DISSERTATION

THREE ESSAYS ON THE ECONOMICS OF UNIVERSITY KNOWLEDGE PRODUCTION AND COMMERCIAL INNOVATION: THE CASE OF COLORADO STATE UNIVERSITY RESEARCH AND TECHNOLOGY TRANSFER

Submitted by

Yoo Hwan Lee

Department of Agricultural and Resource Economics

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Summer 2016

Doctoral Committee:

Advisor: Gregory D. Graff
Co-Advisor: Dana L.K. Hoag

Stephen R. Koontz
David W. Mushinski

ABSTRACT

THREE ESSAYS ON THE ECONOMICS OF UNIVERSITY KNOWLEDGE PRODUCTION AND COMMERCIAL INNOVATION: THE CASE OF COLORADO STATE UNIVERSITY RESEARCH AND TECHNOLOGY TRANSFER

The central aim of this study is to analyze the research production and R&D activities of Colorado State University (CSU) across its different colleges, departments and other research units, and to evaluate how those activities impact the Colorado economy's agriculturally-related sectors. The study consists of three main chapters, to introduce the dynamics of university knowledge transfer to local agricultural economies.

Chapter 1 explores CSU research production and technology-transfer activities, using a unique panel date set for each of 54 academic departments over the period of 1989-2012. In order to estimate the empirical knowledge-production function (KPF), this chapter attempts to build a negative binomial panel regression model with a polynomial distributed lags (PDL) of research expenditures. Three categories of research outputs are modelled, including (1) published journal articles, (2) industry collaboration, and (3) technology transfer mechanisms. In the regression results, publications are clearly the most common research outputs of the university, with a more systematic relationship between research inputs and publications than the other two types of research outputs. Moreover, it appears to exhibit decreasing returns to scale, whereas the collaborative and tech transfer research outputs appear to show increasing returns to scale. In the results of a seemingly unrelated regressions (SUR) model among the three different types of research outputs, publications and the tech transfer mediated research outputs are the primary research outputs in the university and have the maximum impact from past research expenditures.

Furthermore, results indicate that collaboration mediated outputs are substitutional relative to the more formal tech transfer outputs.

Chapter 2 explores the agency of knowledge production, viewing scientific research teams as "quasi-firms" arising as independent knowledge-creating entities within the university context. First, the findings from the ego-centric social network showed that the participation of outside members makes it possible to increase the size of the ego-centric teams and the growth patterns of the percent share are an obvious parallel to the patterns of team size. Particularly, the growth rate of team size is opposite of the percent share of ego's home department co-authors with upward and downward tendencies, respectively. Second, the findings from the regression results showed that the number of CSU departments per team is statistically significant in the team's assembly mechanism for both the article teams and patent teams. Thus, it seems reasonable to conclude that cross-functional team formation is more effective and common in the university research team formation and has a positive impact on the size of research teams. Finally, the quality of research teams' knowledge production tells us that the group with multiple departments per article has a higher research impact than a group using a single department per article. By the same token, larger-sized teams have higher impacts than relatively smaller-sized teams, as well as field variety. The group with multiple references per patent had a higher impact than the groups with a single or no reference per patent. This result tells us that the citation mapping from backward citations to forward citations is a significant factor for testing the research teams' impacts on the economic and social benefits with respect to knowledge spillover.

Chapter 3 has focused on CSU's knowledge spillovers within agriculturally related fields and technologies. The findings indicated that academic knowledge spillovers are geographically bounded, but they are not strictly limited to the regional scale. Crucially, the impact of university spillovers on agriculturally-related industries depends upon which type of knowledge dissemination channel is utilized by university researchers. Broadly speaking this chapter evaluates four types of channel—including the publication mechanism, the industry collaboration and extension mechanism, the technology patenting/licensing mechanism, and the venture creation mechanism—each of which are variously adapted to transmitting different degrees of sticky (tacit) versus slippery (codified) knowledge. The results showed that in both aggregate level of technology and six different technological categories, the spillover impacts of journal publications, are rarely localized within Colorado; rather, the geographic scope of these impacts are national and even global. However, the extent to which the spillover impacts of patented knowledge is localized within Colorado is open to question because it is possible to control permissions for use, but at the same time it is impossible to limit everyone's awareness and use of it, particularly in foreign jurisdictions where patents are not taken out by the university. However, the collaboration mechanism requires closer interaction and greater geographic proximity, which usually prevents global dissemination. Thus, we observe geographic proximity is significantly important for these channels. However, there are even distinctions within these. For example, we find industry coauthorship on articles to be less likely to be localized than privately sponsored grant awards. Nevertheless, the stickiness of these channels might depend also on the different technological categories. As mention as above, the geographic proximity is important only in aggregate level of technology, but it can be varied across the different technological categories, especially the slippery form of knowledge in animal health and nutrition health technology. Finally, university start-ups are highly geographically bounded near universities because in the early stages start-up companies need support from their host university.

ACKNOWLEDGEMENTS

I would like to sincerely thank my advisor, Professor. Gregory D. Graff, for supporting and advising me during my doctoral program years. Dr. Graff showed commitment to and put his heart into my dissertation and long-term career goals. He always encouraged me to improve the most important skills to me as an economist and to grow as a creative and independent researcher. Particularly, I really love his academic leadership and economic intuitions, especially regarding technological innovation and the economics of entrepreneurship. Moreover, his mentorship is paramount in advising his students, especially his international students. Dr. Graff deeply and truly understands his students' problems and concerns, and he always endeavors to solve their problems. Whenever I suffered a hard time in my academic journey, he always cheered me up wholeheartedly and advised me how to resolve the problem. I would like to thank him again and to say that *I could never have achieved my doctoral degree without his advice and encouragement*.

I would also like to thank my co-advisor, Professor. Dana Hoag, for his advice and insightful suggestions on my dissertation and for his letter of recommendation for my job applications. He is the person I met first at Colorado State University, and he hooded me at the commencement ceremonies on behalf of Dr. Graff. Moreover, I am also grateful to my other dissertation committee members at Colorado State University, Professor. David Mushinski and Stephen Koontz, for their time and valuable feedback on a preliminary version of my dissertation, especially relating to their econometric knowledge.

I would like to thank the members of the Department of Agricultural and Resource Economics at Colorado State University, including the professors (Dr. Gregory Perry and Dr. Dawn Thilmany), the administrative staffs (Ms. Denise Davis and Donna Sosna), and my colleagues, especially

Ghulam Samad, Annabelle Berklund, Neama Lariel, Chuba Suntharlingam, and Jakrapun Suksawat (Ton), for their love and support.

I would like to thank the members of CSU Ventures, Dr. Todd Headley and Ms. Sarah Belford, for sharing their valuable dataset and our insightful discussions. I am also grateful to the members of the Office of the Vice President for Research at Colorado State University, Dr. Alan S. Rudolph and Dr. Pam Harrington, for sharing their valuable dataset, discussions, and suggestions.

Finally, I thank my beloved family, my parents in Gumi, parents-in-law in Suji, my sister (Yoo Jin), my nephew (Jun Hyeok), my sister-in-law (Hye Won), for their endless love and commitment. I especially thank my lovely wife (Hye Jun) for her patience, love, and faith. I also thank my little princess: my daughter (Yerin_Manna). In addition, I would like to thank my spiritual family: Pastor Paul J. Gim and Ms. Youngin Lee in the Korean Presbyterian Church of Lawrence for their prayer and love, and the other church family members: Pastor Gi Hyun Park and the other members of the First Korean Church of Fort Collins for also their prayer, love and support.

DEDICATION

I dedicate this dissertation to my Lord and Savior Jesus Christ for the unfailing love and eternal life. “*Surely God is my salvation; I will trust and not be afraid. The LORD, the LORD, is my strength and my song; he has become my salvation*.", (Isaiah 12:2, NIV)

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

OVERALL INTRODUCTION

In recent years, numerous studies have explored the economic impacts of university research and technology transfer on industrial research and development (R&D) and commercial innovation. These studies have contributed to the understanding that university research leads to positive social returns as well as economic growth through stimulating commercial innovation (Adams & Griliches, 1998; Cohen, Nelson, & Walsh, 2002; Jaffe, 1989).

The importance of university technology transfer has been addressed by economists for only a few decades. Of particular importance in this regard is the potential of university knowledge production activities to affect both commercial innovation directly and economic growth indirectly. According to Adams & Griliches (1998), universities account for about 50 percent of basic research conducted in the U.S., and basic research is one of the mainsprings of industrial innovation. If universities' research productivities were to decline overall or were to shift in composition away from basic research, a critical input into industrial research would grow more slowly. Following Cohen, Nelson, & Walsh (2002), university research is of at least moderate importance to a number of industries, including both high technology and more traditional industries. Studies by Mansfield (1991 & 1995), Henderson, Jaffe, & Trajtenberg (1998), Agrawal & Henderson (2002), and others, have sought to model and measure the contributions of academic research to industrial innovation through different knowledge transmission channels and different modes of impact. Moreover, these seminal studies have shown that technological innovation across various industries, including both new products and new processes, have been based on to some extent on academic research. At the same time these studies emphasize the importance of the role of university research in the creation of human capital, arguably the most significant resource in industry and society, through the training of graduate students, many of whom become industrial scientists, engineers, and inventors.

However, these roles and missions of the research university have been transformed by histories of university institutional and national government policy changes. Before the nineteenth century, the function of the university was primarily teaching, i.e. the preservation and dissemination of knowledge. Beginning in the nineteenth century, the university was transformed into the research university, in what is often called the first academic revolution in which a research mission was integrated with the teaching mission (Etzkowitz, 2003). In the early- to mid-twentieth century, the third mission of university emerged, emphasizing outcomes such as economic and social development, in which is called the second academic revolution (Etzkowitz, 2003). In the course of changing their roles and missions, universities have dealt with important policy changes. Although changes in the roles and missions of the university, and changes in policy have been interdependent, with causality flowing in both directions, in general, policy changes have accelerated the speed of university transformation and supported or suggested the dissemination of new missions across the population of universities. In the U.S., universities have encountered several important policy changes, such as the Morrill Land-Grant Act of 1862, the Hatch Act of 1887, the Smith-Lever Act of 1914, the Bayh-Dole Act of 1980, and so on.

Following these policy changes, the Land Grant universities in the U.S. have generally come to embrace three missions: an educational mission, a research mission, and an outreach mission. First, the educational mission is to prepare and educate knowledge workers for statewide, nationwide, and even worldwide labor markets. The university graduates become knowledge workers--whether scientists, engineers, faculty members, or political leaders--contributing the most important factor of production in the knowledge economy, human capital. Second, the research mission is to produce new knowledge, which can be observed as outputs such as journal articles, student theses and dissertations, or even invention disclosures and patent applications. The university output of new knowledge can contribute to industrial innovation and technological progress both indirectly through basic research and directly through development of high technology. Finally, the outreach mission encompasses not only the traditional forms such as public lectures, extension services, and consulting and collaborative R&D with industry partners, but also newer forms such as the licensing of patents and the startup of new companies.

However, in most previous studies of academic knowledge transfer, some important knowledge indicators such as publications, degree awards, private sponsorships, and extension activities are far less fully explored and catalogued than are patents, licenses, and royalties. Furthermore, relatively few studies have been devoted to understanding the multiple knowledge channels of the university, but most of them have been highly skewed toward just one or two knowledge dissemination channels, because of data limitations. This study takes into consideration a full range of potential knowledge dissemination channels, as indicated by such measures as academic journal publications, degrees awarded (master and doctoral), industry co-authorship on publications, private sponsorship of research projects, extension activities, invention disclosures, patent applications, licenses, and startups, all from data collected at the level of the different departments or research units of Colorado State University, from 1989 to 2012. These various measures make it possible to analyze the extent to which the different types of knowledge dissemination channels work.

The concept of multiple knowledge channels in this study can be understood as informed by Samuelson's (1954) classic comparison of excludability and rivalry dimensions in "The Pure Theory of Public Expenditure." As such, four types of knowledge channels are introduced (see

Figure 1). First, what we call the "public domain mechanism" of knowledge dissemination is most appropriate for those outputs of research that have the strongest public good attributes, defined as non-excludable and non-rivalrous. Thus, it is impossible to exclude anyone from accessing this knowledge and to prevent simultaneous use or access, which means full freedom of use and open access.

Second, the "collaboration mechanism" of knowledge dissemination is defined as a common good that is non-excludable and, yet, more rivalrous, due to the tacit or "sticky" nature of the knowledge, with it skills or routines, thus preventing global dissemination via publication or even making simultaneous contact or contract with multiple parties less than effective. Dissemination or transfer of the knowledge requires close interaction, such as apprenticeship or collaboration.

	Non-Rivalrous ("Slippery" information, lower transaction costs, lower capacity requirements)	**Rivalrous** ("Sticky" information, higher transaction costs, higher capacity requirements)
Non-excludable	**Release via Public Domain**	**Collaboration**
Invoked exclusion	**Patenting/Licensing**	**Venture Creation**

Figure 1—The concept of Four Types of Knowledge Dissemination Channels

Third, the utilization of the intellectual property rights and contracts characterize the "patenting/licensing mechanism" of knowledge transfer, which is best suited when a certain degree of excludability is required to create sufficient incentives for follow-on investment in an otherwise non-excludable and non-rivalrous knowledge output. Thus, by virtue of the IP and contracts, it is possible to exclude others from accessing and making use of it. Finally, the "venture creation mechanism" of knowledge dissemination is best for raising private investment in the further development of knowledge treating it most like a private good, by virtue of IP making it relatively excludable and, by virtue of its intrinsic stickiness or context dependence, being relatively non-rivalrous, so it is possible to exclude others from accessing this knowledge and making simultaneous use of it.

The main goal of this study is to analyze the economics of Colorado State University (CSU)'s knowledge production and its spillover activities on the commercial innovation and invention. There are three main objectives: (1) to understand the historical trends and system of university knowledge production and R&D activities across the different colleges, departments, and research units in terms of a full range of potential knowledge dissemination channels, (2) to understand recent changes in the organization of knowledge creation activities, called the rise of the "entrepreneurial university" by Etzkowitz et al (2000), such as university research team structures and assembly mechanisms, and (3) to understand the dynamics of university knowledge spillovers and the geographic scope of their commercial and economic impacts on our society and economy, especially in the agriculturally-related sectors in Colorado and beyond.

This study consists of three main chapters, and each chapter pursues the three objectives of the study in different ways. Chapter 1 analyzes the system of CSU's knowledge production and transfer activities. This chapter introduces several advances to the literature. First, while previous studies have limited their analysis, due to data, to the institutional level, this analysis shows how knowledge production works at the more disaggregated level of the different colleges, departments, and research units within the university. Second, most other studies explore the relationship between research inputs and one type of output, typically research papers or patents. This analysis considers production of several different knowledge outputs that are disseminated through different channels and analyzes how the input output relationship differs across them. This chapter addresses questions of how efficiency of knowledge production can be assessed across the different types of knowledge, why they have different productivity, and what important factors may be related to improving productivity. Finally, this first chapter introduces a novel empirical technique for estimating the knowledge production function that combines a panel count data model with polynomial distributed lags between the inputs and the outputs of the knowledge production function. This approach makes it possible to estimate the input-output relationships systematically across the entire university at the disaggregate level of analysis, and for each of the different types of knowledge outputs.

Chapter 2 explores how knowledge production within the research university works, by analyzing the structure and composition of academic research teams as quasi-firms or the agents of innovation. This is a speculative and preliminary exploration into new concepts about the agency and organization of early-stage innovation, an attempt to look inside the black box of the production function. To this purpose, this study aims to examine what are the main components and structures of research teams, who can be characterized as principals and as agents of the research production process, what kinds of research environments or other factors determine the sizes of scientific research teams, and what sorts of team structures characterizes the relationship between university researchers and industry collaborators.

Chapter 3 examines CSU's knowledge dissemination and its impact on commercial innovation through the several different types of knowledge dissemination channels introduced above. This chapter focuses on the geospatial patterns of CSU's knowledge spillovers and its commercial economic impact especially within Colorado's agriculturally related sectors and industries, but also nationally and globally. This chapter describes CSU's publications, its collaborative activities, such as indicated by data on university-industry co-authorship of articles and private sponsorship of research, its licensing of inventions, and its startup of new ventures in agricultural and food technologies. This chapter traces the locations of private industry innovators impacted by or associated with CSU agricultural research across different types of technology, and a systematic relationship between the types of CSU knowledge transfer mechanisms employed the geographic proximity to CSU of the commercial user of that knowledge.

CHAPTER 1. EMPIRICAL EVIDENCE OF THE UNIVERSITY KNOWLEDEG PRODUCTION FUNCTION AND KNOWLEDGE DISSEMINATION ACTIVITIES: THE CASE OF COLORADO STATE UNIVERSITY

I. Introduction

Knowledge production and transfer activities are an important source of commercial innovation and economic development, such that knowledge spillovers have attracted the attention of many economists and policy makers. Furthermore, university research accounts for 50 percent of basic research in the U.S., and the university research is one of the mainsprings of industrial innovation, and the introduction of new products and processes (Mansfield, 1991, 1995; Adams and Griliches, 1998). Although studies of knowledge production had developed steadily for decades, around the 1980s the area blossomed with seminal papers by Griliches (1979), Pakes & Griliches (1980, 1984), Hausman, Hall, & Griliches (1984, 1986), Jaffe (1989), and Pardey (1989). Since that time, the main conceptual ideas and empirical techniques for analyzing knowledge production have progressed, as the quantity of studies has increased. In most previous studies, including those of both industrial research and development (R&D) and university research, empirical results of the input-output relationship between past research expenditures and current research outputs, respectively, remain ambiguous (Crespi & Geuna, 2008). This is largely due to two major problems: the quality of data and empirical model misspecifications.

In the United States, total R&D spending in 2012[1] was $447 billion of which academic R&D expenditures[2] were $66 billion, nearly 15 percent of the total. According to the National Science

[1] Source: 2014 Global R&D Funding Forecast, Battelle, R&D Magazine, International Monetary Fund, World Bank, CIA Fact Book.
[2] Source: Rankings by total academic R&E expenditures, National Center for Science and Engineering Statistics (NCSES), National Science Foundation (NSF).

Foundation, the industry investment was the largest performance of U.S. total R&D, accounting for 71 percent, and the research universities were the second largest performer (NSF, 2015). Furthermore, from 1996 to 2010, the economic impact of university and non-profit patent licensing was estimated to be $388 billion on the U.S. gross domestic product and $836 billion on the U.S. gross industrial outputs[3]. Since 1980, startup companies affiliated with universities were more than 4,000. In the same period, universities and non-profit entities' patent licensing created more than 3 million jobs in U.S. (AUTM, 2015).

There is a growing body of evidence that academic research and technology transfer activities are an important source of ideas and inputs for commercial research and innovation, including both basic and high technology industries (Cohen et al, 2002). Yet, many of the recent studies of academic knowledge output and its commercial impact focus on patenting and licensing as the newer, more controversial channels of knowledge dissemination by academics. Although patents are a significant output of academic knowledge production and represent one channel of its impact on economic growth and development, other research outputs such as journal articles, degree awards, and more traditional relationships of university-industry collaboration[4] are still the major channels by which research universities affect commercial activity (Agrawal & Henderson, 2002). To create a more accurate holistic picture, analysis of the impacts of academic knowledge production should take into account the full range of potential knowledge dissemination channels. Also, previous studies of academic knowledge production and transfer have focused their analyses at the institutional level, largely because of data limitations. Again, a more accurate picture of academic knowledge production should be possible from a more disaggregated, yet systematic

[3] Source: the AUTM Briefing Book: 2015.
[4] Collaborative channels include consulting to state governments and local industries, seminars and workshops, informal conversation, co-supervising, personnel exchange, and so on.

accounting of research activities within the diverse structure of the contemporary research university.

The main purpose of this chapter is to analyze the disaggregated system of knowledge production and transfer activities within a major research university. In so doing, this chapter introduces a uniquely detailed dataset together with a novel empirical technique for estimating the knowledge production function. First, this chapter introduces and explores data that characterizes the full range of knowledge production inputs and outputs across all of the available accounting units—the different colleges, departments, and research units—of our home institution, Colorado State University (CSU), over more than two decades, accessed from both open access and private sources within the administration of CSU, to explore some basic questions: how does this relationship differ for the different types of knowledge output that impact the economy through different knowledge transfer or dissemination channels; how does the efficiency of knowledge production vary across the different disciplines and research units of the university; and what factors might explain why they vary in output productivity.

Second, this chapter introduces a panel count data model with a polynomial distributed lag scheme to estimate the input-output relationship of knowledge production using the CSU knowledge production and transfer dataset. This unique data allows for more detailed modelling of systematic input-output relationship of academic knowledge production and its relationship with the wider economy, overcoming some of the typical data limitations confronted in prior investigations.

The rest of this chapter is organized into six sections. Section II reviews the literature and describes previous studies of knowledge production. Section III describes a technique for estimating the knowledge production function involving panel count data within a polynomial distributed lag scheme. Section IV describes research inputs-outputs data set. Section V shows results for the four main types of research output, by the fifty-four separate academic departments and research units of Colorado State University from 1989 to 2012, within an adapting a seemingly unrelated regression (SUR) system of knowledge production functions. In addition, this section presents a structural change analysis within this regression framework for one of the three types of research outputs, comparing technology transfer metrics before and after the restructuring of the technology transfer office at CSU. Section VI discusses some the main characteristics of the university departments' knowledge production in terms of classical production theory, including the output elasticity of knowledge production with respect to research inputs and returns to scale of knowledge production across the main research outputs. The section also discusses limitations of this study and future directions for this research. Section VII summarizes the main conclusions and insights of this chapter.

