (1998), however, in order for ordinary least squares analysis to derive the least biased (minimum variance) and maximum likelihood estimators, it must meet three “basic assumptions: normality, common variance, and independence of errors” (1998, p. 325-6). Prior to any model building, the data was tested to insure that it met these assumptions using skewness and kurtosis diagnostics to test the distribution of the data; probability plots to test the homogeneity of error variance; and boxplots to detect any instances of error correlation between cases for the independent variables. Rawlings, Pantula, and Dickey (1998) also list other potential problem areas that should be tested for and, if detected, remedied or at least reduced in order to maximize the adequacy of the model. Specifically, these additional issues include: outliers, collinearity of independent variables, and inadequate specification of the functional form of the model – all of which are also examined in the course of modelling the data for this study.

In specifying the statistical model itself, the key task was to justify which of the independent variables listed above should comprise a final model. As Allen emphasizes, “the specification of a regression models should be based on theoretical considerations rather than empirical ones” (1997, p. 166). As such, the inclusion of some variables in the model were viewed inviolable due to their importance to the study’s conceptual framework as well as the fundamental role that they play in the production of scholarly output. These are faculty count (as incorporated into the measure of the dependent variable), research funding, and student enrollment levels. To state the obvious: without faculty members, no articles would be written; without research funding, no significant research could take place; and without student enrollment, there would be no teaching responsibilities to compete for the researcher’s time. The self-explanatory relationship of these variables to scholarly output would have made their exclusion inconceivable from any quantitative model claiming to measure the determinants of that output. The connection that the remaining independent variables have to scholarly output, though based in logic, cannot be assumed with the same level of certitude.

These remaining independent variables were explored using the enter method to determine whether or not they warranted inclusion in the final model. Some of these variables represent labor-related factors of the production function, which have the potential to either directly influence productivity (research assistants, non-tenured research faculty, post-doctoral fellows, etc.) or to ameliorate the teaching loads (teaching assistants, non-tenured teaching faculty). The remaining independent variables are represented in the literature as characteristics that may have an influence on productivity (e.g. academic rank, library expenditures, department size, percentage of enrollment that are graduate students, etc.). The inclusion of both the labor-related and institutional variables was based on whether or not an empirical justification warranted it. This evidence took the form of their inclusion's impact on R^2 and adjusted R^2 values, the residual means square MS(Res), and the individual significance levels of the independent variables as they were entered or removed from the model.

With a final model in place, the BME programs were ranked in terms of their residual scholarly output and the results can be analyzed in a variety of fashions. These analyses centered on examining the differences between the high-efficiency and low-efficiency programs in terms of the means and standard deviations for each variable in the model by using both descriptive statistics and independent t-tests to identify areas of statistical significant differences between these groups, as well as those programs that fell in the middle between the 10^{th} and 90^{th} percentiles. Further analysis centered on how the method used in this study differs from more tradition rankings of research productivity, such as research expenditures or article counts. The purpose was to explore the extent to which ranking programs by the residual method rearranges the more traditional methods or measures of program rankings. The primary question to be answered was whether or not the programs that occupy each upper and lower ends of ranking the spectrum according to traditional measures are also found at the upper and lower ends of the residual rankings. If there was only a small difference between the two approaches, it would indicate the calculating residuals as a means of rating programs might actually be unnecessary or superfluous.

Finally, the results of the model were reviewed to consider how the qualitative component of the study could be used to better understand the quantitative findings. For instance, for any unexpected or surprising quantitative results, the qualitative approach was tailored to help understand what was driving any unanticipated developments. Essentially, whatever stood out in the empirical results had to be followed up on during the qualitative inquiry.

Qualitative Methodology

As previously stated, the objective of the qualitative component of this study was to identify which characteristics distinguish over-performing BME programs from their under-performing peers and how those characteristics influence faculty productivity. Given the design of the quantitative portion, the expectation upon entering the qualitative stage was that much of what distinguishes these varied BME programs would be related to organizational factors – departmental climate and leadership, access to resources, modes of research support, administration, and other institutional characteristics that might influence productivity in some form or another - but factors that were not directly quantifiable in the model.

Research Questions

R3: Which institutional factors most influence the research efficiency of faculty research productivity efforts in biomedical engineering departments in US doctoral institutions?

R4: How do these institutional factors influence the research efficiency of faculty research productivity efforts in biomedical engineering departments in US doctoral institutions?

Qualitative Research Design

A basic qualitative research design was used for this component of this study, to help "discover and understand" how institutional factors influence research productivity from the perspective of BME researchers and administrators (Merriam, 1998, p. 11). First, data collection and data analysis occurred concurrently. This involved conducting one-on-one interviews with BME faculty and BME affiliated research administrators using an interview guide. Some slight alterations were made data gathering process to flesh out or accommodate new themes as they emerged, but not to the extent that it would qualify as a grounded theory design. Codes and categories were developed based primarily on the data received, not from existing hypotheses. Nevertheless, connections between the data collected and the stated research interests of this study was explored for fit – this included questions about institutional support for research activity (e.g. access to resources), departmental culture and values concerning research, etc. Questions surrounding the role of graduate research assistants and pre- and post-award support were revised and expanded as it became increasingly clear that these two areas were particularly important. The sampling technique was designed to support theory construction and not for representativeness of the overall population. Finally, some post-analysis literature review was necessary to provide additional context for understanding unanticipated results.

Research Population, Sampling and Data Collection

The unit of analysis remained PhD-granting BME programs at US doctoral universities rated "highest research activity," but the sampling frame was based on those programs identified by the quantitative analysis as having the highest and lowest productive efficiency. This maximum variation approach was designed to support the goals of the study rather than to achieve general representativeness of the overall population of BME programs (Sandelowski, 1995; Patton 2002; Givens, 2016). Rather than simply selecting those programs with the largest and smallest residual research efficiency, however, the sample drew from those institutions whose residual fell either above the 90th percentile or below the 10th percentile, excluding outliers as described above. This allowed the study to focus on institutions that better fit the model rather than those that might be outliers and, hence, are not accurately described by the model (Rousseeuw, 1994). The goal was to select three each of high and low research efficiency programs, for a total sample of six. Due to recruitment difficulties, however, the final sample included four institutions – two high and two low – and the number of interviews per institution ranged from a low of three to a high of five.

Using the percentile rankings, the quantitative analysis identified seven programs at either end of the spectrum, but two at the upper end and one at the lower end were excluded as outliers. This left five high and six low productive efficiency programs from which to form the sample for the qualitative component. To draw the sample, each institution was assigned a random number using the =rand() function in excel. The institutions were then sorted by the randomly assigned number and the first three from each end of the spectrum were chosen for the study and recruitment began.

Data collection consisted of one-on-one interviews with BME faculty as well as research administrators with BME affiliations, such an associate dean for research in the school of engineering, a BME researcher who had just stepped down from the role of VP for Research of his institution and two researchers with BME affiliations who also ran research institutes and who has served as deans – and in one case, as president of a research institution – at other institutions. Zoom video conferencing was used to conduct the interviews, which were recorded, saved, and transcribed. Participants were informed that they can end the interview process at any time. Participant confidentiality was maintained by excluding the names of researchers, institutions, and any other potentially identifiable information.

The interviews were intended to provide insight into the research environment surrounding each of the selected BME programs. Of particular interest were the programmatic and institutional characteristics that the researchers believe most help or hinder their productivity. The intent was to allow the researchers to identify the themes and issues that they believe influence their productivity the most, without leading them too much, in a manner that should "add depth, detail, and meaning at a very personal level of experience" (Patton, 2002, p. 17), while also making sure that certain obvious topics were covered consistently across cases. The interview guide's structure was based in part on an instrument in *The Research Productive Department: Strategies for Departments that Excel* (Bland, Weber-Main, Lund, & Finstad, 2005) and can be viewed in appendix A. Although that guide was intended for interviews with department chairs only, it shares many of the same goals. Some of the questions and language from that interview guide are used directly here. Furthermore, as suggested *Qualitative Research and Evaluation Methods* (Patton, 2002, exhibit 7.1) the guide was structured topically in such a way as to allow it to serve as a checklist. If the free flow of the participant responses touches on topics not yet raised by the interviewer, then the interviewer can check the topic on the guide to help maintain a "conversation style but with the focus on particular subject[s] that has been predetermined" (Patton 2002, p. 343). Slight changes made to the original guide during the course of the data gathering period were for the purpose of accommodating new and unanticipated themes as they were identified, or as existing questions either failed to elicit substantive information, or reached a point of data saturation where no further information was needed.

Description of selected institutions

While efforts were made to recruit faculty form all six institutions, ultimately, as is discussed further in the recruitment section, only enough data could be collected from four of the six institutions to be viably included in the study. Two of these programs were high productive efficiency programs, while the other two were low – preserving the originally intended balance. As a general indication of the institutional prestige, the U.S. News and World Report 2018 College Rankings for these schools ranged between the top ten to 60s in the national universities category. Based on the United States Census Bureau's designated regions and division, their geographical disbursement is as follows:

- Northeast Region (1):
  - Mid-Atlantic Division (1)
- South Region (2):
  - South Atlantic Division (2)
- West Region (1):
  - Pacific Division (1)

The two high productive efficiency programs are identified throughout the study as institutions A and B. The two low efficiency programs are identified as institutions C and D. It is important to note that institution A and D are at institutions that have well regarded medical schools. Institution B does not feature a medical school, but they have a formal partnership – including joint appointments – with a top tier medical school at an R1 university that resides very close by. Institution C does not have a full-fledged medical school. As the only university of the four not located in a major metropolitan area, its closest viable partner institution with an established medical school is several hours away.

Recruitment

The recruitment goal was to enlist three tenure-eligible and three tenured faculty members from each of six programs. The rationale for this approach was to stratify the sample between junior and senior level researchers to ensure that both groups were well represented and in order to make a comparison of research experiences by rank if distinct differences emerged (Patton, 2002; Edmonds & Kennedy, 2017; Hesse-Biber, 2017). Once the institutions were identified, the attempt was made to recruit BME researchers directly by sending the IRB approved recruitment email message (see appendix D) using contact information provided on each BME program's website.

Each faculty directory included the rank its members in addition to their contact email addresses. This rank information was used to facilitate the stratified sampling approach. To accomplish this, the name, rank, and email address of each faculty member of each program was copied into an MS Excel file, with one worksheet per program. Each faculty member was assigned a random number using the =rand() function in excel. The information was sorted by the random number and the first three faculty members for each rank were selected for recruitment from each program. As such, the initial recruitment emails were sent to only nine faculty members per department. The logic of this approach was that if emails were sent to everyone simultaneously, it could result in an uneven representation of ranks, particularly for those programs that might have an uneven distribution skewed one way or the other toward junior or senior faculty. If this first wave failed to recruit a sufficient number of respondents – or no respondents – then the idea was to move on to the next faculty member on the list and to repeat the process until enough respondents were successfully recruited.

The first wave of recruitment emails elicited almost no response. The timing of the invitations - which were sent in mid-November 2017 – was much later in the fall semester than the design had envisioned. The delay was caused by unanticipated challenges encountered in the quantitative modelling process, which necessarily had to be worked out before the programs could be selected and recruitment could begin. Coming just before the Thanksgiving break and the run up to exam periods that follows, it is not surprising that the responses were weak heading into a very busy period of the semester. Most faculty did not respond at all and those who did mostly declined. The first round of 54 invitations resulted in just two respondents who were willing to participate in the interview.

Due to the weak response rate, the stratified recruitment strategy was abandoned in favor of sending invitations to all of the remaining tenure and tenure-track faculty in each of the departments. This second round of recruitment emails went out in early December. Being even closer to the exam period for most academic calendars, this was still inopportune timing. To help address this issue, the second email included a timeframe reference which stated that interview could take place "at a time convenient to you between now and early February 2018." The original recruitment email had made no reference as to when the interviews needed to take place, which may have lent the impression that the intent was to conduct interviews in the immediate future. The hope was that the indication of a broad window for interviews, which ran across the winter break and into the early spring semester, would allow a degree of flexibility that would potentially appeal to more faculty members than the original request. This effort netted six more interviews, including enough to make institution A definitely viable for inclusion in the study.

Now that all of the faculty of the selected institutions had been contacted, the determination was made to attempt a follow-up email in mid-January, 2018 to every researcher who had not responded to the early December email. This delay between recruitment efforts was based on the assumption that it would be unproductive to send out the follow up over the winter break. This effort drew a few more promises of interviews, but it was becoming increasingly clear that other steps were needed. Therefore, at the end of January the decision made to give up on recruiting from each institution that had yet to yield any interviews or scheduled interviews. This lead to new recruitment efforts at two schools which garnered three new interview participants. I followed up with these new schools again in early February and scheduled a few more interviews.

By mid-February, one of the key challenges in determining how best to deal with the sluggish recruiting efforts was that interviews had already been collected from five separate institutions. There was no reason to believe that any of the remaining heretofore uncontacted programs would be more responsive than those five. So if attention was shifted toward new programs, that would mean giving up on some of the few existing interviews. The other option was to contact the same schools a third time, which seemed somewhat futile. Ultimately, the decision was made to continue on with the same programs by sending a third recruitment email using a new tactic. Again, anyone who had not yet accepted or declined my invitation was contacted, but the difference was that the third message was sent as a reply to the original recruitment email with additional text. The text of the reply can be found in appendix E, but it essentially indicated that other researchers from their program had been interviewed and, while the data was very interesting, there was not enough to draw reliable conclusions and that just a few more interviews were needed.

This email was more slightly effective and several interviews were scheduled across the rest of February and into mid-March. By that time four interviews had been collected from institution A (including one BME research administrator), five interviews from institution B (including one BME research administrator), four interviews from institution C, and three interviews from institution D (including two BME research administrators). The interview respondent's rank and role in research administration are listed in table 1 by institution.

Table 1

Rank and administrative roles of interview participants at institutions A, B, C, and D

Institution	Researcher Rank	Researcher ID	Admin. Role
A	Associate	A1	No
	Professor	A2	No
	Professor	A3	Yes
	Assistant	A4	No
B	Associate	B1	No
	Professor	B2	Yes
	Associate	B3	No
	Professor	B4	Yes
	Assistant	B5	No
C	Assistant	C1	No
	Assistant	C2	No
	Professor	C3	No
	Professor	C4	No
D	Professor	D1	Yes
	Professor	D2	Yes
	Professor	D3	No

After three attempts over four months, it seemed like there was a limited likelihood that additional researchers could be recruited from these schools; so it was decided that the original goal of six participants per institution was not feasible. The decision seemed reasonable, however, because the existing interviews had each yielded thick, substantive responses that were very illuminating. With the exception of institution D, the data collected seemed to provide a sufficient amount of information to develop a realistic feel for the environment surrounding each program. And while more information from institution D would have been helpful – for instance, none of the respondents were junior or mid-career faculty – the two research administrators interviewed offered broad perspectives and provided the two richest transcripts among the interviews. It just would have been preferable to hear more from researchers on the front lines.

As late March, 2018 approached, a decision had to be made to either attempt to recruit participants from a fifth and sixth school or to move forward with the study based on just four. The former would require extending the study another semester. In consultation with my dissertation advisor, it was decided that the data collected as-is provided enough information to complete the study, even though a wider and larger sample was originally preferred.

Coding and Analysis

A series of codes were developed based on the initial review of transcripts. Additionally, each question from the interview guides had a conceptual underpinning that also served as a code. Each time a code was applied it was accompanied by a qualifier of "positive," "negative", or "neutral," which was subsequently rated as 1, 2, or 3. For example, the response "It is the grant obstruction process, rather than the grant assistance process," would be coded with "grant support services" tag and then qualified as "negative" with a rating of 3 (for strongest). The coding list and scheme can be reviewed in appendix G.

The interview data was loaded into an excel spreadsheet were each participant's response to each question served as a single row. Added to that row was the respondent's unique identifier number, the institution's identifier letter, whether the institution was a high or low productivity environment, and the academic rank of the respondent. Then a table was created that included the list of codes, qualifiers and rating options from the coding scheme. This table was used to create dropdown lists that would allow as many coding tags as necessary to be applied to a row's single question response. This produced a single worksheet that included the full participant information, the data for each question, and the applied coding scheme. This worksheet was then used to power a pivot table. In turn, the pivot table allowed the creation of a series of data slicers based on the coding, institutional affiliations, and questions that could be used to pull any number of related responses in the pivot table. For example, to access all of the data from low productive efficiency programs concerning grant management support services, the user can simply select the buttons for "low productivity" and "grant support" on the respective slicers and all responses with those to attributes were pulled into view on the pivot table. The user could then decide to narrow the analysis further by selecting the "negative" button on the qualifier slicer or, if only responses from an assistant professor at institution C were wanted, those slicer setting could also be activated, and so on.

Trustworthiness

Member checks were conducted with the interview participants from each institution by sharing session transcripts once the preliminary analysis had taken place. Participants were asked to confirm whether or not the transcripts were generally accurate and to note omissions or areas of disagreement. A log of member check comments was maintained using a format exhibited in Merriam's *Qualitative Research: A guide to design and implementation* (2009, p. 217). The log is available in appendix G. Most responding participants simply indicated that the transcript looked accurate to them. None of the responding participates registered any substantive disagreements or areas of omission. Three noted typos and one respondent made minor edits to the transcript to tighten up the wording of his responses in the manner of quote polishing, not in an attempt to change the nature of any of his statements or responses.

Additionally, peer debriefing was employed during the coding and analysis stage to determine if the substance of the data was being handled in a reasonable, accurate, and consistent manner. Lincoln and Guba (1985) suggest peer debriefing as a useful method of supporting the data creditability and in establishing overall trustworthiness of qualitative research studies. As Spall (1998) describes, the point is to discuss general conduct of the qualitative investigation at a handful of points throughout the process in order to help confirm that "the findings and interpretations are worthy, honest, and believable" (p. 280). It involves sharing and discussing the ongoing process of data collection and analysis with a disinterested colleague who is not involved with the project. Janice Baab, a senior research analyst in VCU's Office of Planning and Decision Support, agreed to review the interview guide, coding scheme and a sampling of transcripts. Then, independently, both she and the researcher applied the coding scheme to the transcripts and then compared the results. Both parties generally came to the same conclusions. Our interpretations of the data and applications of the codes were essentially consistent. We also both agreed that the original coding scheme needed to be expanded to account for additional topics and themes. In fact, in the course of coding the sample of transcripts, both parties had independently created the additional codes they felt were needed. When compared, the new codes created by each party were remarkably similar. Each of us agreed that tags were needed for collaboration; that "institutional support" code should be expanded to distinguish “departmental support” from “university support;” and, that a code was needed to label instances of “administrative burden.” Furthermore, Janice also noted a potential inconsistency with the plan to qualify each code as “positive” or “negative.” While agreeing with the concept generally, she noted that some concepts were inherently negative or positive. This could introduce ambiguity to the task of qualifying some data as positive or negative. For instance, “impediments” is clearly a negative concept. If a participant were answer the general question about impediments to their research by indicating that there were none, should the coder apply the “positive” qualifier (because no impediments is a good thing) or should the coder apply the “negative” qualifier (because the participant answer the question in the negative)? We both agreed that it could be problematic, but because I was the only coder and analyst involved, it would not cause any problems in the actual analysis, so long as I was consistent and was aware of the issue.

Pilot Interviews

Prior to beginning work in earnest on this dissertation, pilot interviews were conducted with two senior researchers in Virginia Commonwealth University’s department of biomedical engineering. The interviews were conducted separately using the interview guide found in appendix A. The goal of these interviews was to help develop an understanding of the general nature of BME scholarly research and how external institutional factors might impact research productivity. Each researcher underscored that research productivity in their field was highly dependent on the quality and quantity of graduate assistants available to researchers. They both also perceived an improvement in the quality of services offered by VCU related to grant management, procurement, and other centrally-provided research supports over the years, which they each believed had a positive impact on their productivity. They both suggested that their department's culture values discovery over commercialization of research, although one researcher indicated that departments that do emphasis patents and other commercialized research create pressure on researchers that he believes negatively impacts productivity. Both pilot participants also indicated that newer NIH policies that favor interdisciplinary work between medical researchers conducting clinical trials and BME researchers has led to significant increases in NIH funding directed toward BME, and hence increased productivity.

Researcher Positionality

I am an outsider to biomedical engineering with no prior connection to the field. The same is also true regarding other branches of engineering. I have never contemplated pursuing an engineering degree or career and I have never enrolled in an engineering class of any type. To my knowledge, I had never spoken to a biomedical engineer prior to conducting pilot interviews for this research project. Instead, I came to choose BME as the focus of this dissertation because I am interested in a career in the field of institutional research and, as such, I wanted to choose a topic relevant to institutional effectiveness in a higher education environment. This lead me to focus on the concept of research productivity from an institutional perspective – what can universities do to positively influence the productivity of their researchers? What things should they avoid? In contemplating how to carry out this project, it became evident that I would need to focus on a single field, because comparing research output at the university level would be too large of a scale to effectively analyze. This left me to identify a particular field to study. In going about that task, I worked on the assumption that the more complex the research environment of a particular field, the more opportunity there may be for organizational factors to influence a researcher's productivity. Researchers in every field require adequate time and funding to conduct research, but science and engineering generally require lab space, specialized equipment, and research assistants – all of which leave more room for the institution to have an impact. With BME, these complexities extend to include the commercialization of research which involves managing patent activity, intellectual property rights, and various technology transfer responsibilities. Likewise, BME is also fundamentally interdisciplinary, which is another area where organizational factors could come into play as some institutions may facilitate cross-discipline collaboration better than others. For these reasons BME seemed to offer a particularly complex and multifaceted research environment.

Delimitations

Although it is hoped that the findings of this study can be translated into strategies for improving research productivity in biomedical engineering in particular and STEM-H disciplines in general, it must be noted that the findings of the qualitative component of this study will be drawn from a non-representative sample of BME programs. Likewise, the parameter estimates that are established in the quantitative portion of the study will measure only the correlation between the factors of scholarly production and scholarly output in BME programs. Causation can only be implied. As such, there should be no assumption that the parameter estimates for each input factor represents a value that would marginally add or detract from a program's productivity if changed.

IRB Statement

Prior to the collection of any human research data, this study was submitted for approval by Virginia Commonwealth University Institutional Review Board (IRB). All human research data collection has been carried out in manner consistent with the IRB proposal as approved, and with the *VCU IRB Written Policies and Procedures, Part I: Human Research Protection Program Overview*, available at: http://www.research.vcu.edu/human_research/wpp_guide_part1.htm. Pilot interviews were conducted with two BME researchers at Virginia Commonwealth University prior to the dissertation proposal for the purpose of establishing a basic understanding of the research environment surrounding the field. Data gathered in the interviews were not used in the study.

CHAPTER IV. RESULTS

Pre-modeling data analysis

Once the dataset was assembled, the descriptive statistics for each of the variables were examined to help illustrate the scope and nature of scholarly outputs and the factors of production that define BME programs, as well as to identify potential errors contained in the data. By using the data to describe the population generally, this analysis helped provide context once the subset of high and low productive efficiency programs were identified using the final model.

Dependent variable descriptive statistics

The dependent variable dataset consisted of 19,971 articles published 2,186 different journals from 2014 through 2016 by the 62 institutions included in the study. Approximately one-half of the articles were published in 85 separate journals, while the remaining half were published in the 2,101 separate journals. Table 2 shows the article counts by journal title. The wide variety of journals titles indicates the both vastness and interdisciplinary nature of the field. Of particular note is that three of the top five titles – PLOS One, Scientific Reports, and Proceeding of the National Academy of Sciences – are general scientific journals that do not deal exclusively in either engineering or medical sciences content. Together these three titles represent over 8% of all publications contained in the dependent variable dataset.

Table 2

Distribution of articles by journal title.

Journal Title	Article Count
PLOS ONE	686
SCIENTIFIC REPORTS	594
BIOMATERIALS	408
PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA	348
ACTA BIOMATERIALIA	262
ANNALS OF BIOMEDICAL ENGINEERING	254
JOURNAL OF BIOMEDICAL OPTICS	253
JOURNAL OF BIOMECHANICS	245
BIOMEDICAL OPTICS EXPRESS	242
MAGNETIC RESONANCE IN MEDICINE	223
LAB ON A CHIP	213
NATURE COMMUNICATIONS	213
ACS NANO	183
JOURNAL OF BIOMEDICAL MATERIALS RESEARCH PART A	165
TISSUE ENGINEERING PART A	165
BIOPHYSICAL JOURNAL	150
ACS APPLIED MATERIALS & INTERFACES	148
JOURNAL OF CONTROLLED RELEASE	140
JOURNAL OF BIOMECHANICAL ENGINEERING-TRANSACTIONS OF THE ASME	136
BIOMACROMOLECULES	129
JOURNAL OF MATERIALS CHEMISTRY B	129
MEDICAL PHYSICS	128
JOURNAL OF NEUROPHYSIOLOGY	124
ADVANCED HEALTHCARE MATERIALS	121
OPTICS LETTERS	120
PHYSICS IN MEDICINE AND BIOLOGY	117
ANALYTICAL CHEMISTRY	116
PLOS COMPUTATIONAL BIOLOGY	116
NEUROIMAGE	111
LANGMUIR	104
INVESTIGATIVE OPHTHALMOLOGY & VISUAL SCIENCE	100
JOURNAL OF NEURAL ENGINEERING	99
JOVE-JOURNAL OF VISUALIZED EXPERIMENTS	99
NANOSCALE	98
RSC ADVANCES	97
IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING	95
INTEGRATIVE BIOLOGY	94
ACS BIOMATERIALS SCIENCE & ENGINEERING	89
APPLIED PHYSICS LETTERS	86
JOURNAL OF NEUROSCIENCE	85
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY	85
ADVANCED MATERIALS	84
ACS SYNTHETIC BIOLOGY	82
TOTAL PUBLISHED IN REMAINING JOURNALS	12,435

Table 3 summarizes journal article counts by BME program, before the counts were normalized by the number of tenure and tenure-track faculty in each program. The average number of articles produced over the three-year period was 307 per program, ranging from a low of 40 to a high of 836. Weighting the article counts by impact factor caused each article to be inflated by a factor of 4.64 on average, with the highest factor of inflation for a single program recorded at 9.51 and the lowest being 3.21 times the unweighted article count.

Table 3

Article count summary per BME department, 2014 to 2016

Measure	Article Count	IF Article Count	Inflation Factor
Mean	307	1,569	5.10
Median	276	1,213	4.25
Range, high	836	7,953	9.51
Range, low	40	151	3.21

Next, both the raw article count and the weighted article count were divided by each department's tenured and tenure-track faculty count for the purpose of operationalizing the dependent variable. The resulting descriptive statistics, contained in table 4, indicate that 18.13 articles were produced per tenure or tenure-track faculty member per program from 2014 to 2016, or about six per year. It cannot be stated, however, that the average BME faculty member published at that level as the dependent variable data counted all published articles associated with each program – meaning that articles by graduate students or other researchers associated with a BME department were counted towards the department's total. The most productive department averaged almost 48 articles per faculty member over the same period, which represents about 16 articles per year. Once weighted by impact factor, these number were inflated to 89 weighted articles per faculty member over three years, with the most productive department producing 265 weighted articles per faculty and the lowest producing 23. Approximately 1,000 tenure or tenure track faculty were employed by 62 BME programs represented in the dataset (the average total count over the three-year period).

Table 4

Article count summary per faculty per BME program, 2014 to 2016

Summary per faculty	Article Count	IF Article Count	Inflation Factor
Mean	18	89	4.92
Median	16	70	4.26
Range, high	48	265	5.57
Range, low	6	23	3.88

Independent variable descriptive statistics

Table 5 shows the descriptive statistics produced by the independent variable data. One of the most noticeable aspects is that total research expenditures varied widely among institutions. At $32.7M, the standard deviation associated for research expenditures is actually higher than the mean itself, which totaled $32.6M for the three-year period. Research expenditures over this period ranged from $364,475 by the lowest spending department up to $192,563,662 for the highest spending department. While the student enrollment data also displayed a great deal of variability, the percentage of enrollment that graduate students represented out of all biomedical engineering majors ranged even wider. Some programs consisted as of as little of 8% graduate students, while others registered 100% graduate enrollment. Like research expenditures, the average graduate enrollment of 34% was accompanied by a sizable standard deviation at 23%. Because overall enrollment is assumed to be a constraint to productivity on the one hand and graduate assistants are assumed to boost productivity on the other, the interplay between these two factors presented ramifications in the model building process.

In regard to demographic characteristics, the departments included in the study consisted of 21% female and 36% non-Caucasian tenure and tenure-track faculty members. Departments with the highest values for these two variables were 91% non-Caucasian and 53% female.

Assistant professors made up 29% of the tenure or tenure-track faculty in these departments, ranging from a low of 0% to a high of 78%. Finally, the library expenditures for each institution, which consisted of the average per year from 2013 to 2015, ranged from a low of $7.8M to a high of $80.7M per year. Because all of the institutions represented are large, comprehensive doctoral universities, their libraries are presumably charged with supporting the research of a wide variety of disciplines. As such, it is reasonable to assume that the variability in the size of the library budget from one institutions to the next is a more reflection of institutional wealth and priorities in the distribution of resources, than a reflection of the size of the institution itself.

As table 5 indicates, three of the 62 programs were missing data for the ethnicity variable. These programs either did not respond to that section of the ASEE survey or they reported the ethnicity of their faculty as 100% "other," the latter of which can be characterized as measurement error. This represents the only missing data in the dataset. Several programs were excluded preemptively for lacking key data – such as faculty counts or unreliable article counts – during the data gathering process, before the dataset had been fully assembled. As discussed in chapter 3, this action was based on the critical importance of those data points to the study. For instance, without the ability to derive the dependent variable for a particular program, the only choice is to plainly rule out that program's inclusion in the study. Ethnicity, on the other hand, is not considered theoretically critical enough to warrant the omission of any cases. Because of its relatively low importance to the study, the problem of the missing ethnicity data could have been accommodated, if needed, by conducting pairwise deletion. This is because the other variables have no missing values, so the choice to use pairwise deletion would not have had any other effect on the model. The issue ultimately was ultimately settled, however, when subsequent analysis indicated that the distribution of ethnicity variable was unsuited for inclusion in the model, which is discussed more fully in the data modelling section below.

The variables pertaining to graduate assistants, post-doctoral fellows, and non-tenured or tenure-track faculty are all school-level ratio variables. Each one was derived by dividing the school-wide tenure and tenure-track faculty count into the school-wide counts for graduate assistants, fellows, or non-tenured faculty counts. While the resulting variables describe engineering in general rather than BME program in particular, it is apparent that graduate research assistants number much higher than the other personnel variables that represent additional labor added to the research process or to assist in handling the teaching load. There were approximately 2.5 graduate research assistants to tenure or tenure-track faculty, while ratio was approximately 1 or lower for the other additional personnel variables.

Table 5

Independent variable descriptive statistics

Variables	Descriptive Statistics				
	N	Mean	Std. Deviation	Minimum	Maximum
Research Expenditures	62	32,616,050.05	32,702,461.59	364,475.00	192,563,662.00
Res Exp as % of School	62	0.15	0.12	0.00	0.50
Enrollment	62	422.07	241.06	41.33	1,573.33
% Graduate Enrollment	62	0.34	0.23	0.08	1.00
Non-Tenure Track Teaching Faculty	62	0.22	0.19	-	0.90
Non-Tenure Track Research Faculty	62	0.38	0.40	-	1.67
Fellowships	62	0.95	0.92	-	3.99
Graduate Research Assistantships	62	2.51	1.39	0.55	9.26
Graduate Teaching Assistantships	62	1.04	0.71	-	3.88
% Faculty Assistant Professors	62	0.29	0.16	-	0.78
% Faculty, non-Caucasian	59	0.36	0.17	-	0.91
% Faculty, Female	62	0.21	0.10	-	0.53
Library Expenditures	62	30,302,171.25	15,849,854.33	7,822,750.00	80,661,447.00

Data Modeling

The contents of the master data file were converted into a IBM SPSS Statistics 24 file to conduct statistical analysis and modelling. First, the distribution of the data for each variable was tested to by running descriptive statistics for skewness and kurtosis. This produced a table of skewness and kurtosis statistics (see table 6) which were then interpreted by dividing each variable's statistic by its standard error, the quotient of which was examined to determine if it fell within the range of -1.96 to 1.96, a rule of thumb for normality of distribution in small datasets based on the understanding that 95% of observation in a normal distribution lie within approximately two standard deviations of the mean (Kirkwood & Sterne, 2003; Pandis, 2015; Rawling, Pantula, & Dickey, 1998). The results indicated that all the variables in the dataset were positively skewed. The skew of only one variable, the percentage of non-Caucasian faculty members, fell within the acceptable range. The remainder of the variables fell well outside of that range, often by a factor of two or three times the acceptable level. Similarly, all but three of the variables failed to past the same test regarding kurtosis.

These results indicated that the entire dataset consisted of variables with non-normal distributions. Because the tests for statistical significance and confidence intervals rely data with normal distributions, the variables could not be reliably used for regression analysis in their current state (Allen, 1997). Therefore, data transformation options were explored. Because the data for each variable was skewed positively, or to the right, it suggested the data for most variables likely included a handful of cases with values that greatly exceeded the mean, while much of the remaining cases were concentrated just below the mean. Visual evidence of this pattern was confirmed by generating histograms for each of the variables, which can be found in appendix I. For several variables, the histograms show a majority of cases distributed below the mean while a few cases far exceed the mean, generally lying outside of the distribution curve entirely.

In transforming data distributed in this fashion, it is necessary to use a calculation that reduces the actual values such as a logarithm, cube root, or square root transformation (Allen, 1997, p. 126). The intent is to constrict the actual distance between the high-end outliers and the rest of the cases using a method that, while consistent across all cases, causes a relatively greater reduction for the high-end outliers than it does for the lower value cases. A log natural transformation was attempted first, but it proved too powerful, causing most of the variables to become skewed too far to the left. The less robust method of square root transformation was explored next. This approach was partially effective, causing many of the variables to fall within acceptable parameters, but a crucial exception was the dependent variable, which remained too skewed to the right at 2.34. Although the raw article count did fall within acceptable parameters, the impact factor weighted version was considered the more conceptually relevant of the two, meaning another solution was needed. Furthermore, the key variable for research expenditure remained too skewed as well. As a result, a cube root transformation – weaker than log natural, but stronger than square root – was attempted as a potential middle route. As table 6 demonstrates, the results between the square and cube root methods were mixed, with different variables responding more favorably in each instance.

Table 6

Data transformation results

Indicators falling outside the ±1.96 range of tolerance are highlighted in red.

VARIABLES	Untransformed Data			Square Root Transformation			Cube Root Transformation		
	Skewness	Std. Error	Indicator	Skewness	Std. Error	Indicator	Skewness	Std. Error	Indicator
Article count, weighted	1.386	0.304	4.56	0.711	0.304	2.34	0.494	0.304	1.63
Article count, raw	1.256	0.304	4.13	0.559	0.304	1.84	0.345	0.304	1.13
BME Research Expenditures	2.378	0.304	7.82	0.694	0.304	2.28	0.162	0.304	0.53
BME Res Exp as % of School Res Exp	1.114	0.304	3.66	0.221	0.304	0.73	-0.161	0.304	-0.53
Enrollment	2.012	0.304	6.62	0.492	0.304	1.62	0.015	0.304	0.05
% Enrollment graduate	1.628	0.304	5.36	1.045	0.304	3.44	0.834	0.304	2.74
Library expenditures	1.018	0.304	3.35	0.539	0.304	1.77	0.37	0.304	1.22
Graduate research assistants	2.138	0.304	7.03	0.775	0.304	2.55	0.38	0.304	1.25
Graduate teaching assistants	1.9	0.304	6.25	0.185	0.304	0.61	-0.825	0.304	-2.71
Non-tenure research faculty	1.936	0.304	6.37	0.435	0.304	1.43	-0.344	0.304	-1.13
Non-tenure teaching faculty	1.609	0.304	5.29	0.318	0.304	1.05	-0.729	0.304	-2.40
Fellows	1.498	0.304	4.93	0.429	0.304	1.41	-0.195	0.304	-0.64
% BME faculty female	0.855	0.304	2.81	-0.608	0.304	-2.00	-2.099	0.304	-6.90
% BME faculty non-Caucasian	0.574	0.311	1.85	-0.723	0.311	-2.32	-2.029	0.311	-6.52
% BME faculty assistant professors	1.014	0.304	3.34	-0.153	0.304	-0.50	-1.28	0.304	-4.21

Note: As an indicator variable, the yes/no medical school variable's distribution was not relevant

Although both square and cube root methods each left approximately the same number of variables inadequately transformed – and thus unusable for model building purposes – the cube root method was more successful in terms of bringing the key variables into play. Most importantly, it caused the dependent variable data to fall within acceptable parameters. It was also successful in transforming the critical BME program research expenditure variable into a normal distribution. In regard to graduate assistants, the square method was more successful in transforming the teaching assistant data, but it was not strong enough to bring the research assistant data within acceptable parameters. The cube root produced inverse results, leaving the research assistant data in the acceptable range, which according to the pilot interviews, is the more critical of the two types of graduate assistants in terms of research productivity.

Taking all of these factors into account, the cube root version of the data is clearly the most favorable version to use. Unfortunately, the cube root method was too powerful in regard to the departmental characteristic variables relating to the rank, gender, and ethnicity makeup of the

BME programs, rendering the variables ineligible for inclusion in the model. Other variables that were lost included of the percentage of graduate enrollment, graduate teaching assistants, and non-tenure/tenure track teaching faculty.

Model Specification

Once the distribution of the transformed data was determined to be acceptable, the process of model building began. The weighted article count was selected as the dependent variable and the enter method was used to introduce independent variables into the model. Listwise deletion was used, though this had no particular consequences as there were no cases of missing data. The central task of model specification is to identify which variables should be included or excluded in the process of building a “best fit” model (Allen, 1997). In analyzing different variable selection criteria, Bendel and Afifi (1977) found that the inclusion of all theoretically or potentially relative independent variables available in a dataset generally produced models with inferior estimation power compared to other methods. An exception to this rule can be found in instances involving an exceptionally large sample, but because this study draws on just 62 cases, it must rely on stopping rules to regulate the variable selection process of identifying the best possible subset of variables (Rawlings, Pantula, & Dickey, 1998). Aside from the aforementioned sacrosanct inclusion of research expenditures and enrollment variables, the stopping criteria used to evaluate the different model iterations centered on the coefficient and adjusted coefficient of determination (R^2 and adjusted R^2), the residual means square, and the individual significance levels of independent variables. The R^2 and adjusted R^2 indicate as a measure of the amount of the dependent variable’s variance “explained” by a particular model’s subset of independent variables (Rawlings, Pantula, & Dickey, 1998). The residual mean square MS(Res) “is an estimate of σ^2 [variance] if the model contains all relavent independent variables" (Rawlings, Pantula, & Dickey, 1998, p.222). The MS(Res) value is biased upward when relevant variables are omitted but generally unaffected by the inclusion of unimportant variables, meaning that lower MS(Res) scores indicate a more complete model. Finally, the significance level of the individual variables provides as an important basis for inclusion or exclusion, not only because it indicates a strong correlation with the dependent variable, but also because the coefficients produced in a final model will be used to calculate predicted scholarly output using actual data from each BME program. It would be difficult to justify the inclusion of any variable for which significant uncertainty existed as to whether the sign of its coefficient was positive or negative. If such a variable were included it would present a problem when the resulting regression equation was used to calculate the predicted scholarly output for each department because it may have the effect reducing the predicted value for a factor that may actually aid scholarly research or for increasing the predicted value for a factor that may actually decreases productivity.

As described in chapter 3, an initial model was run which included the research expenditures and average enrollment variables only, as both measures were considered foundational in both the literature and to the conceptual framework that any final model that did not include these fundamental two measures could not be justified. Designated as model 1, its F test produced a p value of .000, indicating that can explain the variance in weighted article counts significantly better than an intercept-only model (Jobson, 1982). It produced an .505 R value, signifying a passable level of multiple correlation between the observed values of dependent variable and the values for weighted article count predicted by these two independent variables (Janke & Tinsley, 2005, p.280). While the model's MS(Res) of .641 is not strong, how this value changes from one model to the next will be more telling than the value itself. The model's R^2 value of .255 suggests that research expenditures and enrollment explained approximately 25% of the variance in weighted article counts from program to program, which is not especially strong. Furthermore, the square of the model's partial and semi-partial coefficient correlation statistics reveal that this explained variance is almost entirely attributable to the research expenditures variable, which accounted for 24.4% of the variance when controlling for enrollment and 24% by itself. Enrollment, on the other hand, uniquely explained only .25% of the variance. Research expenditures were positively and significantly correlated to weighted article counts with a p-value of <.001. As expected, enrollment was negatively correlated with weighted article counts, but with a p-value of .522 the relationship was not statistically significant. With a 95% confidence interval ranging from -.213 to .109, the actual nature of the relationship between enrollment and weighted article counts had no basis for certainty. Both variables fell well within acceptable parameters regarding collinearity, with each producing a tolerance statistic of .856 and a variance inflation factor of 1.168. In terms of standardized coefficients, research expenditures (.530) proved to be a much stronger predictor of scholarly output than enrollment (.078). The full output for model 1 can be found in Appendix J.

<u>Model 1 Expression:</u>

$$\text{IF Article Count/Faculty} = \beta_0 + \beta_1\sqrt[3]{Avg.\ BME\ Res.\ Exp.} + \beta_2\sqrt[3]{Avg. Enrollment} + \varepsilon$$

Expanding the model

The remaining independent variables were introduced systematically into the model along with research expenditures and enrollment, continuing with the enter method and listwise exclusion. This began with introducing the labor-related variable group, followed by the institutional characteristics group, with the former being considered more conceptually relevant as labor is a primary component of the production function. The labor-related variable group, representing additional human resources dedicated to assisting the research process were added to form model 2. This group of variables consisted of graduate research assistants, non-tenure research faculty, and research fellows. As mentioned, the two teaching-related variables that represented additional human resources were both too skewed to be included. The introduction of the variable group almost doubled the R^2 to .454, with an adjusted R^2 of .405. The standard error of the estimate also improved, decreasing from .801 in model 1 to .704 in model 2, indicating that the weighted article counts predicted by the model were moving closer to the actual observed levels of weighted article counts (though the cube root transformation makes the number itself less translatable). The overall model's F test remained significant at .000 and the MS(Res) dropped to .495 suggesting that labor-related group contain important predictors of weight article counts. These variables output also had the effect of improving the significance of the enrollment variable to a p-value of .144 and causing it to register a negative coefficient, matching the original assumption that increases in enrollment relative to the faculty count (encapsulated in the dependent variable) would be negatively associated with scholarly output. Furthermore, the graduate research assistant variable produced a positive coefficient that registered a high level of statistical significance with a p-value of .004. Surprisingly, its positive correlation with weighted article counts generated a beta weight that was slightly stronger than that of research expenditures, at .398 and .391 respectively, suggesting it represented the most influential variable in the model. The partial and semi-partial correlation scores, however, suggested that research expenditures still explained more of the variance in the model, accounting for 17.89% of the variance when holding the other variables constant and 11.9% of the variance by itself. Graduate research assistants, on the other hand, explained 13.84% of the variance while controlling for other variables and 8.82% uniquely. None of the remaining variables could explain more than 2.13% of the variance uniquely. Neither the non-tenure research faculty or fellowship variables demonstrated a statistically significant relationship to weighted article count, producing t-test p-values of .454 and .279 respectively. The non-tenure research faculty had a negative correlation to scholarly output – which contradicted expectations – but the standard error of its coefficient (.325) was greater than the coefficient itself (-.245), suggesting that very little confidence could be placed in the assertion that the relationship was truly negative. The full output for model 2 can be found in appendix K.

Model 2 Expression:

$$\text{IF Article Count/Faculty} = \beta_0 + \beta_1\sqrt[3]{Avg.\ BME\ Res.\ Exp.} + \beta_2\sqrt[3]{Avg. Enrollment} + \beta_3\sqrt[3]{Grad. Res. Asst.} + \beta_4\sqrt[3]{Non - Tenure\ Research\ Faculty} + \beta_5\sqrt[3]{Fellows} + \varepsilon$$

Next, a separate model was explored which included the variables representing institutional or departmental characteristics along with the mainstay research expenditures and enrollment variables to form Model 3. As mentioned, the variables pertaining to demographic characteristics such as rank, ethnicity and gender makeup of each department were too skewed to be included. This was also the case with the variable representing the percentage that graduate students to overall enrollment. This left only the variables representing average library expenditures and the percentage of research funding that each BME department's expenditures represented out of the entire school or college of engineering's research expenditures. The former was intended as both a direct measure library investment as well as a surrogate for institutional investment in research generally, while the latter was concerned with the size of the

BME department's research enterprise in relation to the overall school of engineering's research enterprise.

The introduction of these two variables produced slightly weaker R^2 and adjusted R^2 values than model 2, at .422 and .382 respectively. The MS(Res) score increased to .514, suggesting that important variables had been removed from model 2 to model 3. Research expenditures continued to demonstrate a strongly positive and statistically significant relationship to scholarly output, but the p-value of the enrollment variable was only .616, which was much weaker than in the previous model based on the human resource-related variables. Likewise, the p-value for percentage of BME research expenditures of school expenditures also produced a p-value that was far from statistically significant at .484. Library expenditures, however, produced a highly significant p-value of .001. Its relationship with scholarly output was strongly positive, with a beta weight (.390) the was second only to research expenditures (.485) in terms of importance to the overall model. The squares of the partial and semi-partial correlations attributed approximately the same amount of variance to both research expenditures and library expenditures, at about 18% each when controlling for other variables and with about 13% being uniquely attributed to both. The full SPSS output for model 3 can be found in appendix L.

<u>Model 3 Expression:</u>

$$\text{IF Article Count/Faculty} = \beta_0 + \beta_1\sqrt[3]{Avg.\ BME\ Res.\ Exp.} + \beta_2\sqrt[3]{Avg. Enrollment} + \beta_3\sqrt[3]{\%\ BME\ Res. Exp.} + \beta_4\sqrt[3]{Library\ Exp.} + \varepsilon$$

Model 4 explored the inclusion of all the independent variables. This model produced the highest R^2 and adjusted R^2 values thus far at .547 and .488 respectively. It also produced a lower standard error of the estimate at .652. The p-value of the overall remained significant at .000 and, at .426 the MS(Res) value suggested that this model included most important variables yet. As demonstrated in table 7, research expenditures, graduate research assistants, and library expenditures all maintained a strong, positive and statistically significant correlations with weighted article counts. The remaining variables – enrollment, percentage of BME to school research expenditures, non-tenure research faculty, and fellowships – each failed to demonstrate statistical significance. Of these, enrollment had the lowest p-value (.303) and remained negatively correlated with scholarly output. At a 95% confidence level, the variables for percentage of BME-to-school research expenditures, non-tenure track research faculty and number of post-doctoral fellowships all had confidence intervals with lower and upper bounds that spanned from negative to positive. The bounds of the enrollment variable also spanned zero, but the portion of upper bound that was above zero represented only about 25% of the entire interval. The full SPSS output for model 4 can be found in Appendix M.

<u>Model 4 Expression:</u>

$$\text{IF Article Count/Faculty} = \beta_0 + \beta_1\sqrt[3]{Avg.\ BME\ Res.\ Exp.} + \beta_2\sqrt[3]{Avg. Enrollment} + \beta_3\sqrt[3]{Grad. Res. Asst.} + \beta_4\sqrt[3]{Non - Tenure\ Research\ Faculty} + \beta_5\sqrt[3]{Fellows} + \beta_6\sqrt[3]{\%\ BME\ Res. Exp.} + \beta_7\sqrt[3]{Library\ Exp.} + \varepsilon$$

In order to produce a final model with trustworthy coefficients, model 5 incorporated only those variables that had consistently demonstrated, with certainty, that the sign of the coefficient was reliable – research expenditures, library expenditures and graduate research assistant variables. The exception to this rule was the enrollment variable, which was also included despite the uncertainty of its coefficient. Its inclusion was based on both its conceptual importance to model and because, though not statistically significant, its confidence interval suggested that its relationship to scholarly output was much more likely to be negative than positive. As such, model 5 consists of research expenditures, enrollment, graduate teaching assistants and library expenditures as it predictors. It produced R^2 and adjusted R^2 values of .534 and .501, respectively, suggesting that these predictors explained approximately 50% of the variance in the observed weighted article count variable, the most favorable value yet produced. Likewise, the MS(Res) of .415 indicated that model 5 consisted of the best combination of relevant variables compared to the other iterations. The p-value for F test remained statistically significant at .000.

<u>*Model 5 Expression:*</u>

IF Article Count/Faculty = $\beta_0 + \beta_1\sqrt[3]{Avg.\ BME\ Res.\ Exp.} + \beta_2\sqrt[3]{Avg. Enrollment} + \beta_3\sqrt[3]{Grad. Res. Asst.} + \beta_4\sqrt[3]{Library\ Exp.} + \varepsilon$

Research expenditures, library expenditures, and graduate research assistants all remained positively correlated with scholarly output and each was statistically significant well below the <.05 level. The coefficient beta weight for graduate teaching assistants was strongest at .388, followed by research expenditures (.334), library expenditures (.327), and enrollment (-.125). The square of the partial correlation statistics suggests that variance in the ratio of graduate research assistants to tenure or tenure-track faculty accounts for 20% of the variance in a weighted articles counts. Library expenditures and research expenditures accounted for 17% and 16% respectively. The significance of the enrollment variable’s coefficient improved to a p-value of .221. Despite not achieving the standard p-value of <.05 for statistical significance, the results for enrollment were the stronger in model 5 than the others, with the exception of model 2 where it produced a p-value of .144. While its confidence interval spanned zero, 78% of the interval was negative. The inclusion of the variable measuring the number of graduate research assistantships appears to impact the enrollment variable, in that the significance of enrollment improved dramatically in the models that featured graduate research assistants (2, 4, and 5) compared those that did not (1 and 3). This interaction has a logical explanation, in that graduate research assistants, as students, are a component of enrollment that apparently have a significantly positive impact on scholarly productivity. Meanwhile, as hypothesized, enrollment generally appears to have a negative impact on scholarly impact because, as it increases in relation to the faculty count, it represents an increase in non-research responsibilities of the departments faculty. It is possible that this conflicting relationship is clarified when graduate research assistants are accounted for in the model. Finally, the independent variables all passed the collinearity diagnostics, with tolerance values well above 0.10 and variance inflation factors well below 10, the standard rules of thumb for both statistics (Rawling, Pantula, & Dickey, 1998). The low collinearity statistics lend confidence in regard to the general accuracy of the regression coefficients. The full SPSS regression output for model 5 is available in Appendix N.