II. Literature Review

A. The Knowledge Production Function: Theory and Empirical Analysis

The knowledge production function is based on the concept of the neo-classical production function, and it is useful for describing the unobservable, yet valuable, additions that research contributes to the stock of knowledge capital. The concept of a knowledge production function was first introduced by Griliches (1979) and further developed by Pakes & Griliches (1980, 1984). These initial papers analyzed the rate of "production" of patents, considered a useful indicator of unobservable knowledge increments, resulting from the research and development (R&D) activities of U.S. industry, and set up an empirical model of the relationship between past R&D expenditures and the output of patented inventions.

In one of the early applications of the Griliches knowledge production function to academic research, Jaffe (1989) modeled a production function of commercial spillovers from university R&D, examining the importance of geographical proximity between university research and commercial innovation. The same year, Pardey (1989) analyzed the input-output relationship of knowledge production at the state agricultural experiment stations (SAES) of a number of major state universities in the U.S. using a panel data set. The model was also based on the Griliches knowledge production function to explore the relationship between past research expenditures and research outputs, as indicated by research publications and their citations. This paper provides that the quality of the publications depends on the number of citations and increases the research output elasticity. Moreover, levels of research expenditure showed little evidence of systematic short-run or point-to-point influence in the input-output relationship, but only a long-run or average influence.

Adams & Griliches (1998) explored research performance and productivity within a set of U.S. universities, with a model based on the Griliches knowledge production function. They show that, at the aggregate level, research outputs such as publications and citations tend to follow constant returns to scale, but, at the individual institutional level, they follow diminishing returns to scale. Moreover, the paper analyzes variations in the relationship between research outputs and R&D expenditures across different fields, types of universities (public versus private), funding sources, and degree awards.

Subsequently, few other studies have extended and developed Griliches' knowledge production function, especially empirically. According to Crespi and Geuna (2008), previous studies provide little systematic evidence that such investments lead to increased levels of knowledge outputs, such as scientific publications, patenting, or other measures of innovation, and ultimately to better economic performance. Crespi and Geuna critique several common methods used in estimating Griliches' knowledge production function, including the misspecifications of a Cobb-Douglas functional form and problems with Koyck lags. They instead introduce a polynomial distributed lags (PDL) model for this input-output relationship.

This paper seeks to develop a new empirical approach, utilizing both a more detailed dataset and a polynomial distributed lags model to analyze how R&D inputs relate to knowledge outputs within the university context.

B. Different Types of Knowledge Dissemination Channels

First, we begin with a recognition that the type of university knowledge outputs (and the type of knowledge dissemination channels on which they tend to rely) can vary across the many different research environments and disciplines found within the university. In general, basic knowledge outputs in the university can be measured by scientific publications, citations made to those publications, and graduate degrees awarded. There is a long tradition of "unpublished" or "prepublication" collaborative exchange of new knowledge with industry, via a range of formal and informal contacts, such as joint conduct of research, workshops and seminars, extension activities, personnel exchanges, and consulting relationships. Newer types of academic knowledge outputs include invention disclosures, patents, and even, occasionally, hi-tech startup ventures founded by university researchers. Many previous studies of academic knowledge production have tended to focus on just one type of research output in their empirical models, particularly favoring journal publications, citations, or patents (see Jaffe, 1989; Johnes & Johnes, 1995; Mansfield, 1995; Narin et al., 1997; Adams & Griliches, 1998; Henderson et al., 1998; Crespi & Geuna, 2008; Whalley, 2013).

Because of data limitations, it has been challenging to take multiple types of knowledge output systematically into account. Agrawal & Henderson (2002) attempted to consider various types of knowledge outputs. They introduced relative importance of university knowledge channels such as publications, conferences, consulting, informal conversations, collaborative research, patents & licenses, recruiting of graduates, and co-supervising. These channels were not analyzed within empirical or other statistics methods, but simply summarized as percentages and ratios, relative to one another, across the entire university at an institutional-wide level of analysis.

This study exploits an extensive data set of knowledge inputs—including grant and contract awards, research expenditures, researcher FTEs, laboratory space, equipment—as well as outputs—including publications, graduate degrees awarded, proxies for university-industry collaboration and extension activities, invention disclosures, patent applications and grants, license agreements, and startup companies. All data are collected at the lowest possible institutional level of analysis, that of the academic department or research unit, as annual counts, over more than twenty years for many of the series.

III. Model Framework

A. Empirical Model Framework

1. Functional Forms for Modelling Knowledge Production

In classical production theory, various functional forms are used to represent the relationship between inputs and outputs: linear, log-log, quadratic, Cobb-Douglas, constant elasticity of substitution (CES), transcendental, von Liebig, Mitscherlich-Baule, translog, *etc*. However, most previous empirical studies of knowledge production have utilized one of most common, the Cobb-Douglas production function, because of its amenability to econometric techniques but also because of its suitable representation of some of the inherent characteristics of knowledge production.

Although Griliches' concept of knowledge production originates in neo-classical production theory, the production of knowledge differs somewhat from that of normal economic goods. There are two major differences. First, the profit maximization problem is rarely applied to the knowledge production problem due to the lack of a stable, appropriated market price of research outputs. In other words, markets fail to form for most knowledge outputs, for a variety of reasons that will be discussed later. Second, the units or increments of actual or "underlying" economically valuable technological knowledge are often unobservable. According to Pardey (1989), empirical studies of knowledge production is limited in large part because of the difficulties of obtaining suitable indicators of research outputs. However, we can still be confident that there exists a systematic input-output relationship between research inputs and new knowledge outputs.

The university knowledge production function developed here is based on previous models (Griliches, 1979; Pakes & Griliches, 1984; Jaffe, 1989; Pardey, 1989; Adams & Griliches, 1998) using classical microeconomic assumptions of production theory. The knowledge production function, describing the technical relationship between research inputs and outputs, is structurally analogous to the neoclassical production function. Consider an accounting center within the university—a college, department, or research unit–within which known amounts of inputs, such as salaries of R&D staff, equipment, laboratory space, are used and within which research outputs—such as datasets, software, research publications, inventions—are produced. The knowledge production function defines the technical relationship by which the observed research inputs are transformed into the observed research outputs.

Equation (1) represents the empirical functional form for panel data analysis of the university knowledge production that relates research inputs to outputs.

$$Y_{i,t} = \alpha + \sum_{j=0}^{k} \beta_j R_{i,t-j} + \beta_L L_{i,t} + \beta_E EQ_{i,t} + \beta_C C_{i,t} + \varepsilon_{i,t} \qquad (1)$$

where i stands for the i^{th} department or research unit and t stands for the t^{th} time period. Y is the vector of university research outputs. R represents research expenditures in the current and the k lagged time periods. L is the count of full-time-equivalent (FTE) researchers, which consists of faculty, research staff, and graduate research assistants employed in time period t, as human capital[5]. EQ represents the value of research equipment, or physical capital, employed in research. C are other control variables. The variable $\varepsilon_{i,t}$ is an independent and identically distributed panel disturbance term. This term can be comprised of a group-variant but time-invariant error term, u_i, and a group and time-variant idiosyncratic error term, $e_{i,t}$, in which case it is decomposed to the following:

$$\varepsilon_{i,t} = u_i + e_{i,t} \qquad (2)$$

Assuming that these are characterized by mean zero and homoscedastic variance, such that $u_i \sim \left(0,\ \sigma_u^2\right)$, $e_{i,t} \sim \left(0,\ \sigma_e^2\right)$ and $\operatorname{cov}\left(u_i,\ e_{i,t}\right) = 0$, and assuming no serial correlation, such that $\operatorname{cov}\left(e_{i,t},\ e_{i,s}\right),\ \forall t \neq s$, then equation (3) follows from equations (1) and (2),

$$Y_{i,t} = \left(\alpha + u_i\right) + \sum_{j=0}^{k} \beta_j R_{i,t-j} + \beta_L L_{i,t} + \beta_E EQ_{i,t} + \beta_C C_{i,t} + e_{i,t} \qquad (3)$$

[5] There exists a positive relationship between FTEs and research expenditures, because research expenditures include the salary of faculty members, GRAs and research staffs. This can cause a collinearity problem, which we can control, if it is detected, by adopting an alternative model.

where the slope coefficients $\beta_0,\cdots,\beta_k,\beta_L,\beta_E,$ *and* β_C are constant in every group, but $(\alpha+u_i)$ vary by group, such that this latter term measures the heterogeneity of research environments and inherent characteristics across colleges and research units, sometimes called a varying coefficient model.

In contrast, another useful panel disturbance term consists of two components, where one is time-variant but group-invariant, μ_t, and the other is both time and group-variant, as depicted in equation (4),

$$\varepsilon_{i,t}=\mu_t+e_{i,t} \tag{4}$$

Again, it is assumed these components are characterized by mean zero, with homoscedastic variance and no serial correlation. Equation (5) is the modified model from equations (1) and (4):

$$Y_{i,t}=(\alpha+\mu_t)+\sum_{j=0}^{k}\beta_j R_{i,t-j}+\beta_L L_{i,t}+\beta_E EQ_{i,t}+\beta_C C_{i,t}+e_{i,t} \tag{5}$$

So, these two variants control only one or the other source of variation, either group or time heterogeneity, and are thus considered a one-way error component model.[6]

2. Count Data as Research Output

Most academic research outputs are measured by count data,[7] such as the number of articles published per year, the number of graduate degrees conferred per year, the number of articles with co-authors from industry per year, the number of invention disclosures or patent applications per year, and the number of startups created per year. The fundamental variable in each of these measures is discrete, taking only a finite number of non-negative, integer values. For count data,

---

[6] A two-way error component model would control both unobserved group and time heterogeneities at the same time. However, in this two-way method many degrees of freedom would be lost, so our analysis will not use the method.

[7] A type of data in which the observations can take only the non-negative integer values {0, 1, 2, 3, …} and where these integers arise from counting rather than ranking.

the Poisson and negative binomial maximum likelihood regression models are well established. First, when a Poisson model is appropriate for research outputs with count data, the probability distribution function is given by

$$f\left(Y_i \middle| \theta\right)=\begin{cases}\dfrac{\theta^{Y_i}e^{-\theta}}{Y_i\,!}; & Y=0,1,2,\ldots\\ 0 & ;\ \mathit{Otherwise}\end{cases} \tag{6}$$

where θ is known as the population rate parameter, and the Poisson PDF has an equal-dispersion property where $E\left(Y_i\right)=\theta$ and $\sigma^2=\theta$.

Second, the negative binomial is an extension of the Poisson with a probability distribution given by

$$f\left(Y_i \middle| \theta\right)=\begin{cases}\left(\dfrac{r}{r+\theta}\right)^r \dfrac{\Gamma\left(r+Y_i\right)}{Y_i\,!\Gamma\left(r\right)}\left(\dfrac{\theta}{r+\theta}\right)^{Y_i}; & Y_i=0,1,2,\ldots\\ 0 & ;\ \mathit{Otherwise}\end{cases} \tag{7}$$

where r is the dispersion parameter and Γ is the gamma function. The negative binomial does not have the equal-dispersion property, and different values for mean and variance allow for over-dispersion, with $E\left(Y_i\right)=\theta$ and $\sigma^2=\theta+\dfrac{\theta^2}{r}$.

In many cases of count data, the value of the sample variance is, indeed, greater than the value of the sample mean, indicating over-dispersion; yet, the Poisson distribution does not allow for variance to be adjusted independently of the mean. In such cases, the negative binomial distribution can be used. In general, in testing for over-dispersion it is not sufficient merely to compare sample mean and variance, but it is necessary to proceed with a hypothesis test with a null hypothesis, $H_o: \mathit{Dependent\ variable} \sim \mathit{Poisson\ distribution}$. However, if such a test is not feasible, alternatively, information criteria, such as the Akaike information criteria (AIC) and the

Schwarz Bayesian information criteria (SBIC), can be used, as the Poisson distribution is a special case of the negative binomial distribution, with simply a zero value of the dispersion parameter, $r = 0$ in equation (7) above.

The first step of negative binomial maximum likelihood estimation is to derive the log likelihood function with independently and identically distributed disturbances from equation (7), the negative binomial probability distribution:

$$L\left(\beta \middle| Y_{i,t}, \theta_{i,t}\right) = \prod_{i=1}^{N} f\left(Y_{i,t} \middle| \theta_{i,t}\right) = \prod_{i=1}^{N} \left[\frac{\Gamma\left(Y_{i,t} + r\right)}{Y_{i,t}! \cdot \Gamma\left(r\right)} \cdot \left(\frac{r}{r + \theta_{i,t}}\right)^{r} \left(\frac{\theta_{i,t}}{r + \theta_{i,t}}\right)^{Y_{i,t}} \right] \tag{8}$$

Next, taking the logarithm on the both sides of equation (8), we obtain the log-likelihood function for negative binomial maximum likelihood estimation,

$$\begin{aligned} \ln L\left(\beta \middle| Y_{i,t}, \theta_{i,t}\right) &= \sum_{i=1}^{N} \ln \left[\frac{\Gamma\left(Y_{i,t} + r\right)}{Y_{i,t}! \cdot \Gamma\left(r\right)} \cdot \left(\frac{r}{r + \theta_{i,t}}\right)^{r} \left(\frac{\theta_{i,t}}{r + \theta_{i}}\right)^{Y_{i,t}} \right] \\ &= \sum_{i=1}^{N} \left[\ln \Gamma\left(Y_{i,t} + r\right) - \ln\left(Y_{i,t}!\right) - \ln \Gamma\left(r\right) \right] \\ &\quad + r \cdot \sum_{i=1}^{N} \left[\ln\left(r\right) - \ln\left(r + \theta_{i,t}\right) \right] + Y_{i,t} \cdot \sum_{i=1}^{N} \left[\ln\left(\theta_{i,t}\right) - \ln\left(r + \theta_{i,t}\right) \right] \end{aligned} \tag{9}$$

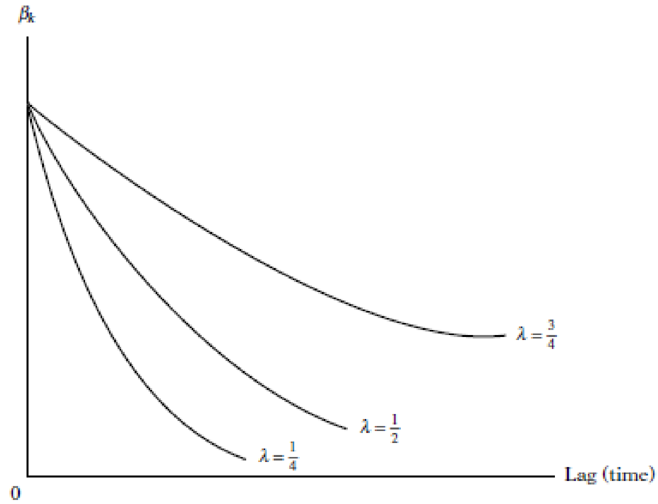
where $\theta_{i,t} = E\left[Y_{i,t} \middle| X_{i,t}\right] = \exp\left(X_{i,t}{}'\beta\right) = \exp\left(\sum_{j=0}^{k} \beta_j R_{i,t-j} + \beta_L L_{i,t} + \beta_E EQ_{i,t} + \beta_C C_{i,t} \right)$

Equation (9) is maximized with respect to the unknown parameter, $\theta_{i,t}$, the mean of the negative binomial process.

3. Polynomial Distributed Lags (PDL) Model

Much of the literature has shown that past research expenditures are the main input of knowledge production. A number of previous studies have thus adopted research expenditures as the primary research input and have used a distributed-lag model with a finite lag of *k* time periods.

A few have adopted the Koyck approach to modelling the distributed lag, which is based on the assumption that coefficients decline geometrically as the lag lengthens. See the Figure 2.



Source: D.N. Gujarati (2004) Basic Econometrics 4th edition

Figure 2—Koyck Scheme (Declining Geometric Distribution)

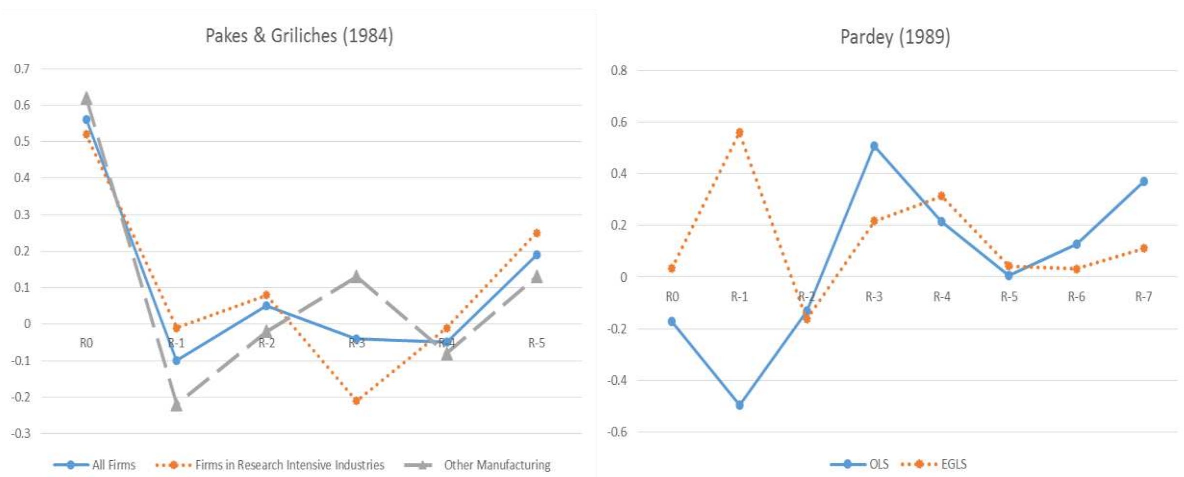


Figure 3—Coefficient Patterns of Pakes & Griliches (1984) and Pardey (1989)

However, the assumption has limitations in estimating a knowledge production function, and in particular in estimating lagged research expenditures. In Figure 3, the slope coefficients of Pakes & Griliches (1984) and Pardey (1989) do not seem to follow a geometrically decreasing pattern, but appear to follow something more like a fourth-degree polynomial pattern, such that the Koyck scheme of distributed-lags does not appear to work as well in these cases as an *ad hoc* distributed lag scheme.

In the current study, Almon's (1965) scheme, with a polynomial of degree *m* and a lag length *k* of past inputs will be adapted to the knowledge production function model. The advantages are an increased goodness-of-fit of the regression model and more accurate verification of the regression results. However, there are few examples in the literature, see Crespi & Geuna (2008), that employ this methodology in the knowledge production context, which is a disadvantage, along with the risk of model misspecification. However, in light of the failure of the Koyck assumption, the polynomial distributed lag (PDL) model is more precise than an *ad hoc* distributed lag model.

A panel count polynomial distributed-lag model (PDL) is derived from equation (1) and the negative binomial maximum likelihood equation. First, equation (10) is a polynomial distributed lag scheme generated by different degrees of polynomial. In this the maximum degree, *m*, of the polynomial, $p = 0,1,2,\cdots,m$, must be smaller than the maximum lag, $j = 0,1,2,\cdots,k \ (\forall m < k)$:

$$\beta_j = \omega_0 + \omega_1 \cdot j + \omega_2 \cdot j^2 + \cdots + \omega_m \cdot j^m = \sum_{p=0}^{m} \omega_p \cdot j^p \tag{10}$$

Second the corresponding equation of *m*-degree and k lags of unrestricted polynomial distributed lag is

$$Y_{i,t} = \alpha + \sum_{p=0}^{m} \omega_p Z_{p,i,t} + \beta_L L_{i,t} + \beta_E EQ_{i,t} + \beta_C C_{i,t} + \varepsilon_{i,t} \tag{11}$$

where $z_{0,i,t} = \sum_{j=0}^{k} j^0 R_{i,t-j} = R_{i,t} + R_{i,t-1} + R_{i,t-2} + R_{i,t-3} + \cdots + R_{i,t-k}$

$$z_{1,i,t} = \sum_{j=1}^{k} j \cdot R_{i,t-j} = R_{i,t-1} + 2R_{i,t-2} + 3R_{i,t-3} + \cdots + kR_{i,t-k}$$

$$\vdots$$

$$z_{m,i,t} = \sum_{j=0}^{k} j^m R_{i,t-j} = R_{i,t-1} + 2^m R_{i,t-2} + 3^m R_{i,t-3} + \cdots + k^m R_{i,t-k}$$

The degree of the polynomial and the length of the lag can be tested by two common measures for comparing maximum likelihood models, the Akaike information criterion (AIC) and the Schwarz Bayesian information criterion (SBIC) .[8] However, the slope coefficients, ω_p, from equation (11) are not true values of the model estimators, but the true slope coefficients need to be recovered. The formulae for recovered slope coefficients are the following:

$$
\begin{aligned}
\tilde{\beta}_0 &= \hat{\omega}_0 \\
\tilde{\beta}_1 &= \hat{\omega}_0 + \hat{\omega}_1 + \hat{\omega}_2 + \cdots + \hat{\omega}_m \\
\tilde{\beta}_2 &= \hat{\omega}_0 + 2\hat{\omega}_1 + 4\hat{\omega}_2 + \cdots + 2^m \hat{\omega}_m \\
&\vdots \\
\tilde{\beta}_k &= \hat{\omega}_0 + k\hat{\omega}_1 + k^2\hat{\omega}_2 + \cdots + k^m \hat{\omega}_m
\end{aligned}
\tag{12}
$$

Thus, the $\tilde{\beta}$ s from equation (12) are the estimated slope coefficients, and the unrestricted PDL model is

$$Y_{i,t}^u = \alpha + \sum_{j=0}^{6} \tilde{\beta}_j^u R_{i,t-j} + \beta_L L_{i,t} + \beta_E EQ_{i,t} + \beta_C C_{i,t}\,. \tag{13}$$

The unrestricted PDL model has no *a priori* restrictions, but a restricted PDL model can be limited by endpoint restrictions on the current and the k^{th} or k+1$^{\text{st}}$ lagged coefficients, which are zero. If imposing this restriction on the coefficient of the current time period's input variable, it is called a left restriction or a near-end restriction. If imposing the restriction on the k^{th} or $k+1^{\text{th}}$ lagged coefficient, it is called a right restriction or a far-end restriction. Following Gujarati (2004), such restrictions are explained by psychological, institutional, or technical reasons. In this paper, we assume that unobservable inputs made beyond the k^{th} lag year no longer impact the current research outputs, but that research expenditures in the current year do have impact.