Table 7

Model comparison

VARIABLES	Model 1		Model 2		Model 3		Model 4		Model 5	
	Beta	Sig	Beta	Sig	Beta	Sig	Beta	Sig	Beta	Sig
BME Research Expenditures	0.530	0.000	0.391	0.001	0.485	0.001	0.304	0.026	0.334	0.002
BME Res Exp as % of School Res Exp					-0.087	0.484	0.050	0.676		
Enrollment	-0.078	0.522	-0.163	0.144	-0.057	0.616	-0.109	0.303	-0.125	0.221
Library expenditures					0.390	0.001	0.331	0.002	0.327	0.001
Graduate research assistants			0.398	0.004			0.344	0.010	0.388	0.000
Non-tenure research faculty			-0.076	0.454			-0.079	0.407		
Fellows			0.136	0.279			0.113	0.331		
Overall model sig.	0.000		0.000		0.000		0.000		0.000	
Std. Error of the Estimate	0.801		0.704		0.717		0.652		0.644	
R^2	0.255		0.454		0.422		0.547		0.534	
Adjusted R^2	0.230		0.405		0.382		0.488		0.501	
MS (Res)	0.641		0.495		0.514		0.426		0.415	
n	62		62		62		62		62	

Dependent variable: weighted article count

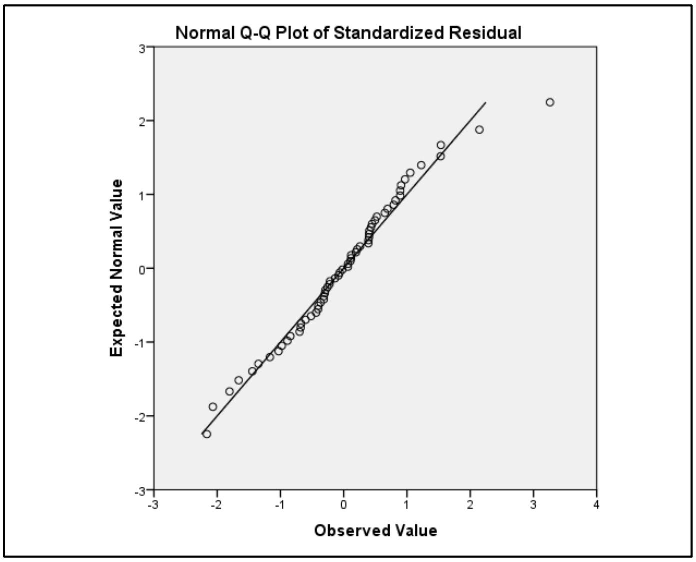
Regression model diagnostics

Before moving forward with model 5 to run the predicted scholarly output values and calculate residual productivity for each BME program, the model needed to be tested to determine that it met basic assumptions of least squares regression: the normality of residual distributions and homogeneity of variance (Rawlings, Pantula, & Dickey, 1998, p. 325; Fahrmeir, Kneib, Land & Marx, 2013, p. 74), recalling that collinearity diagnostics were run throughout the various model iterations and no problems were detected in terms of independent variable correlations.

The residual values represent the difference between the observed dependent variable data and the values predicted by the model's regression equation, or the error term of the model. The assumption is that these error terms should have a mean of zero and are normally distributed (Rawlings, Pantula, & Dickey, 1998, p. 342). If this assumption is violated, it is possible that p-values for the t-tests – and any subsequent inference based on those p-values – are not actually valid (Allen, 1997). To test the distribution, a Q-Q plot was generated by saving the standardized residuals for model 5 as separate a variable (ZRE_1). Figure 1 shows that the model's residual values mostly fall along the expected line of a normal distribution, with exception of a couple values at the high end of the range. Had the path of the values strayed further from the line or if the path had the intersected line at one or more points rather than lie along it, that would have indicated an non-normal distribution of residuals. Figure 1, however, illustrates a reasonably normal distribution.

Figure 1

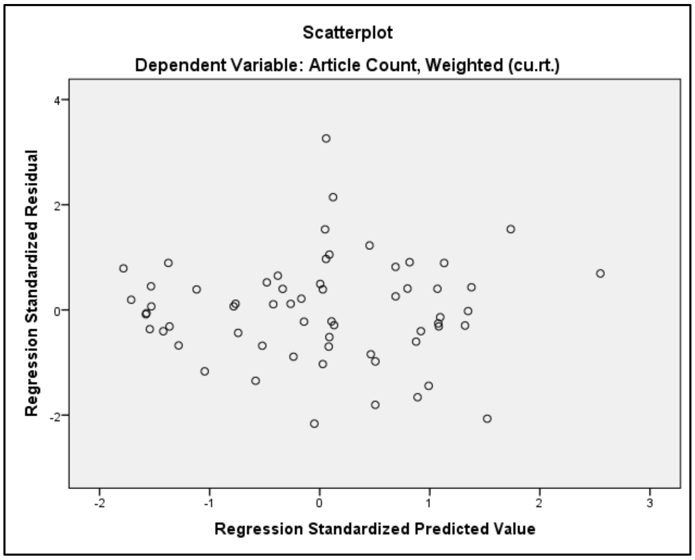
Q-Q plot of standardized residuals for model 5



Continuing with the ZRE_1 variable of the model's residual values, the homogeneity of variance in the model was tested visually by using a scatterplot of the standardized residual values charted against standardized predicted values, seen in figure 2. For a homogeneous variance the resulting pattern should appear evenly distributed across all ranges of the of the predicted values. If the variance fluctuates across the range of predicted values, then it would indicate that the data is homoscedastic. This would suggest that model is better at predicting values within some ranges than others or, stated differently, that some ranges contain more explanatory information than others (Rawlings, Pantula, & Dickey, 1998, p. 328). In cases of heterogeneous variances, most typically the variance in the residuals increases at the higher end of the predicted values, creating a fan or cone shaped pattern. Some degree of heterogeneous variance can be detected in the scatter plot in figure 2 in the middle of the range of predicted values, but nothing so strong as to resemble the telltale cone shape.

Figure 2

Scatterplot of residual and predicted values for Model 5



Applying and analyzing the model

Based on the analysis discussed above, model 5 was considered to be the best fit for use in running predicted values of scholarly output for each of the BME departments in order to determine their scholarly output residual for the purpose of identifying both low- and high-efficiency departments. The SPSS data was exported into an excel spreadsheet where that actual values of each BME department were multiplied by the model's coefficients then summed along with the regression equation's constant to provide a predicted weighted article count value for each department. This value was then subtracted from the department's actual weighted article count to calculate its residual scholarly output. The departments were then ranked in descending order by their residual scholarly output, whereby those departments with the higher residual values were considered high-efficiency than those with lower values.

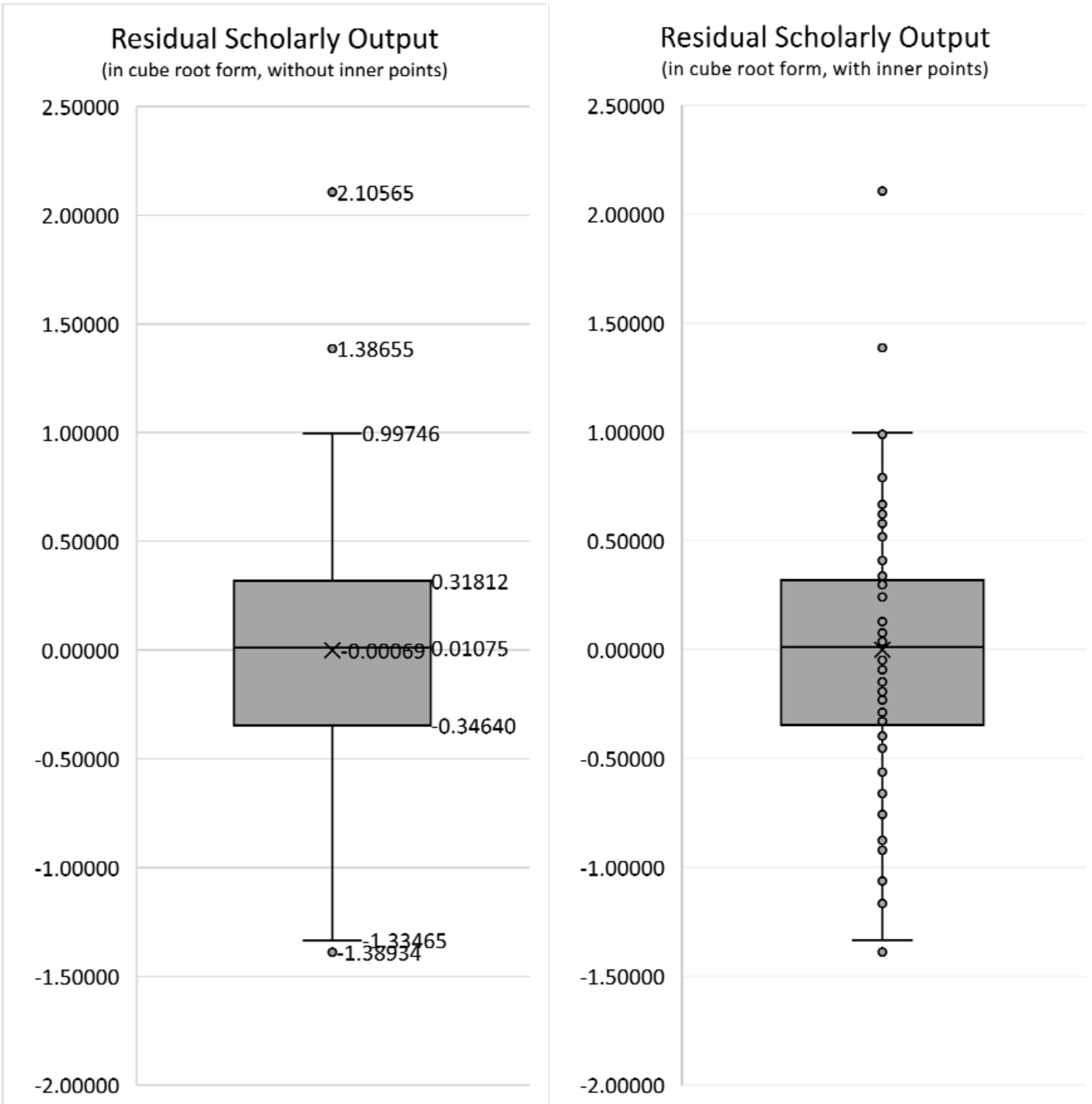
According to the approach detailed in chapter 3, those departments whose residual scholarly output fell between three and four standard deviations from the mean, either above or below, would be considered high- and low-efficiency departments respectively, from which the sample would be drawn to identify those institutions that would serve as the subjects of the qualitative component of the study. Only a single institution, however, fell outside three standard deviations from the mean and all but four institutions fell within of two standard deviations of the mean. The compactness of this range meant that standard deviation was not a reasonable measure of dispersion for distinguishing high- and low-efficiency performance, as 93.5% of the departments fell between within two standard deviations of the mean.

Therefore, percentiles were used as an alternative method of identification. Those programs with a residual scholarly output residing above the 90^{th} percentile of the entire range were designated as high-efficiency (.6190 and above), while those that fell below the 10^{th} percentile were designated low-efficiency programs (-.7480 and below). This approach identified seven departments each as high- or low-productive efficiency. As demonstrated in figure 3, however, three of the institutions fall outside of the box's upper or lower whiskers, with two above the top and one below the bottom, indicating that these programs are outliers that do not fit the model well. These departments were eliminated from eligibility for sampling based on that assumption, leaving five high-efficiency and six low-efficiency departments from which to draw a sample.

A sample was drawn by dividing the remaining eligible departments into groups based on high- or low-productive efficiency. The departments in each group were assigned a random number using the =rand() equation function in excel. Each group was then sorted by its random number into ascending order. The three institutions whose assigned random number sorted to the top of each group were then selected for qualitative analysis which is described later in this chapter.

Figure 3.

Box plot of residual productivity where values are in cube root form.



Comparative analysis of model 5 variables

While the beta weight coefficients from model 5 provide evidence as to which variables are most strongly associated with scholarly output, the this does not provide any indication as to how these factors vary between high and low productive efficiency departments. If scholarly output alone were the focus of the study, it would be logical to assume that low productivity programs would typically exhibit lower levels of the variables positively correlated with output. Because the calculation of residual output accounts for the level of inputs, however, it cannot be assumed that the low productive efficiency departments should exhibit lower input levels. Therefore, the untransformed data of each of the model's variables was explored to determine what characteristics, if any, distinguished the departments designated as high productive efficiency research environments from those exhibiting low productive efficiency. The 48 departments that fell in between these designations – within the 10th to 90th percentiles – were also included for additional context. Table 7 shows the mean and standard deviation for each of model 5's variables broken out by productive efficiency groupings.

Table 7.

Means comparison of model 5 variables by productive efficiency group.

Productive Efficiency Grouping	Statistic	Article Count/Faculty (weighted)	Article Count/Faculty (raw)	Research Expenditures	Enrollment	Graduate Research Asst.	Library Expenditures
High (90th percentile, n=7)	Mean	181.21	30.76	$ 65,034,376	563	2.60	$ 32,836,305
	Std. Deviation	61.20	9.01	$ 64,198,516	476	1.27	$ 19,008,293
Low (10th percentile, n=7)	Mean	42.27	11.12	$ 39,568,625	468	2.82	$ 29,373,724
	Std. Deviation	18.29	5.06	$ 18,413,794	306	1.49	$ 13,562,757
Middling (10-90th percentile, n=48)	Mean	82.55	17.30	$ 26,874,460	395	2.46	$ 30,068,009
	Std. Deviation	48.15	7.64	$ 24,923,732	175	1.41	$ 15,978,104
Total (n=62)	Mean	89.14	89.14	$ 32,616,050	422	2.51	$ 30,302,171
	Std. Deviation	58.71	58.71	$ 32,702,462	241	1.39	$ 15,849,854

The most noticeable aspect of this comparison is that difference in articles counts from one group to next suggests that high productive research efficiency is strongly associated with high overall research productivity. Programs above the 90th percentile in productive efficiency produced much higher overall article counts than their peers in the lower and middling percentile rankings, whether measured by weighted or raw counts. In raw counts – the actual number of articles published – the 90th percentile group produced almost twice as many articles per tenure or tenure-track faculty member as the middling group and almost three times more than the low efficiency group. This is despite having higher enrollment levels and roughly the same rate of graduate research assistantships. Likewise, the high productive efficiency group also had noticeably higher research expenditures, though it is interesting to note that the low productive efficiency programs had noticeably higher research expenditures than the middling group. Library expenditures were roughly even across groups.

To further illustrate the descriptive statistics by productive efficiency rankings, table 8 shows the statistics by each individual institution at the high and low ends of the spectrum.

Table 8.

Descriptive statistics of high- and low-productive efficiency programs.

Productive Efficiency Ranking	Article Count/Faculty (weighted)	Article Count/Faculty (raw)	Research Expenditures	Enrollment	Graduate Research Asst.	Library Expenditures
1	47	264	$ 28,960,570	113	2.3	$ 18,971,335
2	38	189	$ 38,988,201	442	0.8	$ 62,396,620
3	23	149	$ 1,391,817	537	3.0	$ 55,984,751
4	28	265	$ 192,563,662	588	4.3	$ 23,335,927
5	27	155	$ 102,447,457	1573	4.1	$ 15,534,054
6	23	126	$ 57,011,917	238	1.9	$ 19,609,086
7	30	121	$ 33,877,007	452	1.8	$ 34,022,364
56	6	23	$ 10,512,112	196	1.5	$ 14,340,237
57	7	28	$ 32,323,646	437	2.0	$ 13,545,640
58	17	64	$ 63,737,583	593	2.8	$ 44,270,670
59	13	55	$ 35,129,287	430	2.9	$ 47,365,287
60	11	41	$ 34,094,101	76	2.5	$ 24,676,033
61	18	62	$ 39,172,590	516	6.0	$ 35,996,168
62	6	23	$ 62,011,055	1028	2.1	$ 25,422,036

The results displayed in tables 7 and 8 present a number of significant implications regarding the efficacy of this study's approach toward measuring research productivity based on productive efficiency. These implications will be more fully explored in the discussion chapter, but the most immediate takeaway is the appearance that the top BME programs could simply be identified by the more traditional measures of research expenditures or article counts alone. For all the effort of collecting other data and running regression analysis, table 8 creates the impression that the programs ranked highest by this method seem to be those that had the most scholarly output and research funding to begin with. At the same time, however, table 8 also suggests that the model did have the ability to bring light to some under-producing programs that would have been ranked higher if looking at research expenditures alone, as evidenced by the above average level of research expenditures recorded by the bottom 10th percentile. What complicates both notions, however, is that the standard deviations associated with each variable's mean tended to be high – often more than one-half the mean – suggesting that the means were not overly representational of the individual cases. In particular, it should be noted that the standard deviation for research expenditures for the high-efficiency group was almost as high as the mean itself. To determine if the observed differences in the means were statistically significant, two-sided t-test were produced in SPSS. The top portion of table 8 shows the means for each variable by percentile grouping, while the bottom portion shows where statistically significant differences occurred between groups, relative to the high productive efficiency departments denoted as "A". A significance difference was registered at the .05 level between the high productive efficiency groups and the other two groups for the weighted articles counts.

Table 9.

Results of means comparison for model 5 variables

Untransformed Variables	Variable Means by Productivity Percentile		
	90th & above	10-90th	10th & below
Article Count, weighted	181.21	82.55	42.27
Research Expenditures	65,034,376	26,874,460	39,568,625
Enrollment	563.24	394.80	467.86
Graduate Research Asst.	2.60	2.46	2.82
Library Expenditures	32,836,305	30,068,009	29,373,724

Comparisons of Column Means	Statisically Significant Differences		
	90th & above	10-90th	10th & below
	(A)	(B)	(C)
Article Count, weighted	B C		
Research Expenditures	B		
Enrollment			
Graduate Research Asst.			
Library Expenditures			

Results are based on two-sided tests assuming equal variances. For each significant pair, the key of the smaller category appears in the category with the larger mean.

Significance level for upper case letters (A, B, C): .05

Tests are adjusted for all pairwise comparisons within a row of each innermost subtable using the Bonferroni correction.

Aside, from that, the only other statistically significant difference between means detected was for the difference in research expenditures between the high productive efficiency group and those departments that fell between the 10th and 90th percentiles. The results indicate that the resources associated with high productivity efficiency environments do not vary significantly in comparison to the low-efficiency environments, despite the contrast in mean values. At seven institutions each, however, the total number comprising these two groups are low enough that it may hamper the ability to demonstrate a statistically significance difference. But as it stands in this analysis, scholarly output is the only measure that constitutes an empirically discernable significant difference between the high and low productive efficiency departments.

Comparative analysis of other program variables

Some variables failed to demonstrate a statistically significant correlation to scholarly output during the regression modelling process and many others were excluded in advance because the data did not meet the necessary assumptions of that method. Clearly, these variables cannot be assumed to be drivers of scholarly output based on the data gathered for this study and the modelling technique it employs. On the other hand, the grouping of BME programs by productive efficiency, based on the model's regression equation, does afford the opportunity to compare the means of these remaining variables by those groupings by using the same sort of analysis as presented in tables 7 through 9. This analysis can help identify other characteristic differences that exist between the three productive efficiency groupings. Despite the fact that there is no quantitative basis for suggesting these extra-model variables have an impact on scholarly output, identifying any differences between these program-types is worth noting in the event that the qualitative inquiry produces additional data suggesting these characteristics matter.

Table 10 contains the means and standard deviation for those independent variables not included model 5. Because more variables are included here than in table 7, the format is transposed to display the variables in rows and the productivity efficiency groupings in columns. Some noticeable differences that distinguish the high productive efficiency group from the others include higher percentages of graduate student enrollment, higher rates of post-doctoral fellows and higher proportion of overall school research expenditures. The percentage of assistant professors increases inversely with the productive efficiency ranking, which matches the findings of much of the general literature on research productivity. On the other hand, the percentage of female tenure or tenure-track faculty and non-tenure teaching faculty were remarkably consistent across groupings. Non-tenure research faculty differed only for the low productive efficiency group, which average only half that of the high and middling groups.

Table 10.

Means comparison of other variables by productive efficiency group.

Varible	Statistic	High (90th percentile, n=7)	Middling (10-90th percentile, n=48)	Low (10th percentile, n=7)	Total
Research Exp as % of School*	Mean	20.1%	14.7%	14.2%	15.2%
	SD	17.9%	12.0%	6.9%	12.3%
% Grad Enrollment	Mean	47.3%	31.2%	36.1%	33.5%
	SD	17.3%	22.8%	30.4%	23.4%
NT Teaching Faculty to School Faculty	Mean	0.23	0.22	0.23	0.22
	SD	0.27	0.19	0.10	0.19
NT Research Faculty to School Faculty*	Mean	0.41	0.41	0.17	0.38
	SD	0.55	0.39	0.17	0.40
Fellows to School Faculty*	Mean	1.41	0.92	0.71	0.95
	SD	1.19	0.92	0.60	0.92
GTA's to School Faculty	Mean	0.90	1.01	1.43	1.04
	SD	0.36	0.70	0.98	0.71
% Assistant Professors in Program	Mean	20.4%	28.9%	34.7%	28.6%
	SD	12.0%	15.1%	22.0%	15.8%
% non-Caucasian Professors in Program	Mean	29.1%	37.3%	35.8%	36.2%
	SD	12.2%	17.7%	13.2%	16.7%
% Female Professors in Program	Mean	21.7%	21.5%	20.1%	21.4%
	SD	7.0%	10.6%	5.8%	9.7%

When the means for these variables were compared using two-sided t-tests, however, none of the differences were found to be statistically significant of the .05 level. Again, that result is not surprising given the low number of cases (n = 7) represented in the high and low productive efficiency grouping and given the fairly high standard deviations.

Exploring the Effect of Residual Scholarly Output Ranking

Table 11 contains the rankings for each of the 62 programs included in the study, sorted by their residual scholarly output ranking on the far left column. Additionally, each department's ranking for article count, impact factor weighted article count, and research expenditures are displayed across the rows. This includes rankings on a per faculty basis as well as total articles counts and research dollars per department without regard to the number of faculty. The programs that were the focus of the qualitative study have been highlighted in bolded red font. Institution A is ranked 7th, institution B is 5th, institution C is 60th, and institution D is 61st.

A compelling reason for using the residual scholarly output as a means of ranking effectiveness is that the rankings are relative to the program's resources. As such, it was expected that this method would upend more traditional rankings based on grant dollars or journal article counts because programs with less grant funding or fewer researchers would be expected to produce less scholarship. If the residual-based rankings proved to be highly correlated to the more traditional rankings, then it would indicate that the model and residual calculation might be unnecessary steps, and that ranking output alone could serve as the basis for identifying high- and low-efficiency research environments. Following the total research expenditures as an example, however, it is clear that the residual ranking has little or no correlation to expenditure rankings of the programs. The highest ranked program according to residual was only 28th in terms of total research expenditures, while the lowest residual ranked program was ranked 10th in terms of research spending.

Table 11.

Residual scholarly output rankings compared to rankings by research expenditures and article counts.

RESIDUAL SCHOLARLY OUTPUT	RANKING PER FACULTY			RANKING TOTAL		
	Article Count	Weighted Article Count	Research Expenditures	Article Count	Weighted Article Count	Research Expenditures
1	2	3	5	23	18	28
2	3	4	21	3	5	16
3	15	9	61	13	8	60
4	7	2	1	1	1	1

5	**9**	**8**	**4**	**5**	**3**	**2**
6	18	14	8	18	14	12
7	**6**	**16**	**34**	**4**	**11**	**22**
8	10	7	25	15	12	27
9	5	5	32	20	16	37
10	41	39	59	40	39	58
11	11	10	30	9	9	25
12	49	46	46	48	49	46
13	1	1	24	14	13	41
14	22	25	27	16	19	17
15	34	28	3	49	44	18
16	16	22	43	19	23	38
17	45	48	48	54	52	48
18	13	6	6	7	4	5
19	19	26	23	45	48	45
20	21	11	41	25	15	40
21	8	13	19	11	17	19
22	39	23	20	10	7	7
23	35	44	60	44	46	61
24	12	18	17	30	30	32
25	55	53	52	61	61	55
26	25	27	47	17	22	39
27	31	31	56	33	31	54
28	30	32	29	41	42	42
29	40	41	28	39	40	30
30	50	52	62	53	55	62
31	38	43	51	56	54	53
32	17	12	10	2	2	3
33	54	54	50	59	58	49
34	58	55	54	50	51	50
35	14	17	9	8	10	6
36	33	29	36	36	34	35
37	32	34	42	58	53	51
38	20	19	16	6	6	8
39	37	30	44	32	28	36
40	4	15	2	12	20	4
41	23	20	11	28	21	14
42	56	56	58	57	57	59
43	57	59	57	52	56	57
44	53	57	35	55	59	43
45	24	21	18	31	26	26
46	47	49	49	26	29	31
47	36	36	45	24	27	33

48	26	24	39	22	24	29
49	51	50	55	51	50	56
50	59	60	53	60	60	52
51	42	40	15	21	25	11
52	28	33	14	27	32	13
53	52	47	31	38	37	23
54	46	37	40	47	41	44
55	43	45	26	42	45	34
56	61	62	33	62	62	47
57	60	58	37	43	43	24
58	29	35	7	29	33	9
59	44	42	22	37	36	20
60	**48**	**51**	**38**	**34**	**35**	**21**
61	**27**	**38**	**12**	**35**	**38**	**15**
62	62	61	13	46	47	10

Informing qualitative inquiry

The quantitative analysis above was principally devised as a means of calculating residual scholarly output for the purpose of detecting high and low productive efficiency programs as suitable candidates for qualitative inquiry. It also served the purpose of identifying parameter estimates and predictive capacity that help confirm, reject, or otherwise provide measureable scales to assumptions about the factors that drive research productivity for the field biomedical engineering. But it serves another important function by eliciting as many questions as it answers. In doing so, it does more than point the qualitative inquiry in the direction of one program or another, it interactively informs the process itself by presenting questions to be resolved and results to be validated.