[8] $AIC = -2 \times \ln(\text{Likelihood Function}) + 2 \times p$ and $SBIC = -2 \times \ln(\text{Likelihood Function}) + \ln(N) \times p$ where p is number of parameters estimated and N is number of observations.

For example, setting m=4, k=6, and far endpoint=7, β_7 in equation (14) is the far endpoint restriction for the regression model:

$$\begin{aligned}&\beta_7 = \omega_0 + 7\omega_1 + 49\omega_2 + 343\omega_3 + 2401\omega_4 = 0 \\ &\Rightarrow \ \omega_0 = -7\omega_1 - 49\omega_2 - 343\omega_3 - 2401\omega_4\end{aligned} \tag{14}$$

Equation (14) is substituted into equation (11) and then the result is:

$$Y_{i,t} = \alpha + \omega_1\left(z_{1i,t} - 7z_{0i,t}\right) + \omega_2\left(z_{2i,t} - 49z_{0i,t}\right) + \omega_3\left(z_{3i,t} - 343z_{0i,t}\right) + \omega_4\left(z_{4i,t} - 2401z_{0i,t}\right) + \varepsilon_{i,t} \tag{15}$$

Then, recovering the true estimated slope coefficients from equation (15) by the same manner as in the unrestricted model, equation (16) is the resulting restricted PDL model:

$$Y_{i,t}^r = \alpha + \sum_{j=0}^{6} \tilde{\beta}_j^r R_{i,t-j} + \beta_L L_{i,t} + \beta_E EQ_{i,t} + \beta_C C_{i,t} \tag{16}$$

In knowledge production, there is no doubt that past research expenditures impact current research outputs, but there are some ambiguities between near-endpoint, far-endpoint, and both-endpoint restrictions. In general, these depend on the type of research output, the inherent characteristics of the research environments, and the different purposes of R&D projects across the different departments or research units.

4. Effective Labor in Knowledge Production Function

In classical production theory, the primary input variables in the production function are capital and labor. In a similar way, research expenditures and full time equivalent (FTE) researchers can be considered the main input variables in the knowledge production function. Since, most of research expenditures go toward the salaries of faculty members, research staff, and graduate students, these two input variables in the knowledge production function interact with each other. Thus, this section introduces one of the alternative ways for avoiding the resulting problems. In Solow's (1956) model, human knowledge cannot be removed from the labor input, so knowledge and labor enter multiplicatively, which is referred to as an effective labor, labor-augmenting, or Harrod-neutral model.[9] Moreover, according to Romer (1986, 1990), the stock of human or knowledge capital determines the rate of growth and is a non-rival and semi-excludable resource, in his alternative specification known as endogenous technological change. Even though these growth models are macroeconomic, they are based on the notion of the production function, sometimes called "*the microfoundations of macroeconomics.*"[10] In general, aggregate data of production functions assume constant returns to scale instead of diminishing returns to scale, as in individual firm data. However, if we follow this and assume constant returns to scales with respect to labor in the knowledge production function for the university, we can generate equation (17) below

$$\begin{aligned} & f(\mu R,\ \mu L,\ \mu EQ) = \mu \cdot f(R,\ L,\ EQ) \\ \Rightarrow\ & f\left(\frac{R}{L}, 1, \frac{EQ}{L}\right) = \frac{1}{L} \cdot f(R,\ L,\ EQ),\ if\ \mu = \frac{1}{L} \end{aligned} \tag{17}$$

By assuming a Cobb-Douglas functional form, and taking the logarithm on both sides of equation (17), we can denote research output per unit of effective labor (in our case, FTEs) as a function of research expenditures per unit of effective labor (or FTEs) and the value of research equipment per unit of effective labor (or FTEs), as in equation (18):

$$\begin{aligned} & \ln\left(\frac{Y_{i,t}}{L_{i,t}}\right) = \alpha + \sum_{j=0}^{k} \beta_j \ln\left(\frac{R_{i,t-j}}{L_{i,t}}\right) + \beta_E \ln\left(\frac{EQ_{i,t}}{L_{i,t}}\right) + \varepsilon_{i,t} \\ & \Leftrightarrow \\ & y_{i,t} = (\alpha + u_i) + \sum_{j=0}^{k} \beta_j \cdot r_{i,t-j} + \beta_E \cdot eq_{i,t} + e_{i,t} \end{aligned} \tag{18}$$

[9] Romer (2001) "Advanced Macroeconomics" 2nd edition.
[10] The microeconomics of individual agents' behavior such as households or firms that underpins a macroeconomic theory.

It should be noted that equation (18) is still a panel group fixed-effects model with polynomial distributed lags of research expenditures, but it is not a negative binomial maximum likelihood equation. At the department level, there are many zero values that causes the estimation routine to drop the observations due to the characteristics of the logarithm.

IV. Data Descriptions

The data describing both research inputs and research outputs was collected from a variety of sources, and was denominated at the smallest common organizational unit that proved feasible across the entire set of data sources, which was the department or analogous research unit (e.g. various semi-autonomous research centers or services) within the university.

A. Research Input Data

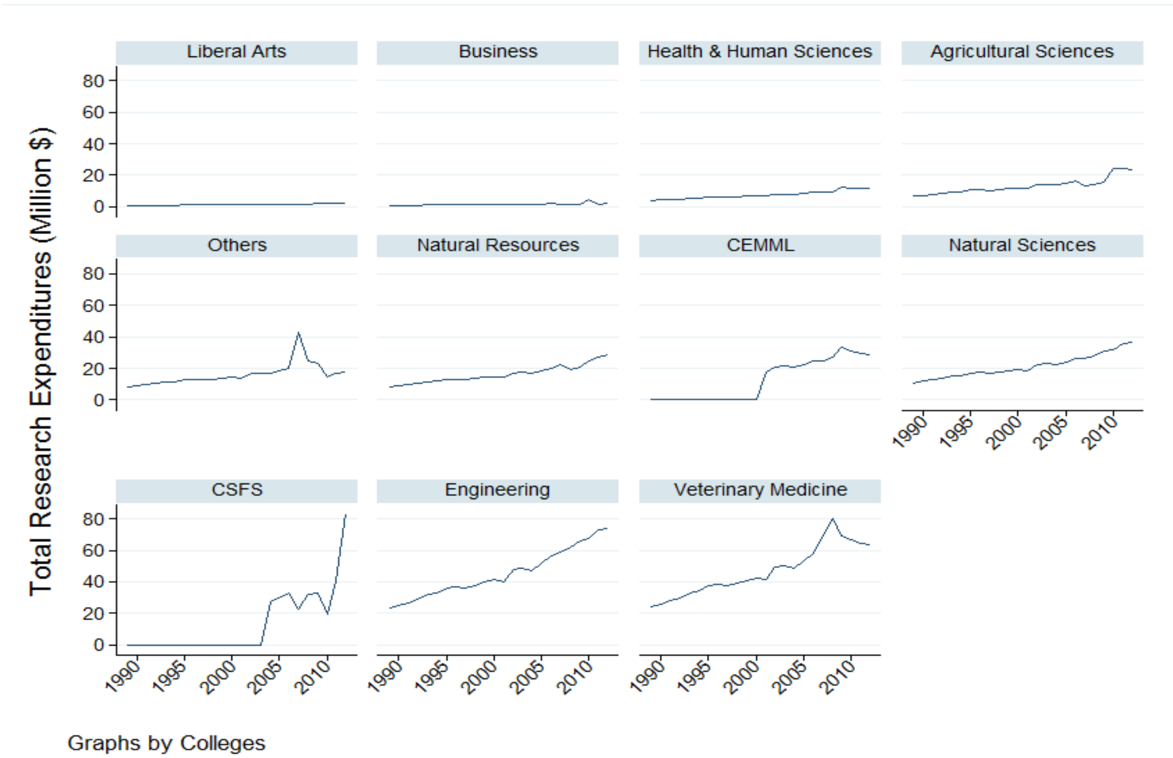
Research input data collected for the university consist of three different categories, broadly representing financial capital, human capital, and physical capital. Total research expenditures and awards of grants and contracts are the financial capital variables collected. Full-time equivalents (FTEs) of researchers represent the human capital input. Office and laboratory space and the value of research equipment are the physical capital variables collected.

1. Financial Capital

Total Research Expenditures: CSU research expenditures are reported from 1989 to 2012 for the university overall, and from 2003-2012 at two more disaggregate levels, including colleges as well as departments and other research units that, combined, make up the colleges.[11] These data come from the Office of Vice President of Research and the University Fact Book Online compiled by the Institutional Research office at CSU. In 2012, the university's overall level of total research

[11] Given the more limited data reporting at the departmental level, estimates for departmental research expenditures were "back cast" for the period from 1989 to 2002, based on the total university research expenditures for those years, subdivided according to the average share of total university research expenditures observed for each department during the period of 2003-2012.

expenditure was $371.6 million. Among colleges, the College of Veterinary Medicine and Biomedical Sciences was the top ranked in 2012 in terms of total research expenditures, at $63.3 million (see Figure 4).



Note: CEMML is Center of Environmental Management of Military Lands, and CSFS is the Colorado State Forest Service. These distinct programs' research budgets are of sufficient size to be counted alongside the eight academic colleges. "Others" include the Vice President for Research (VPR) Office, the libraries, and other central functions that have research budgets.

Figure 4—Annual trends of total research expenditures, by college, (1989-2012)

Grant and Contract Awards: Another measure of financial input is the announced awarding of research funding from external sources via grants and contracts, the records of which are characterized by investigator names, project titles, funding sources, award amounts, college and department affiliations of investigators, and the date and year when the award was announced, from 1989 to 2012. The total number of awards announced over this time period was 38,792 and the total amount dollar was $3,951.3 million. The data comes from the online Proposal and Award

Search at the Research Data Center of the CSU Office of Vice President for Research (http://web.research.colostate.edu/datacenter/). In 2012, the value of overall grant awards announced totaled $257.4 million, and the top ranked college was College of Engineering, at $61.2 million. The disaggregated grant and contract awards and total research expenditures at the departmental level are highly correlated. Moreover, grant awards do not represent actual expenditures, so total research expenditures are considered a better measure of financial inputs to research. The additional information on individual investigators and funding sources, and the fact that it is at a more disaggregate level can make grant and contract award data useful for other purposes.

2. Human Capital

Full-time equivalent (FTE) researchers: The research human capital within the university consists of faculty, research staff, postdoctoral fellows, and graduate research assistants (GRAs).

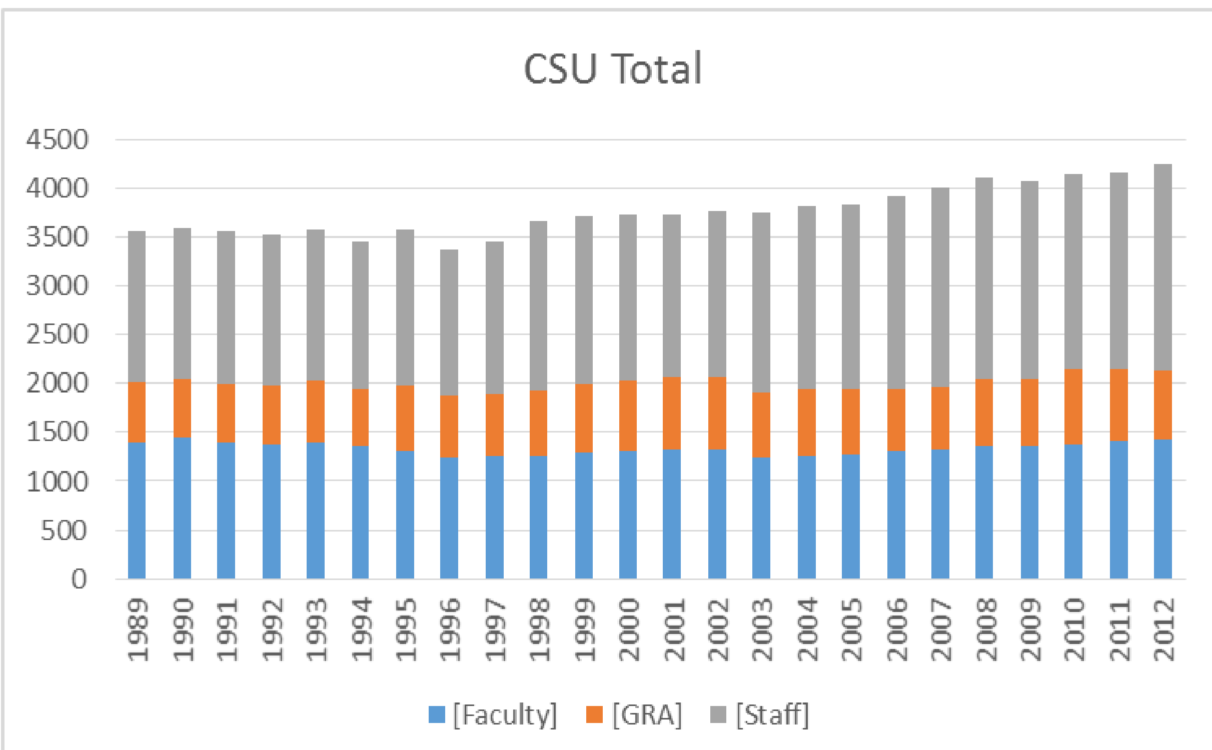


Figure 5—Full-time equivalent (FTE) researchers for the university overall, (1989-2012)

The data on university research staff is tabulated for the full university, as well as for individual colleges and departments. The data comes from the Office of Institutional Research at CSU, but the time period is only 10 years, from 2003 to 2012 rather than 24 years. Totals for other years within the range of this study, from 1989 to 2002, are collected from the CSU Fact Book report.

3. Physical Capital

Laboratory Space: Physical research capital is accounted for across the full university, as well as the level of individual colleges and departments. The data source is the Department of Business and Financial Services, Cost Accounting Division, at CSU.

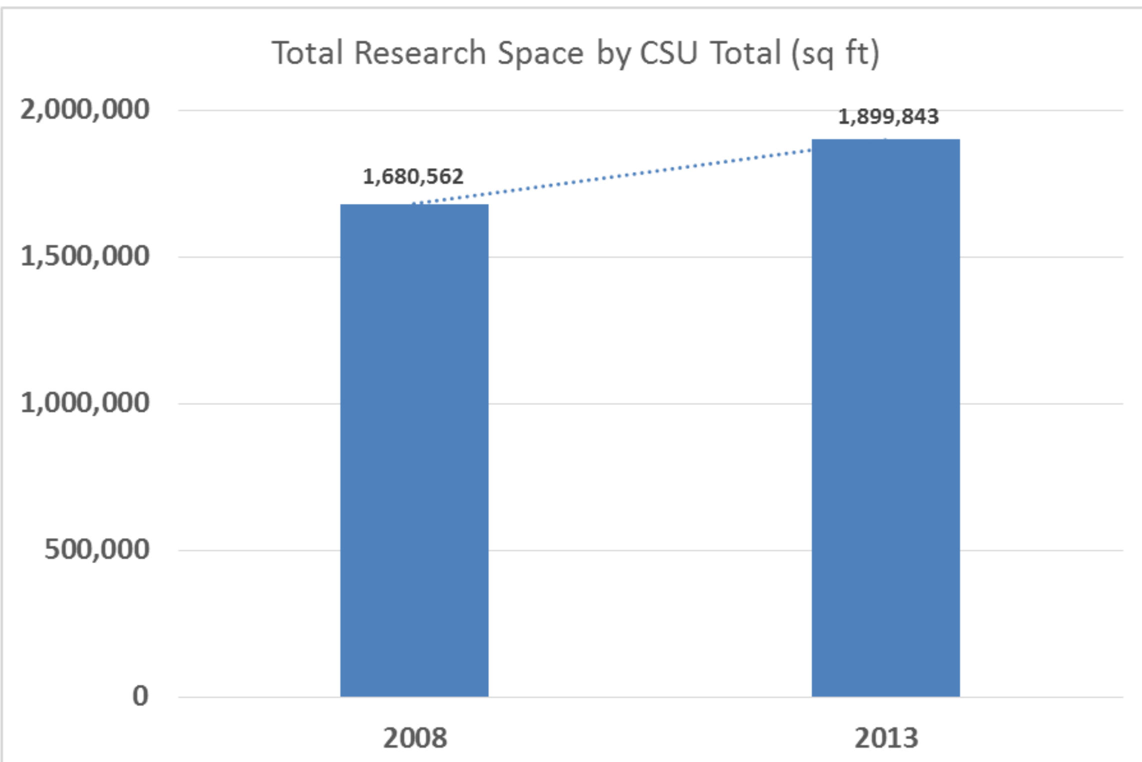


Figure 6—Three different types of research space for the university overall, (2008 and 2013)

However, laboratory space has only been inventoried at two points in time during the range of this study, in 2008 and again in 2013. Even though it does not provide an annual time series, it is still valuable for investigating heterogeneity of physical research capital across the university. In

Figure 4, the total research space for CSU overall in 2008 was 1,680,562 square ft. and in 2013 was 1,899,843 square ft., considering laboratory space, office space, and other research facility space (see Figure 6). By the mean value between 2008 and 2013, the College of Veterinary Medicine was the top ranked, with 29% of the university total, and the College of Business was the lowest, with 2% of total.

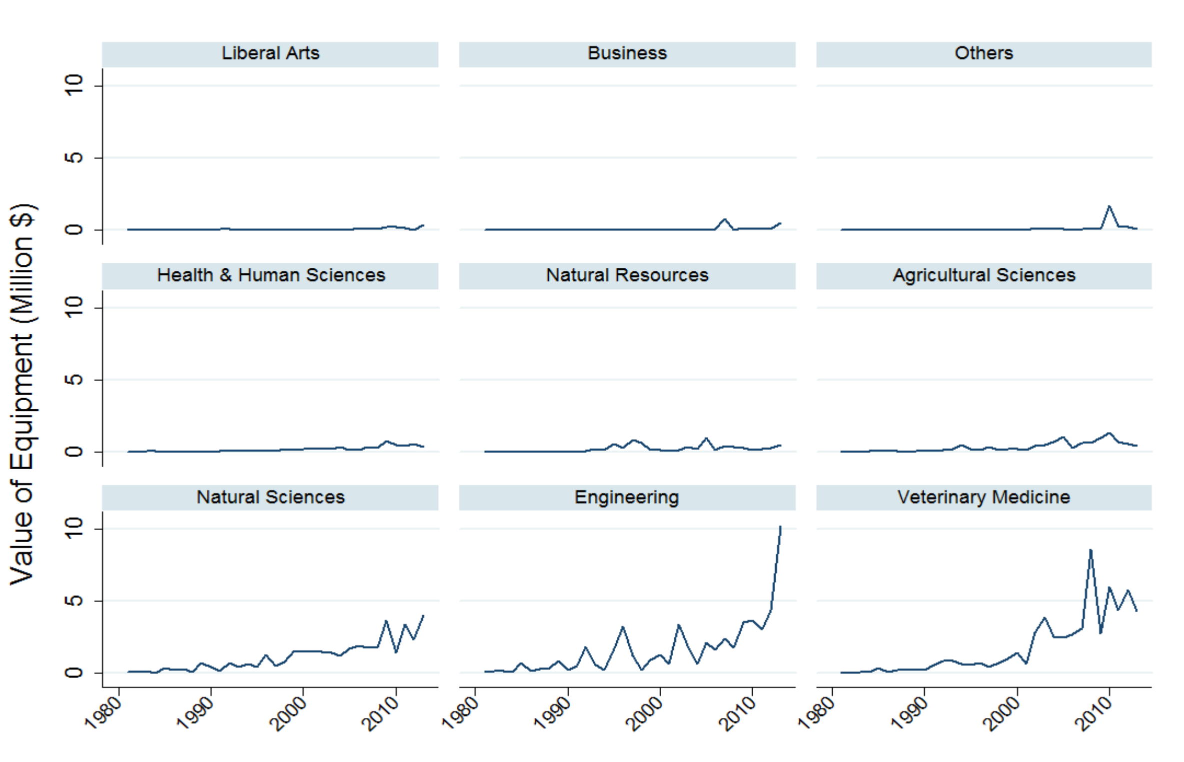


Figure 7—Value of research equipment acquisition, by college, (1981-2013)

Research Equipment: Finally, value of equipment as physical capital is characterized for the university, colleges, and departments from 1981 to 2013. The data comes from the Department of Business and Financial Services, Cost Accounting Division, at CSU. The value of equipment is recorded by acquisition date. Technically, for the value of research equipment to be suitable as an annualized capital input variable for the empirical model, each single piece of equipment would need to be depreciated independently in estimating the value of the stock of equipment. However, even just acquisition values are valuable for examining the heterogeneity across the university; and moreover, we find that annual acquisition values are highly correlated with our preliminary calculation of the value of equipment stock using uniform depreciation rates. In the empirical model, to capture some of this, the equipment acquisitions variable is transformed by a moving window, the sum of the value of research equipment acquisition in the current year and two or three previous years' values.

B. Research Output Data

Research output data were categorized according to the main mechanism or channel of impact they were most likely to represent. We develop a simple taxonomy of three mechanisms: the public domain mechanism; the mechanism of direct collaboration with industry users; or the mechanism of technology transfer mediated by formal intellectual property, including licensing of patents or the creation of startup ventures.