The most glaring question is why graduate research assistants are so strongly correlated with research productivity. That they should aid research productivity is not only consistent with the study's conceptual framework and the information gleaned during the pilot interviews, it is also what assistantships are precisely intended to do. So while it does not require a leap of logic to grasp why there should be a positive correlation, what is unexpected is that relationship between graduate assistants and scholarly output was both stronger and more significant than that of research expenditures – the measure by which BME researchers and programs are most typically judged. The interview guide already contained general questions about how graduate assistants might aid research productivity, but those questions were subsequently expanded to help add clarity and specifics about their direct role in conducting the research and in writing papers. The guide also asks researchers about their ability to support a sufficient number of graduate research assistant and the general availability and quality graduate research assistants in their program. When the responses to these questions are compared across cases, it will help determine whether the weight that assistantship variable carries is actually true reflection of the importance of their role in the production of scholarly output or more of an aberration of the model itself.

Additionally, developing a better understanding about the role that graduates play should also help inform why the negative correlation between enrollment and scholarly output was not as significant as anticipated. The expectation was that enrollment would have a fundamental impact on scholarly output, second only to research expenditures. Yet, not only was the correlation weaker and less significant than expected, but the programs that registered the highest productive efficiency generally had higher enrollment. The study's measure for enrollment includes all BME students, whether undergraduate, masters, or doctoral level. If doctoral students prove to a key ingredient in producing scholarship, then it becomes less surprising that the results regarding the enrollment variable where less certain that anticipated. The interview guide asks researchers if they feel like they have sufficient uninterrupted time to dedicate to scholarly activities, without asking them directly about their teaching loads. So the responses to that question, in combination to those pertaining to graduate research assistants, should help ascertain if the weaker-than-expected results for the enrollment variable is because teaching responsibilities are not as much of a constraint to scholarly output as originally anticipated or if the variable simply has been operationalized in less than optimal fashion. Furthermore, the analysis of the extra-model variables indicated that high productive efficiency programs had higher percentages of graduate students, yet these programs had roughly the same ratio of graduate research assistants to faculty members as the average across all programs. So if high efficiency programs do not have more graduate assistants, do they have a higher quality of graduate assistants due to the size of the pool available to them? Or is there some other aspect about the graduate curriculum that is fundamentally symbiotic to scholarly output?

The robust results of the library expenditures variable should also benefit from the substance of the interviews. From the start, the inclusion of that variable was based on two separate – though not necessarily competing – assumptions. The first assumption is that a more comprehensive and well-organized collection of engineering-related library materials might help researchers accomplish their literature reviews in a timelier and more effective manner and, perhaps, may even offer them direct support in the lab through the provision online reference materials. The second assumption, which is based in part on the fact that overall library expenditures do not offer any direct indication of the quality of the collections related to engineering, is that the amount that a university invests in its library may be a reflection of it overall commitment, emphasis, and capacity to support research generally. Because the library expenditures variable produced a standardized coefficient that was essentially as strong as research expenditures and graduate assistants, it is hard to imagine that the quality of the collections alone could have produced such a large impact. This lends more credence to the theory of library expenditure as a surrogate for overall institutional support, but the interview responses should help flesh the issue out further because it asks about other institutional-level research support, such as the efficacy of the Office of Research or availability and usefulness of core research facilities.

Qualitative Analysis

The qualitative analysis begins by following the format of the interview guide, which divided the questions between those relating to research support services and resources, on one hand, and those related to department climate and workplace culture on the other. The results for each section are discussed below, largely using the interview guide's questions as topical section headings. Then, moving on from the structure of the interview guide, the overarching themes that emerged in the course of analysis are discussed. The preponderance of substantive information regarding how institutional factors influence research productivity came decidedly from the research support services and resources section. While climate appeared to be an issue at one of the two low productive efficiency departments, the disparity of results concerning climate were less clear and less cleanly split between high and low productive efficiency departments. Furthermore, the qualitative data and analysis offered considerable insight into the questions raised by the quantitative model regarding the surprisingly weak negative correlation between faculty-to-student ratios and scholarly output as well as the surprisingly strong weighting given to graduate assistants and library expenditures.

Most sections begin with two or more quotes that typify the responses related to the topics they covered. Where relevant, there are counts of the number of respondents from each type of institution who replied positively or negatively to certain topics. For example, if a particular section were to note that 7 of 9 respondents from a high productive efficiency environment offered positive comments regarding that particular topic, it does not necessarily mean that there were only seven positive quotes related to that topic, because each respondent could have offered multiple points of data on the topic. Instead it is just to serve as a head count for the purpose of providing a general sense of how researchers in each type of environment reacted to a particular topic or question.

The high productive efficiency programs are identified throughout the study as institutions A and B. The low productive efficiency programs are identified as institutions C and D.

Research Support Services and Resources

Each heading in this section represents research support services and resources that BME researchers were asked about specifically in the interview guide. Additionally, at the beginning and end of the interview guide, participants also had the opportunity to provide open comments on resource or service factors they believe influenced their productivity either positively or negatively. The data collected from those responses did not reveal any additional categories to be added beyond what was already covered by the specific questions in the interview guide; therefore, the data elicited from open questions were simply incorporated under the existing headings. However, it is worth noting that allowing respondents to list and discuss the resources and services they felt were most important, prior to be prodded by questions seeking details on a specific topic, provided a better sense of what issues the researcher felt most impacted their success.

Graduate Research Assistants

"They are the ones who are doing everything. From any sort of effort in terms of collection and analysis to dissemination of the work."

-*Assistant professor, high productive efficiency institution*

High productive efficiency programs: 9 of 9 respondents offered positive comments.

Low productive efficiency programs: 7 of 7 respondents offered positive comments.

Of the results of the quantitative analysis, nothing was more curious than the finding that graduate research assistants were more strongly associated with scholarly output for BME programs than grant funding. Of the results of the qualitative analysis, however, no finding was more consistent across cases than how prominently graduate students fit into the process of conducting research and producing scholarship. Although grant funding is unquestionably vital to facilitating research, according to the respondents, graduate assistants do almost all of the work. Simply put, faculty researchers are not the primary labor component of scholarly production in biomedical engineering. Graduate students – and to a lesser degree postdocs and, to an even lesser degree, undergraduate assistants – run the experiments, collect and analyze the data and write the papers. All of this is done with varying degrees of guidance from the BME faculty member, whose primary role in the enterprise is to find new or continued grant funding. One professor summarized it as “my job is to bring in the money and their job is to carry out the work.” In keeping with the production function that served as the basis of the quantitative model, students are the labor function and the research funding, space, and equipment are the capital. Presumably this leaves the BME faculty member to serve the executive functions of the enterprise - overseeing the production process and attracting investment.

While the interviews elicited a wealth of data on the importance of graduate students, it yielded no clear suggestion of a difference in regard to the experiences faced by either low or high productive efficiency environments. Instead, the discussion surrounding graduate research assistants were remarkably similar for both types of programs. A typical account of how they fit into the research process is explained by a professor at institution A: “They’re the ones that actually do the bench work in the laboratory. Also, in a lot of projects I’ll have a clinical collaborator and they [GRA’s] serve, really, as the bridge between myself – the scientist – and the clinical collaborator. And they do that quite effectively. I would say that I mostly help with conception of ideas as well as dissemination of those ideas. Making sure they are well communicated. But everything else is them.” Similarly, a senior researcher at institution C indicated: “They run all of the experiments for me. We might troubleshoot some stuff together, but they are critical.” These sorts of comment are repeated throughout the data.

Across all cases, researchers also offered a near uniform description of the role that graduate research assistants play in the writing and publishing of journal articles. An assistant professor at university A briefly describes how her graduate research assistants fit into the publication process in a way that typifies what most respondents had to say: “They are part of the entire process. So they write a draft of the different parts, and it is true that I end up re-writing a lot of that. I try to go over the changes with them so that there is some learning involved and so that they can do better the next time.” An associate professor at institution B explained how the process actually initiates: “In terms of manuscript preparation, the expectation is that the student who is the primary intellectual contributor and/or the primary data generator takes a lead role in the actual writing of the manuscript and oversees its submission to an appropriate journal.” To be clear, the graduate research assistants receive full credit for their contributions. Several researchers indicated that the common practice is for students to be recognized as the first author on any papers resulting from his or her projects as well as secondary authorships for other students’ projects for which the student may provide significant assistance.

While the role that graduate research assistants play in conducting research and publishing findings does not offer any sort of obvious distinction between high and low productive research environments, the manner in which it was described by the researchers does help make sense of the quantitative results – they are equally as important to the production of output as the grant itself. It is a chicken-and-egg type of relationship, where one is not possible without the other. There could be a difference in the quality of the graduate assistants from high to low productive efficiency programs, which would presumably translate into the publication of more articles in higher impact journals. However, it is impossible to determine from either the qualitative or quantitative data if this is the case. In terms of the qualitative data, each of the researchers were probed about the quality of their students, but none offered anything that could be characterized as a negative assessment of the students' quality. In terms of the quantitative data, of course, one graduate research assistant carries the exact same weight as the next, regardless of ability.

Quality of Grant Support

"I think it is fantastic. I just tell them that I am going to write this grant and, if I know in advance that it is going to be due in a month or two months, then they will have it in their calendar and they will know what to do."

-*Assistant professor, high productive efficiency institution*

"It is the grant obstruction process, rather than the grant assistance process."

-*Professor, low productive efficiency institution*

Quality of pre- and post-award grant management was the single most obvious difference between high and low productive efficiency programs. And while high productive efficiency programs seem to have found better, more proactive ways of supporting their faculty throughout the grant process, it is a topic that elicited negative comments from every respondent that was interviewed, regardless of institution type. Pre-award grant activities include locating funding opportunities, writing proposals, assembling the supporting documentation (e.g. budget statements, investigator CVs, etc.) and the actual submission of the grant application package. Post award grant activities include budgeting, accounting, procuring, reporting and the closing out of research grants.

Pre-Award Support

> High productive efficiency programs: 3 of 9 respondents offered negative comments, all moderate in severity
>
> Low productive efficiency programs: 7 of 7 respondents offered negative comments, 4 high and 2 moderate in severity

A theme that emerged clearly from the data is that faculty contentment with the pre-award process is heavily tied to the extent to which their home department or school chooses to be earnestly engaged in that process. At one end of the spectrum, some departments generally leave it to the faculty researchers themselves to work directly with the university's office of sponsored research to get most grant proposals submitted to funders. At institution C, for instance, an assistant professor indicated that "how you do the budget is you go to the Office of Sponsored Programs and you go back and forth with them, but for all of the other stuff like bio-sketches, facilities descriptions, the aims…that's all on you or you might get peers to look at it." One of his senior colleagues acknowledged that it is an area that needs improvement. "We talked about getting someone who can help you with those forms – not the whole grant – but some pieces of paper," he explained. A major consequence of providing little or limited pre-award support can be the length of time it takes to get a grant submitted. "I have to be careful that I don't sound too negative, in the sense that I think that the university really does try to be supportive in identifying funding opportunities for faculty members," explained the same senior researcher, "but the reality of the situation is that the faculty members almost have to know the opportunities that are coming down the pipeline before they are ever publically announced [or] they most likely aren't going to be in a position to present a competitive proposal – particularly for major activities." Viewing the pre-award process as something "we are weak on," another senior researcher's strategy was to rely on the medical school of a separate institution with which institution C frequently collaborates. "I submitted center grant with [the partner institution] and I let them lead because they have the infrastructure that was so much more supportive in terms of getting it submitted," he explained, "they can pull it all together, make it synergistic, and make sure everything lines up." In comparison, he noted that "our sponsored program office will help us…things like they will review your budget justification," he explained, "you have to write your budget justification, but they will review it and tell you if you have any typos or errors…but they are not going to read through any of your other documents." In characterizing the quality of department-level pre-award support at institution D, a senior researcher and research administrator described it as such: "Department-wise there is the standard sort of helping you with the budget sort of thing, but nothing exceptional…nothing that I would right a testimonial about." This theme is not limited to low productive efficiency environments. An assistant professor at institution A also referenced pre-award challenges she faces related to assistance in assembling and reviewing grant submissions. "My main complaint is that the administrative people are not willing to review or to help you with those documents (budget, etc.) until after you have a complete proposal," she explained. "That is kind of bad, because the science part always takes you longer to figure out," she reasoned, "so in my mind it would be easier if I could just say: 'Here are all of the administrative documents that need to be filled out, while I work on my part at the same time.'" The associate dean for research in her engineering school indicated "[t]here is not as much effort on the pre-award side at the department level, typically at our school of engineering, that is centralized more at the school level." Nonetheless, he explained that "it is not very good" and that they "need to upgrade."

The other high productive efficiency program at institution B, however, does largely handle the ancillary proposal requirements that are at the heart of pre-award support frustrations expressed by faculty at the other institutions. As an assistant professor there explained:

> "Like NIH awards, for example, they will send me templates for what the budget is supposed to look like and I just fill in the gaps for the needs of my individual research. They will do the same for personnel – typical language that is used for describing personnel. [They handle] those additional parts which don't really deal with the research plan, but that are more just functional for the NIH or the NSF or others to see 'okay, who is going to be here? So being a new, fairly new, professor it is good for me to have these things. It also helps make it easier and just less time consuming for me. So that it pretty well done."

As a result of this type of support, institution B researchers registered only positive comments about the pre-award process. "The pre-award grant application and administration is important," an associate professor at institution B explained, "so that is someone who collects bio sketches, reads the actual call for proposals, who assures that – for whatever we are applying for – the application is complete and consistent with the application instructions, shuttles the application through the office of sponsored programs or sponsored research in a timely and appropriate way…that kind of thing." Other researchers at institution B offered similar descriptions of the department lead grant support. According to a senior faculty member "locally in the department the grant support and the administration is very good, but I think at the institutional level it has gone through challenges." He elaborated: "I think that is mainly due to personnel issues. You know, sometimes you have good personnel and they leave and they may not be replaced by someone who is as good. And that lead to an unevenness in the quality of support."

The experiences of each of the institutions suggest that the school or department cannot expect to rely on solid pre-award support from central units that are not solely responsible or invested in engineering alone and that, even with best intentions, are likely be too overstretched to provide much in the way of individual attention.

Post-Award Support

> <u>High productive efficiency programs</u>: 1 of 9 respondents offered negative comments, 4 of 9 offered positive comments
> <u>Low productive efficiency programs</u>: 5 of 7 respondents offered negative comments, 1 of 7 offered positive comments

"We know not only how much have we spent, but what it is going to look like a year out or two years out. So they are really good at helping us manage our money so that we don't end up in hole a couple three years from now."

-Professor, High productive efficiency program

"The funder will call me and be annoyed and say: 'I am just trying to give you guys some money, how hard is that?'"

-*Professor, Low productive efficiency program*

Comments about post-award grant support centered on establishing the grant in the university's accounting system once the award was announced, the budgeting and accounting of the funding over the life of the grant, and, to a much lesser degree, procurement of equipment, supplies and services. Positive comments centered on the degree to which these responsibilities were well-handled. Negative comments centered on how the mishandling of these responsibilities could lead to significant problems that distract from research and impede progress.

One 'war story', offered by an assistant professor at institution C, illustrates how the pre- to post-award hand off – a process which seems like it should be routine, but which apparently causes trouble for most institutions – can have palpably negative impact on research productivity:

> "One grant took a full year from being accepted to getting the money. It took probably multiple hours working with multiple people in my department to figure out what codes to use for this or that, while simultaneously taking multiple tens of hours sending emails and making phone calls to the [funder], trying to figure out what was going on and what can I do to expedite this. If you put all of the time it took from getting the award to getting the funds – like all of the hours that I put in between those, because I hadn't had the funds just given to me – we are talking like probably in excess of 25-30 hours of lost time. That's not trivial. If I didn't put my nose in with the [funder], it was just would have stayed stuck. But even more than that, even if I wasn't poking and just waited, most of the time was not doing the poking for the money, it was doing the shuffling for the students I was planning on paying with these funds. I had to figure out how to put all of that on to a consortium of other funds. And then, once the money did come, I had to reimburse all of that. It was just a mess. And each of those decisions took talking to multiple people over iterations to get ironed out. So that was really frustrating. It was at least a week of my entire year lost to just managing that situation – maybe even two weeks."

The issues and potential potholes surrounding the transition to post award was a theme mentioned by each of the research administrators interviewed. "The baton-pass is badly done at most schools that I know about," explained an administrator at institution D, "and it is something that can really get in the way of the work, because it is right as you are trying to start the work that you are not getting the forward movement that you need - so it is an area that is a common stumbling block." Comparing experiences between engineering and medicine, another researcher at institution D who has dual appointments indicated that the medical school has a strategy for dealing with this problem that he would like to see employed by engineering. "It is absolutely critical, that you have a pre-award office and post-award management to grants that are integrated together. In other words, staff that both know how to help you submit your grants and, at the same time, those same staff can actually help you with post-award budgeting and management of those same grants," he indicated, adding, "I would be telling BME they need to do that if that was my primary place." A professor at institution C indicates that he routinely encounters an inverted version of this problem whereby when he is re-awarded the same grant from the same funder he encounters a lack of continuity:

> "I feel like I have to start all over again. I have to jump through all these hoops, and then I get another grant from them and it like there is no history – and you would think that it would be easier because you have gotten this one before – but it is like 'no, you have start over new,' he explained, adding "it is very frustrating [because] the funder will call me and be annoyed and say: 'I am just trying to give you guys some money, how hard is that?'"

The problem of transition from pre- and post-award is not an issue that came up in any of the interviews with researchers at high productive efficiency programs.

The single negative comment on post-award grant support originating from a high productive efficiency program came from an associate professor at institution A and was related to the financial tracking of the grant. "In our department we don't get a lot of updated information unless we specifically request it," he explained, "I know that's something that they are trying to achieve, but you know, I think it is rarely accomplished." Researchers at the low productive efficiency programs also mentioned similar issues. According to each of the research administrators interviewed, a major challenge in this area is that university accounting systems are largely incongruent with post-award tracking and reporting needs, making it difficult for financial officers and principle investigators alike to have confidence in budget projections or even available balances. Regardless of whether the issue was related to accounting, reporting, or other post-award administrative necessities, the general consensus of post-award concerns was that too much of the burden fell on the researcher and not the administrative infrastructure of the department or the university. As a researcher at institution D summarized the situation: "Post-award grant management has been good, but it varies of course, even from subgroup to subgroup from within the department. So for me, it has been adequate, but some of the work has been supported by people from within my lab rather than people within the department. And I think it is something that should be expanded and better supported."

Research Facilities

The interview guide asked researchers specifically about the adequacy of the research space provided by their departments or affiliated university institutes. The responses suggest that, like graduate research assistants, this is also not an area of significant difference between high and low productive efficiency programs. Most BME researchers interviewed indicated that they had the space and foundational equipment that they needed to be productive, with the exception of institution C. While most of the researchers there expressed that they were content with their current space, there were a few mentions of complications.

A senior researcher at institution C who feels "fortunate" that he currently has adequate space, characterized his university as "a space-starved institution." He explained that "we have a number of older buildings that we continue to use for research purposes, so as a result, they are serving functions that they were not necessarily designed to serve." This limited lab space initially affected an assistant professor at institution C. "I didn't [have adequate space] the first three years," he explained, "they gave me one lab, but I grew pretty fast for a new professor. At one point I had five graduate students and 12 undergrads, so 17 people who were all working in this one small to midsize lab space. It was just insane." He applied for space in an adjacent lab when it became vacant, but was turned down by the department because he did not "have the stature to justify" the additional space as he had yet to secure significant research funding.

Though outside of the department's control, the situation suggests an inability to properly invest in the success of a junior faculty member. Certainly, the situation represented a hindrance to his productivity potential and, presumably, to other colleagues of his who may have faced similar space challenges. The senior researcher did indicate that the university recognizes the problem and has made moves to improve the situation. He currently resides in a research institute that has grown considerably in the past few years where he enjoys access to the space he needs and a variety of shared equipment laboratories.

Core Lab Facilities

"I think that [at my institution] we tend to have a very archaic core facility structure where some tired, burnt out PhD who never quite made it as a faculty member is looking after a confocal microscope that's five years old. That can cause real varying experiences for people who need to get things done."

-*Professor, Low productive efficiency institution*

"I think that at the institutional level we have done much better recently at providing shared laboratory sorts of equipment. In the past were really didn't have institutionally supported technicians. So while we might have pieces of equipment distributed across the campus, people were reluctant to let graduate students outside of their research groups use the equipment out of fear that the equipment could become damaged if they weren't adequately trained and so on. So I think having the technical support is essential."

-*Professor, Low productive efficiency institution*

"So there are people who are running these facilities who know what they are doing, so it makes it easy."

-*Professor, High productive efficiency institution*

"Yes, so the staff that manages our microscopy course are especially helpful and…I would say, generally speaking, all of the core equipment staff persons are very, very helpful. And I would say that the quality of the animal care, the biosafety processes, and the training manuals that they provide are also quite useful."

- *Professor, High productive efficiency institution*

"There are all sorts of resources that we have available from imaging to animal, machine shops, manufacturing, nano-manufacturing, and even other resources that are used for teaching. So there is a lot and it is fairly flat ground and there aren't any significant hurdles to getting access to those spaces."

-*Assistant Professor, High productive efficiency institution*

High productive efficiency programs: 2 of 9 respondents offered negative comments, 6 of 9 offered positive comments
Low productive efficiency programs: 6 of 7 respondents offered negative comments, 3 of 7 offered positive comments

In addition to asking about the researcher's own lab space, the interview guide also included probes to specifically ask about their use of core laboratory facilities at their institution. By providing access to shared equipment and capabilities beyond what many of the researchers can expect to furnish in their own labs, cores represent an institutional factor with the potential to significantly impact productivity. The type and quality of core facilities can both set boundaries to the kind of research that can be accomplished as well as the time that it takes to complete important research tasks. Respondents were encouraged to indicate whether or not they had access to the types of spaces and equipment that they needed to conduct their research as well as to describe their experiences gaining access to such facilities and the quality of the equipment and technical support provided.

Researchers from high productive efficiency programs offered a more positive picture of the cores generally, but researchers as both types of institutions had good things to say about their ability to access the equipment and spaces they needed. The substantive difference between the two groups centered on the quality of support and expertise that they and their students received in the use and handling of the complex machinery and systems that the cores offer.

In terms of gaining access to the types of facilities and equipment needed, even though both types of programs were generally positive, low productive efficiency programs respondents did mention some challenges, whereas researchers at high productive efficiency programs mentioned none. At institution C, for instance, a researcher uses a clean room at non-university facility that is a two-hour drive away. His university does offer a clean room, but he is able to use the other one for no cost. All of the institutions studied typically charge researchers' grants for the use of core facilities, so that is not it unusual. It could be the researcher is just being thrifty, but, generally speaking, it is not an efficient use of his time. The experience of one of his departmental colleagues is perhaps more troublesome. Part of his research involves the sorting of human cells, a task for which he relied on access to a particular core facility. An administrative decision was made, without warning or consultation, that the facility would only allow animal cell research moving forward. He has an active grant in place that requires the sorting of human cell tissue, so he was left to determine how to move forward on his own.

> "We don't have a good solution yet," he said. "We can go to [a city that is a 45-minute drive away] and do some work down there to sort the cell, but if you are trying to do the experiments on your scope and you've got everything set up, then you don't want to go and sort the cell and bring them over to do your work [somewhere else] because the cells are…it's just not a good answer. It is our best answer right now, but we are still working on it."

So in this case, the institution has the capacity needed, but the researcher clearly does not have reasonable access to the tools he needs to conduct his research.

Regarding the core facilities used by the BME faculty at institution D, a research administrator and active BME researcher there indicated that "the tools and the amount of space is not as great as it needs to be – just because the need is growing and starting to exceed capacity." In this case, however, his institution proved to be responsive to the situation and is about to open a new facility to meet the need increasing demand. What continues to concern him, however, is that a better support structure is needed in terms of expertise to really leverage the capacity of the core facilities. "I guess the issue for me, and probably more for other users beside me, is that the cores are more than just equipment, the cores are expertise," he explained, "and so to be able to use the equipment right and have that sort of personal handholding …I think is very important." This is the aspect that seemed to divide high and low productive efficiency programs. One of his colleagues, who has served in the role of dean at other institutions as well as president of a major, stand-alone research center, characterized the current management of core facilities at his institution as one where "the research office is more focused on the system it put it place for recharge [payment by researchers for use of the facilities], than it is on keeping the facility cutting edge." By this, he explained that he means support, more so than equipment or facilities, by offering that "if you want a core facility to be really cutting edge, really enabling, keeping up with fast moving changes in that particular technology area, then you are going to need a faculty member [who is a domain expert] to sort of oversee that." This vision, if it is realized, would probably exceed the current quality of support described by the high productive efficiency programs. But in a blunt assessment of the current quality of core facility management and support at institution D, a senior researcher there lamented that "it is just the state of the facilities and the responsibilities, it is just complete lack of management…the kind of thing where you say 'I am paying 65% indirect costs' and you shake your head and say 'what in the hell for.'"

By comparison, of the nine researchers and research administrators at high productive efficiency programs, one offered a minor complaint about the online reservation systems for a particular core, and a researcher at the other institution suggested that is sometime took too long for damaged equipment to be repaired. No other negative comments were registered.

Library Resources

"Primarily it is helping me have access to literature that is relevant to my research."
-Associate professor, High productive efficiency program

"To be honest, it is just [about] having the proper online subscriptions so that when I look online I can get the actual pdf's of the papers."
-Assistant professor, Low productive efficiency program

"I always tell me students that two years of research will help save a half an hour in the library."
-Professor, Low productive efficiency program

Another of the surprises of the quantitative model was that the variable for the total library expenditures produced a beta weight that was roughly equal to BME research expenditures in its correlation to scholarly output. As such, one of the goals of the qualitative study was to develop a better understanding of what was driving these results. Presumably, every research project touches on library resources at some point in the research cycle and, if nothing else, the literature review is a crucial component of that cycle. At the same time, it is difficult to conceive of those resources impacting the scholarly output of a single department on the same level of magnitude as research funding or graduate research assistants. To help better understand what was behind these results, each researcher was asked to describe the role that library resources play in aiding their research productivity. They were also probed specifically whether library resources played any sort of direct role in the lab itself. The objective was to find out if they were using library resources for anything beyond the literature search itself.