1. Outputs Disseminated via the Public Domain Mechanism

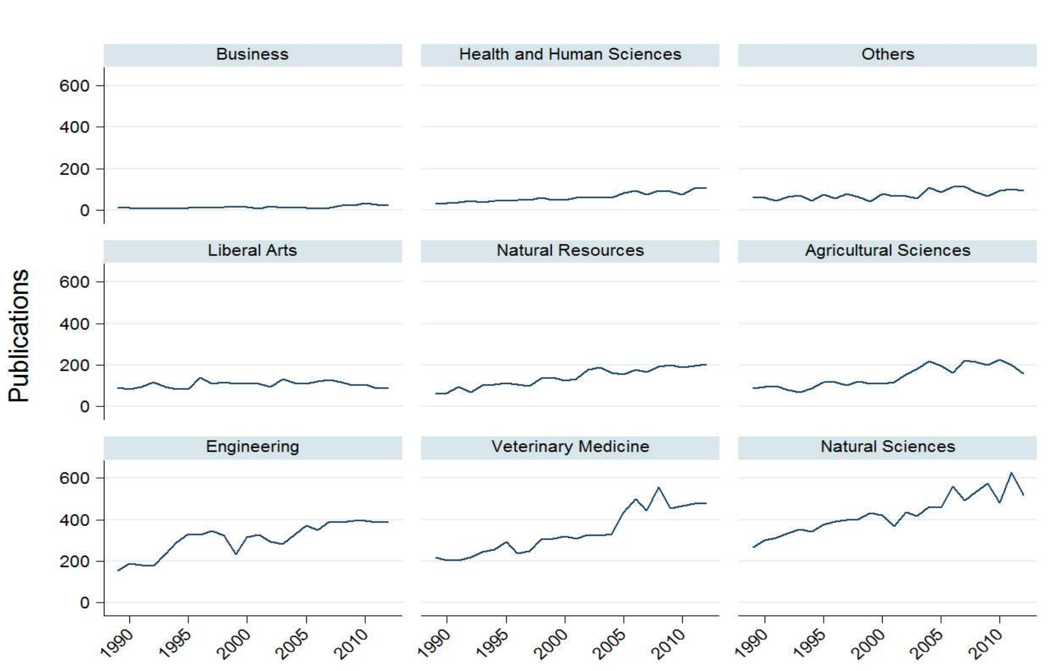


Figure 8—Annual published academic journal articles, by college, (1989-2012)

Published academic articles: This primary proxy of research output is characterized by journal names, author names, publication dates, keywords, citation counts, numbers of references, and organizational details such as college, department, and academic disciplinary field. The data was collected from the Web of Knowledge (Thomson Reuters) accessed via CSU Libraries. The total count of CSU affiliated publication, from 1989 to 2012 is 38,916, within which the highest annual count was 2,211 in 2011. Moreover, among colleges, the College of Natural Sciences had the largest annual count of publication, with 603, in 2011, and at the departmental level, the Department of Chemistry had the highest annual count, at 252, in 2006.

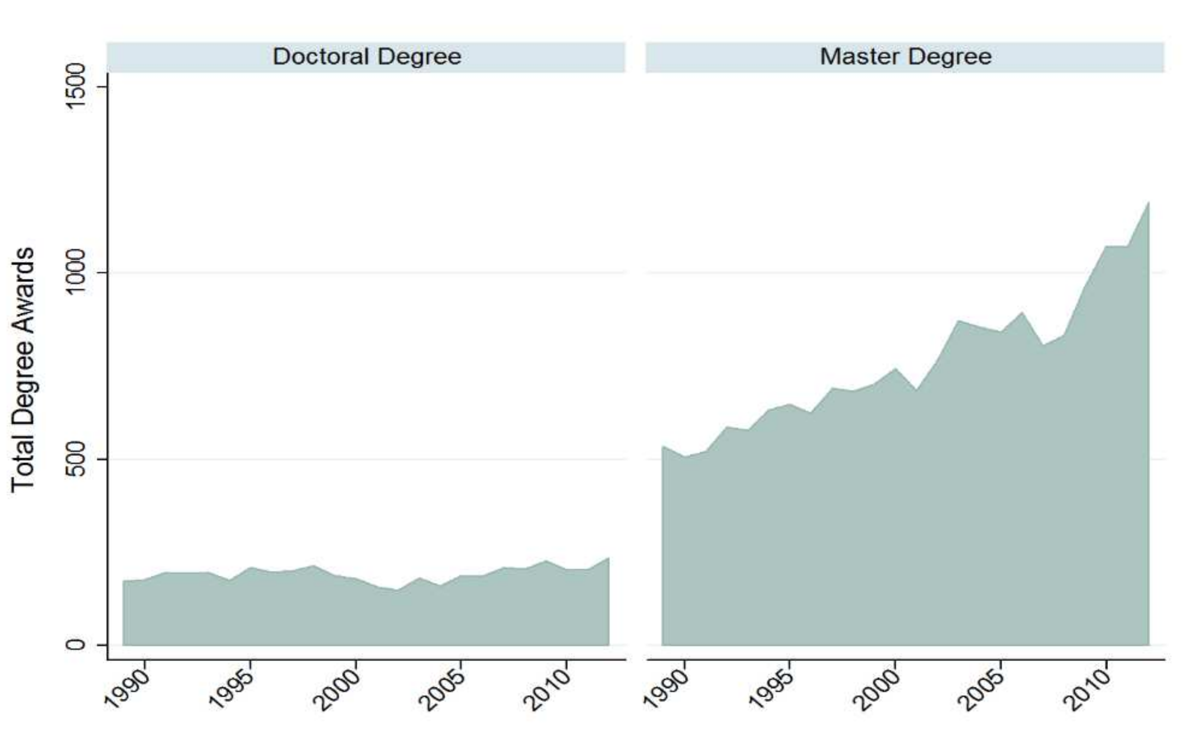


Figure 9—Total (research based) masters and doctoral degree awards, by the university overall, (1989-2012)

Masters and doctoral degree awards: Since not all research conducted by graduate students is published as articles, this additional measure of research output is characterized by colleges and departments, and the years of coverage is from 1989 to 2012. The data source is Institutional

Research Online at CSU. Note that professional degrees, that are not considered research based, such as MBA and Doctor of Veterinary Medicine are excluded. Total numbers of master's degrees and doctoral degrees awarded from 1989 to 2012 were 16,839 and 4,609, respectively.

2. Outputs Disseminated via the Collaboration Mechanism

Given the challenge of measuring directly the results of collaboration between university researchers and industry R&D, several indirect measures are devised to ascertain the research output that has impact via the channel of collaboration.

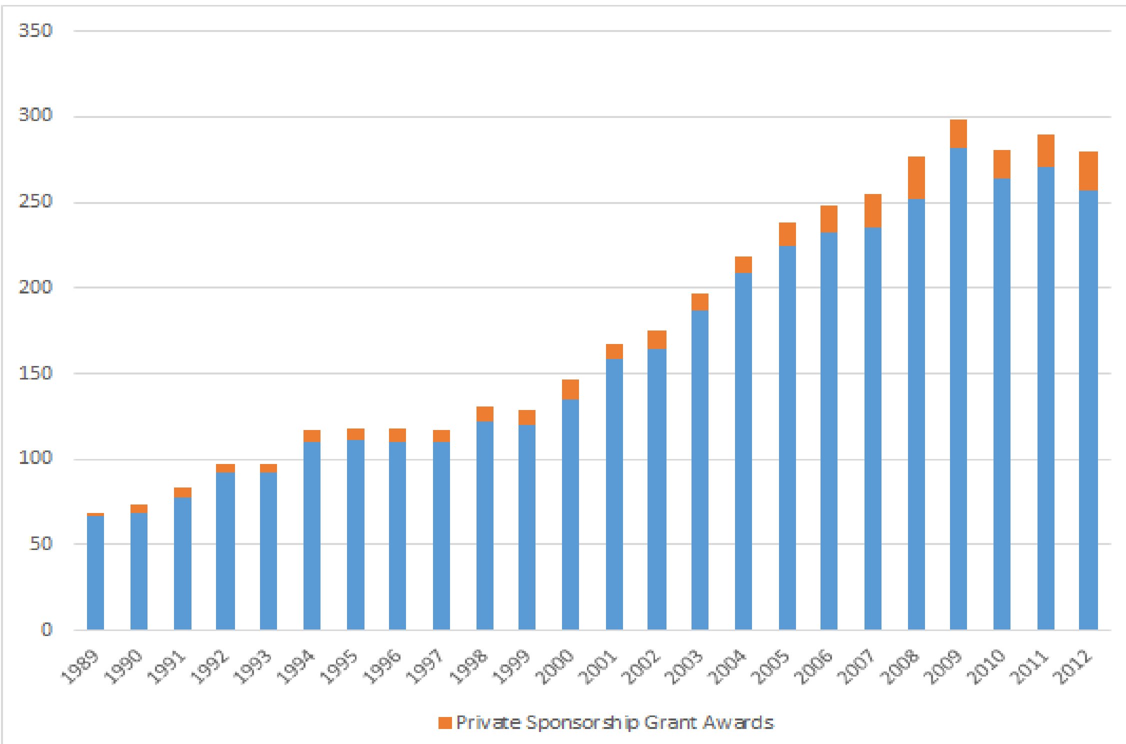


Figure 10—Contract and grant awards from industry sponsors as a share of the total contract and grant awards to the university overall, (1989-2012), million $

Grant and Contract Awards from Industry Sponsors: Even though, technically speaking, grants or contracts are considered an input to research, the extent of private sector sponsorship of grants to a given department can be considered an indication of the amount of research conducted in that department of which the impact is realized, at least in part, via collaborative interaction with industry sponsors. Grant awards data is characterized by principle investigator names, project titles, funding source, college and department of the investigator, award amount, and award dates from 1989 to 2012.

The source is the online Research Data Center (http://web.research.colostate.edu/datacenter/) of the Vice President for Research at CSU, under Proposal and Award Search. The nature of the sponsor, if it was a public sector (federal or state) entity or an industry sponsor, was ascertained by hand. The total number of industry sponsored awards made from 1989 to 2012 is 4,387. Over the 24 years, on average, the value of industry sponsored contract and grant awards to the university makes up 7% percent of total value of awards. In 2012, the value of industry sponsored awards was $21.8 million. The largest single year's industry award to an individual college was to the College of Veterinary Medicine in 2007, at $8.5 million, which was largely driven by the largest single award to a single department, to the Department of Clinical Sciences in 2007, at $5.2 million.

Published academic articles with industry co-authors: Co-authorship on research articles with someone from industry is another indication of the collaboration mechanism at work. While the article itself indicates an impact via the public domain, the fact that the research was conducted and the publication was authored jointly with industry R&D personnel indicates that likely other results of the research, such as tacit skills, and possibly technical findings of relevance to the industry partner, could have arisen and been exchanged more informally. The data source for academic articles is the Web of Knowledge, accessed via CSU Libraries (as already detailed above). All authors on the academic articles associated with CSU were labelled based on their affiliation, being either "CSU", "other public sector", or "private sector" affiliation. Total number of the articles published from 1989 to 2012 with at least one CSU author and at least one private sector author is 2,960 (or 7.6% of total articles published during this time period). The single largest observation for a single college in a single year was the College of Veterinary Medicine with 94 articles with a private sector co-author in 2006.

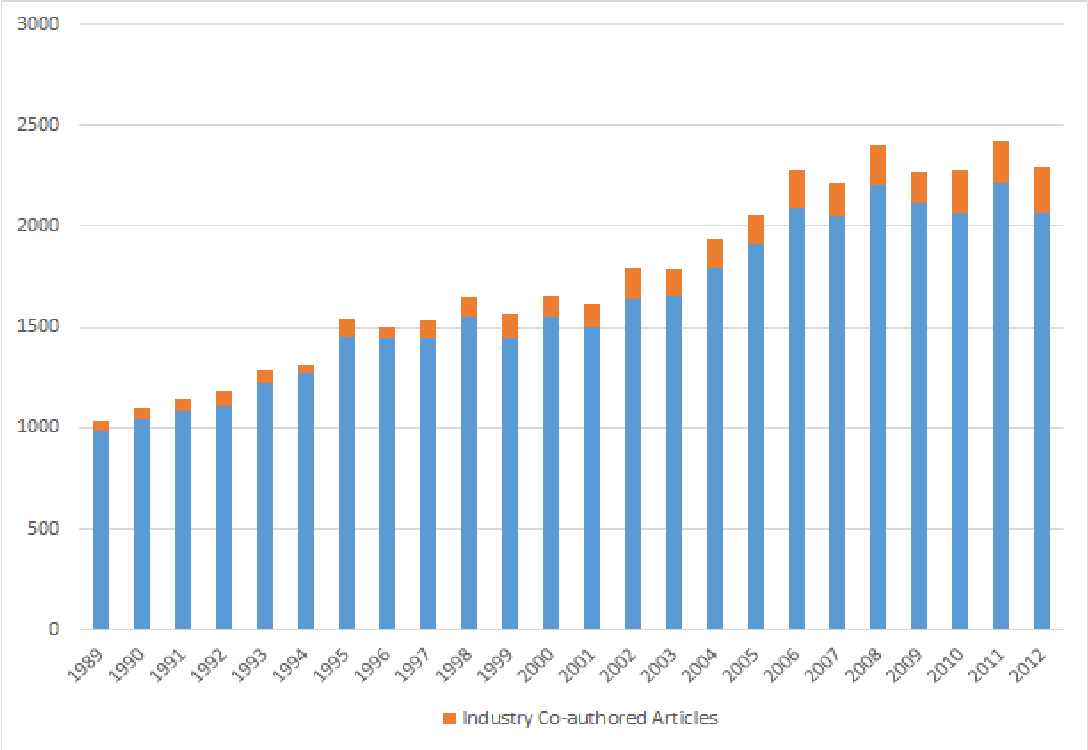


Figure 11—Total published academic articles with industry co-authors, for the university overall, (1989-2012)

Departmental level expenditures on Extension: Extension activities, especially when involving appointments in academic departments, is considered another form of the collaboration mechanism of disseminating research impact. Faculty with extension appointments tend to be engaged with private sector stakeholders throughout the state, communicating their own research results as well as those of their colleagues. Other indicators of extension activity, such as contact hours, numbers of consultations, etc., were not found to be systematically collected across all departments.

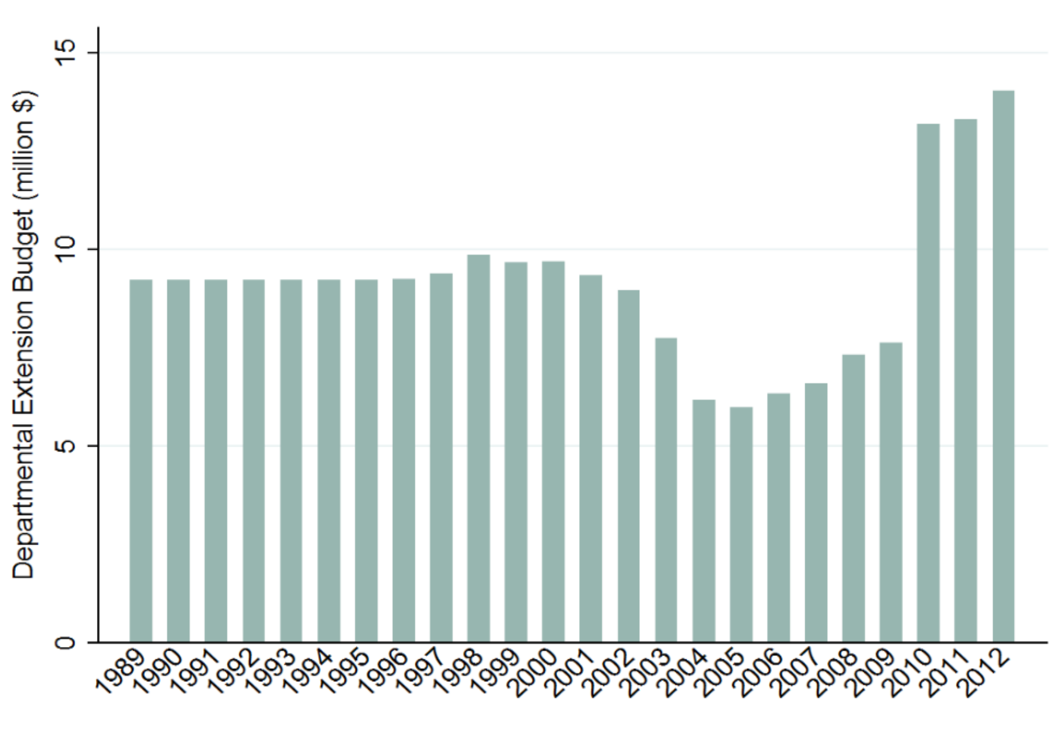


Figure 12—Total annual departmentally-administered expenditures on extension activities, (1989-2012), million $

Again even though expenditures on extension appointments may be technically considered an "input" variable, the level of department level extension expenditures is likely to indicate, in a reasonably systematic and comparable way, across departments and across years, the amounts of engagement with stakeholders outside the university made by extension appointed faculty members of a department. Extension budgets are characterized both at the college and the department level, in dollars per year, with coverage from 1989 to 2012. The data for the college and the department level was provided by accountants in the Agricultural Business Center, of the College of Agricultural Sciences at CSU, from 2003 to 2012. Total extension budget for the remaining period, from 1989 to 2002, was collected from the CSU Fact Book, and the share of expenditures for individual departments over these years was estimated from the annual total, based on average departmental shares in the observed data from 2003-2012. Naturally, some colleges and departments have no extension budget or activity.

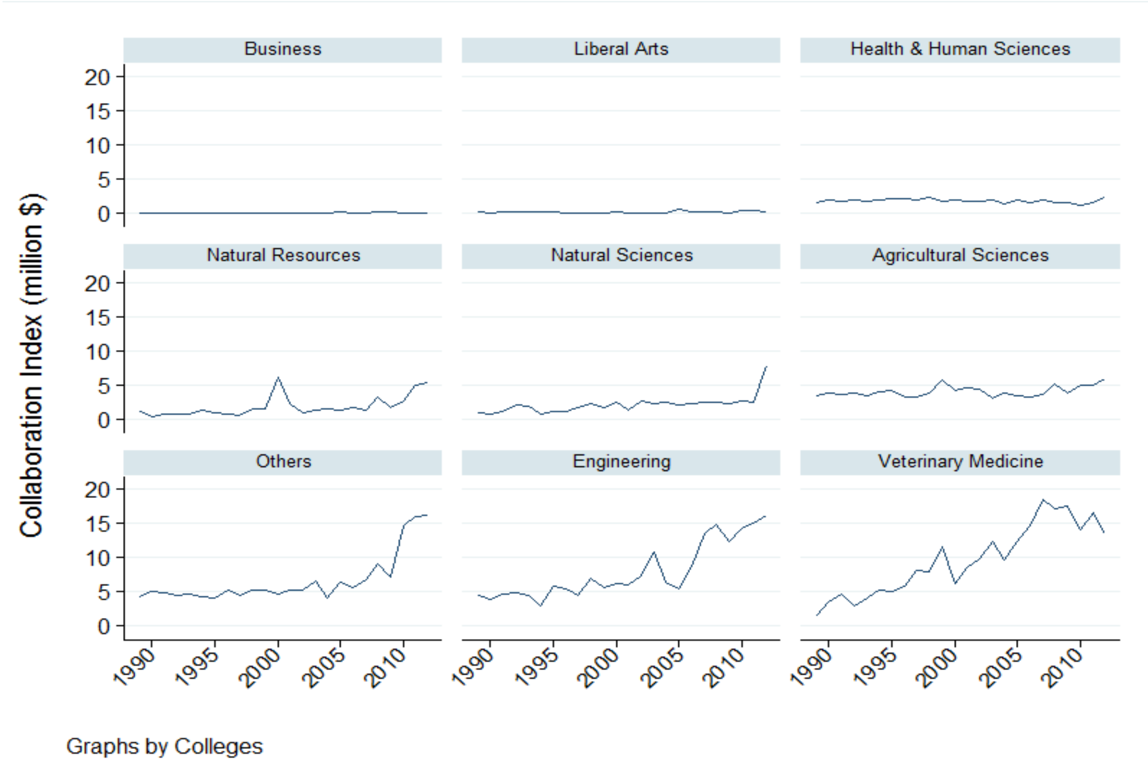


Figure 13—Annual combined industry research collaboration index, by college, (1989-2012), million $

Collaboration Index: The three different collaboration proxies—industry co-authorship on academic articles, research grant awards from industry sponsors, and departmental extension expenditures—were found to be relatively uncorrelated and thus independent of one another. Therefore, a combination of these three into a single variable could best represent the level of industry collaboration activities. In constructing such an index, however, in order to compare industry co-authorship on articles, which is count data, with the other two variables which are financial data, we transformed the publications count data into financial-type data based on average overall research expenditures observed per published article for each respective department or research unit.

3. Outputs Disseminated via Technology Transfer Mechanism

Invention Disclosures: The first indicator of research results that have impact via the technology transfer mechanisms mediated by formal intellectual property are inventions that result from university research formally disclosed to the university's technology transfer office. Invention disclosures are characterized by names of inventor, funding sources that supported the research that led to the disclosed invention, resulting patents related to the invention, percent contribution, as well as the college and department affiliation of the inventor(s). This data spans from 1989 to 2012. The data was provided by CSU Ventures, which serves as the technology transfer office, located within the CSU Research Foundation, on behalf of the university. The total number of CSU invention disclosures was 1,564 from 1989 to 2012. The highest single year, by college, was by the College of Engineering, with 49 invention disclosures in 2009.

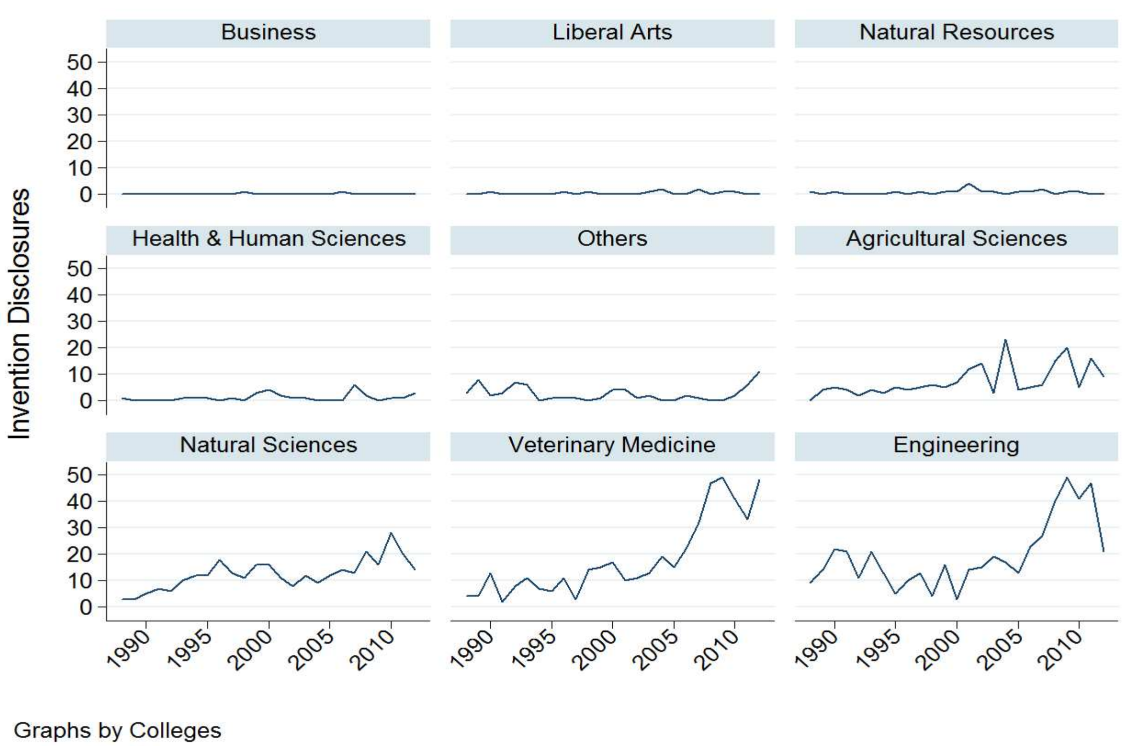


Figure 14—Annual summation of invention disclosures by college level, (1989-2012)

Published Patent Applications and Granted Patents: A second indicator of the patenting/licensing mechanism of impact is patent data. Patent data is characterized by inventor names, published or issued dates, assignee, technology classifications, numbers of citations, and were augmented with the college and department of the CSU affiliated inventors, from 1990 to 2011. Patent data was collected from the Thompson Innovation database by Thompson Reuters. The total number of CSU patents application between 1990 and 2011 was 195. The single largest annual count was from the College of Engineering, with 17 in 2008.

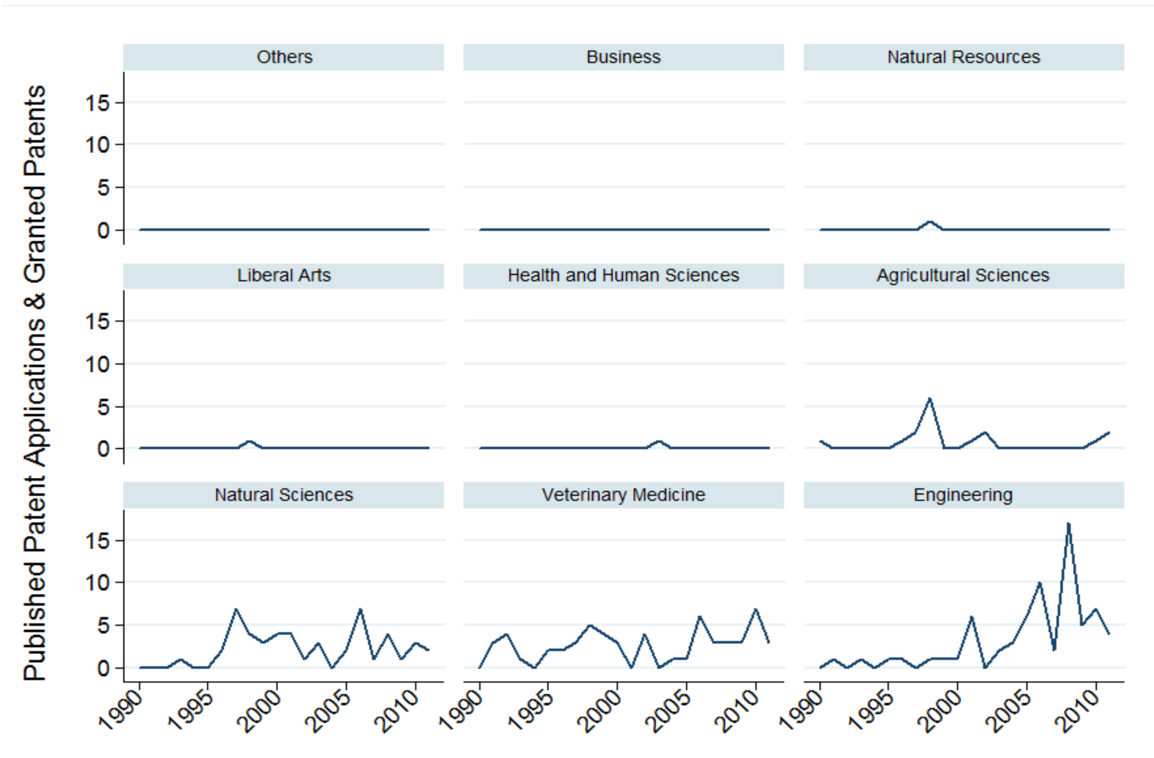


Figure 15—Annual summation of published patent applications and granted patents by college level, (1989-2012)

Startup companies arising from university research are the main indicator of the venture creation mechanism of impact. Data on CSU startup companies is characterized by names of the companies, names of founders, incorporation dates, current employee count, capital raised, as well as college and department of the CSU founder, spanning from 1989 to 2012. During this time CSU research has led to 41 startup companies. The data source is again the CSU Ventures database.

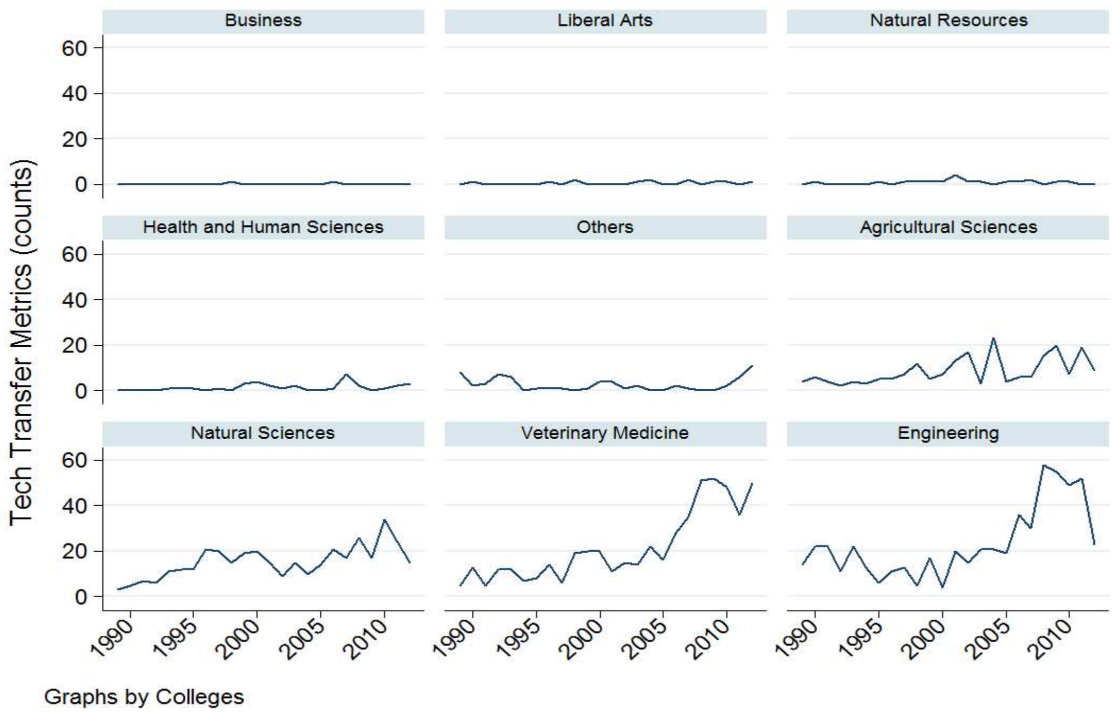


Figure 16—Annual combined tech transfer metric, by college, from 1989 to 2012

Combined Tech Transfer Metrics Index: This numerical value measures the combination of invention disclosures, patent applications and issued or granted patents, and the number of startup companies. These three variables represent publicly observable university tech transfer activities. The combined tech transfer metrics index is conceptually distinguished from the previously introduced collaboration index, as it is more based on intellectual property rights and the "patenting/licensing mechanism" and the "venture creation" mechanism of research output, rather than the "collaboration mechanism" of research output.

V. Empirical Results

In order to assess the efficiency of knowledge production across the different disciplines and research units of the university and to test the effects of changes in research inputs on the various knowledge outputs that impact the economy through different channels, several empirical relationships were estimated. This section consists of two parts of regression models results. One involves regression models which show the results of different knowledge dissemination channels independently. The other involves a system of equations, which provides a test of the interrelationships among the three sets of independent regression models. In addition, we examine the structural changes in the model around the time period of the creation of CSU Ventures, as the technology transfer office (TTO) for Colorado State University.

Table 1 provides summary statistics of the research input and output variables for 54 research units. As detailed in the previous section, we have collected data on several major research inputs, including total annual research expenditures, research FTEs employed, value of research equipment acquisitions, and other control variables. Two of the input variables introduced in Section III, contract and grant awards and research space, were not employed in the regression models, due to the fact that they were found to be highly correlated with some of the other input variables[12] and issues with the quality of data[13], respectively.

In addition, we have eight research output variables, most of which occur as count data, but some of which, including contract and grant awards from industry sponsors and departmental level expenditures on extension, are financial data[14]. Moreover, some of the data collected are used as

---

[12] Total research expenditures and total grant awards are highly correlated each other about 90.4%. In addition, total research expenditure is actual spending and total grant award is an announcement of winning grants.

[13] We have only two data points within the time period, in 2008 and 2013, for laboratory space.

[14] Thus, these research output models will not be measured in the same manner as other research output models, negative binomial maximum likelihood estimation (MLE). This problem will be explained in a later section. In the collaboration index, the index values were constructed as counts (in order to employ negative binomial MLE) by the combination of annual counts of industry co-authored articles, dollars awarded in grants & contacts from private

control variables, including graduate student enrollment by department, which consists of both masters and doctoral degrees. Some departments have clearly identified their master's and doctoral enrollment, but others are not as clearly specified because of different policies of their academic programs.

Table 1—Summary statistics: research inputs and outputs at department level, (1989-2012)

Summary statistics at department level: research input-output variables from 1989 to 2012

	Mean	Std. Dev.	Min	Max	Sum	Group	Obs
Research inputs							
Research expenditures (million $)	2.7	4.6	0	37.9	3,546	54	1,296
Full-time equivalent researchers (FTEs)	56.8	46.1	0	282.4	73,592	54	1,296
Value of equipment (3) (million $)	0.3	0.7	0	7.2	380	54	1,296
Value of equipment (2) (million $)	0.2	0.5	0	4.2	264	54	1,296
Research outputs							
Published academic articles (counts)	28.6	35.9	0	252.0	37,029	54	1,296
Doctoral degree awards (counts)	3.4	5.4	0	54.0	4,415	54	1,296
Industry co-authored articles (counts)	2.2	5.0	0	58.0	2,872	54	1,296
Private grant awards (million $)	0.2	0.5	0	5.4	230	54	1,296
Extension budget (million $)	0.1	0.2	0	1.3	106	54	1,296
Invention disclosures (counts)	1.1	3.0	0	28.0	1,470	54	1,296
Patent apps and granted patents (counts)	0.12	0.5	0	9.0	160	54	1,296
Startup companies (counts)	0.03	0.2	0	3.0	40	54	1,296
Research output index							
Collaboration index (counts)	44.8	98.6	0	812.0	58,096	54	1,296
Tech Transfer metrics (counts)	1.3	3.4	0	29.0	1,670	54	1,296
Other control variables							
Enrollment (counts)	124.1	132.0	0	1,075.0	160,717	54	1,296

Note: 1. The number is window width: (3) is the summation of current value and two previous years that is called three years of window, similarly (2). 2. Graduate level enrollment.

Given the inherent characteristics of panel data analysis, given that it contains both cross-section and time-series, two significant issues that must be controlled for are heteroscedasticity and autocorrelation, respectively.

sponsors, and departmental level of extension expenditures. Similarly, the combined tech transfer metrics index was also constructed as a count variable by the combination of counts of patenting/licensing and venture creation research outputs. The formulas and methodologies are explained in a later section.

Table 2—The results of augmented Dickey-Fuller unit root test of all research inputs and outputs

	Augmented Dickey-Fuller Unit Root Test			
	Inverse chi-squared (108)	Inverse normal	Inverse logit t(274)	Modified inv. chi-squared
Research Inputs				
Total research expenditures	206.58***	-1.93**	-3.15***	6.71***
FTE researchers	289.72***	-9.48***	-9.87***	12.36***
Value of equipment (2)	165.05***	-3.66***	-3.80***	3.88***
Value of equipment (3)	209.21***	-3.92***	-4.71***	6.89***
Research Outputs				
Publications	281.27***	-9.89***	-9.95***	11.79***
Degree (PhD)	366.71***	-12.82***	-13.49***	17.60***
Co-authorship articles	307.25***	-9.42***	-10.29***	13.56***
Private grant awards	251.73***	-8.33***	-8.57***	9.78***
Extension budget	108.97	-1.91**	-1.78**	0.07
Collaboration index	178.08***	-3.46***	-3.72***	4.77***
Patent applications & grants	266.29***	-8.53***	-8.89***	10.77***
Invention disclosures	258.15***	-8.96***	-8.92***	10.22***
Startup companies	185.27***	-4.41***	-4.75***	5.26***
Tech transfer metrics	258.95***	-9.13***	-9.02***	10.27***

*Note: *** 1%, ** 5%, and * 10% levels of statistically significant*

First, time-series analysis assumes that the data represents a stationary process, that the statistical properties of the distribution from which the observations are drawn does not change over time. If, however, a non-stationary process is detected by a unit root test, the analysis should be estimated by a first difference, period to period change. A well-established unit root test is the augmented Dickey-Fuller (ADF) test. Table 2 displays the results of ADF unit root tests for the variables summarized in Table 1. These test statistics indicate a stationary process for most of the input and output variables at a 1 percent level of statistical significance, by four different methodologies. In considering potential heteroscedasticity, the negative binomial maximum likelihood estimation has the advantage of controlling for heteroscedasticity problems with the Huber White sandwich variance-covariance estimation, bootstrap estimation, Jackknife estimation, and others.

A. Research hypotheses

1. Identifying research inputs on the knowledge production function

According to the Pakes and Griliches (1984), a knowledge production function translates past research expenditures and a disturbance term, encompassing unobserved inputs, into inventions. However most previous studies use only past research expenditures because of the quality of data. In this study, we attempt to decompose the disturbance term and extract some important research input variables from it. In particular, we seek to test the statistical reliability of the stock of human capital, as indicated as FTEs, and stock of physical research capital, as indicated by the value of research equipment, on the knowledge production function. Although adding more independent variables into the objective function has both advantages and disadvantages, these two variables makes it possible to derive the best linear unbiased estimators of a knowledge production function[15], and improve on the connection to classical production theory[16]. Therefore, we hypothesize that both FTEs and the value of equipment are positively related to the measures of university research outputs.

2. Model specification on the lags of research expenditures

Crespi and Geuna (2008) point out that only a few studies have focused on the relationship between investment in science and measures of research outputs. As shown as previous section, we want to build an empirical model of knowledge production function, employing polynomial distributed lags (PDL) of research expenditures, and to test the comparison between the PDL model and previous studies' model which employ *ad hoc* lags of research expenditures on the knowledge production function. We hypothesize that the PDL model should have statistically

[15] The error term is to converge into zero and the model has the lowest variance of the estimate. It is also called Gauss–Markov theorem.
[16] In classical production function, the major inputs consist of labor, financial capital, and physical capital.

better results than ad hoc lag models and better explains unobservable effects of past research expenditures.

3. Research productivity and long-run effect

In order to assess university research productivity, there are some important indicators, which came from the classical production theory. Each slope coefficient on the knowledge production function model represents the elasticity of production, which is an indicator of short-run productivity. The sum of all independent variables' slope coefficients is the returns to scale of production, which is an indicator of long-run productivity. Basically, we expect that the productivity of university knowledge can be varied across the different types of knowledge dissemination channels because of their own inherent characteristics. First, we hypothesize that the public domain mechanisms of knowledge such as the published journal articles and doctoral degree awards exhibit degreasing returns to scale. According to Adams and Griliches (1998), at the aggregate data, the publication as research outputs follows constant returns to scale, but at the individual university level, the result follows decreasing returns to scale.

Second, we hypothesize that research outputs based on the invention and innovation, such as patenting/licensing mechanisms of knowledge exhibit increasing returns to scale. Artz et al (2010) study the innovation productivity of firms, examining the relationship between a firm's commitment to R&D and its outcomes in terms of innovations and inventions, and the impact of these outcomes on firm performance. Their findings show that the output of patents exhibits increasing returns to scale with respect to R&D spending, which contradicts previous studies with decreasing returns to scale, but which is consistent with an economic argument for advantages of scale in innovation, similar to findings in Henderson and Cockburn (1996) and Bettencourt et al, (2007).

B. Independent Regression Model Results

This section shows the results of independent regression models across the different knowledge outputs, in which we take into account a full range of potential knowledge dissemination channels: (1) published journal articles, (2) doctoral degree awards, (3) the collaboration index, and (4) the technology transfer index. The output measures associated with each of the knowledge dissemination channel are estimated with four different types of knowledge production function model, which are introduced in Section III, at the department level. We examine in this section the results of the knowledge production function models across the different knowledge dissemination channels, and how the heterogeneous characteristics of each knowledge channel may explain economic intuitions of research production in the university.

1. Estimating Output of Published Journal Articles

Published journal articles are the primary research output of knowledge production across the different colleges, departments, and research units of the university. Both Adams & Griliches (1998) and Pardey (1989) found that journal articles as an output of knowledge production were increasing in research expenditures, with some lag.

The total count of published journal articles by CSU authors from 1989 to 2012 was 38,916, compared to just 1,470 invention disclosures over the same time period. 10,275 or 26 percent of the university's total publications were from authors in the College of Natural Sciences, followed by 22 percent from the College of Veterinary Medicine, and 19 percent from the College of Engineering. Thus, these three colleges account for 67 percent, or two thirds, of the university's total research publications. Adding two more colleges, the College of Agricultural Sciences and the College of Natural Resources, altogether accounts for almost 85 percent (see Figure 7). Similarly in the department level, a relatively small number of departments account for a relatively large percentage of the university. Thus, the distributions of the publication production are highly skewed toward a few colleges or departments. In other worlds, a few number of colleges, departments, or research units shares of the total capacity or production of publications in the university (Scherer and Harhoff, 2000).

Table 3 displays the results of the panel estimation of the production function of published journal articles as a function of lagged research expenditures for the fifty-four departments of the university from 1989 to 2012. The model structures are based on Pakes and Griliches (1980, 1984), and the empirical technique follows closely from Crespi and Geuna (2008). However, it is distinguished from those earlier papers in that we have adopted a negative binomial maximum likelihood estimation (MLE) with polynomial distributed lags (PDL). The four different model approaches: model 1 is an unrestricted polynomial distributed lags (PDL) model; model 2 is a restricted PDL model with end-point restriction (k+1=0); model 3 is an ordinary distributed lag model or *ad hoc* distributed lag model similarly with previous studies; model 4 is a log-log PDL model with White's robust standard error, effective labor model, but not negative binomial MLE.

$$H = \left(\tilde{\beta}_{FE} - \tilde{\beta}_{RE}\right)' \left[\operatorname{var}\left(\tilde{\beta}_{FE}\right) - \operatorname{var}\left(\tilde{\beta}_{RE}\right)\right]^{-1} \left(\tilde{\beta}_{FE} - \tilde{\beta}_{RE}\right) \tag{19}$$

All four models assumed a group fixed effect. Equation (19) is the Hausman test, of which the result for model 1 rejects the null hypothesis of a random effect model at the 1 percent level of statistical significance; the Chi-square statistics is 25.88, and the p-value is 0.0005. Thus, the assumption of group fixed effects can be used in model 1. The test results of other models are comparable. There are several criteria for choosing the maximum length of the lag, *k*, and the degree of the polynomial, *m*. The preferred model is the one with minimum values of the AIC and SBIC. Further, we assumed that the end of the lag window, the k^{th} time period, must be statistically significant at least 5% level and that the maximum length of the lag cannot be greater than 10 years to prevent the loss of degrees of freedom. Using a top-down approach, we choose the lag length, *k*, and, in a similar manner, degree of polynomial, *m*.