A handful of questions on the interview guide tended to elicit uniform responses across all respondents, and the library question was one of these. The first two quotes at the start of this section summarize the extent researchers felt that library resources impacted their research. Largely, they just want easy online access to the journals and articles necessary for their work. Some expressed appreciation for the library's efforts in developing their students' research skills, relating to locating journal articles, while others questioned the library's reason for existence beyond paying for the subscriptions and making the articles available electronically.

With the exception of institution B, all researchers felt that they had adequate and current access to the journal titles that they required. Three of the five researchers at institution B, however, raised mild concerns about the breadth of titles to which their library subscribed, with the consensus being "good, but not great." No one characterized this as an impediment or significant barrier to the productivity. Furthermore, as a high productive efficiency program, it seems doubtful that whatever is lacking in the library's journal offerings had a significant negative impact on research productivity. As one of the researchers, who complained about the issue of title availability, said: "I think as being part of the BME program which is partnered with [a neighboring institution's medical school] we have access to their library – their digital library – but I haven't tried to fully explore that avenue yet." Logic would suggest that if library resources were capable of impacting his research productivity on a scale nearing that of research expenditures or graduate assistants, he would have explored what the other institution had to offer.

The general uniformity of the qualitative data collected on this topic indicates that the quantitative results of the library expenditures variable must be serving, at least partially, to measure the effect of something beyond the library itself in impacting scholarly output. It is reasonable to suggest, as did Dundar and Lewis (1998), that library expenditures provide an indirect measure of a university's investment in its own research enterprise and infrastructure. Furthermore, it might also be a reflection of a university's wealth, which by extension might represent its capacity to invest in its research infrastructure. Are high levels of investment in the library correlated to high levels of investment in the Office of Research or other centralized components of a university's research infrastructure? Do more librarians and subscriptions suggest more office of sponsored program support personnel? While the actual driver of this variable remains unclear, the qualitative data – something that Dundar and Lewis' (1998) study did not include – provides enough information to conclude that it is not just a measure of the library's direct impact. If library resources were impacting research output on the same scale as research funding and graduate research assistants, it would not make sense that none of the researchers interviewed brought up the topic of the library until they were specifically asked about it. Conversely, when answering the opening questions on the interview guide, which asked respondents to list out the factors that aided or hindered their research productivity without any specific prompts, two of the most frequently mentioned topics were grants and research assistants.

Department Climate and Research Environment

The impact of department climate on research productivity or efficiency was much less clear than that of research resources or support services. Many of the themes are common across both high and low productive efficiency departments, and even were there were disparities between the two, it was difficult to discern a path from those disparities that led clearly to scholarly output.

General Departmental Climate

"Oh, we have a great environment, with a high level of esprit de corps, lots of enthusiastic people at every level from junior to senior with a willingness to work together."
-*Professor, Low productive efficiency program*

"I would say [morale] it is quite high. That's one of the reasons that I continue to stay here. One of the reasons that I was attracted to starting my career here, because I knew about the climate here." -*Associate professor, High productive efficiency program*

Questions in this section of the interview guide were intended to develop an understanding of how researchers in both environments viewed their departments' group identity, the level of participative governance, and the nature of interrelations between department personnel. In response to the question about morale and interrelations, all of the respondents expressed a positive view of their colleagues and suggested that interrelations were not a problem. The terms 'collegial' and 'collegiality' were most frequently used to describe department atmosphere. Any difference between high and low productive efficiency programs was minor to non-existent.

There was also a high level of commonality in the views expressed about participative departmental governance. It should be noted that by participative governance, most respondents took it to mean taking part in the recruitment and hiring process – both in terming what types of position to hire and in selecting the candidates. A BME researcher and research administrator at a high productive efficiency program offered a description that typifies many of the responses: "They are welcomed to be part of the decision process," he explained, "but, as you know, a lot of faculty don't care if it doesn't relate directly to them." This sentiment is repeated throughout the data, emphasizing that faculty have the ability to engage in decision making to the extent that they choose to do so.

Some variations on this theme were present, however, which did suggest that a disparity exists between two of the programs – one high and one low productive efficiency departments each – in terms of general climate. Respondents from Institution B, which has the largest number of tenure or tenure-track faculty of any of the programs studied, typically expressed a relatively high level of active faculty engagement. For example, a BME researcher and associate dean for research commented that: "The department has a well-attended annual retreat…to distill where we think our research needs are and why. Last year attendance was over 90%, which is phenomenal." While most of the respondents across programs were generally positive in describing their colleagues and departmental work environment, those from institution B offered much more elaborate descriptions of what made their work environment great. Characterizing what made his department unique, an associate professor at B said:

> “[R]egardless of what our self-perception is like in regard to critical mass or national or international competitiveness – there is a – whatever this quality is, I am calling it entrepreneurship – there is a creative energy and a grit that characterizes the thinking about and the execution on big ideas and opportunities that I think pays off.”

In terms of thinking big and taking on high profile projects, the associate dean for research suggested that “the BME faculty are aware that we are considered a leader amongst our peers… they kind of know that other schools look to us for how we do things and where some of the most cutting-edge research is.” In expanding on institution B’s group identity, he added: “I don’t know how we got here, but we have some of the lowest institutional barriers to collaborate compared to any of our peers. And people who spend any amount of time here comment on that.”

A senior researcher at institution C, on the other hand, described departmental governance as follows: “[My institution] is different from other institutions, where a lot of other institutions it is more like everyone votes on everything and [mine] is more top down. It is more like what it is at a company.” A significant issue at institution C was created when two, somewhat disparate, departments were merged a few years ago. The merger happened without any consultation or advance warning to the faculty. A senior faculty member – who heard of the merger from an email from a colleague at another institution – commented on the awkwardness of the fit between the two programs. “Some people who weren’t doing anything [BME] are saying ‘hey, do I fit anymore?’ And the people in [BME] are saying, ‘hey, I am not really a [other discipline] person per se, do I fit in?’” he explained, in a sentiment that was expressed by other researchers from institution C. While each of the respondents indicated that interrelations between faculty members in the two departments is not an issue, there is the suggestion that this move – aside from violating shared governance norms – clearly offers lower opportunities for inter-department collaboration because the two disciplines are too far apart from one another.

Furthermore, the sort of half-baked nature of a department formed by fitting two disparate disciplines together presents a recruiting challenge as well. Researchers from either discipline likely find it confusing and less ideal of a fit than other institutions. As an assistant professor indicated: "For our identity or long term viability, I don't think anybody can communicate their way out of inherent structural problems we have with the department."

Climate for Research

"I think the focus is on research. I think being at a public institution with somewhere in the realm of one thousand undergraduate students, some will tell you that we are also very focused on teaching. We care. And we have people who really, really care about teaching who are teaching faculty. But from the research faculty who make up the core of the department, our research is pretty much our lives."

-*Assistant professor, High productivity efficiency program*

"We strongly emphasis research. It aspires and it has been told to aspire to have a high level of ranking in the research environment."

-*Professor. Low productive efficiency program*

The interview guide intended to measure each program's climate for research with questions that asked how strongly the department emphasizes research over other priorities; if the department set collective goals; and if the researchers felt the productivity of their peers impacted their own productivity. As interviews progressed, additional questions were added to ask if and how the department measure research success and what types of research expectations were communicated to the faculty.

Unsurprisingly, no one suggested that any other priority was ahead of research in terms of importance. Some respondents took the opportunity to express the importance of their education mission, but only suggested that education was equally important in emphasis to research. So while little was learned with this particular question, it is worth noting for the record that none of the respondents – particularly those that were measured as low productive efficiency – suggested that research output of their program might be lower because the department is more concerned with its teaching mission. Researchers also suggested that, for the most part, their own productivity was not influence by that of their peers. In most cases, the respondents indicated that they were self-motivated and that competitive pressures were not a factor in their behavior.

Similarly, no respondents indicated that their departments set collective research goals (e.g. number of publication, total research grants, achieving a particular ranking, etc.), except to the extent that some researchers interpreted the collective goal question broadly enough to include the recruitment of new faculty into existing research clusters. This was not the type of information that the question was designed to elicit – it was more concerned with whether benchmarks were being set – but this information was useful for informing other sections of the study. In regard to benchmarks or other collective goals, the consensus was that faculty are individually motivated and individually judged on their performance through evaluations and the promotion and tenure process. Furthermore, other suggested that BME is too broad across sub-disciplines to set quantitative standards. This raises concerns about the validity of the quantitative phase of this study, which are addressed in Chapter 5.

Once it was clear that the questions surrounding research climate were yielding repetitive, non-substantive responses with each successive interview, the questions were revised to ask more about how research success is quantified by their programs, and whether or not faculty members faced clear expectations about what they should achieve. These questions sparked a good deal of discussion about quantitative and qualitative merits of assessing productivity that were helpful in developing a better understanding of the BME environment. It is questionable whether enough data was collected on this topic to draw firm conclusions about whether or not clear differences exist between high and low productive efficiency programs in respect to their approaches to measuring research success. However, the respondents at high productive efficiency programs all indicated that the administrators in their programs (keeping in mind that some of the respondents were those administrators themselves) understood the challenge of relying on quantitative measures of success. As a senior researcher at institution A explained: “I think that it will be something that chairs will ultimately have to make the qualitative decisions. You know, I don't think we'll get rid of those metrics immediately, but now the chairs are trying to rank faculty within their departments more broadly, rather than relying on these more traditional academic measures.” Likewise, an assistant professor at institution B acknowledged that, “there is no one single outline that works for everybody, and everybody [here] is very much aware of that situation.” Contrastingly, an assistant professor at institution C characterized his view of his department’s expectations as follows: “The sense I am getting is that as long as you get your $1M and you are graduating people, that’s what we are looking for. I don’t think they care as much if I had put out 10 passable papers as opposed to 10 outstanding papers. It seems like that is just a bonus. It seems like legitimately innovative scholarship is like a bonus, but if you can get the money and graduate the students and just publish so-so papers, that is generally enough.” One of his senior colleagues elaborated further on how research success is discussed at institution C: “The deans will say: ‘We want to increase our research revenue and our trajectory and our number of PhD students,’ and all this stuff, but then they just show you the graph. They don’t show you what makes those numbers go up.” Again, the data surrounding research metrics and success measurement are limited, but it is worth noting the high productive efficiency programs respondents who were asked all noted that their departments were aware of the intricacies surrounding the task, and none of them offered anything like the two respondents from institution C.

Leadership

While a good deal of the literature suggests that quality of department leadership can impact the productivity of a department, the three questions related to leadership drew very little useful information for the purposes of this study – except to the extent that the data collected suggests that this is not a distinguishing factor between high and low productive efficiency programs in this area. Based largely on Bland et al.'s (2005) study on research on productive departments, researchers were asked whether they viewed the chair as an accomplished researcher, if the chair uses his/her expertise to support the productivity of other researchers, and if the chair communicates an overall vision for the department in terms of research goals. These questions elicited some of the briefest and least substantive responses of any section of the interview guide. Everyone felt their chair was an accomplished researcher, but there was very little evidence to suggest that chairs either contribute directly to the productivity of others or that they communicate concrete research goals for the department as a whole. Often when the question of the chair's contributions towards productivity of others came up, the common response was that BME was too broad of a discipline for a chair to be able to help others directly. Likewise, the vastness of BME was commonly cited as a reason why collective or department research goals did not make much sense.

Emergent Themes

After analyzing the data collected from the interviews topically, it was clear that a few transcendent themes spanned across the topics structured by the interview guide. First, despite the literature's emphasis on teaching loads as the primary factor negatively impacting research productivity, none of the respondents described their teaching responsibilities as impeding on their research. In terms of impediments, they were much more inclined to identify general administrative burdens as getting in the way of their progress in terms of research. Finally, the importance of collaboration and challenges and barriers to facilitating interdisciplinary research came up frequently in responses to a variety of the questions contained in the interview guide.

Teaching Load as a Non-finding

The literature surrounding research productivity focuses largely on teaching loads and, to a lesser degree, service. These other two legs of the academic stool are often represented as the primary competitors for the faculty member's time which can often confound their productivity as a researcher. The interview guide contained multiple questions that provided faculty with an opportunity to indicate the extent to which they believed that their teaching or service responsibilities were impeding on their productivity as researchers. These included:

- Do you find that you have sufficient uninterrupted time to dedicate to scholarly activities?
- What do you consider to be the biggest barriers to your productivity?
- How would you describe your department's emphasis on research, compared to other priorities?
- Is there anything left that we have not discussed that you think is relevant to your productivity – either negatively or positively?

Although none of these questions ask specifically about teaching or service loads, the expectation was that they would elicit responses related to those topics to the extent that the respondents viewed teaching loads or service as significant impediments. Only one respondent, an associate professor at institution B, brought up the issue of teaching loads as an impediment to his research productivity, describing it as follows:

> "I don't want to use the word barrier – but we have [a very large] undergraduate population compared to our peers. So that obviously puts a level of pressure on our faculty to be able to teach at a high level in a way that maintains a reasonable student-to-faculty ratio. So maintaining the administrative support for our teaching mission in a way that doesn't over burden our commitment to scholarship and innovation can be a challenge when budgets at institution of higher education are strained."

In response to the question regarding 'sufficient time to dedicate to scholarly activities' some respondents did acknowledge that teaching represented a commitment of their time, but no else one characterized their teaching loads a significant impediment or source of frustration as might have been anticipated. The response of a full professor and former research administrator at a high productive efficiency program was more typical of the types of descriptions that were offered on the subject to teaching:

> "I would not say it is teaching. I think at our institution our teaching loads are not onerous. I don't feel that they impede my productivity at all. I could see at other institutions how that could be a burden, because there is only so much time in the day."

Despite that last sentence, none of researchers from other institutions expressed this sentiment, aside from the lone negative comment above. This strongly suggests, as this quote implies, that the BME researchers in these programs do not generally face onerous teaching loads. In interpreting these results, it bears repeating that all of the respondents were tenure or tenure track faculty by design. No lecturers, term faculty, adjuncts, professors of practice, or others were included in the study – although each of these programs avail themselves of the services of these types of faculty, along with doctoral students, to assist in the delivery of instruction. Given that every respondent indicated that their department emphasizes research over other priorities and that BME is a particularly research-intensive field, it is not surprising if all of the programs recognize that ensuring a low teaching load is a fundamental feature for keeping their tenure and tenure track faculty viable as researchers. Nonetheless, the pre-analysis assumption was that subject of teaching loads would play a larger role than the data gathered suggests. To confound pre-analysis assumptions further, the one lone negative comment came from a high productive efficiency program. In contrast, multiple professors at one of the low productive efficiency programs counted the low teaching loads there as among the most significant aids to their productivity as researchers. These results offer further evidence as to why the enrollment variable failed to produce more significant results in the quantitative study. They also strongly suggest that enrollment is not key factor that distinguishes low productive efficiency programs from high ones.

Administrative Burden

"Administrative type stuff, I mean I know it is not science, but all of that administrative stuff that you need to deal with on a day-to-day basis..."
-Assistant professor, high productive efficiency institution

"[W]e should have proactive staff who are lowering the barrier for entry for faculty to get these things done in a diligent and efficient way."
-Professor, low productive efficiency institution

"I think, for me, that it is a kind of a general communication overwhelm, meaning too many inputs from too many people getting in the way of prioritizing actual research."
-Associate professor, high productive efficiency institution

"Basically, our bureaucracy, for a number of historical reasons, believes that rather than it being a service structure, that we the faculty are there to keep them in full employment."
-Professor, low productive efficiency institution

"Of course, we have to champion safe work environments and we have to take our responsibilities very seriously as the recipients of federal, state or other types of research awards, but I really think that [my institution] really falls down in helping their faculty manage their compliance burden."
-Professor, low productive efficiency institution

As the most common impediment cited by researchers, the administrative burden faced by faculty was present across both high and low productive efficiency programs. This often pertained to administrative duties arising from the research itself, but in other instances respondents complained about administrative responsibilities more broadly. While some focused on the tasks requiring attention themselves, most respondents focused on the extent and the quality of support they receive in dealing with general administrative burdens. Chief among these concerns was the support and attention received during both pre- and post-award grant management – discussed in detail above – but other common areas that were also mentioned include dealing with an array of compliance concerns, the often unresponsive bureaucracy of central university support units, and the ever-present deluge of emails requiring action. Most respondents seemed to accept that these types of burdens are inevitable, ubiquitous, and often primarily the responsibility of the researcher. No one appeared to harbor any unrealistic expectation that these burdens could be eliminated completely, but many of them expected better support and wondered why they did not have it. Researchers in the low productive efficiency programs clearly had more complaints, war stories, and generally negative comments regarding the adequacy of the administrative support, but the theme also manifested itself in the high productive efficiency programs as well – just with less frequency and less intensity.

An example that illustrates how this plays out in a manner of degrees between high and low productive efficiency environments concerns procurement. An assistant professor in a high productive efficiency program listed procurement as the first thing that came to his mind as an organizational impediment to his research productivity. He characterized his experiences as follows: "All of the red tape around ordering equipment has to be done through this one vendor-approved website – and while it is pretty easy to use – if the vendors that I want to purchase things from are not listed, then we just have to go through this longer process." Addressing essentially the same issue, a professor in low productive efficiency program choose to describe the situation in harsher terms: "Purchasing has a vendor management system that is so hostile that some of our key suppliers have refused to use it. I am unable to obtain certain key supplies." While this single example could be interpreted as potentially saying more about the differences in disposition of the two respondents than differences in the realities of their surroundings, the greater severity of this type of problem in low productive efficiency programs was a theme that was reinforced frequently. The same pattern can be found in the discussion of administrative support personnel assigned to faculty members, which in all four programs consisted of a single assistant assigned to a handful of faculty members. Researchers at both ends of the program spectrum cited an unevenness in the quality of support provided from one assistant to the next. An associate professor at a high productive efficiency program characterized the inconsistency by noting that "although the department provides administrative support, the kind of level or the quality of that support and the appropriateness of the support background can vary a lot. I have had administrative assistants that have been incredible to administrative assistants that don't necessarily have the skillset or experience to help me." But again, faculty in low productive efficiency programs expressed more strident concerns. A professor at Institution C indicated that "I think one area that I would say is not a positive as it used to be: When I started as faculty member there was probably more support staff that were available to assist faculty members. Over time, in response to some of the state budget priorities shifting, as well as institutional priorities, we have reduced the number of staff members who were available to help faculty members." A junior colleague of his at the same institution expressed further frustration that the administrative assistant assigned to support him was located in a building across campus, making interaction inefficient. As a result, he seeks support for only routine matters like travel reimbursements, because "it just isn't worth the effort of walking over there" and, otherwise, to utilize the assistant more would effectively require more email correspondence, just increasing another administrative burden. In regard to the situation, he suggested that "I get the sense that there's more she can do [to help out]." Summarizing support services generally, a researcher at institution D indicated that "all of the services are necessary, but I can't think of any of them that are done particularly well."

Facilitating Collaboration

"That is the big elephant in the room. So I work with different medical schools around the country, which is what I have to do. But once you start needing patient samples or things like that, that's when it gets tricky."
-*Professor, Low productive efficiency program, when asked about his institution's lack of an established medical school nearby with which to collaborate.*

"The department helped a lot. And there is also an institute here that encompasses not only the people in bioengineering, but the people interested in bioengineering from mechanical engineering, electrical engineering, and other departments. And they have a connection with a number of clinicians. They also support a number of core facilities, so I made a number of connections through that institute. And in general, there are even some faculty in BME who have clinical appointments at [a nearby medical school] and I have made connections through them as well."
- *Professor, High productive efficiency program, when asked how he established collaborative connections after transferring to his current institution recently.*

From its name alone it is evident that biomedical engineering inherently involves interdisciplinary research and collaboration. The interview guide did not ask respondents about these topics specifically, but the guide's open questions offered an opportunity for researchers and administrators to bring these issues up on their own. Furthermore, when responding to specific questions about core facilities, grant support, or the role of graduate students, the topics of collaboration and interdisciplinary connections also arose because, in addition to aiding research productivity, these factors are also seen as potential catalysts in forging interdisciplinary relationships.

It was clear that every institution studied values collaboration – interdisciplinary as well as intra-departmental – and that they are all taking earnest steps in the attempt to facilitate it as best they can. Each program features some type of multidisciplinary workspace that combines people from across other areas of engineering as well as the sciences in a single setting. Each program also focuses heavily on the nature of their relationship with either their own institution's school of medicine – or the nearest practical partner in the event there is no medical program – with an eye toward maximizing interaction with clinical researchers in the attempt to foster more translational research. Accordingly, the difference between the high and low productive efficiency programs in this area does not appear to be a lack of emphasis or effort in fostering collaboration, but rather the low productivity programs just seem to face more obstacles.

In addition to facing the challenge to intra-departmental collaboration caused by merging incongruent departments together, as mentioned previously, institution C faces the greater challenge of not having a full-fledged medical school within its metropolitan area with which to collaborate. Instead, they have a relationship with another institution's medical school in a separate state that is a two-hour drive away. Although this relationship garnered some positive responses, it makes the department naturally more conducive for basic, less translational research. As one researcher indicated, "we have a lot of people doing more fundamental research [and] if you are doing fundamental research, like small animal studies or cell studies, it works out fine. But once you start needing patient samples or things like that, that's when it gets tricky." As an example of the challenge this presents to clinical efforts, he mentioned: "So, today one of my students is at [an out of state university, other than the partner institution] working with one doctor collecting some samples. That's a pain, right, because that is six hours away. Then two of my other students are at [another out-of-state] medical center to do some work about three hours away." So even if the department was able to overcome the bridge-building barrier created by distance, despite whatever collaborative synergy the faculty and students might be able foster, the research productivity is still hampered by the inefficiency of that distance.

Institution D seems to have a productive collaborative relationship with their medical school, but an area that they would like to improve is both the capacity and quality of their core facilities. Both research administrators interviewed there see the cores not only as important resources for facilitating the tasks of the researcher, but as vital cross-pollinators drawing a cross-section of their STEM-H community together in ways that would not happen otherwise. “One of our challenges is to make it so that people can interact in the most fruitful way,” one research administrator explained, “for me the thing that has turned out to be the best Rosetta Stone are the cores.” He explained that he helped foster this sort of environment at another top tier school and had been brought to institution D to do the same. He added “because the people within the cores are helpful, and the equipment is helpful,

then the tools and the people end up being great meeting grounds.” However, as each of the respondents from institution D mentioned above, currently the cores there do not have the capacity – either in terms of physical space and equipment or expertise – to leverage the sort of outcomes they are looking for. With new core space coming on line this year and leaders with the kind of perspective of the two BME researchers/administrators that were interviewed, institution D may be on the road to resolving these issues. Nonetheless, it is possible that the current situation helped contribute to their under-performance in terms of residual scholarly output detected in the quantitative model.

Neither institution A nor B appears to have the same impediments to collaboration faced by C and D. Respondents at both institutions suggest that they enjoy strong, synergistic relationships with their associated medical programs that are devoid of any significant hurdles. Likewise, both seem to have high-functioning cores. Additionally, the size of Institution B, with approximately 70 tenure and tenure track faculty members in the BME department, offers a scale that allows for richer intra-department collaboration than the other programs that were studied. As the associate dean for research explains, “all [of] these faculty [members] represent clusters of research strength that actively work together.” Other departments with 25 or 35 faculty members would obviously have a difficult time finding the same level of overlapping research interests while still being able to reasonably provide the breadth of experience needed to adequately cover the curriculum needs of their degree programs.

CHAPTER V. DISCUSSION

Introduction

This chapter discusses the implications of the finding of the study. Specifically, it explores whether the application of the production function to the scholarly environment provided a proper fit; whether total factor productivity and the calculation of residual scholarly output can be used as a viable methodology for indirectly measuring institutional effectiveness in research productivity; and it also introduces the concept of business models and the facilitation of business processes as a theoretical framework for describing the characteristics that differentiate high and low productive efficiency programs for the purpose of better understanding how institutional effectiveness can be increased in fostering faculty research productivity. The study's research questions are answered within the context of discussing these three topics. Additionally, the chapter will also address the limitations of the study and offer suggestion for future strands of research in this area.

When this study commenced, the application of the production function and total factor productivity to an academic environment presented concerns – at least in the mind of the researcher – as to the appropriateness of the theoretical fit. Is it suitable to compare the academic and commercial spheres? The creation of scholarly content in an academic department does have somethings in common with the commercial production of goods: both commit resources toward the production of output. And while business and academic enterprises both presumably want to perform effectively in the process of production, the fit seemed roughly analogous at best and a bit crude at worst. Commercial environments are profit driven. The microeconomic and business literature tells us that commercial enterprises mostly care about the quality of their output or its impact on society to the extent that those factors impact revenue and profits (Kichmer, 2017). The academic enterprise, on the other hand, should have a purer focus that is driven by the value and contribution of their output in and of itself, by producing quality scholarship that impacts, if not society broadly, at least the knowledge that surrounds a given discipline. Of course, in recent years a growing body of literature suggests that higher education has become increasingly corporatized (Morrisey, 2012; Engwall, 2016). Yet, even if the view is taken that academic institutions are motivated by reputation and prestige more so than actual impact – and that such reputation seeking is analogous to the corporate world's profit seeking in explaining behavior – the translation of scholarly output into reputation is still not as clear or as clean to measure as the translation of labor and capital into goods that are sold for a price in the marketplace. The value added and profits accrued through those processes can be measured with a high degree of precision. If a faculty member publishes a high-profile paper, it may add credence to the assertion that a program is high quality, but how much reputation is derived from that article? Or from a department's collective scholarly output, for that matter? Clearly, the drivers of reputation for academic excellence or reputation are more manifold than the highly manifest exercise turning labor and materials into profits.

Despite the concerns of applying an econometric approach to the less manifest production environment of scholarship, the findings suggest that not only was the attempt not off-base, but that BME programs function more like businesses than anticipated. First, in terms of the application of the production function itself, findings show that when the data surrounding the BME scholarly environment is organized and arranged in terms of input and outputs in the fashion of the production function, it produced a model capable of describing that environment that both meets modelling standards and is logically consistent with how research and scholarship are carried out in the BME field. Second, when the concept of total factor productivity was applied by using the results of the production function to calculate residual scholarly production, it led to qualitative findings that strongly reinforced the notion that this quantitative approach proved to be effective in empirically distinguishing programs that were well run from those that were facing more challenges. Because of this, the study is able to meet its primary objective, which was both to identify institutional effectiveness in supporting research productivity and to inform its practice. In doing so, the study also accomplished the objective of establishing the use of residual scholarly output as viable methodology for measuring institutional effectiveness in research administration, an approach that offers broad potential for future research and application. Finally, the qualitative findings can be interpreted to suggest that the BME scholarly environment is akin to business or commercial operations, and the key factor that most divides programs in terms high and low productive efficiency is how they assign responsibility for the facilitation of key business processes.

The Production Function and the Scholarly Environment

The first issue to address is whether or not a convincing model can be derived by arranging the data surrounding the BME scholarly environment into the form of the production function. The findings of chapter 4 indicate that it is both possible and reasonable, as the model that was produced both meets the requirements of ordinary least squares regression analysis and consisted of components that are conceptually consistent with the production function. Recalling that the production function, in its most basic form, indicates that output is a function of the combination of capital and labor (Robinson, 1955; Hulten, 2001; Fioretti, 2007), the model's components all conform to this notion. BME research expenditures and library expenditures represent capital investment into the enterprise and graduate research assistants, along with the tenure and tenure-track faculty count imbedded in the dependent variable, represent the labor. While the enrollment variable does not represent labor or capital, the constraint that it represents to the researcher's time is still a factor that can influence output in a fashion similar to how tax structures are incorporated into production function to the extent that it can influence the prices and demand for the labor or materials used in the production process (Hall & Jorgenson, 1967; Mertens & Ravn, 2012).

If the production function concept and its relation of inputs to outputs had been a poor fit for the scholarly environment, it is unlikely that the data would have produced as significant of a model. Furthermore, the only major factor that was absent from the model was research space and equipment. This would represent a capital component akin to the factory and machinery needed to manufacture output in a production function in an industrial setting. This is important to note because it is reasonable to assume that this omission would be noticeable to anyone with some familiarity of the science and engineering research environment equally as much as the omission would be troublesome from an industrial production function perspective. Unfortunately, no data was available that could serve as a measure for this factor, but it is assumed had such data been available the model would have achieved higher significance and would have been capable of explaining more of the variability in output among programs. The qualitative data certainly suggested that adequate access to the right equipment and space was a distinguishing characteristic between high and low productive efficiency programs.

The second issue to address is what was learned about the relation of each of these components to scholarly production during the course of the study. Some of the discussion points were covered more fully in chapter 4, but they are repeated here for the purpose of providing a complete summary in one place. By virtue of their inclusion in the model, the factors discussed below combine to answer the first research question:

> R1: What factors of faculty research productivity are most strongly correlated with scholarly output as measured by weighted article count per tenured or tenure track faculty at PhD-granting biomedical engineering programs in Carnegie-classified "highest research activity" doctoral universities?

Graduate Students

The literature and pilot interviews both suggested that graduate research assistants play a significant role in aiding research productivity in BME. The quantitative results suggested their role was roughly as significant as that of research expenditures and the qualitative data offered a clear explanation as to why. Each BME researcher interviewed indicated that the assistants handle almost all of the research and initiate almost all of the publications. A few researchers did indicate that they rely more on postdocs fellows to fill essentially the same role, though the quantitative data did not show a significant relationship between scholarly output and postdoctoral fellowships, perhaps because they were fewer in numbers than graduate research assistants. But in either case, they fit into the process the same. BME faculty member's primary role is to attract funding and provide oversight for the lab while graduate assistants, postdocs, or a combination of both do the actual work. Hence graduate research assistants are *the* labor component, rather than serving some sort of auxiliary labor role as the study originally assumed. The extent of this reliance on graduate assistants was evidenced by the explanatory strength of that variable during the model building process. Each time it was included in an iteration, regardless of which other variables were present, its relationship to scholarly output consistently produced among the highest beta weights of any variable and the relationship was always statistically significant at the highest levels.

The clear implication that needs to be emphasized, particularly to an audience outside of BME, is how fundamentally integral graduate research assistants are to the research process in this field. Clearly, the qualitative data suggests that their importance to the process is well known to those familiar with the BME research environment. Furthermore, as the descriptive statistics indicate, the entire range of R1 PhD-granting BME programs employs roughly the same number of graduate research assistants to tenure or tenure-track faculty (about 2.5 GRA's per researcher). As such, there is no evidence to suggest that the number of graduate research assistants plays a role in distinguishing high productive efficiency programs from the less efficient peers. While it is possible that the quality of graduate research assistance may impact the productive efficiency of a program, there was no measure of quality available to introduce into the model and the qualitative data turned up no results to suggest a disparity.

This study's finding concerning graduate students also provides parameter estimates of the of the magnitude by which they impact scholarly output. This empirical evidence of their effect offers some potential benefit, even to those who already understand it, in demonstrating the value that graduate research assistants provide to research productivity in BME and possibly engineering in general. Anyone who is interested in making an evidence-based decision or argument in favor of investment in graduate research assistantships can support the claim that they are as fundamental to the research and publication process as research expenditures themselves. While it must be noted that these results measure correlation, not causation, the straightforwardness of the role that graduate research assistants play in conducting research and producing scholarship offers strong theoretical support to the claim of a cause-and-effect relationship.

Research Expenditures

The quantitative analysis confirmed just how significant a role research expenditures play in the production of scholarly output. The variable performed as expected in the production function, with the only surprise result being that graduate research assistants and library expenditures variables registered similarly strong beta weights and levels of statistical significance. As discussed above, it is worth noting that the qualitative analysis indicated that BME faculty's primary role in the research process is the securing of research dollars, not the actual research itself. Furthermore, several interview participants indicated that in the current environment of sequestration and with other reductions of funding they are spending more time – reportedly by a factor of two or three – to be able to locate and secure the same levels of research funding than they did in the past. This presents implications that are addressed below in the discussion of BME as a business process.

Library Expenditures

The library expenditures also achieved results that were on par with research expenditures in terms of beta weight and statistical significance. Such findings are not without precedence, as Toutkoushian and Porter (2005) registered similar results with their library expenditures variable when they applied the concept of the production function to the scholarly environment. They thought it was unlikely that library expenditures were actually having so strong of an effect on scholarly output and, instead, they hypothesized that the library variable was likely serving as an indirect measure of the university's capacity and propensity to invest in its research infrastructure. The qualitative data and analysis in this study certainly supports Toutkoushian and Porter's assumption that library variable was indeed measuring something more than just the direct impact of the library on research productivity. The responses to the question on the interview guide regarding library resources and services offered no evidence that would suggest that library resources or services were as important in the research process as funding or graduate assistants.

Enrollment

The correlation between enrollment levels and scholarly productivity was expected to produce a stronger, more negative coefficient than the final model indicated. One reason for this assumption was an under-appreciation of the important role that graduate research assistants play in the research process. With graduate students representing more than one-third of total BME enrollment, this variable was simultaneously measuring both a constraint and an aid to productivity. Furthermore, researchers at three of the four institutions indicated relying on undergraduate students in the lab to some degree as well as graduate students. So, while relatively high enrollment levels do cause teaching responsibilities to compete with research activities – as evidenced by the negative correlation – they also represent assistance in the research process. It could be that the relationship between enrollment and scholarly output might be better expressed with a different measure of enrollment than was used in this study. The variable for percentage of graduate enrollment was intended to help mitigate this potential problem, but unfortunately, it was too skewed to use in the model. Perhaps a variable that measures only undergraduate enrollment, rather than the total enrollment variable used in this study, would represent a truer measure of the effects of teaching load. Another reason for the weaker-than-anticipated finding is the likelihood that BME is distinct from many other disciplines in terms of the size of the faculty teaching load. While the literature consistently describes the balancing of teaching and research responsibilities as a fundamental challenge to research productivity, the general absence of its mention in the interview data suggests that BME researchers likely face lower teaching loads than researchers from many other disciplines.

The results of the production function model itself is not profound and does not offer too much information about the BME research environment that was not already understood, beyond adding parameter estimates to known factors. That grant funding can be transformed into scholarship by faculty and graduate students does not offer a challenge to conventional wisdom, but it is important because it suggests that this first step, from which other more important steps follow, is a valid way to conceive of and measure scholarly production.

Total Factor Productivity as a Measure of Institutional Effectiveness

Much more critical and much less certain than whether or not the production function fits the scholarly environment, was the question of whether or not the concept of total factor productivity (TFP) and residual production could be subsequently used to detect institutional effectiveness in the support of research productivity by measuring productive efficiency. There was no challenge in using the production function equation to calculate each program’s level of residual production by comparing actual to predicted scholarly output. Claiming that some portion of that residual is truly a measure of productive efficiency, however, was much less certain. As the centerpiece of the study’s conceptual framework, if residual scholarly output had failed to adequately detect institutional effectiveness in supporting research productivity, then study itself would have been unable to accomplish its objectives.

The inherent problem is that every regression model produces random error, whereby the actual value of most case’s dependent variable will lie above or below the predicted value, meaning that a residual value is produced regardless of its cause. Because of this, the quantitative data is insufficient to the task of determining whether the residual approach actually distinguished high from low productive efficiency environments. The answer had to come from the qualitative data. When the BME program residuals were first calculated and were demonstrated to have a very compressed range (93.5% of observation fell within two standard deviations of the mean) there was some initial concern that the programs at either end of the spectrum would not be too different from one another. After all, interpreted from a production function perspective, the compactness of that range would suggest that the BME field is fairly consistent in terms of its efficiency in producing scholarship. To the extent to which the limited number of programs represented in this study could offer confirmation, however, the qualitative findings strongly suggest that productive efficiency in the production of scholarly output was being measured, and that this efficiency was tied to institutional effectiveness. There was a palpable contrast in the nature of responses between participants in those programs designated as low productivity efficiency and that of their peers in programs with the high designation.

Researchers from programs with the lowest residual production levels mentioned more challenges and described less effective support structures. Further research is needed before the value or success of this methodology could be declared with a high degree of certitude, but its application in this study achieved what it was intended to do.

In doing so, it answered the second research question:

> R2: Which Ph.D.-granting biomedical engineering programs in Carnegie-classified "highest research activity" doctoral universities exhibit the highest and the lowest levels of productive efficiency in creating scholarly output?

For the purpose of confidentiality of the research participants, the names of the institutions have been excluded. This research question was never intended to produce a public ranking or identification of individual institutions, but rather to serve a means of facilitating the sampling of the qualitative analysis.

Research Support as a Business Model

"Our administrative structure, our university's administrative structure, is the same as that of a flea market: It provides a roof over ours head and everyone runs their own stall."
-Professor, Low productivity efficiency program

As the qualitative findings suggest, the characteristics that most differentiate the two types of programs were that the researchers in high productive efficiency departments faced fewer administrative burdens in both grant management and in general; had better access to well-equipped, institutionally provided core facilities; and had core facilities that were well run and staffed with personnel who had the expertise to make the most of the equipment and research tools. These findings answered the third research question:

> R3: Which institutional factors most influence the research efficiency of faculty research productivity efforts in biomedical engineering departments in US doctoral institutions?

In the quote above, a BME researcher from institution D was referring to the fact that neither his school nor department provides technical or machinist support for his lab. Instead, researchers are left to contract for services on a lab-by-lab basis. Like most labs, he does not need full time support. Therefore, he is left to identify the appropriate contractor and then use his institution's "Byzantine" procurement system to secure services and pay invoices for assistance in things such as research computer support or the other specialized tasks that are needed to keep his projects moving forward and his lab running effectively. He does have access to an office manager who is shared with other faculty, but despite that assistance, he is largely left to play the role of business manager. Aside from the cost of the services themselves, there is the cost of his time and focus as he is pulled away from research or grant writing to deal with business matters.

The arrangement that this BME researcher describes can be characterized as a business model, which in this case he has labelled as a flea market, whether his institutional leadership conceives of it in those terms or not. The idea of the business model offers a heuristic approach towards conceptualizing the key differences that distinguish high productive efficiency departments from lower ones. It is the provision of these resources and services, and perhaps even more so the effective quality of those services and resources, that allow some researchers to move forward with minimum distractions while others face an accumulation of obstacles and diversions from progress. This discussion provides an answer to the final research question:

> R4: How do these institutional factors influence the research efficiency of faculty research productivity efforts in biomedical engineering departments in US doctoral institutions?

The literature surrounding business models is conceptually broad, evolving and is far from having a single definition (Bankvall, Dubois, & Lind, 2017). A general description that fits the purpose of this discussion, however, is that a business model is the "link between business strategy, processes, and information technology" (Veit et al., 2014). Focusing on the business processes component of the business model, Kirchmer (2017) offers a hierarchy consisting of "governance processes," "management processes," and "operational processes" (p. 7). In this hierarchy, governance processes give shape to the overall business model by ensuring that stakeholder expectations, mega-trends, and technological developments are incorporated into the operations of the business. The management processes are responsible for ensuring efficient and effective performance of all work processes. Finally, the operational processes are the means by which the actual work gets done.

Applying this concept to the BME research environment, the operational processes that facilitate the creation of scholarly output – aside from conducting the actual research and writing of manuscripts – can be viewed as consisting of pre- and post-award grant management (including budget, procurement, accounting as well as grant submission support) and the technical support and expertise needed for the both the researcher's lab and the cores facilities. The management processes can be thought of the extent to which performance of the operational processes are supervised and evaluated to ensure effectiveness. The governance processes consist of strategic provision of resources, such as core facilities or shared equipment, and the investment in and proper the organization development of support services personnel.

The key distinction in business models between high and low productive efficiency programs in the generation of scholarly output is how much of the business processes fall on researchers. Clearly, the researcher must be responsible for grant writing and the direct oversight of the research and publication activities. They are highly trained and skilled to be productive in those central tasks of the research enterprise. The question is: why have them be responsible for anything else needed to facilitate the research process? By default, however, any operational effort or managerial oversite that is not institutionally provided falls on the researcher to provide. This is because the onus rests with the researcher, by virtue of their faculty contract and performance evaluation, to do whatever is necessary to secure grants and translate those funds into publications. In the absence of a more sophisticated support structure that diffuses any ancillary responsibilities effectively, the faculty member is left to fill the gaps.

This is not to suggest that the concept that researchers benefit from support is not universally and inherently understood by university and departmental leadership or the research community generally. It is a basic and obvious concept that is reflected in the fact that every university makes considerable efforts at considerable expense to centrally support sponsored research and most departments – to the extent that they can – make additional localized efforts to ease the burdens their researchers face. Regardless of how basic a concept it is, however, it clearly receives various levels of emphasis or achieves various results in execution by the programs that were explored in the qualitative component of this study. The level of supports that were offered and the quality of those supports varied so greatly that at some point it can be suggested that they constitute separate business models: one where the researcher maximally facilitates the operational and managerial processes of producing scholarship and one where the researcher minimally facilitates those processes. The extent to which each of the institutions in the study offers support, and the extent to which that support is effective, is summarized below program-by-program in the attempt to illustrate this further.

The qualitative data suggests that institution D leaves more operational and managerial responsibilities to the researcher than the other programs studied. As discussed above, researchers are left to both manage and pay for technical support contracts with outside vendors. They receive limited pre-award support in terms of handling the non-scientific portion of grant submission process. Furthermore, indications are that the university's centralized office of sponsored programs is largely unhelpful in the submission process as well. There is no integration of pre- and post-award efforts, and most of the post-award management support is in the form of a single office manager who is shared with other faculty. Finally, the core facilities appear to be insufficient in terms of capacity to meet demand and are operated without an adequate staffing model to provide the expertise needed to ensure that those facilities function optimally. Institution D is a top tier university with an overall ranking higher than the other institutions discussed here. It is not public and it is not financially limited. To its credit, the university recognized the need to invest in better core facilities and they responded by constructing a significant multidisciplinary research facility housing shared equipment which is about to come on line at the time of this writing. Nevertheless, doubts were still expressed about the adequacy of the staffing model in place for the new facility. With the accumulation of impediments and responsibilities faced by researchers at institution D, it is not surprising that their collective scholarly output was so much lower than their input of resources suggested it should be. In terms of the dependent variable data that was assembled for the quantitative analysis, institution D was ranked 38 out of the 62 programs in weighted article count per faculty member. But when factoring their resources to calculate their residual production, they fell to 61 of 62 in productive efficiency ranking. It is easy to imagine that the BME program would produce a level scholarly output more in line with their level of research funding and number of graduate research assistants if the faculty were not busy acting as business managers and navigating institutional bureaucracy.

Institution C appears to have better access to core facilities and shared equipment, though some difficulties clearly exist. Institutional funding for technical support does appear to be on the rise. But they offer very limited pre- and post-award support in comparison to the high productive efficiency programs, and perhaps even less than Institution D in that regard. Researchers are left to construct their budgets on their own, assemble the various supporting documents, and work directly with the central office of sponsored programs to facilitate submission. Furthermore, several researchers suggested that difficulties can arise between award and funding that are much more extreme than the experiences that researchers described at other institutions.

Institution A appears to have excellent core facilities that are accessible, well managed, and staffed with expert personnel. The researchers all mentioned enjoying departmentally provided technical support in maintaining their own lab spaces as well as the availability of machinists and fabricators to help in the performance specialized tasks. The pre- and post-award support that they described could be categorized as reasonable, but some researchers indicated there was room for improvement in both phases.

BME researchers at institution B clearly offered the most positive account of institutional support for their research. They described a very systematic approach to pre-award grant management where support personnel were very familiar and proactive in providing the type of documentation that most funding entities required. The manner in which BME researchers at Institution B describe the support they receive suggest that the grant support personnel take a certain level of ownership that was not reported by researchers at other institutions. Like Institution A, there were only positive comment concerning the availability of core facilities and the personnel who run them. Likewise, they seemed to enjoy adequate institutionally-provided technical support for their own labs.

As the quantitative data indicated, these programs are not too different in terms of resources. In 2015 research dollars, for example, institutions A, C, and D were all within a $1.5M dollars of each other, with A being the lowest of the three. Institution B did record a little more than twice as much in research expenditures as the others, but the size of their faculty was also more than twice as much as the other three programs. They all had similar ratios of graduate research assistants to faculty members. But institutions A and B have clearly found better ways to support their researchers.

If institutions like C and D wish to become more effective in terms of research productivity, they need to concentrate on maximizing the support they provide to their researchers by investing more in their support infrastructure and making it a highly visible department priority. The findings of this study suggest that departments, generally, need to determine how much support infrastructure – in terms of the numbers and type of personnel – are needed per researcher to effectively support their efforts. Respondents from all institutions relayed stories of departmental meetings where faculty eagerly discussed expanding into new area through strategic hiring of new faculty as a means of strengthening departmental research clusters or expanding into new research domains. For institutions that are not performing to the level of their existing resources, like C and D, those discussions should focus on growth of support infrastructure first. They need to determine what types and what numbers of personnel can best provide the substantive support their faculty need in order to stay focused on the research itself. Once achieved, they need to make sure the quality of that support remains a priority, and that the infrastructure expands as their department grows in order to make sure that their productive efficiency remains sustainable.

Limitations

There are various limitations to this study. The most important to acknowledge it that this study applies only to Ph.D.-granting biomedical engineering programs in Carnegie-classified "highest research activity" institutions. As a result, the extent to which these results are generalizable to other types of programs, disciplines, or institutions is questionable and should only be done with caution. Additionally, the number of respondents that participated in the qualitative phase of the study is not as high as originally intended, though there was a sense of data saturation in regard to the items on the interview guide. In terms of the quantitative aspects of the study, some of the measures were imprecise because they were measured at the school level rather than the BME program level (this is discussed more fully in Chapter 3).

Future Research

It is hoped that the concept of the production function and TFP's residual scholarly output could be tested further in other disciplines and followed up with qualitative inquiry in order to determine whether this approach is widely applicable for detecting institutional effectiveness in facilitating research productivity. Furthermore, while the approach of selecting the institutions at the high and low ends of the residual production spectrum seemed to produce useful results, an alternative method might be to select programs based on how much their rankings shifted from the weighted article count per faculty member to the residual ranking. By comparing those programs with the greatest upward shift to those with the greatest downward shift, it is possible be that an even greater contrast in institutional effectiveness could be found.

Perhaps most importantly, while this study suggests that the level and quality of support services is the most important factor in distinguishing high from low productive efficiency programs, it did not dig deeply into how each institution's support structure was designed and maintained. It would be useful to better understand the types of personnel employed, their professional training and backgrounds, and how their responsibilities are divvied and/or integrated from one program to the next. This would help produce more specific recommendations as to what works best.

Another important dimension in BME research which was addressed somewhat in the findings, but not as fully as would have been preferred, was the intersection of collaboration and research productivity. What organizational strategies best foster interdisciplinary, intra-departmental, and multi-institutional research? How does this impact the collection research productivity of a BME department? Does it introduce economies of scale or efficiency into the research process? While this topic did come in several of the interviews, there simply was not enough data from enough participants or institutions to develop a solid theory on this important aspect.

References

Aiston, S.J. and Jung, J. (2015). Women academics and research productivity: an international comparison. *Gender and Education, 27(3)*, pp. 205-220. DOI:10.1080/09540253.2015.1024617

Allen, M.P. (1997*). Understanding Regression Analysis*. New York, NY: Plenum Press.

Baldwin R. G., Lunceford, C. J., and Vanderlinden, K. E. (2005). Faculty in the Middle Years: Illuminating an Overlooked Phase of Academic Life. *The Review of Higher Education, 29(1)*, pp. 97-118. DOI: 10.1353/rhe.2005.0055

Bankvall, L., Dubois, A. & Lind, F. (2017). Conceptualizing business models in industrial networks. *Industrial Marketing Management, 60(1)*, pp. 196-203. https://doi.org/10.1016/j.indmarman.2016.04.006

Bauer, K. (2001). "Assessment for Institutional Research: Guidelines and Resources." In W. E. Knight (ed.), *The Primer for Institutional Research*. Tallahassee, Fla.: Association for Institutional Research.

Bellas, M.L. and Toutkoushian, R.K. (1999). Faculty Time Allocations and Research Productivity: Gender, Race, and Family Effects. *The Review of Higher Education, 22(4)*, pp. 367-390.

Betancourt, R.R. (1986). A generalization of modern production theory. *Applied Economic, 18*.

Biddle, J. (2012). Retrospectives: The Introduction of the Cobb-Douglas Regression. *Journal of Economic Perspectives*, 26(2): 223-36.

Black, J., Hasimzade, N., and Myles, G. (2009). *A Dictionary of Economics* (4th ed.). Oxford, UK: Oxford University Press.

Blackburn, R. T. & Bentley, R. J. (1993). Faculty Research Productivity: Some Moderators of Associated Stressors. *Research in Higher Education, 34(6)*, pp. 725-745. http://www.jstor.org/stable/40196121

Bland, C.J. and Ruffin, M.T. (1992). Characteristics of a Productive Research Environment: Literature Review. *Academic Medicine, 67(6)*, pp. 385-396.

Bland, C.J., Weber-Main, A.M., Lund, S. & Finstad, D. (2005). *The Research-Productive Department: Strategies from Departments that Excel*. Boston: Anker Publishing, Inc.

Britt-Fries, S. L., Rowan-Kenyon, H. T., Perna, L. W., Milem, J. F., & Howard, D. G. (2011). Underrepresentation in the academy and the institutional climate for faculty diversity. Journal of the Professoriate, 5(1), pp. 1-34.

Charmaz, K. (2004). Grounded Theory. In S.N. Hesse-Biber & P. Leavy (Eds.), *Approaches to qualitative research: A reader on theory and practice*. New York: Oxford University Press.

Charmaz, K. (2015). Grounded Theory: Methodology and Theory Construction. In J.D. Wright (Ed.) *International Encyclopedia of Social and Behavioral Sciences, 2nd Ed*. London: Elsevier RELX Group

Cole, S. (2007). Research Administration as a Living System. *Journal of Research Administration (38)*, pp. 139-152.

Corbin, J.M. and Strauss, A.L. (2008). *Basics of qualitative research: techniques and procedures for developing grounded theory*. Los Angeles, CA: Sage Publications.

Diamond, A. M. (1986) The Life-Cycle Research Productivity of Mathematicians and Scientists. *Journal of Gerontology, 41(4)*, pp. 520-525.

Dundar, H. & Lewis, D. R. (1998). Determinants of Research Productivity in Higher Education. *Research in Higher Education, 39(6),* pp. 601-631. http://www.jstor.org/stable/40196313

Eagan, M. K. & Garvey, J. C. (2015). Stressing Out: Connecting Race, Gender, and Stress with Faculty Productivity. *The Journal of Higher Education, 86(6)*, pp. 923-954. DOI: 10.1353/jhe.2015.0034

Engwall, L. (2016). The Internationalization of Higher Education. *European Review, 24(2),* pp. 221-231.

Fioretti, G. (2007). The production function. *Physica A, 374*. p. 708

Garfield, E. (2005, September). *The Agony and the Ecstasy: The History and Meaning of the Journal Impact Factor.* Paper presented at the International Congress on Peer Review and Biomedical Publication, Philadelphia, PA.

Given, L.M. (2016). *100 Questions and answers about qualitative research*. Los Angeles: Sage.

Glaser, B.G., and Strauss, A.L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Chicago: Aldine.

Golden, J., and Carstensen, F. V. (1992b). Academic research productivity, department size and organization: Further results, comment. *Economics of Education Review, 11(2)*, pp. 153-160.