Table 3—Regression results: Published journal articles at the department level, (1989-2012)

Published Journal Articles (1989-2012)

	Unrestricted PDL[3] (k=6, m=2)[1]	**Restricted PDL[4] (k=6, m=2)[1]**	**ODL[2] (k=6)[1]**	**Effective labor PDL[5] (k=6, m=2)[1]**
	[1]	[2]	[3]	[4]
Negative Binomial	YES	YES	YES	NO
Fixed Effect	YES	YES	YES	YES
Expenditure (t-0)	0.0375***	0.0269***	0.0478***	0.1073***
	(0.0071)	(0.0068)	(0.0120)	(0.0411)
_t-1	0.0072*	0.0152***	-0.0142	0.1632***
	(0.0037)	(0.0030)	(0.0164)	(0.0374)
_t-2	-0.0111**	0.0061***	-0.0099	0.1906***
	(0.0050)	(0.0013)	(0.0160)	(0.0455)
_t-3	-0.0174***	-0.0004	-0.0100	0.1895***
	(0.0053)	(0.0026)	(0.0166)	(0.0431)
_t-4	-0.0117***	-0.0042	-0.0041	0.1600***
	(0.0038)	(0.0034)	(0.0179)	(0.0339)
_t-5	0.0061	-0.0054*	-0.0335*	0.1020**
	(0.0045)	(0.0033)	(0.0180)	(0.0506)
_t-6	0.0360***	-0.0040*	0.0754***	0.0155
	(0.0112)	(0.0021)	(0.0202)	(0.1045)
Total lag multipliers	0.0466***	0.0342***	0.0514***	0.9280***
	(0.0080)	(0.0075)	(0.0084)	(0.1712)
Mean lag	2.9511	1.5521	3.5935	2.5385
FTEs (t_5)	0.0044***	0.0044***	0.0045***	—
	(0.0011)	(0.0011)	(0.0011)	
Equipment (3)[6]	0.0497**	0.0495**	0.0526***	0.0069
	(0.0199)	(0.0202)	(0.0201)	(0.0286)
Cons	2.0221***	2.0339***	2.0275***	2.3774***
	(0.1082)	(0.1069)	(0.1054)	(0.5715)
AIC[7]	5747.4	5758.7	5748.8	528.6
SBIC[7]	5776.6	5783.0	5797.4	546.2
log-likelihood	-2867.7	-2874.4	-2864.4	-260.3
Obs	954	954	954	589
Group	53	53	53	43

Note: ***1***. *k is the length of lags and m is the degree of polynomial,* ***2***. *Ordinary distributed lag or ad-hoc model, not PDL,* ***3***. *Unrestricted PDL model,* ***4***. *End-point restriction PDL model (k+1=0),* ***5***. *Output per unit of FTE with PDL model,* ***6***. *Window width is three,* ***7***. *Akaike Information Criterion and Schwarz' Bayesian Information Criterion. *** at 1%, ** at 5%, and * at 10% level of statistically significant. Parentheses are standard errors, but model 4 has White's robust standard errors*

In Table 3, the maximum length of lag in the publication models is 6 years and the better fit is obtained with a second-degree polynomial. In particular, the information criteria for the PDL models are much smaller than for the ODL model, with the one exception of the AIC for the restricted PDL model[17]. Moreover, the ODL model is not statistically significant in the middle time periods, from years 1 to 4, but the results of the PDL models are statistically significant at least at the 5 percent level, particularly in the unrestricted PDL model and the effective labor model.

In model 2, the restricted PDL model, the end-point restriction assumes that there is no impact beyond six years of lagged research expenditures on current year publications, yet the estimation result tells us that only three years[18]—the current, and the first and second lagged years' expenditures—affect strongly and positively the current year's publications. Similarly, model 4 finds that multiple years of past research expenditures per unit of FTE positively and significantly affect current publication counts, also measured per unit of FTE. This fourth model finds no negative effects of past research expenditures and finds the 2nd lagged time period's research expenditures to have the maximum impact. The other three models do not have a maximum point. However, model 1 does indicate small negative impacts for several years of lagged research expenditures.

The total lag multiplier represents a long-run or total impact of past and current research expenditures on current year publications. It measures how the log count of publications at department *i* change in response to changes in research expenditures. All four models have statistical significance, while model 3 has a higher magnitude of the coefficients than the others. Thus, research expenditures have a positive and significant long-run impact on publication counts

[17] The interpretation of AIC & SBIC is that the difference between 0 and 2 is only weakly significant, between 2 and 6 is significant, between 6 and 10 is strongly significant, and over 10 is very strongly significant.

[18] This is likely due to the characteristics of 2nd degree of polynomial PDL.

according to all models. The results also shed some light on the nature of the time lag between expenditures and resulting publications. The mean lag represents a weighted average of time in the short-run, and the formula for calculating it is shown in equation (20) below. This corresponds, for example, to the average lag between a research project's inception and completion. (In practice, actual expenditures typically begin some time after project inception, due to the delay in applying for and receiving funding. While at the same time, publications occur some time after project completion, due to the similar delay in submitting and publishing articles.) Model 1 tells us that, on average, research teams spend 2.95 years, or almost three years of their time on average in generating a publication: similarly, model 2 finds a mean lag of 1.55 years; model 3, 3.59 years; and model 4, 2.53 years.

By comparison, in previous studies, Pardey's mean lag of citation adjusted publications is 3.87 years in his OLS model, 3.30 years in his "within" model, 4.64 years in his "between" model, and 3.62 years in an EGLS model. In Crespi and Geuna's results, the unrestricted PDL model shows a mean lag of 3.48 years and the restricted PDL model, 4.16 years. Thus, this analysis finds a somewhat smaller mean lag compared to previous studies. The difference may be related to the different research settings: while this study uses data at a university department level, Pardey analyzed state agricultural experiment station (SAES) data at an institutional level; Crespi and Geuna used 14 countries of higher education research and development (HERD) data at a national level.

$$Mean\,lag = \frac{\sum_0^k k \cdot |\beta_k|}{\sum_0^k |\beta_k|} \qquad (20)$$

The full-time equivalent (FTE) researchers, lagged 5 years[19], as measure of human capital research input has positive and statistically significant (at the 1 percent level) impact on publication counts. This variable represents that *ceteris paribus*, for every unit increase in researcher FTE, the log count of publications five years later is increases by approximately 0.0044 in model 1 and 2 and by 0.0045 in model 3. In model 4, the FTE variable was controlled for as effective labor.

Next, we include the value of research equipment as a physical input to the research production process. There are several ways that equipment could be represented: as a fixed or stock variable that changes over time with the addition of new pieces of equipment and the depreciation of existing ones, or as a stream of services provided by the existing stock of equipment in any given year. The collected data represents the value of equipment at time of acquisition, but does not include the (highly heterogeneous) depreciation rates for each of those, making depreciation calculations impractical.

Understanding that the production of research outputs in a given year should be affected by equipment acquired in that same year as well as previous years cumulatively, we assume that this can be reasonably approximated by an average value of acquisitions within a moving window. In these models, the width of the window used is three years, which was selected by model specification tests, information criteria (AIC and SBIC), and the statistical significance of their estimated coefficients relative to windows of other sizes. The three-year moving average of annual values of equipment acquisitions in model 1 to model 3 have statistical significance at least at the 5 percent level, but the estimated coefficient on the variable in model 4 is insignificant. Thus, the value of research equipment per FTE, does not affect publications per FTE by department. Since model 1 to model 3 are negative binomial MLEs and thus their slope coefficients do not directly

[19] The lag length is also chosen by the smallest value of AIC and SBIC with comparing other lagged years and by stronger statistically significant of their estimated slope coefficients.

reveal the marginal effect, it needs to be calculated[20]. Instead, an alternative option is to report an incident rate ratio (IRR) for coefficients estimated in the negative binomial MLE model. The IRR displays the change in the research output, in this case publications, in terms of a percentage increase or decrease, with the precise percentage determined by the amount by which the IRR, deviates above or below one. Table 4 represents incident rate ratio of the three negative binomial MLE models, models 1 to 3, with the IRR values transformed by taking the exponent of the slope coefficients in Table 3.

Table 4—Negative binomial incident rate ratio (IRR): Published journal articles

Incident Rate Ratio (IRR): Published Journal Articles (1989-2012)			
	Unrestricted PDL (k=6, m=2)	**Restricted PDL (k=6, m=2)**	**ODL (k=6)**
	[1]	[2]	[3]
Expenditures (t-0)	1.0382***	1.0272***	1.0489***
_t-1	1.0072*	1.0153***	0.9859
_t-2	0.9889**	1.0061***	0.9901
_t-3	0.9827***	0.9996	0.9900
_t-4	0.9884***	0.9958	0.9959
_t-5	1.0062	0.9946*	0.9670*
_t-6	1.0366***	0.9960*	1.0784***
Total lag multipliers	1.0477***	1.0348***	1.0527***
FTEs (t_5)	1.0044***	1.0044***	1.0046***
Equipment (3)	1.0510**	1.0508**	1.0540***

In order to interpret the IRR values and take into account the marginal effects of the independent variable, the conversion $\left(IRR-1\right)\cdot 100=\%$ is more clear: effectively, the marginal value of the independent variable's impact on the dependent variable is the value by which its IRR value exceeds or falls below 1. For example, in the unrestricted PDL, the IRR value suggests that, *ceteris paribus*, for every one percent increase in lagged year 1 research expenditures, the number

20 $\partial E\left(Y_{i,t}\right)/\partial X_{i,t}=\partial \exp\left(X_{i,t}\beta\right)/\partial X_{i,t}=\beta \exp\left(X_{i,t}\beta\right)$

of publications are expected to increase by approximately 0.72 percent, with a statistical significance at the 10 percent level. In contrast, for every one percent increase in lagged year 2 research expenditures, publications decrease by 1.11 percent. Over the long-run, publications are increased by approximately 4.7 percent for every percent increase in research expenditures, with 1 percent level of statistical significance.

At the margin, FTEs lagged 5 years have a positive impact on current publication counts, with 1 percent level of statistical significance in all models. *Ceteris paribus*, for every one percent increase in FTE, lagged 5 years, current year publications can be expected to increase approximately 0.44 percent. Considering the value of research equipment acquisitions, one percent increase in the three year moving average value of research equipment acquisitions can be expected to increase current year publications by approximately 5.1 percent.

2. Estimating Output of Doctoral Degree Awards

Similar to published articles, doctoral degree awards represent an important output of the knowledge production process at the university. Each doctoral dissertation represents a body of research conducted at the university. In this section, we estimated doctoral degree awards as a function of research expenditures by department. We do not estimate master's degree awards, because the counts contain not only those master's degrees from research-based programs but also professional degree programs such as MBA, masters of social work, fine arts performance, and so on. In the same vein, we removed Doctor of Veterinary Medicine (DVM) degree awards from the count of doctoral degree awards. The departments of veterinary medicine as well as animal science do offer research-based PhD degrees, which were included.

Table 5 presents the regression results of doctoral degree awards at the department level from 1989 to 2012. The set of regression models is the same as in the previous section: unrestricted

PDL, restricted PDL, ordinary distributed lag, and effective labor models. All four have same lag length, k=8, and the three PDL models have the same degree of polynomial, m=4. Results of the PDL models generally show greater statistical significance than the ODL model, except for the restricted PDL model. Model 1, the unrestricted PDL, shows that research expenditures lagged 1, 5, 6, and 7 years affect positively the current year's number of doctoral degree awards. In particular, we might interpret the strong significance of research expenditures lagged 5 and 6 years to reflect the average total time to completion of a doctoral program. The investment decision for admitting and funding new Ph.D. students likely proceeds or corresponds to the first year of the program.

The slope coefficients on the middle range of lagged research expenditures, from years 2 to 4, indicate a negative relationship to the current year's count of doctoral degree awards; however, these statistically insignificant. However, these lags correspond to the time periods in the degree cycle when most students are taking core classes as well as preliminary or qualifying exams. Mean lags show that the average lag times between research expenditures and awarded degrees is 6.40 years in model 1: similarly, it is 4.50 years in model 2; 5.99 years in model 3; and 4.38 years in model 4. These results generally correspond with the reflections above on average time periods for the completion of doctoral degree programs.

In these regressions, FTEs and value of research equipment are insignificant. We expect, naturally, that the number of FTEs and the annual number of doctoral degree awards are highly correlated with each other, as the FTEs variable contains graduate research assistants as well as faculty who serve as Ph.D. advisors. Model 2, the PDL with endpoint restriction, k+1=0, shows that all estimated slope coefficients of lagged research expenditures are insignificant as well as FTEs and value of equipment variables. Only graduate enrollment is significant.

Table 5—Regression results: Doctoral degree awards at the department level, (1989-2012)

	Doctoral Degree Awards (1989-2012)			
	Unrestricted PDL[3] (k=8, m=4)[1]	Restricted PDL[4] (k=8, m=4)[1]	ODL[2] (k=8)[1]	Effective labor PDL[5] (k=8, m=4)[1]
	[1]	[2]	[3]	[4]
Negative Binomial	YES	YES	YES	NO
Fixed Effect	YES	YES	YES	YES
Expenditures (t-0)	-0.0043	0.0208	-0.0171	0.3835**
	(0.0247)	(0.0235)	(0.0287)	(0.1689)
_t-1	0.0270*	-0.0144	0.0607	0.3635***
	(0.0164)	(0.0144)	(0.0375)	(0.1318)
_t-2	-0.0033	-0.0079	-0.0477	0.0251
	(0.0148)	(0.0149)	(0.0381)	(0.1200)
_t-3	-0.0251	0.0111	0.0031	-0.2595***
	(0.0156)	(0.0128)	(0.0404)	(0.0736)
_t-4	-0.0058	0.0235	-0.0125	-0.3026***
	(0.0158)	(0.0149)	(0.0392)	(0.1070)
_t-5	0.0490***	0.0198	0.0496	-0.1013
	(0.0142)	(0.0131)	(0.0396)	(0.1104)
_t-6	0.0965***	0.0004	0.1082*	0.1627
	(0.0260)	(0.0132)	(0.0616)	(0.1214)
_t-7	0.0559*	-0.0243	0.0622	0.1232
	(0.0297)	(0.0237)	(0.0643)	(0.1232)
_t-8	-0.1915***	-0.0345	-0.2062***	-0.7709***
	(0.0463)	(0.0266)	(0.0551)	(0.2770)
Total lag multipliers	-0.0015	-0.0055	0.0002	-0.3763*
	(0.0270)	(0.0278)	(0.0274)	(0.2175)
Mean lag	6.4006	4.5001	5.9983	4.3794
FTEs (t_2)	0.0023	0.0025	0.0023	—
	(0.0027)	(0.0027)	(0.0027)	
Equipment (2)[6]	-0.0361	-0.0019	-0.0394	-0.0319
	(0.0477)	(0.0472)	(0.0477)	(0.0359)
Enrollment (t_4)[7]	0.0022***	0.0022***	0.0022***	0.4540***
	(0.0003)	(0.0003)	(0.0003)	(0.1104)
Cons	4.5634**	3.6873***	4.8101	-4.6806***
	(2.3020)	(1.0116)	(2.9509)	(0.6112)
AIC[8]	1947.7	1963.3	1953.9	537.5
SBIC[8]	1985.8	1997.2	2009.0	565.1
log-likelihood	-964.8	-973.6	-963.9	-261.8
Obs	512	512	512	383
Group	32	32	32	30

Note: ***1***. *k is the length of lags and m is the degree of polynomial,* ***2***. *Ordinary distributed lag or ad-hoc model, not PDL,* ***3***. *Unrestricted PDL model,* ***4***. *End-point restriction PDL model (k+1=0),* ***5***. *Output per unit of FTE with PDL model,* ***6***. *Window width is two,* ***7***. *Graduate enrollments,* ***8***. *Akaike Information Criterion and Schwarz' Bayesian Information Criterion. *** at 1%, ** at 5%, and * at 10% level of statistically significant. Parentheses are standard errors, but model 4 has White's robust standard errors*

There are two possible explanations for this result: one would be model misspecification and another would be the characteristics of the imposed restriction. First of all, if there were model misspecification, then the unrestricted PDL model should also be insignificant, because both models are strongly related to each other. Considering the second possibility, endpoint restrictions are generally imposed to cut off the effect of longer lagged time periods on the current output variable. As this seems a more likely explanation, we expect that doctoral degree awards are connected with longer and cumulative lagged years of research expenditures as well as investments in graduate education. To the extent that doctoral degree awards reflect at least as much about the educational attributes of a department, as the research attributes, it is not surprising to find that doctoral degree awards are observed to have a less systematic relationship with research inputs.

Finally, model 4 in Table 5 shows that only two years' of research expenditures per FTE, the current and one lagged year, affect positively the department's doctoral degree awards per research FTE. With regard to the negative relationship observed for research expenditures lagged 3 to 5 years, this may again reflect the doctoral program's gestation period, that of taking core classes and qualifying exams. However, we expect that the current year and the one prior year to a doctoral degree grant are generally the time periods of doctoral students doing their dissertation, such that the research expenditures per unit of FTE during those years affect positively doctoral degree awards per unit of FTEs.

Table 6 displays the incident rate ratio (IRR) of three negative binomial MLE models. First, model 1's IRR shows that, ceteris paribus, for every one percent increase in research expenditures, the log count of doctoral degree awards is expected to increase by 5.03 percent five years later and by 10.14 percent 6 years later, with 1 percent level of statistical significance.

Table 6—Negative binomial incident rate ratio (IRR): Doctoral degree award

Incident Rate Ratio (IRR): Doctoral Degree Awards (1989-2012)			
	Unrestricted PDL (k=8, m=4)	**Restricted PDL (k=8, m=4)**	**ODL (k=8)**
	[1]	[2]	[3]
Expenditures (t-0)	0.9957	1.0210	0.9830
_t-1	1.0274*	0.9857	1.0625
_t-2	0.9967	0.9921	0.9534
_t-3	0.9752	1.0112	1.0032
_t-4	0.9942	1.0238	0.9876
_t-5	1.0503***	1.0200	1.0508
_t-6	1.1014***	1.0004	1.1142*
_t-7	1.0575*	0.9759	1.0642
_t-8	0.8257***	0.9661	0.8137***
Total lag multipliers	0.9985	0.9945	1.0002
FTEs (t_2)	1.0023	1.0025	1.0023
Equipment (2)	0.9645	0.9981	0.9614
Enrollment (t_4)	1.0022***	1.0022***	1.0022***

3. Estimating Output captured in the Collaboration Index

For many years, the role of the university has included not only basic research and education, but also outreach and informal knowledge transfer activities, sometimes referred to as the "third mission" of the university. Faculty members in the university have long worked with industry and private sectors recipients of their knowledge through a variety of relationships, including consulting, conferences, informal conversations, practitioner networks, collaborative research, and co-supervising or internship programs for training R&D staff. Such activities emerged early in the history of the research university in the 19th and early 20th centuries and became embedded as important knowledge dissemination channels from the university to industry.

However, such collaborative activities by university researchers are often hard to detect in terms of their magnitude, size, and scope, particularly in a systematically consistent way across the entire institution. In a regression analysis, this problem can cause empirical model specification errors and distort estimation results. To overcome this, we propose three separate variables as reasonable proxies of knowledge outputs disseminated through channels based on collaboration or close interaction between university and industry: (1) industry co-authorship on academic journal articles, (2) the value of grants and contracts awarded from private sector sponsors, and (3) the departmental level expenditures on cooperative extension appointments and activities. Again, a major advantage is that each of these can be collected systematically for all departments and research units across the university.

First, the total count of journal articles with at least one industry co-author was 2,960 from 1989 to 2012. The College of Veterinary Medicine and the College of Engineering have published 1,059 and 810, respectively representing about 63 percent of the university total. At the department level, the Department of Clinical Sciences (in the College of Veterinary Medicine) and the Department of Electrical and Computer Engineering (in the College of Engineering) have published 523 and 286 articles with industry co-authors, respectively.

Second, across all of CSU, the total amount of private grant and contract awards was $268.7 million, which represents about 6.8 percent of total grant awards the university total of $3,95 billion from 1989 to 2012. The College of Veterinary Medicine and College of Engineering have received $93.1 and $66.8 million in private contract and grant awards, respectively.

Third, CSU's total amount of departmental level expenditures on extension was $108.5 million from 2003 to 2014, with the College of Agricultural Sciences accounting for the largest college level share at $24.4 million. One problem with this data is that most departments have no extension expenditures.

In this section, we estimate the collaboration-based activities of the university departments as a single research output variable by combining these three proxy variables into one single count or index. First of all, since three different variables have different data types, including both count and financial data, we transform the count data of industry co-authored articles into a comparable financial-type metric, according to equation (21),

$$DIA_{i,t} = \left[\left(\frac{TPubs_{i,t}}{R_{i,t}} \right)^{-1} \right] \cdot IA_{i,t} \tag{21}$$

where *DIA* is the dollar-equivalent value attributed to industry co-authored articles, *TPubs* is the total publication count, *IA* is industry co-authored articles, a subset of total publications, and R is research expenditures, all of which are specific to department *i* and time period *t*. In effect, in equation (21), the ratio between research expenditures and total publications per year represents the amount of (non-lagged) research expenditures attributed to a single paper in that year. Thus, *DIA* is effectively the research expenditure value that can be associated with the department's sum of articles co-authored with an industry collaborator in that year, in millions of dollars.

Then, as described in equation (22) we create a linear combination of the three proxy variables for collaboration activities, denoting this the *collaboration index*.

$$Coll_{i,t} = DIA_{i,t} + PGrant_{i,t} + Extension_{i,t} \tag{22}$$

where *PGrant* is the value of contract and grant awards from private sector sponsors and *Extension* is the extension budget[21] for department *i* in time period *t*. This collaboration index, *Coll*, is essentially financial data type (million \$). In order to make appropriate for use within the negative binomial MLE and consistent with other research outputs, including publications, doctoral degree awards, and our tech transfer index (see next section), we simply transform the collaboration index into a count data type by rounding off to units of ten thousand dollar increments: see equation (23).

[21] The time coverage of extension budget is from 2003 to 2012 comparing the other two variables from 1989 to 2012.

$$Collu_{i,t} = \left(Coll_{i,t}\right) \cdot 100\Big|_{Round\ off\ all\ decimal\ points} \tag{23}$$

The correlation between equation (22), the continuous variable, and (23), the rounded-off count variable, is 0.9999, almost perfectly correlated.

Table 7 presents regression results of the collaboration index, at the department level, from 1989 to 2012. Models 1 to model 3 are negative binomial MLE models and model 4 is log-log PDL model with White's robust standard error, an effective labor model. As in previous sections, all of the PDL models have better fit than the ordinary distributed lag model. Model 1, the unrestricted PDL, shows that the 2nd and 3rd years of lagged research expenditures are positively related to a given year's collaboration activities, as well as, interestingly, the 9th year's lagged research expenditures. However, research expenditures from lagged years 4 to 8 years have a negative relationship with a given year's level of collaborations.

The cumulative effect of the positive coefficients, at 0.2835, is somewhat greater than that of the negative coefficients, at -0.2464 (not considering their statistical significance). All three of the count models (i.e. except for model 4), have certain continuous strings of years indicating a negative impact of research expenditures on the collaboration index in the current year t=0. There may be several explanations. One possibility is that this may reflect a search cost for collaborating with private sector partners. Or, it may indicate that private sector collaboration is a strategy adopted in the wake of several 'lean' years of research funding. In other words, the negative relationship is effectively inverse: when research expenditures are down, the collaboration index goes up. Or, conversely, this result may indicate that in the years following overall increased research expenditures, the large bulk of which comes from federal and state sources, industry collaboration activities tend to get "crowded out".

Table 7—Regression results: Collaboration count at the department level, (1989-2012)

	Collaboration count (1989-2012)			
	Unrestricted PDL[3] (k=9, m=4)[1]	Restricted PDL[4] (k=9, m=4)[1]	ODL2 (k=9)[1]	Effective labor PDL[5] (k=9, m=2)[1]
	[1]	[2]	[3]	[4]
Negative Binomial	YES	YES	YES	NO
Fixed Effect	YES	YES	YES	YES
Expenditures (t-0)	0.0326	-0.0173	0.0459	0.3181***
	(0.0214)	(0.0210)	(0.0288)	(0.1061)
_t-1	0.0206	0.0365***	-0.0031	0.2248***
	(0.0128)	(0.0123)	(0.0435)	(0.0862)
_t-2	0.0249*	0.0298**	0.0476	0.1570
	(0.0132)	(0.0143)	(0.0427)	(0.0989)
_t-3	0.0204**	-0.0006	0.0199	0.1146
	(0.0101)	(0.0098)	(0.0382)	(0.1111)
_t-4	-0.0037	-0.0287***	0.0020	0.0977
	(0.0109)	(0.0098)	(0.0328)	(0.1099)
_t-5	-0.0444***	-0.0388***	-0.0593*	0.1061
	(0.0122)	(0.0125)	(0.0357)	(0.0938)
_t-6	-0.0842***	-0.0256**	-0.1250*	0.1400*
	(0.0194)	(0.0123)	(0.0661)	(0.0723)
_t-7	-0.0920***	0.0057	-0.0404	0.1994**
	(0.0275)	(0.0126)	(0.0652)	(0.0831)
_t-8	-0.0221	0.0396**	-0.0501	0.2841*
	(0.0218)	(0.0167)	(0.0677)	(0.1494)
_t-9	0.1850***	0.0498***	0.1999***	0.3943
	(0.0403)	(0.0167)	(0.0560)	(0.2506)
Total lag multipliers	0.0370*	0.0503**	0.0375*	2.0362***
	(0.0226)	(0.0232)	(0.0218)	(0.4669)
Mean lag	6.3400	5.0117	6.2293	4.8431
FTEs (t_6)	0.0116***	0.0102***	0.0121***	—
	(0.0019)	(0.0019)	(0.0019)	
Equipment (2)[6]	0.1049**	0.1627***	0.1171**	-0.0123
	(0.0530)	(0.0505)	(0.0571)	(0.0608)
Cons	-1.0137***	-0.9856***	-1.0507***	6.0247***
	(0.1111)	(0.1122)	(0.1136)	(1.3813)
AIC[7]	4922.8	4936.5	4928.8	1025.4
SBIC[7]	4959.4	4968.5	4988.3	1041.6
log-likelihood	-2453.4	-2461.2	-2451.4	-508.7
Obs	720	720	720	427
Group	48	48	48	41

Note: ***1***. *k is the length of lags and m is the degree of polynomial,* ***2***. *Ordinary distributed lag or ad-hoc model, not PDL,* ***3***. *Unrestricted PDL model,* ***4***. *End-point restriction PDL model (k+1=0),* ***5***. *Output per unit of FTE with PDL model,* ***6***. *Window width is two,* ***7***. *Akaike Information Criterion and Schwarz' Bayesian Information Criterion. *** at 1%, ** at 5%, and * at 10% level of statistically significant. Parentheses are standard errors, but model 4 has White's robust standard errors*

The total lag multipliers, reflecting the total or long-run impact of research expenditures, indicate they have a positive impact on the current collaboration count, with at least a 10 percent level of statistical significance across all four models. Mean lags in each model can be interpreted to represent the average lag between effective inputs and the measured output, or the so-called gestation period. These values show that changes in research expenditures seem to affect collaboration activities 6.34 years later in model 1: similarly, 5.011 years later model 2; 6.23 years in model 3; and 4.84 years in model 4. The lags here are two or three years longer than those observed in the publications equations and similar to the years observed with doctoral degree awards equations.

The number of FTEs, lagged 6 years, impacts the current collaboration index, at the 1 percent level of confidence, in models 1 to 3. *Ceteris paribus*, for every one unit increase in FTEs, the log of the collaboration index six years later increases by 0.0116 in model 1, by 0.0102 in model 2, and by 0.0121 in model 3. This may reflect that growth in the number of researchers in a department will, over time, expand capacity or opportunities for diversification by members of that department into engagement with industry.

The two-year moving window of the value of research equipment acquisitions has a positive and statistically significant relationship, at the 5 percent level, with the collaboration index. Interestingly, the magnitude of the estimated slope coefficient for equipment acquisitions is much higher than those of other input variables in these models. This might reflect that collaborative research is often associated with specific laboratory and engineering equipment. In general, we find that in the publication models and the doctoral degree award models, the magnitudes of the parameter estimates on the equipment variable are much smaller than in the collaboration models and the tech transfer models.

Table 8—Negative binomial incident rate ratio (IRR): Collaboration count

Incident Rate Ratio (IRR): Collaboration count (1989-2012)			
	Unrestricted PDL (k=9, m=4)	Restricted PDL (k=9, m=4)	ODL (k=9)
	[1]	[2]	[3]
Expenditures (t-0)	1.0331	0.9828	1.0470
_t-1	1.0208	1.0372***	0.9969
_t-2	1.0252*	1.0302**	1.0487
_t-3	1.0206**	0.9994	1.0201
_t-4	0.9963	0.9717***	1.0020
_t-5	0.9566***	0.9619***	0.9425*
_t-6	0.9192***	0.9747**	0.8825*
_t-7	0.9121***	1.0057	0.9604
_t-8	0.9781	1.0404**	0.9512
_t-9	1.2032***	1.0511***	1.2213***
Total lag multipliers	1.0377*	1.0516**	1.0382*
FTEs (t_6)	1.0116***	1.0102***	1.0121***
Equipment (2)	1.1106**	1.1767***	1.1242**

Table 8 shows the incident rate ratio (IRR) of the collaboration index models, derived from Table 7 by taking the exponents of the slope coefficients. These show us that, for example, in model 1, *ceteris paribus*, for every one percent increase in research expenditures, the log count of the collaboration index after a two year lag will be increased by 2.52 percent, with 90 percent confidence and after a three year lag will be increased by 2.06 percent with 99 percent confidence.

However, research expenditures tend to have a decreasing effect on the collaboration index, lagged four to eight years. As discussed above, this negative relationship between research expenditures and the collaboration index may represent either a search cost, reliance on industry collaboration as an alternative when public funding sources are lacking, or a crowding out of industry collaboration by publicly funded research, over the time period 4 to 8 years earlier. *Ceteris paribus*, for every one percent increase in research expenditures over this time period, the collaboration index will be subsequently decreased by approximately 23.7%.

4. Estimating Output of Combined Tech Transfer Metrics

In terms of university knowledge dissemination and impact, traditional collaborations, as introduced in the previous section, have limited scope of activities, subject to constraints such as geographical proximity, non-divisibility of time and persons, higher transaction costs, informality, strategic and political considerations, and so on, even though they represent one of the major knowledge dissemination channels available in the university research context. Alternatively, patenting and licensing, and the creation of startup companies, represent a more recently emphasized mode of dissemination and impact, which, as such, have received much attention and analysis, particularly regarding the role and use of intellectual property rights (IPRs) in the commercialization of university knowledge. However, these IPR-mediated technology transfer activities, when compared with the other knowledge dissemination channels of the university, are still a minor mechanism. Yet, it may have potential for growth as a mechanism for managing knowledge outputs and as well as generating revenues for the university in the future.

In this section, we estimate CSU's production of research outputs disseminated via such tech transfer activities, using the same empirical regression models and methodology employed in the previous sections for research outputs disseminated via the public domain and research outputs disseminated via collaboration, even though the overall magnitude of tech transfer activities is considerably smaller.

CSU's total invention disclosures from 1989 to 2012 was 1,564, and of these 488 were made by the College of Engineering and 455 were from the College Veterinary Medicine. Thus, overall, 60 percent came from just these two colleges. At the department level, the Department of Electrical and Computer Engineering has the highest count of invention disclosures, at 229. Following the disclosure of an invention comes the decision to make a patent application. The university's total count of patent applications filed and on which patents were granted were 195 from 1989 to 2012. The College of Engineering, College Veterinary Medicine, and College of Natural Sciences accounted for 69, 58, and 49 patent filings, respectively. The university's total count of startup companies was 41 from 1989 to 2012.

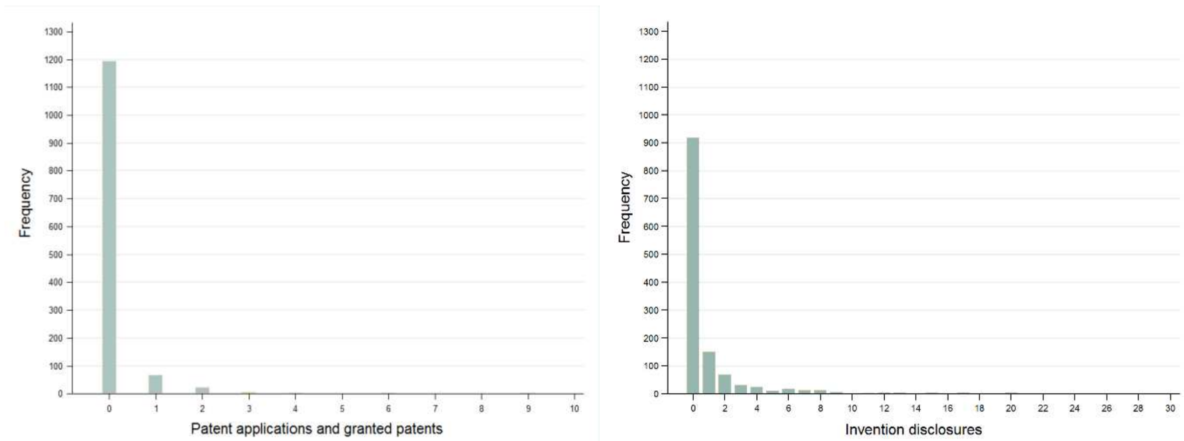


Figure 17—Frequency of invention disclosures and patent filings at 24 years overall, (1989-2012)

In the empirical model, which considers invention disclosures, patent filings, and startups made by each department each year, these three variables have very few observations (often zero, or just one or two inventions by department per year), as shown in Figure 17. Thus, these variables cannot be estimated without a zero inflated count model regression procedure. Another possibility is to create a linear combination of these three variables into a single variable. This, in turn, may introduce multi-counting of single inventions. The typical technology transfer process can consist of invention disclosures, patent applications, issued patents, licensing contracts, and even startup companies, all around a single invention. Although, it might be interpreted that a linear combination of the three count variables thus gives more weight to those inventions that proceed further through the typical technology transfer process.

Equation (24) represents the combined variable, which we call the *tech transfer metrics* (TTM) index, which consists of the sum of invention disclosures (*invention*), patent applications & granted patents, (*patent*), and the number of startup companies by CSU inventors, (*startups*), for department *i* in time period *t*.

$$TTM_{i,t} = invension_{i,t} + patent_{i,t} + startups_{i,t} \tag{24}$$

As shown Table 9, we run four different models, the same as we have done for the other research output models in previous sections. All three PDL models have greater statistical significance and better goodness of fit than the ODL model. In model 1, the unrestricted PDL has positive and significant lag coefficients of research expenditures from the current year to the 4th lagged year. Similarly, model 2, the endpoint restricted PDL (k+1=0), shows positive and significant impacts of research expenditures from the current year through year 3, but this model also has a longer period of zero or negative relationships between research expenditures and tech transfer activities, from lagged years 4 to 9.

According to Heher (2007), the tech transfer process typically takes six to ten years from the moment of invention disclosure to the time when significant income can be generated from a license. In this regard, the negative relationship with research expenditures over longer time periods could be connected with the gestational or provisional periods of innovation, proof of concepts, and such. It also may be an artifact of the data: as before a major restructuring in the university's tech transfer office in 2007, invention disclosures, patent filings, and startups were relatively quiet, even though research expenditures had been growing rapidly. After 2007 the positive relationship may have been restored. The possibility of a structural break in this relationship is explored further below.

Table 9—Regression results: Combined tech transfer metrics at the department level, (1989-2012)

Combined Tech Transfer Metrics (1989-2012)				
	Unrestricted PDL[3] (k=9, m=2)[1]	**Restricted PDL[4] (k=9, m=2)[1]**	**ODL[2] (k=9)[1]**	**Effective labor PDL[5] (k=9, m=2)[1]**
	[1]	[2]	[3]	[4]
Negative Binomial	YES	YES	YES	NO
Fixed Effect	YES	YES	YES	YES
Expenditures (t-0)	0.0414**	0.0470***	0.0766**	0.0326
	(0.0179)	(0.0177)	(0.0385)	(0.1476)
_t-1	0.0415***	0.0303***	-0.0240	0.2075**
	(0.0112)	(0.0103)	(0.0591)	(0.0939)
_t-2	0.0380***	0.0162***	0.1021*	0.3263***
	(0.0102)	(0.0046)	(0.0609)	(0.0786)
_t-3	0.0308***	0.0049*	-0.0716	0.3891***
	(0.0114)	(0.0029)	(0.0599)	(0.0877)
_t-4	0.0200*	-0.0038	0.0994*	0.3958***
	(0.0117)	(0.0055)	(0.0528)	(0.0964)
_t-5	0.0054	-0.0099	-0.0265	0.3463***
	(0.0102)	(0.0076)	(0.0526)	(0.0956)
_t-6	-0.0128	-0.0132	0.0162	0.2408***
	(0.0087)	(0.0086)	(0.0919)	(0.0867)
_t-7	-0.0347***	-0.0139*	-0.0082	0.0792
	(0.0120)	(0.0084)	(0.0898)	(0.0831)
_t-8	-0.0603***	-0.0120*	0.0412	-0.1385
	(0.0213)	(0.0069)	(0.0954)	(0.1094)
_t-9	-0.0896***	-0.0073*	-0.2206***	-0.4122**
	(0.0349)	(0.0041)	(0.0821)	(0.1719)
Total lag multipliers	-0.0203	0.0381*	-0.0155	1.4670***
	(0.0325)	(0.0227)	(0.0352)	(0.3735)
Mean lag	5.1411	5.1520	5.0149	4.7344
FTEs (t_6)	0.0071**	0.0053*	0.0076**	—
	(0.0032)	(0.0032)	(0.0035)	
Equipment (2)	0.1365**	0.1260*	0.1609**	-0.0491
	(0.0683)	(0.0704)	(0.0769)	(0.0410)
Cons	0.0896	0.0679	0.0041	0.7170
	(0.2963)	(0.2948)	(0.3143)	(1.1022)
AIC	1351.7	1355.3	1359.3	480.9
SBIC	1377.9	1377.1	1416.1	495.2
log-likelihood	-669.8	-672.6	-667.0	-236.5
Obs	585	585	585	260
Group	39	39	39	33

Note: ***1***. *k is the length of lags and m is the degree of polynomial,* ***2***. *Ordinary distributed lag or ad-hoc model, not PDL,* ***3***. *Unrestricted PDL model,* ***4***. *End-point restriction PDL model (k+1=0),* ***5***. *Output per unit of FTE with PDL model,* ***6***. *Window width is two,* ***7***. *Akaike Information Criterion and Schwarz' Bayesian Information Criterion. *** at 1%, ** at 5%, and * at 10% level of statistically significant. Parentheses are standard errors, but model 4 has White's robust standard errors*

Overall, the total or long-run impact of research expenditures on the tech transfer metric is not significant in models 1 & 3, but it is positive and significant in models 2 & 4. Mean lags indicate that the average lag time between research expenditures and the observation of an invention disclosure or other tech transfer event is 5.14 years in model 1: similarly, 5.15 years in model 2; 5.01 years in model 3; and 4.73 years in model 4. These time periods seem to be shorter than those calculated by Heher, but his measure includes the time required to realize positive royalty incomes, which typically involves a significantly longer wait. In Heher's result, the average time between the invention disclosure and final patent granted is 6 years, which is sufficiently close to the mean lag estimates here.

The human capital input variable, research FTEs, lagged 6 years, is also found to affect the combined tech transfer metric index positively, with a 5 percent level of statistical significance in models 1 & 3, and a 10 percent level of significance in model 2. This time lag of the influence of changes in the size of a department's research FTEs on its combined tech transfer metrics index value is similar to the time lag of its influences on the collaboration index in the previous section. Whether in the more traditional or the newer mode of knowledge transfer to industry, the timing of impact from increases in FTEs are similar.

The physical capital input variable, a two year moving average of the value of research equipment acquisitions, is significant in all except model 4, and it has a much higher magnitude of the estimated coefficient than do the other input variables. Again, this result is very similar to the results in the collaboration index model, in the previous section. Therefore, it appears that research equipment is important to the kinds of research that lead to both our measures of collaboration index and the combined tech transfer metrics.

Table 10 displays the incident rate ratio (IRR) of the tech transfer models. Model 1 indicates that, *ceteris paribus*, for every one percent increase in research expenditures, the log of the combined tech transfer metrics index is expected to increase by approximately 4.22 percent, similarly the log of tech transfer outputs increase by 4.24 percent from an increase of research expenditures lagged 1 year, by 3.88 percent from an increase in research expenditures lagged 2 years, by 3.31 percent, from an increase in expenditures lagged 3 years, and by 2.02 percent from an increase lagged 4 years. Model 1 has an (insignificant) negative long-run relationship of research expenditures to tech transfer metrics, but model 2 indicates the opposite.

Table 10—Negative binomial incident rate ratio (IRR): Combined tech transfer metrics

Incident Rate Ratio (IRR): Combined Tech Transfer Metrics (1989-2012)			
	Unrestricted PDL (k=9, m=2)	**Restricted PDL (k=9, m=2)**	**ODL (k=9)**
	[1]	[2]	[3]
Expenditures (t-0)	1.0422**	1.0481***	1.0797**
_t-1	1.0424***	1.0307***	0.9763
_t-2	1.0388***	1.0164***	1.1075*
_t-3	1.0313***	1.0049*	0.9309
_t-4	1.0202*	0.9962	1.1045*
_t-5	1.0054	0.9902	0.9738
_t-6	0.9873	0.9868	1.0163
_t-7	0.9659***	0.9862*	0.9918
_t-8	0.9415***	0.9881*	1.0421
_t-9	0.9143***	0.9927*	0.8020***
Total lag multipliers	0.9799	1.0389*	0.9847
FTEs (t_6)	1.0071**	1.0053*	1.0076**
Equipment (2)	1.1463**	1.1343*	1.1746**

5. Returns to scale and long-run productivity

In this section, we consider the university's research productivity, across the different knowledge dissemination channels in the long-run. In the production of knowledge measured by published journal articles, the interpretation of estimated slope coefficients in the effective labor or log-log PDL model (model 4 in Table 3) are marginal effects directly, due to the logarithm of the Cobb-Douglas production function. They also represent output elasticities with respect to each lagged value of research expenditures. In the previous section, it was hypothesized that university knowledge production should exhibit decreasing returns to scale, which would mean the sum of all estimated slope coefficients in this model equals one in absolute value, but in the estimates of the model, is 0.9349, or slightly below one, which indicates instead "very slightly" decreasing returns to scale (DRTS) in the publications output model. We expect that publications have significant public good attributes and represent the most basic knowledge output within the university. Therefore, it is more closely associated with classical production theory's production function with the law of diminishing marginal returns.