Grunig, S. D. (1997). Research, Reputation, and Resources: The Effect of Research Activity on Perception of Undergraduate Education and Institutional Resource Acquisition. *The Journal of Higher Education, 68(1)*, pp. 17-52. http://www.jstor.org/stable/2959935

Guarino, C.M. and Borden, V.M.H. (2017). Faculty Service Loads and Gender: Are Women Taking Care of the Academic Family? *Research in Higher Education 58(1)*, pp. 1-23

Hall, R.E. and Jorgenson, D.W. (1967). Tax policy and investment behavior. *American Economic Review, 57* (3), pp. 391-414

Head, R.B. (2011). The Evolution of Institutional Effectiveness in the Community College. *New Directions in Community Colleges, 153*, pp. 5-11.

Helsi, V.L. and Lee, J.M. (2011) Faculty Research Productivity: Why Do Some of Our Colleagues Publish More than Others? *Political Science and Politics, 44(2)*, pp. 393-408. http://www.jstor.org/stable/41319926

Hirschfield, L.E. and Joseph, T.D. (2012). We need a woman, we need a black woman: gender, race, and identity taxation in the academy. *Gender and Education, 24(2)*, pp. 213-227. DOI: 10.1080/09540253.2011.606208

Horn, A.S. and Lee, G. (2016). The Reliability and Validity of Using Regression Residuals to Measure Institutional Effectiveness in Promoting Degree Completion. *Research in Higher Education (57)*, pp. 469-496. DOI 10.1007/s11162-015-9394-7

Hu, Q. and Gill, T. G. (2000). IS Faculty Research Productivity: Influential Factors and Implications. *Information Resources Management Journal, 13(2)*, pp. 15-26.

Hulten, C.R. (2007). Total Factor Productivity: A Short Biography. In Hulten, C.R., Dean, E.R., and Harper, M.J. (Eds). *New Developments in Productivity Analysis*. Chicago, IL: University of Chicago Press.

Jordan, J. M., Meador, M., and Walters, S. J. K. (1988). Effects of departmental size and organization on the research productivity of academic economists. *Economics of Education Review 7(2)*, pp. 251-255.

Jordan, J. M., Meador, M., and Walters, S. J. K. (1989). Academic research productivity, department size, and organization: Further results. *Economics of Education Review 8(24),* pp. 345-352.

Kirchner, M. (2017). *High Performance Through Business Process Management: Strategy Execution in a Digital World. Third Edition*. Cham, Switzerland: Springer International Publishing AG, 2017.

Kirkwood, B.R and Sterne, J.C (2003). *Essential Medical Statistics. 2nd ed*. Oxford, UK: Blackwell, 2003.

Kyrillidou, M., Potter, G.P., Franklin, B., Cook, C. and Thompson, B. (2012). The ARL Investment Index Revisited: In Search of Reliable and Valid Indicators of Extensiveness of Research Libraries in the 21st Century. In Hiller, S. et al. Eds. *Proceeding of the 2012 Library Assessment Conference: Building Effective, Sustainable, Practical Assessment*.

Paper presented at the 2012 Library Assessment Conference, Charlottesville, VA (p. 61-66). Washington DC: Association of Research Libraries.

Kyvik, S. (1990). Motherhood & Scientific Productivity. *Social Studies of Science, 20(1)*, pp. 149-160. http://www.jstor.org/stable/285104

Kyvik, S. (1995). Are big university departments better than small ones? *Higher Education 30(3)*, pp. 295-304.

Levin, S.G. & Stephan, P.E. (1991). Research Productivity Over the Life Cycle: Evidence for Academic Scientists. *The American Economic Review, 81(1)*, pp. 114-132. http://www.jstor.org/stable/2006790

Long, J. S. (1978). Productivity and academic positions in the scientific career. *American Sociological Review, 43(4)*, pp. 889-908.

Long, R.G. (2008). The Crux of the Method: Assumptions in Ordinary Least Squares and Logistic Regression. *Psychological Reports, 103*, p.431-434.

Maske, K.L., Durden, G.C., and Gaynor, P.E. (2003). Determinants of Scholarly Productivity Among Male and Female Economists. *Economic Inquiry, 41(4)*, pp. 555-564. DOI: 10.1093/ei/cbg027.

Massy, W.F. and Wilger, A.K. (1995). Improving Productivity. *Change 27 (4)*. pp. 10-26.

Mason, M.A., Wolfinger, N.H., and Goulden, M. (2013) *Do Babies Matter?: Gender and Family in the Ivory Tower* . New Brunswick, US: Rutgers University Press

McLeod, M.W. and Atwell, C.A. (1992). A Mini Dictionary of Institutional Effectiveness Terms. *Community College Review 20(2)*, p. 30-38.

Merriam, S. B. (1998). *Qualitative research and case study applications in education*. San Francisco: Jossey-Bass.

Merriam, S. B. (2009). *Qualitative research: A guide to design and implementation*. San Francisco: Jossey-Bass.

Mertens, K. & Ravn, M. (2012). Empirical evidence on the aggregate effects of anticipated and unanticipated US tax policy shocks. *American Economic Journal: Economic Policy, 4(2),* pp. 145-181.

Merton, R.K., Fiske, M. and, Kendall, P.L. (1990). *The Focused Interview: A Manual of Problems and Procedures*. New York, NY: The Free Press.

Meyer, K. A. (1998). *Faculty workload studies: Perspectives, needs, and future directions*. ASHE-ERIC Higher Education Report No. 1. Washington, DC: George Washington University, Graduate School of Education and Human Development.

Miller, J. C., Coble, K. H., and Lusk, J. L. (2013). Evaluating top faculty researchers and the incentives that motivate them. *Scientometrics, 97*, pp. 519-533. DOI: 10.1007/s11192-013-0987-7

Morrissey, P. (2012). Higher education and the imperative to build reputations. *Asian Education and Development Studies, 2012*, 1(2), pp.112-123

Mueller, B. (2014, April 28). Administrators Call for Rewarding Professors' Patents with Tenure. The Chronicle of Higher Education. Retrieved from: http://chronicle.com/article/Administrators-Call-for/146245/

O'Meara, K.A. (2005). Encouraging Multiple Forms of Scholarship in Faculty Reward Systems: Does it Make a Difference? *Research in Higher Education, 46(5)*, p. 479-510). DOI: 10.1007/s11162-005-3362-6

Padilla, A.M. (1994). Ethnic Minority Scholars, Research, and Mentoring: Current and Future Issues. *Educational Researcher (23),4*, pp. 24-27. http://www.jstor.org/stable/1176259

Pandis, N. (2015). The Sampling Distribution. *Statistics and Research Design 147*(4), 517-19.

Patton, M.Q. (2002). *Qualitative Research & Evaluation Methods, 3rd ed.* Thousand Oaks, CA: Sage Publishing, Inc.

Pike, G., Kuh, G., McCormick, A., Ethington, C., & Smart, J. (2011). If and When Money Matters: The Relationships Among Educational Expenditures, Student Engagement and Students' Learning Outcomes. *Research in Higher Education, 52*(1), 81-106. Retrieved from http://www.jstor.org/stable/41483772

Powell, B.A., Gilleland, D.S., and Pearson, L.C. (2012). Expenditures, Efficiency, and Effectiveness in U.S. Undergraduate Higher Education: A National Benchmark Model. The Journal of Higher Education, 83(1). P. 102-27.

Production Function. (1998). In *Encyclopedia Britannica online*. Retrieved from: http://www.britannica.com/topic/production-function

Ransome, P. (2013) Ethics and Values in Social Science Research. New York: Palgrave MacMillan

Rawlings, J.O., Pantula, S.G, & Dickey, D.A. (1998). *Applied Regression Analysis: A Research Tool*. New York: Spring-Verlag.

Research Information Network (2010). *Research Support Services in UK Universities*. London, UK: Author. Retrieved from: http://www.rin.ac.uk/our-work/using-and-accessing-information-resources/research-support-services-what-services-do-research

Rogers, J. T. "Assessment in Accreditation: Has It Made a Difference?" *Assessment Update*, 1997, *9*(4), pp. 1–2, 15.

Rousseeuw, P; Leroy, A. (1996). *Robust Regression and Outlier Detection (3rd ed.)* New York: John Wiley & Sons

Samprit, C. & Simonoff, J.S. (2013). *Handbook of Regression Analysis*. Hoboken, NJ: John Wiley and Sons.

Sax, L.J., Hagedorn, L.S., Arredondo, M. and Dicrisi, F.A. (2002). Faculty Research Productivity: Exploring the Role of Gender and Family-Related Factors. *Research in Higher Education, 43(4)*, pp. 423-426.

Stack, S. (2004). Gender, Children, and Research Productivity. *Research in Higher Education, 45(8)*, pp. 891-920.

Streuly, C.A. & Maranto, C.L. (1994). Accounting faculty research productivity and citations: Are there gender differences? *Issues in Accounting Education, 9(2)*, p. 247.

Tierney, W. (1999). Faculty Productivity and Academic Culture. In W. Tierney (Ed.) *Faculty Productivity: Facts, Fictions and Issues*. New York: Falmer Press.

Toutkoushian R.K. and Porter, S.R. (2005). Institutional research productivity and the connection to average student quality and overall reputation. *Economics of Education Review 25*, pp. 605-617

Veit, D., Clemons, E. Benlian, A., Buxmann, P. Hess, T., Kundisch, D., Leimeister, J.M., Loos, P. and Spann, M. (2014). Business Models: An Information Systems Research Agenda. *Business and Information Systems Engineering, 6(1),* pp. 45-53.

Volkwein, F.J. (2011). A Model for Assessing Institutional Effectiveness. *New Directions for Institutional Research 2011(149)*, p. 13-28.

Welsh, J.F. and Metcalf, J. (2003). Cultivating Faculty Support for Institutional Effectiveness Activities: Benchmarking best practices. Assessment and Evaluation in Higher Education 28(1), p. 33-45. DOI: 10.1080/02602930301682

Wergin, J.F. (2004). *Departments that work: Building and sustaining cultures of excellence in academic programs*. Boston: Anker Publishing Co. Inc.

Zuckerman, H. (1991). The careers of men and women scientists: A review of current research. In H. Zuckerman, J. Cole, and J. Bruer (eds.), *The Outer Circle: Women's Position in the Scientific Community*. New York: W.W. Norton.

Appendix A: Interview Guide for BME Researchers

Thank you for your willingness to participate in this study. Before we begin, there are a few point I want to make sure that you are aware of:

- You may end the conversation at any time during the interview for any reason.
- Your responses are confidential. Neither you nor your institution will be mentioned by name anywhere in the resulting dissertation and every effort will be made to exclude any potentially identifying information that you may provide.
- You will have the opportunity to review a transcript of this conversation for accuracy.
- You will also have the opportunity to confirm whether or not the substance of my analysis and interpretation of this conversation is generally accurate and to note omissions or areas of disagreement.

I am interested in identifying how various institutional factors, largely outside the control of an individual researcher, might influence the collective research productivity of biomedical engineering departments.

Before we start, do you have any questions about the study or the interview process?

RESEARCH RESOURCES AND SUPPORT SERVICES

1. Which resources or services offered by your institution, school, or department do you find the most helpful in aiding your productivity as a researcher?
2. What do you consider to be the biggest barriers to your productivity?
3. Do have adequate access to the lab space and equipment that you need to conduct your research?
    - Probe: Do you have adequate departmental lab space (e.g. your own lab space)?
    - Probe: Do you have adequate access to the university's core facilities?
4. Please describe quality of grant management support services that are offered by your department and your Office of Research? (locating/writing/applying/accounting/procuring/reporting)
5. How important are graduate assistants in aiding your research productivity?
    - Probe: Please describe how they aid your research.
    - Probe: Are there a sufficient number of assistants to aid you in your research?
6. Are there other personnel in your department who provide significant assistance to your research efforts? (e.g. non-student research assistants, other support personnel, etc.)
7. Do library resources play a role in aiding your research productivity?
    - Probe: Do you find general quality of the engineering library materials acceptable?
    - Probe: Does the library support your research directly? (e.g. materials data data curation)
8. Are there any other resources or supports services that we haven't discussed that you believe have an impact on your research productivity, either positively or negatively? (e.g. technology transfer, patents application, facilitation of peer networking opportunities, etc.)

DEPARTMENTAL RESEARCH ENVIRONMENT

[*Note: The bracketed/italicized language will not read to the interview participant. It describes the goals and concepts that each question is designed to address. It is based loosely on Bland, Weber-Main, Lund, & Finstad (2005).*]

1. <u>General Departmental Culture</u> [*Does department have a distinctive organizational culture that provides a group identity?*]
    - o How would you describe your department's identity as a group?

- Do all feel that all departmental members contribute to decision making in a substantive manner? [*Assertive Participative Governance*]
- How would you describe department morale and interrelations between individual faculty members? [*Group climate*]

2. Climate for research [*Does departmental research mission serve to focus faculty on high research potential?]*
   - How would you describe your department's emphasis on research, compared to other priorities? [*Research Emphasis*]
   - Does your department establish collective research goals or strategies? [*Coordination Function*]
   - Do you find that the productivity of you peers influences your own productivity? If so, how? [*Research Productive Peers*]
   - Do you find that you have sufficient uninterrupted time to dedicate to scholarly activities? [*Sufficient Work Time*]
3. Leadership [*Extent to which department chair is accomplished researcher, provides professional support faculty, communicates vision and shared goals, and contributes to maintaining a positive work environment*]. Note: The questions will be re-phrased accordingly when interviewing the department chair.
   - Do you view your chair an accomplished researcher? [*Accomplished Researcher*]
   - Does your chair use his/her expertise to directly support your research or the research efforts of your colleagues (e.g. technical support, professional contacts, mentoring, etc.)? [*Provides Professional Support*]
   - Does your chair clearly communicate an overall vision for your department in terms of research goals? If so, is that vision frequently reinforced? Do you find it authentic and aligned with your own values? [*Communicate Vision*]

CLOSING QUESTIONS

1) Are there any important contributors to research productivity in your department that we haven't addressed?

2) Are there any barriers to you research productivity that we haven't discussed?

Appendix B: Interview Guide for Office of Research Administrators

Thank you for your willingness to participate in this study. Before we begin, there are a few point I want to make sure that you are aware of:

- You may end the conversation at any time during the interview for any reason.
- Your responses are confidential. Neither you nor your institution will be mentioned by name anywhere in the resulting dissertation and every effort will be made to exclude any potentially identifying information that you may provide.
- You will have the opportunity to review a transcript of this conversation for accuracy.
- You will also have the opportunity to confirm whether or not the substance of my analysis and interpretation of this conversation is generally accurate and to note omissions or areas of disagreement.

I am interested in identifying how various institutional factors, largely outside the control of an individual researcher, might influence the collective research productivity of biomedical engineering departments.

Before we start, do you have any questions about the study or the interview process?

<u>GENERAL LEAD QUESTIONS</u>

1) What is the most important thing that your office does to support the research productivity of your institution's faculty?

2) Conversely, what factor(s) at your institution – whether under your office's purview or not – do you think hinders faculty research productivity the most?

<u>MAIN QUESTIONS</u>

Goals and Priorities

1) How do you summarize your office's mission?

2) How would characterize an ideally functioning Office of Research?

3) What is currently the greatest challenge in realizing that vision at your institution?

Perceptions of the Research Community

1) Do you formally assess how your institution's research community views of the quality of your department's services and resources?

2) If so, in which areas are you are doing the best job?

3) Which areas present the greatest concern?

Core research facilities

1) Do you feel that your university's core research facilities adequately meet the demands of your research community in terms of capacity and in quality?

1) What are the main challenges you face in operating core research facilities?

2) Does your office assist in patent applications, technology transfer, and the management of intellectual property rights?

FINAL QUESTIONS

1) Are there any important contributions that your office makes toward support research productivity that we haven't address?

2) Are there any barriers to you research productivity at your institution that we haven't discussed?

Appendix C: Pilot Interview Guide

Goal 1: To understand the general nature of BME scholarship and its place within the greater context of BME research.

Goal 2: To understand how organizational resources, policies, values, and other characteristics external to the researcher may impact research productivity.

Project goals and disclaimers:

- Why study BME research productivity?
  - Interest in IR – topic of research productivity in general is germane to institutional effectiveness
  - Public policy component – prevalence of STEM-H in policy and political discussion over higher education.
  - More prolific in scholarship than other engineering disciplines.
  - Most moving pieces (lab requirements, technology transfer, commercialization of research, etc.)
- Disclaimers:
  - This is preliminary research and the collected will not be included in my actual dissertation, the research design for which is not yet developed.
  - No part of this discussion well be used in the dissertation. It may be used in discussion with committee members.

Goal 1 Questions:

1. Where does scholarship (mainly publication of journal articles) fit in with other research goals of BME (e.g. developing commercially useful technology, patentable inventions, etc.)?
2. How does the field value scholarship in comparison with these other research goals?
3. What is the typical content of a BME journal article? And at what point in the research cycle might you share your findings?
4. Is there conflict between the sharing of new knowledge via scholarship and the proprietary nature of invention commercialization?
5. Do graduate students typically share in scholarship?
6. Is BME fundamentally interdisciplinary? Is research frequently conducted with those in the medical or health sciences disciplines?

Goal 2 Questions:

1. I would like to go through a list of institutional factors that might potentially influence (either help or hinder) your research or scholarly productivity. I would like for you to share your own experiences and observation about each item:

    a. Support for the grant process (e.g. locating, securing, managing grants)
    b. Access to adequate research facilities and equipment
        i. Does this influence the competitiveness of your grant applications?
    c. Availability of research and/or graduate assistants
    d. Tenure and promotion policy requirements
    e. Teaching loads
    f. Institutional support for technology transfer, management of intellectual property rights, review of invention disclosure, application for patents, etc.
2. Can you think of other types of organizational factors (departmental or university-wide) that influence your ability to conduct research?
3. Are there institutional efforts to support research that you find of little or no value?
4. Can you think of ways that institutional supports have changed (for better or worse) throughout your career? Or from institution to institution (if you have been employed by other institutions)
5. Do your colleagues from other institutions discuss/report organizational challenges to their research?
6. Do you think it makes a difference for a BME program if it is part of an institution that includes a school of medicine (from a research perspective as opposed to attracting students)?
7. Is your research productivity impacted in anyway by the School of Medicine or other health sciences units? (e.g. access to research space or equipment, collaboration with health sciences faculty and researchers, etc.)

Appendix D: Text of Recruitment Email

Dr. [INSERT NAME],

I would like to request the opportunity to interview you regarding your experiences conducting research at your institution. [When being referred by a participating colleague] Dr. _______ at [INSERT INSTITUTION] has recommended your name as potential participant. [And/or when being referred by a VCU engineering faculty member(s)] Dr(s). ________ are familiar with my study and mentioned your name as a potential participant.

I am Ph.D. candidate in Public Policy and Administration at Virginia Commonwealth University with a research interest in higher education administration. For my dissertation I am studying how institutional factors can influence the research productivity of faculty members in your field. My goal is to better understand how institutions both help and hinder the productivity of their faculty for the purpose of informing institutional effectiveness in the area of research administration.

I plan to interview a handful of tenured/tenure-eligible researchers and research administrators at six different universities, including your own. The interviews will consist of one-on-one video and/or audio conference sessions using Zoom Desktop Conferencing that should take about 45 minutes. Your identity, responses and institutional affiliation will be treated confidentially. You can also terminate your participation in the study at any time and for any reason.

I hope you will consider participating in this study. Please let me know if you have any questions.

If you are willing to grant an interview or if you have any questions, please contact me at rawlsmm@vcu.edu.

Thank you for your consideration. I hope you will choose to participate in this study.

Sincerely,

Michael M. Rawls

Ph.D. Candidate
L. Douglas Wilder School of Government and Public Affairs
Virginia Commonwealth University

Appendix E: Text of Reply Email

Dr. [NAME]:

Just following up on my interview request from last month. I have had the opportunity to interview some of your BME colleagues with the [NAME OF ENGINEERING SCHOOL], but I could really benefit from two or three more interviews in order to have a sufficient amount of data to keep [NAME OF UNIVERSITY] viable as one of the six institutions in my study. That's very important to me because the [NAME OF UNIVERSITY] interviews I've conducted so far have provided excellent insight into how institutional factors can impact faculty productivity, but I just don't have quite the sample size to feel confident about those results. I would really appreciate the opportunity to learn about your experiences in order to add to my understanding.

I hope you will consider about 30 minutes of your time at some point over the next couple of weeks to participate in an online interview. Thank you for your consideration.

Best,

Mike

Appendix F: Consent Form

RESEARCH SUBJECT INFORMATION AND CONSENT FORM

TITLE: Assessing Research Productivity from an Institutional Effectiveness Perspective: How Universities Influence Faculty Research

VCU IRB NO.: HM20010747

INVESTIGATOR: Dr. Maike Philipsen

If any information contained in this consent form is not clear, please ask the study staff to explain any information that you do not fully understand. You may take home an unsigned copy of this consent form to think about or discuss with family or friends before making your decision.

PURPOSE OF THE STUDY

The purpose of this study is to identify institutional factors that most influence the research productivity of biomedical engineering faculty at U.S. doctoral institutions.

You have been asked to participate in this study because you are either a faculty researcher in the field of biomedical engineering or a research administrator at one of six universities that was selected to participate in this study.

DESCRIPTION OF THE STUDY AND YOUR INVOLVEMENT

If you decide to be in this research study, you will be asked to sign this consent form after you have had all your questions answered and understand what will happen to you.

This study is concerned with assessing performance in research administration, rather than studying faculty performance or accomplishments in conducting research. You will be asked about various aspects of conducting research at your institution, such as: “What do you consider to be the biggest barriers to your productivity?” OR “Please describe quality of grant management support services that are offered by your department and your Office of Research.”

The interview will take place using Zoom Desktop Conferencing and will last approximately 45 minutes. Audio and video of the conversation will be recorded. After your interview is transcribed and analyzed, the researcher will submit a copy of the transcript and any identified themes and/or interpretations for to review for accuracy and agreement. Once this process is completed, the video/audio recording of the interview will be deleted.

BENEFITS TO YOU AND OTHERS

You may not get any direct benefit from this study, but, the information we learn from people in this study may help add to the literature in the field of research administration.

COSTS

There are no costs for participating in this study other than the time you will spend in the interview and the optional review of the resulting transcript and themes.

RISKS

There are no risks above a minimal risk. It is possible, though unlikely, that you may become upset talking about research pressures or lack of institutional support. In this event, a list of references to the higher education literature suggesting strategies for dealing with ever increasing productivity standards can be made available. There is a small risk for loss of confidentiality, which will be minimized by securely storing your data.

CONFIDENTIALITY

All personal identifying information will be kept in password protected files and these files will be deleted during or after the completion of the study, which is scheduled to be completed by May, 2018. Potentially identifiable information about you will consist of recorded audio and video images of the interview, your name and your institutional affiliation. Data is being collected only for research purposes. The audio and video recordings will not be included directly in the dissertation in any form, nor will any participants or institutions be named. No administrators or other individuals at your institutions will access to your responses.

There will be an individual Zoom Desktop Conferencing session for each interview that only the student researcher and the participant will have access. The participant will receive a link the hangout session via email. The recorded file will be stored on a password-protected secured VCU network drive. Once the interview has been transcribed and the member checking is complete, the files will be destroyed. The transcription will include a coded identifier (e.g. researcher 2 at university 6). The member check log will be stored in a password-protected excel file that will include the researcher's first name as well as his/her transcription code identifier. This log will also be destroyed after the project is completed.

VOLUNTARY PARTICIPATION AND WITHDRAWAL

Your participation in this study is voluntary. You may decide to not participate in this study. If you do participate, you may freely withdraw from the study at any time. Your decision to not participate or to withdraw will involve no penalty or loss of benefits to which you are otherwise entitled. If you choose to withdraw, the video/audio recording of your interview and any subsequent transcripts will be deleted.

QUESTIONS

If you have any questions, complaints, or concerns about your participation in this research, contact:

Michael M. Rawls
Virginia Commonwealth University
P.O. Box 842033
Richmond, VA
Telephone: (804) 828-1275 (work) (804) 291-7162 (cell)
Email:rawlsmm@vcu.edu

The researcher/study staff named above is the best person(s) to call for questions about your participation in this study.

If you have any general questions about your rights as a participant in this or any other research, you may contact:

Office of Research
Virginia Commonwealth University
800 East Leigh Street, Suite 3000
P.O. Box 980568
Richmond, VA 23298
Telephone: (804) 827-2157

Contact this number to ask general questions, to obtain information or offer input, and to express concerns or complaints about research. You may also call this number if you cannot reach the research team or if you wish to talk with someone else. General information about participation in research studies can also be found at http://www.research.vcu.edu/human_research/volunteers.htm.

Appendix G: Coding Scheme

Coding Rules:
Any portion of a respondent's reply that fits within the definition of the codes below should be labelled with the corresponding code.
Each time a code label is applied, it will be accompanied by one of the following qualifier coding labels: "Positive", "Negative", or "Neutral"
A response or part of a response can be tagged with as many code labels as necessary.

Code List	**Definitions**
Admin Burden - General	General administrative burdens that get in the way of research
Advantage	Any institutional support or service that relieves ancillary burdens of research - particularly the administrative processes
Budget	Related to pre-award budget construction
Bureaucracy	Burdensome rules or compliance requirements that create seemingly unnecessary complications
Champions	Any instance where university staff or administrators are cast in the light of someone who has the faculty's interests in mind and are capable of easing research-related burdens
Collaborators	Mention of collaboration
Compliance	Any mention - good or bad - of handling research compliance issues (COI, IRB, IACUC, etc.)
Core facility lab space	Lab space the research users that belong to the university or other colleagues
Department provided research support	To be used whenever a resource or service is provided by the department or school
Email overload	Mention of emails as major part of daily tasks
Faculty Recruitment	Any statement about the process of recruiting new faculty to the department
GRA Recruitment	Process and/or success of recruiting research assistants
GRA Reliance	Reliance on graduate research assistant in the process of conducting research.
GRAs and Publishing	Any indication of graduate research assistants collaborating in the authorship of journal articles

GRAs Funding	Any statement about how research assistants are paid in terms of the source of funds.
GRAs Quality	Any statement assessing the quality of the graduate student assistant pool or the abilities of existing graduate students.
Great Place	Any general statement about the program being a great place to work (do not include statements regarding collegiality among faculty, b/c everyone says that)
Impediment	Any institutional factor in the research process that hinders, slows, complicates, distracts or leads to faculty frustration
Indirects	Any mention of indirect fund distribution or use
Lab Space	Researchers personal or primary lab only
Library - Journal Articles	Any mention of the exent and accessibility of online journal articles
Library - Other Resources	Any other mention of library resources or services
Medical School	Med school interaction
Mentoring	Any mention of mentorship, or lack thereof, within a program
Online Resources	The quality, availability, and usability of online resources and systems that support research
Organizational Problems	Problems with leadership or climate that hinder productivity
Other support personnel	Any mention of other staff who support of research activity (e.g. IT, Equipment Specialists, admin support, etc.)
Post-award support	Any mention of support in the process of administering active grants (e.g. budgeting, reporting, etc.)
Postdoc	Any mention of the use of post-doctoral fellowship in support of research activity
Pre-award support	Any mention of support in the process of applying for grants.
Procurement	Related to procurement of equipment and supplies for the lab
Res. Metrics - Authentic	Goals, metrics, or expectations related to research that the subject finds authentic and aligned with their values
Res. Metrics - Inauthentic	Goals, metrics, or expectations related to research that the subject finds inauthentic and unaligned with their values
Seed Money	Internal funds to determine viability and then go after big money
Silo-breaking	Any policy, practice or organizational feature that encourages interdisciplinary research and collaboration
Silo-building	Any organizational barriers or attitudes that inhibit interdisciplinary research or collaboration

Strategy	Any university or department policy or practice designed to facilitate research productivity
Teaching load	Any mention of teaching responsibilities within the context competing with research and scholarship responsibilities
Tech Transfer	Any discussion about tech transfer, patenting, and commercialization of research
Translational	Translational values, goals, or activities, not the tech transfer itself
University provided research support	To be used whenever a resource or service is provided by the University

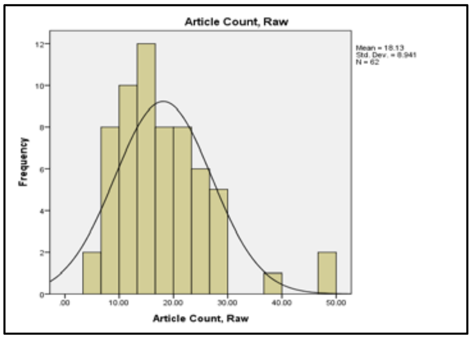
Question ID	Question Text	Question Category	Concept
Q_1	Please list some of the resources or services offered by your institution, school, or department that you find most helpful in aiding your productivity as a researcher?	Resources and support services	Advantages
Q_2	What do you consider to be the biggest barriers to your productivity?	Resources and support services	Impediments
Q_3	Do have adequate access to the lab space and equipment that you need to conduct your research?	Resources and support services	Lab Space
Q_4	Please describe quality of grant management support services that are offered by either your department and your Office of Research? (locating/ writing/ applying/ accounting/ procuring/ reporting)	Resources and support services	Grant Support
Q_5	How important are graduate assistants in aiding your research productivity?	Resources and support services	Graduate Assistants
Q_6	Are there other personnel in your department who provide significant assistance to your research efforts? (e.g. non-student research assistants, other support personnel, etc.)	Resources and support services	Other personnel
Q_7	Describe the role library resources play in aiding your research productivity?	Resources and support services	Library resources
Q_8	Are there any other resources or supports services that we haven't discussed that you believe have an impact on your research productivity, either positively or negatively? (e.g. technology transfer, patents application, facilitation of peer networking opportunities, etc.)	Resources and support services	n/a
Q_9_1	How would you describe your department's identity as a group?	Department Culture	Group Identity
Q_9_2	Do all feel that all departmental members contribute to decision making in a substantive manner?	Department Culture	Assertive Participative Governance
Q_9_3	How would you describe department morale and interrelations between individual faculty members?	Department Culture	Group Climate
Q_10_1	How would you describe your department's emphasis on research, compared to other priorities?	Climate for Research	Research Emphasis
Q_10_2	Does your department establish collective research goals or strategies?	Climate for Research	Coordination Function
Q_10_3	Do you find that the productivity of you peers influences your own productivity? If so, how?	Climate for Research	Research Productive Peers
Q_10_4	Do you find that you have sufficient uninterrupted time to dedicate to scholarly activities?	Climate - General	Sufficient Work Time
Q_11_1	Do you view your chair an accomplished researcher?	Climate - Leadership	Leadership - Accomplished
Q_11_2	Does your chair use his/her expertise to directly support your research or the research efforts of your colleagues (e.g. technical support, professional contacts, mentoring, etc.)?	Climate - Leadership	Leadership - Mentoring
Q_11_3	Does your chair clearly communicate an overall vision for your department in terms of research goals? If so, is that vision frequently reinforced? Do you find it authentic and aligned with your own values?	Climate - Leadership	Leadership - Authentic Vision
Q_12	Are there any important contributors to research productivity in your department that we haven't addressed?	Closing	Advantages
Q_13	Are there any barriers to you research productivity that we haven't discussed?	Closing	Impediments

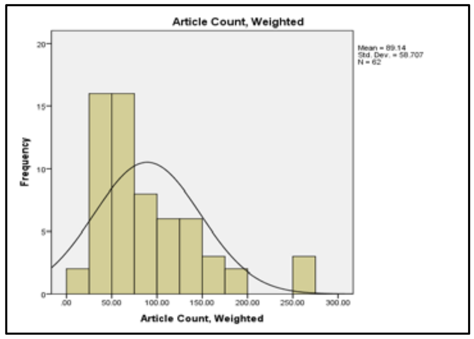
Appendix H: Member Check Log

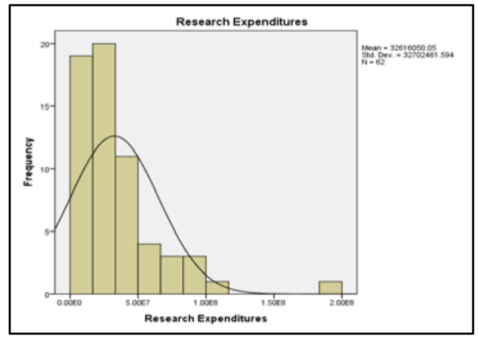
Research Participant	Transcript Sent for Review	Participant Comments	Actions/Notes
A1	Yes	My pleasure. Looks good to me. Please let me know if you need any clarification on anything. Good luck with your dissertation!	
A2	Yes	NO REPLY	
A3	Yes	I went through the transcript and edited to make sense and convey what I was trying to say. Let me know if you have any questions. Good luck with the study. I would be interested in reading it when completed.	Reviewed edits to determine if any substance changes had been made. None were found. Made sure to use edited version of transcript for any direct quotes
A4	Yes	NO REPLY	
B1	Yes	NO REPLY	
B2	Yes	NO REPLY	
B3	Yes	NO REPLY	
B4	Yes	NO REPLY	
B5	Yes	Mike: I have gone through this and I have taken out parts that I think will enable others to identify the institution. I have also made some other comments/edits.	Made sure to use edited version of transcript for any direct quotes
C1	Yes	NO REPLY	
C2	Yes	There were indeed some minor typos, but the heart of the conversation seems intact. So as long as my name and institution aren't being used, I have no problem with this transcript being used for your study.	
C3	Yes	NO REPLY	
C4	Yes	boy am I good at complaining :) This is actually pretty funny to read. I guess if anything, can you change BME to "department"	I indicated that I would use BME generically to describe all of the departments and that the study clearly identified BME as the discipline of focus, so there was no way to conceal the type of department that he was in.

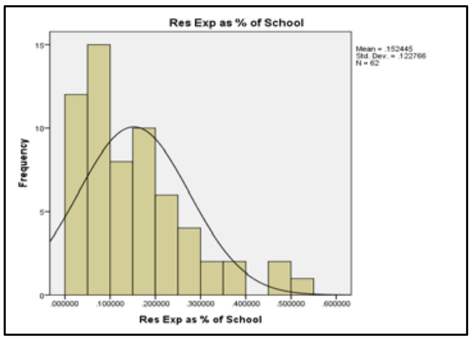
D1	Yes	no problems	Participant also corrected the name of a center at his instiutution, but the center is not named in the dissertation and the portion of his reply that offered the correction was not included here as it should be considered identifiable data.
D2	Yes	NO REPLY	
D3	Yes	The attached has some minor corrections, but generally accurate.	Made sure to use edited version of transcript for any direct quotes
E1	No	n/a	Institution E was excluded due to inability to recruit a sufficient number of participants.

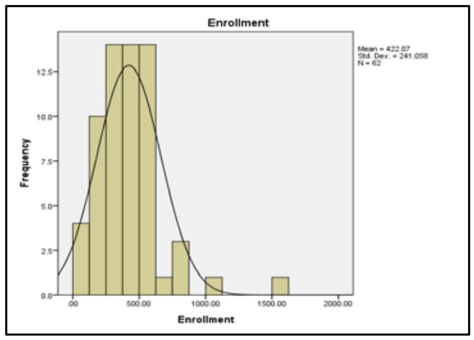
Appendix I: Histograms of Untransformed Variables

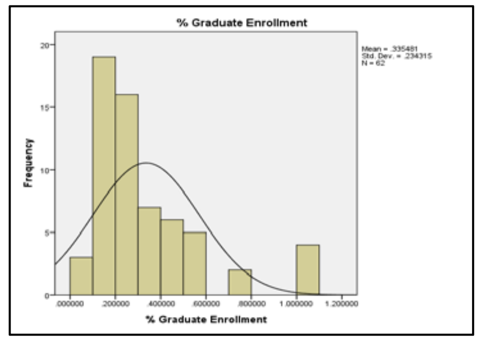


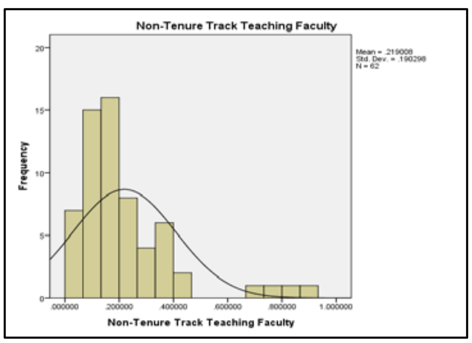


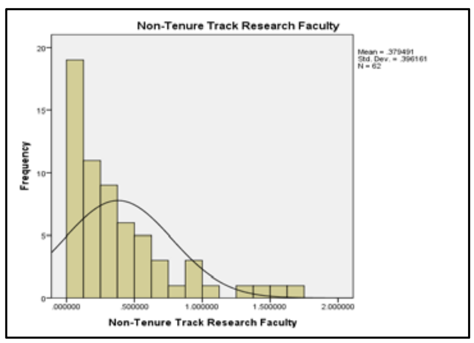


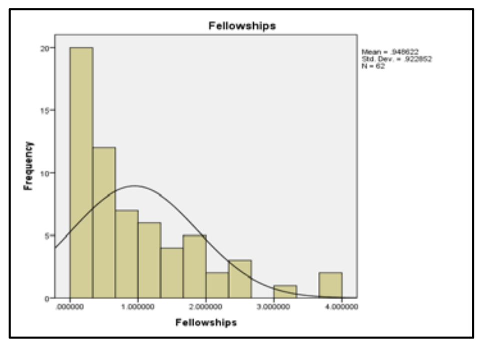


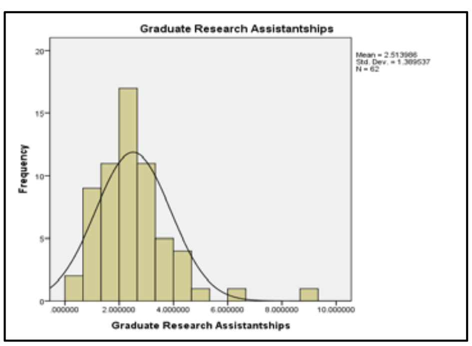


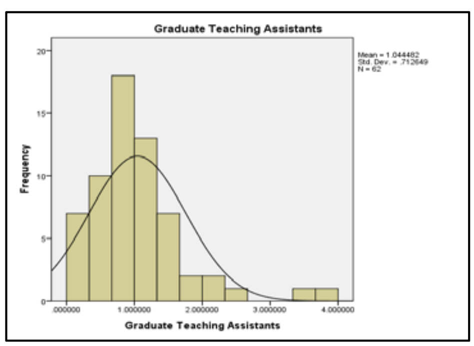


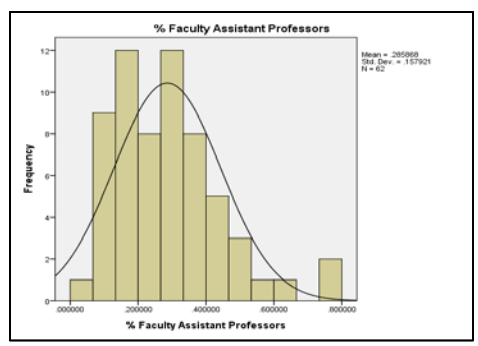


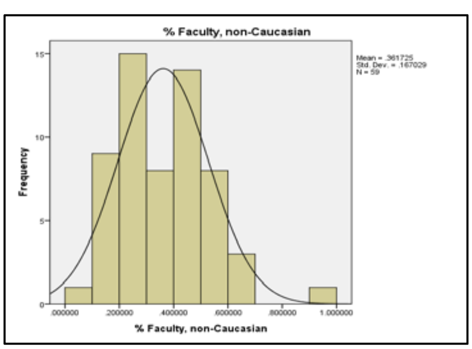


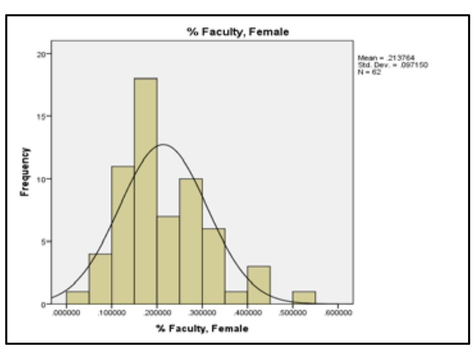


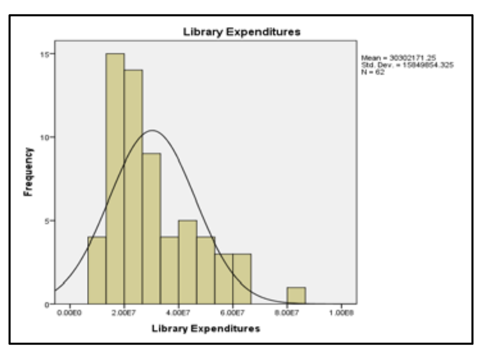












Appendix J: Regression Output for Model 1

```
REGRESSION
 /MISSING LISTWISE
 /STATISTICS COEFF OUTS R ANOVA COLLIN TOL ZPP
 /CRITERIA=PIN(.05) POUT(.10)
 /NOORIGIN
 /DEPENDENT CURT_AC_FC_IF
 /METHOD=ENTER CURT_BME_RE CURT_Avg_Enrollment.
```

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	Enrollment (cu.rt.), Research Expenditures (cu.rt.)[b]	.	Enter

a. Dependent Variable: Article Count, Weighted (cu.rt.)

b. All requested variables entered.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.505[a]	.255	.230	.80054

a. Predictors: (Constant), Enrollment (cu.rt.), Research Expenditures (cu.rt.)

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	12.951	2	6.475	10.104	.000[b]
	Residual	37.811	59	.641		
	Total	50.762	61			

a. Dependent Variable: Article Count, Weighted (cu.rt.)

b. Predictors: (Constant), Enrollment (cu.rt.), Research Expenditures (cu.rt.)

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients			Correlations			Collinearity Statistics	
		B	Std. Error	Beta	t	Sig.	Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	3.292	.558		5.903	.000					
	Research Expenditures (cu.rt.)	.005	.001	.530	4.361	.000	.500	.494	.490	.856	1.168
	Enrollment (cu.rt.)	-.052	.081	-.078	-.645	.522	.122	-.084	-.072	.856	1.168

a. Dependent Variable: Article Count, Weighted (cu.rt.)

Appendix K: Regression Output for Model 2

```
REGRESSION
 /MISSING LISTWISE
 /STATISTICS COEFF OUTS R ANOVA COLLIN TOL ZPP
 /CRITERIA=PIN(.05) POUT(.10)
 /NOORIGIN
 /DEPENDENT CURT_AC_FC_IF
 /METHOD=ENTER CURT_BME_RE CURT_Avg_Enrollment CURT_New_GRA CURT_PC_NT_RFAC CURT_PC_Fellows.
```

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	Fellows (cu. rt.), Enrollment (cu.rt.), Non-Tenure Research Faculty (cu. rt.), Research Expenditures (cu.rt.), Graduate Research Assistants (cu.rt.)[b]	.	Enter

a. Dependent Variable: Article Count, Weighted (cu.rt.)

b. All requested variables entered.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.674[a]	.454	.405	.70371

a. Predictors: (Constant), Fellows (cu. rt.), Enrollment (cu.rt.), Non-Tenure Research Faculty (cu.rt.), Research Expenditures (cu.rt.), Graduate Research Assistants (cu.rt.)

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	23.030	5	4.606	9.301	.000[b]
	Residual	27.731	56	.495		
	Total	50.762	61			

a. Dependent Variable: Article Count, Weighted (cu.rt.)

b. Predictors: (Constant), Fellows (cu. rt.), Enrollment (cu.rt.), Non-Tenure Research Faculty (cu.rt.), Research Expenditures (cu.rt.), Graduate Research Assistants (cu.rt.)

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients			Correlations			Collinearity Statistics	
		B	Std. Error	Beta	t	Sig.	Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	1.810	.639		2.834	.006					
	Research Expenditures (cu.rt.)	.004	.001	.391	3.497	.001	.500	.423	.345	.781	1.281
	Enrollment (cu.rt.)	-.108	.073	-.163	-1.482	.144	.122	-.194	-.146	.808	1.237
	Graduate Research Assistants (cu.rt.)	1.578	.526	.398	3.003	.004	.563	.372	.297	.554	1.805
	Non-Tenure Research Faculty (cu.rt.)	-.245	.325	-.076	-.753	.454	.041	-.100	-.074	.957	1.045
	Fellows (cu. rt.)	.361	.331	.136	1.092	.279	.420	.144	.108	.632	1.582

a. Dependent Variable: Article Count, Weighted (cu.rt.)

Appendix L: Regression Output for Model 3

REGRESSION

/MISSING LISTWISE

/STATISTICS COEFF OUTS R ANOVA COLLIN TOL ZPP

/CRITERIA=PIN(.05) POUT(.10)

/NOORIGIN

/DEPENDENT CURT_AC_FC_IF

/METHOD=ENTER CURT_BME_RE CURT_Avg_Enrollment CURT_BME_RE_All CURT_Avg_Tot_Lib_Exp.

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	Library Expenditures (cu.rt.), Enrollment (cu.rt.), Res Exp as % of School (cu. rt.), Research Expenditures (cu.rt.)[b]	.	Enter

a. Dependent Variable: Article Count, Weighted (cu.rt.)

b. All requested variables entered.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.650[a]	.422	.382	.71728

a. Predictors: (Constant), Library Expenditures (cu.rt.), Enrollment (cu.rt.), Res Exp as % of School (cu.rt.), Research Expenditures (cu.rt.)

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	21.435	4	5.359	10.416	.000[b]
	Residual	29.326	57	.514		
	Total	50.762	61			

a. Dependent Variable: Article Count, Weighted (cu.rt.)

b. Predictors: (Constant), Library Expenditures (cu.rt.), Enrollment (cu.rt.), Res Exp as % of School (cu.rt.), Research Expenditures (cu.rt.)

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations			Collinearity Statistics	
		B	Std. Error	Beta			Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	1.489	.866		1.720	.091					
	Research Expenditures (cu.rt.)	.004	.001	.485	3.559	.001	.500	.426	.358	.546	1.831
	Enrollment (cu.rt.)	-.038	.075	-.057	-.505	.616	.122	-.067	-.051	.798	1.253
	Res Exp as % of School (cu.rt.)	-.517	.734	-.087	-.705	.484	.098	-.093	-.071	.659	1.518
	Library Expenditures (cu. rt.)	.007	.002	.390	3.599	.001	.501	.430	.362	.863	1.158

a. Dependent Variable: Article Count, Weighted (cu.rt.)

Appendix M: Regression Output for Model 4

```
REGRESSION
 /MISSING LISTWISE
 /STATISTICS COEFF OUTS R ANOVA COLLIN TOL ZPP
 /CRITERIA=PIN(.05) POUT(.10)
 /NOORIGIN
 /DEPENDENT CURT_AC_FC_IF
 /METHOD=ENTER CURT_BME_RE CURT_Avg_Enrollment CURT_New_GRA CURT_PC_NT_RFAC CURT_PC_Fellows
  CURT_BME_RE_All CURT_Avg_Tot_Lib_Exp.
```

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	Library Expenditures (cu.rt.), Enrollment (cu.rt.), Non-Tenure Research Faculty (cu. rt.), Res Exp as % of School (cu. rt.), Fellows (cu. rt.), Graduate Research Assistants (cu.rt.), Research Expenditures (cu.rt.)[b]	.	Enter

a. Dependent Variable: Article Count, Weighted (cu.rt.)

b. All requested variables entered.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.740[a]	.547	.488	.65244

a. Predictors: (Constant), Library Expenditures (cu.rt.), Enrollment (cu.rt.), Non-Tenure Research Faculty (cu.rt.), Res Exp as % of School (cu.rt.), Fellows (cu. rt.), Graduate Research Assistants (cu.rt.), Research Expenditures (cu. rt.)

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	27.775	7	3.968	9.321	.000[b]
	Residual	22.987	54	.426		
	Total	50.762	61			

a. Dependent Variable: Article Count, Weighted (cu.rt.)

b. Predictors: (Constant), Library Expenditures (cu.rt.), Enrollment (cu.rt.), Non-Tenure Research Faculty (cu.rt.), Res Exp as % of School (cu.rt.), Fellows (cu. rt.), Graduate Research Assistants (cu.rt.), Research Expenditures (cu.rt.)

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.	Correlations Zero-order	Correlations Partial	Correlations Part	Collinearity Statistics Tolerance	Collinearity Statistics VIF
1	(Constant)	.217	.881		.246	.807					
	Research Expenditures (cu.rt.)	.003	.001	.304	2.284	.026	.500	.297	.209	.474	2.111
	Enrollment (cu.rt.)	-.072	.069	-.109	-1.041	.303	.122	-.140	-.095	.768	1.303
	Graduate Research Assistants (cu.rt.)	1.361	.510	.344	2.668	.010	.563	.341	.244	.505	1.978
	Non-Tenure Research Faculty (cu.rt.)	-.254	.304	-.079	-.835	.407	.041	-.113	-.076	.939	1.065
	Fellows (cu. rt.)	.301	.307	.113	.980	.331	.420	.132	.090	.630	1.588
	Res Exp as % of School (cu.rt.)	.296	.703	.050	.421	.676	.098	.057	.039	.593	1.686
	Library Expenditures (cu. rt.)	.006	.002	.331	3.315	.002	.501	.411	.304	.840	1.191

a. Dependent Variable: Article Count, Weighted (cu.rt.)

Appendix N: Regression Output for Model 5

```
REGRESSION
 /MISSING LISTWISE
 /STATISTICS COEFF OUTS R ANOVA COLLIN TOL ZPP
 /CRITERIA=PIN(.05) POUT(.10)
 /NOORIGIN
 /DEPENDENT CURT_AC_FC_IF
 /METHOD=ENTER CURT_BME_RE CURT_Avg_Enrollment CURT_New_GRA CURT_Avg_Tot_Lib_Exp.
```

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	Library Expenditures (cu.rt.), Enrollment (cu.rt.), Graduate Research Assistants (cu.rt.), Research Expenditures (cu.rt.)[b]	.	Enter

a. Dependent Variable: Article Count, Weighted (cu.rt.)

b. All requested variables entered.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.731[a]	.534	.501	.64410

a. Predictors: (Constant), Library Expenditures (cu.rt.), Enrollment (cu.rt.), Graduate Research Assistants (cu.rt.), Research Expenditures (cu.rt.)

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	27.114	4	6.779	16.339	.000[b]
	Residual	23.647	57	.415		
	Total	50.762	61			

a. Dependent Variable: Article Count, Weighted (cu.rt.)

b. Predictors: (Constant), Library Expenditures (cu.rt.), Enrollment (cu.rt.), Graduate Research Assistants (cu.rt.), Research Expenditures (cu.rt.)

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations			Collinearity Statistics	
		B	Std. Error	Beta			Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	.255	.698		.365	.716					
	Research Expenditures (cu.rt.)	.003	.001	.334	3.235	.002	.500	.394	.292	.765	1.307
	Enrollment (cu.rt.)	-.083	.067	-.125	-1.238	.221	.122	-.162	-.112	.804	1.243
	Graduate Research Assistants (cu.rt.)	1.537	.406	.388	3.782	.000	.563	.448	.342	.776	1.289
	Library Expenditures (cu. rt.)	.006	.002	.327	3.421	.001	.501	.413	.309	.896	1.116

a. Dependent Variable: Article Count, Weighted (cu.rt.)