In the production of knowledge measured by doctoral degree awards, the sum of estimated slope coefficients (in model 4 in Table 5) is 0.0458; as it is below one that can be interpreted to indicate decreasing returns to scale. As in the case of publication (see model 4 in Table 3), doctoral degree awards are understood to be a proxy for a kind of knowledge output that has public-good attributes; the result of doctoral research is, predominantly, one of the basic knowledge outputs within the university context. However, the results of long-run productivity in the publication and doctoral degree awards are different each other even though they exhibit decreasing returns to scales with respect to all input variables, as is consistent with classical production theory. The doctoral degree awards have a much lower productivity in the long-run than the publications, because of the different characteristics. The doctoral degree awards are more likely to be close to the educational mission in the university, so it is distinguished from the general research outputs.

In contrast, the sum of all the estimated slope coefficients in the effective labor model of the collaboration index (in model 4 in Table 7) is 2.023, which is interpreted as indicating that the collaborative types of research outputs exhibit *increasing returns to scale* with respect to all input variables. Similarly, Model 4 in the regressions (Table 9) shows that, similarly, the combined tech transfer metrics index per FTE increases significantly from greater research expenditures per FTEs in the time periods one to six years earlier. Again, the sum of all slope coefficients, at 1.4179, is greater than one, indicating increasing returns to scale. Generally interpreted this means that a doubling in R&D spending would lead to more than a doubling of the tech transfer outputs, and, thus in long-run, it exhibits an economies of scale.

In terms of production of knowledge associated with inventions and innovations, the assessment of productivity is complicated, especially in an empirical study, because of prior conditions and variation in propensity to patent across different technologies and fields. However, the result of increasing returns to scale is still meaningful in the knowledge production. Another possibility of the results of increasing returns to scale in the collaboration and formal tech transfer outputs might be an empirical bias introduced from the linear combination of different proxy variables or the related skewed distribution of observations across the different departments and research units. Nevertheless, it is important that we test the hypothesis, that in the long-run, the productivity of university knowledge production is varied across the different types of knowledge dissemination channels. Our findings indicate that the publication and doctoral degree award models of knowledge disseminated via the public domain mechanism exhibit diminishing returns to scale with respect to all input variables. Meanwhile, the collaboration and tech transfer models exhibit increasing returns to scale of knowledge production.

C. System of Equation Model Results

So far, we have assumed that the different types of knowledge creation are independent of one another. However, it is certainly reasonable to assume that the various types of research outputs are interrelated as co-products of the underlying knowledge production efforts of the university.

Moreover, if the error terms among the models of different types of research outputs are correlated for a given department, but are uncorrelated across departments, we cannot derive efficient and best linear unbiased estimators from independent regression models of each research output, as we have attempted in the previous sections. A better approach is a seemingly unrelated regression (SUR) model (Zellner, 1962), a system of linear equations that consists of j linear regression equations for i departments and t time periods.

In the previous sections, we utilized four sets of independent regression models, one each for publications, doctoral degree awards, collaborative outputs, and tech transfer metrics, but in this section, we use only three output measures--publications, the collaboration index, and the combined tech transfer metrics--within a system of equations. Moreover, because of the non-linear regression[22], we use the effective labor model, a log-log PDL model with White's robust standard errors. Equation (25) represents a matrix version of the seemingly unrelated regression with group fixed effect panel polynomial distributed lag (PDL) model for department i and time period t, and equation (26) is a compact version of the matrix:

$$\begin{bmatrix} y_{i,t,1} \\ y_{i,t,2} \\ y_{i,t,3} \end{bmatrix} = \begin{bmatrix} (\alpha_1 + u_{i,1}) & r_{i,t,1} & \cdots & r_{i,t-(k-3),1} & 0 & 0 & 0 & eq_{i,t} \\ (\alpha_2 + u_{i,2}) & r_{i,t,2} & \cdots & r_{i,t-(k-3),1} & r_{i,t-(k-2),2} & r_{i,t-(k-1),2} & r_{i,t-k,2} & eq_{i,t} \\ (\alpha_3 + u_{i,3}) & r_{i,t,3} & \cdots & r_{i,t-(k-3),1} & r_{i,t-(k-2),3} & r_{i,t-(k-1),3} & r_{i,t-k,3} & eq_{i,t} \end{bmatrix} \cdot \begin{bmatrix} 1 \\ \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \\ \beta_E \end{bmatrix} + \begin{bmatrix} e_{i,t,1} \\ e_{i,t,2} \\ e_{i,t,3} \end{bmatrix} \tag{25}$$

$$\mathbf{Y} = \mathbf{X} \cdot \mathbf{B} + \mathbf{E} \tag{26}$$

Where the third subscript represents the output type: 1=publication, 2=collaboration, and 3= tech transfer. y is a logarithm of the respective research output per unit of labor (per research FTE),

[22] There exists a non-linear SUR model technically, see Winkelmann (2000).

r is the logarithm of research expenditures per unit of labor (per research FTE), and eq is the logarithm of the two year moving window of value of research equipment acquisitions per unit of labor (per research FTE). u_i is a time-invariant error term and $e_{i,t}$ is a group and time-variant idiosyncratic error term for each equation. $\mathbf{Y}$ is 3×1 matrix that consists of a 3×12 matrix of regressors included in a constant term, $\mathbf{X}$, a 12×1 matrix of coefficients, $\mathbf{B}$, and a 3×1 matrix of error terms, $\mathbf{E}$. β_0 to β_k are slope coefficients of the calculated and recovered lag scheme of the polynomial distributed lag (PDL) model with lag length k and degree m. (For more detail, see Section III.)

Table 11 presents regression results of the estimated SUR model, with bootstrapped standard errors, that consists of four different sets of equations at the department level from 1989 to 2012 as well as polynomial distributed lag (PDL) slope coefficients of research expenditures. First is a system of (1) research publications and (2) the collaboration index as measures of output; second is a system of (1) research publications and (3) the combined tech transfer metrics as measures of output; third is (2) the collaboration index and (3) combined tech transfer metrics; and fourth is the system of all three together, (1), (2), and (3).

First, considering the research publications regression without considering correlations of the residuals, it is evident that the magnitude of estimated slope coefficients in model 1 in Table 11 are larger and their standard errors are smaller than in model 2 and 4. However, all three SUR models that include publications indicate a negative impact in the last lagged year of research expenditures, comparable to some of the results of the independent regression models in the previous section. In general, comparing between the independent regression models and the three SUR models that include publications as a research output, there are no notable changes or surprises.

Table 11—The regression results of SUR model with PDL of research expenditures at department level, (1989-2012)

	System of Equations: Seemingly Unrelated Regression PDL (1989-2012)								
	[1]		[2]		[3]		[4]		
	Publications (k=6, m=2)[1]	Collaboration Index (k=9, m=2)[1]	Publication (k=6, m=2)[1]	Tech Transfer Metrics (k=9, m=2)[1]	Collaboration Index (k=9, m=2)[1]	Tech Transfer Metrics (k=9, m=2)[1]	Publications (k=6, m=2)[1]	Collaboration Index (k=9, m=2)[1]	Tech Transfer Metrics (k=9, m=2)[1]
Width of Equip[2]	3	5	3	4	5	4	3	5	4
Expend (t_0)	0.1039*	0.3180***	0.1710*	0.0359	0.3621*	0.0149	0.1569**	0.3641*	0.0192
	(0.0594)	(0.1141)	(0.0878)	(0.2679)	(0.1981)	(0.1680)	(0.0737)	(0.1986)	(0.1683)
t_1	0.1827***	0.2253***	0.1741***	0.2142	0.1893	0.1960*	0.1897***	0.1935	0.2051*
	(0.0306)	(0.0768)	(0.0495)	(0.1877)	(0.1222)	(0.1084)	(0.0451)	(0.1229)	(0.1098)
t_2	0.2157***	0.1579*	0.1629***	0.3347**	0.0686	0.3206***	0.1910***	0.0740	0.3321***
	(0.0465)	(0.0862)	(0.0499)	(0.1396)	(0.1073)	(0.0894)	(0.0607)	(0.1079)	(0.0911)
t_3	0.2030***	0.1157	0.1376***	0.3975***	0.0000	0.3886***	0.1608**	0.0055	0.4002***
	(0.0528)	(0.1026)	(0.0511)	(0.1202)	(0.1225)	(0.0959)	(0.0672)	(0.1231)	(0.0969)
t_4	0.1444***	0.0987	0.0980**	0.4025***	-0.0164	0.4000***	0.0990*	-0.0120	0.4095***
	(0.0420)	(0.1058)	(0.0471)	(0.1152)	(0.1303)	(0.1018)	(0.0576)	(0.1309)	(0.1022)
t_5	0.0402	0.1071	0.0441	0.3499***	0.0192	0.3548***	0.0056	0.0216	0.3598***
	(0.0387)	(0.0922)	(0.0731)	(0.1111)	(0.1205)	(0.0965)	(0.0596)	(0.1210)	(0.0967)
t_6	-0.1099	0.1406**	-0.0239	0.2395**	0.1070	0.2530***	-0.1193	0.1062	0.2513***
	(0.0883)	(0.0681)	(0.1437)	(0.1059)	(0.1013)	(0.0820)	(0.1155)	(0.1007)	(0.0825)
t_7	—	0.1994***	—	0.0714	0.2469**	0.0947	—	0.2419**	0.0839
		(0.0697)		(0.1125)	(0.1110)	(0.0779)		(0.1076)	(0.0794)
t_8	—	0.2835**	—	-0.1544	0.4389**	-0.1203	—	0.4286**	-0.1424
		(0.1323)		(0.1503)	(0.1854)	(0.1168)		(0.1800)	(0.1189)
t_9	—	0.3928*	—	-0.4379	0.6831**	-0.3918**	—	0.6663**	-0.4276**
		(0.2324)		(0.2241)	(0.3091)	(0.1964)		(0.3024)	(0.1984)
Equipment	0.0274	-0.0127	0.0577	-0.0465	-0.0162	-0.0421	0.0616	-0.0145	-0.0383
	(0.0242)	(0.0539)	(0.0441)	(0.0602)	(0.0807)	(0.0548)	(0.0389)	(0.0796)	(0.0539)
cons	2.0269***	5.9367***	2.3217***	0.8504	7.4098***	1.0913	2.0684***	7.3840***	1.0361
	(0.3427)	(1.0779)	(0.6285)	(1.4448)	(1.3881)	(1.0427)	(0.5759)	(1.3595)	(1.0516)
R-sq	0.8054	0.5397	0.7196	0.5274	0.6073	0.5313	0.6977	0.6073	0.5313
Chi2	1763.28	499.51	667.47	290.59	383.49	281.17	573.14	383.06	281.66
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Obs	426	426	260	260	248	248	248	248	248

Note: ***1***. *k is lag length and m is degree of polynomial,* ***2***. *Window width. All models are unrestricted PDL. .*** at 1%, ** at 5%, and * at 10% level. Parentheses are Bootstrap standard errors.*

The SUR models indicate a lower returns to scale than the independent model, even though all still evidence diminishing returns to scale. This is consistent with the idea that publications as a research output are the most fundamental knowledge product, widely generated across all of the departments and research units of the university.

Second, considering the collaboration index as a measure of research output, without considering correlations of residuals for the moment, we see that the regression results of model 1 in Table 11 are similar to the independent models in the previous section; however, the collaboration index results in models 3 and model 4, including it in systems with the combined tech transfer metrics and with all other research outputs, respectively, are significantly different from the results of the independent models, including smaller slope coefficients and larger standard errors as well as differences in the statistical significance and the signs of each coefficient.

These results might be due to the characteristics of tradition collaborations between university and private sectors, in that there are many kinds of informal relationships and most of them are not measured here. We speculate that industry collaborations may be complements, or preliminary activities leading to other research outputs involving private sector partners in those research activities picked up in the combined tech transfer metrics.

Finally, looking at the combined tech transfer metrics models, without considering correlations of residuals, we see that the statistical significance of models 3 and 4 in Table 11 are somewhat better than model 2. Moreover, the magnitudes and signs of the slope coefficients of the SUR models are quite similar to those in the independent regression models in the previous section. This research output was infrequently occurring relative to the other research outputs within each of the respective systems of equations. Thus, we expect that although the tech transfer outputs—which consist of invention disclosures, patent applications, granted or issued patents, and startup companies—are still at an early stage of development and smaller than other research outputs, seem to have potential as a channel of knowledge dissemination in the university.

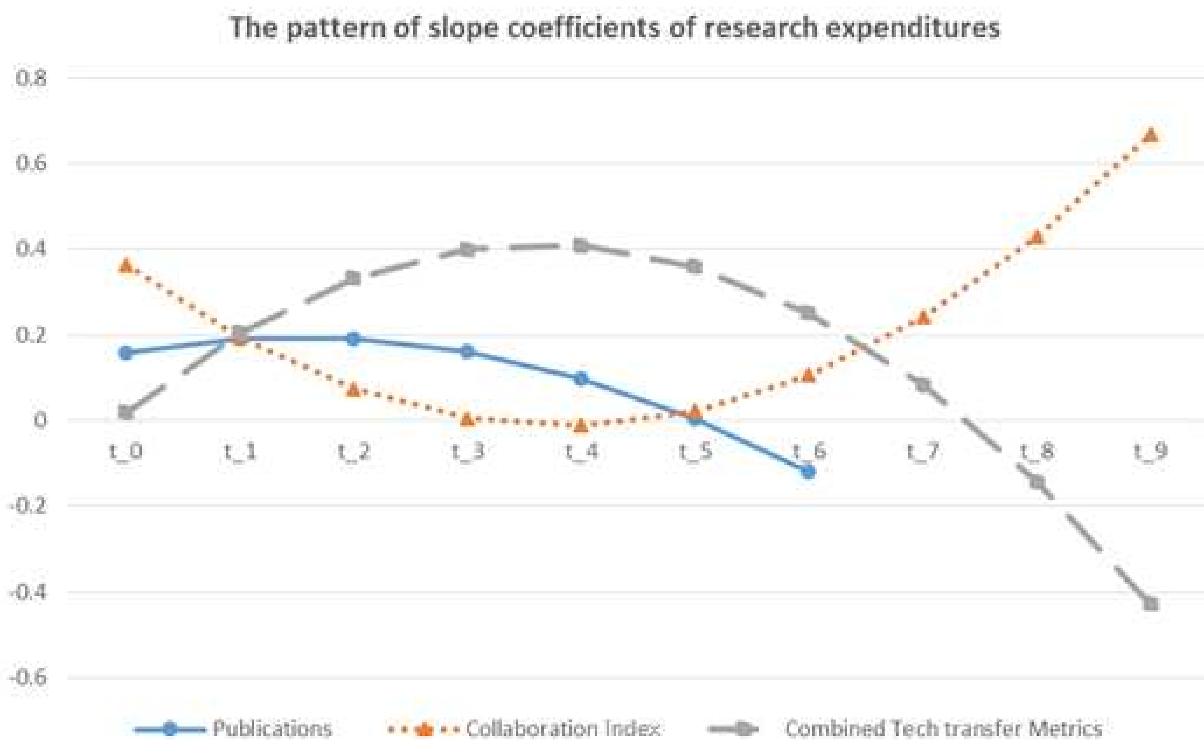


Figure 18—Comparison of estimated slope coefficients of three research outputs in model 4

Figure 18 represents the comparison of estimated slope coefficients of lagged research expenditures from the full system of equations, model 4 in Table 11. As shown in the figure, publications and combined tech transfer metrics have similar patterns of slope coefficients across lagged research expenditures, each following an inverted U-shape or "projectile" trajectory. In contrast, the pattern of coefficients on lagged research expenditures for the collaboration index equation shows a U-shape. From the figure, we can interpret that, holding the degree of polynomial constant, publications and combined tech transfer metrics as research outputs have a maximum impact, in terms of time lagged research expenditures per FTE, between the second and fourth year, respectively, after research expenditures, but the collaboration index exhibits a minimum in terms of impact from research expenditures made within roughly the same lagged time frame.

Moreover, interestingly, the minimum impact of lagged research expenditures on the collaboration index corresponds exactly to the maximum impact of lagged research expenditures on the combined tech transfer metrics, which may indicate that more traditional mechanisms of collaboration and newer tech transfer mechanisms have a substitutional relationship, at least within the time period of this analysis. In other words, department *i*'s research expenditures are tending to result in tech transfer outputs rather than the more traditional collaboration activities.

Table 12—The test results of correlation of residuals

	Publications & collaboration index	Publications & Combined tech transfer metrics	Collaboration index & Combined tech transfer metrics	All research outputs
	[1]	[2]	[3]	[4]
Correlation of residuals	-0.0041	0.0599	-0.0662	0.0336 for pub&coll[1] 0.0824 for pub&ttm[2] -0.0662 for coll&ttm[3]
Breusch-Pagan test of independence	chi2(1)=0.007 (0.9333)	chi2(1)= 0.932 (0.3344)	chi2(1)=1.087 (0.2971)	chi2(3)=3.052 (0.3837)

Note: ***1****. Correlation of residuals between publications & collaboration index,* ***2****. Correlation of residuals between publications & combined tech transfer metrics,* ***3.*** *Correlation of residuals between collaboration index & combined tech transfer metrics. The Parentheses are p-values.*

Finally, Table 12 displays the correlation of residuals in each system of equations model. In model 1, correlation between the errors in the two equations of publications and the collaboration index is -0.0041, and the Breusch-Pagan Lagrange multiplier test for error independence indicates this is not significantly different from 0. In other words, there is no significant correlation. Model 2 indicates a slightly positive, but insignificant correlation; and model 3, a slightly negative, but insignificant correlation. Model 4, the system of all three research output equations, has positive correlation of the residuals between publications and the collaboration index, in contrast to model 1, but the other two combinations of research outputs have correlations in the same directions as the comparable models 2 and 3, respectively.

Therefore, we reject the null hypothesis that the different types of research outputs are related with each other and their error terms are correlated. Instead it is appropriate to estimate independent regressions for each research output. However, without considering the correlation of residuals in SUR models, it would not be possible to support the intuitions and results we consider in the overall results and conclusions of this chapter.

D. Structural Changes

Panel data is based on both time series and cross sectional data. In time series data, there can be structural changes in the relationship between the dependent variable, in this case the research output, and one or more of the independent variables, in this case the research inputs. By structural change, we mean that the value of the parameters of the model do not remain the same through the entire time period, from 1989 to 2012. Structural change can occur at some point in time as a result of changes in policy, consumer preferences, new institutions and technical systems, or some other external shock. In particular, we want to test for structural changes in the model around the time of the creation of CSU Ventures as the technology transfer office (TTO) for Colorado State University.

In 1963, Colorado State University Research Foundation (CSURF) was officially established as an independent non-profit entity to hold and manage certain capital assets on behalf of the university, which, in time, came to include management of intellectual property arising from inventions made at the university. In a major restructuring of the technology transfer operations of CSURF in 2007, CSU Ventures was founded as a wholly-owned subsidiary of CSURF. Since CSU

Ventures specializes in the commercialization of university inventions, including the management of patenting and licensing, as well as the launch of startups, we want to test only the combined tech transfer metrics research output variable, which consists of the combination of the department level count of invention disclosures, patent applications and granted patents, and startup companies founded.

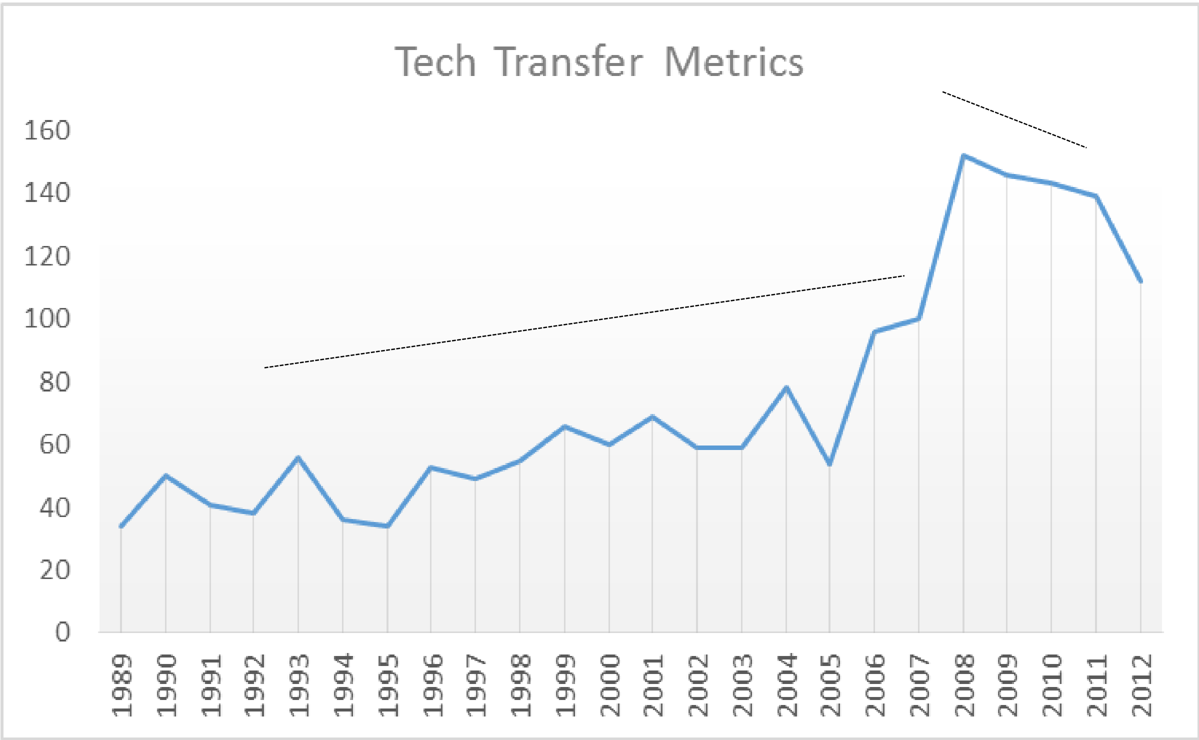


Figure 19—The time trend of the combined tech transfer metrics at CSU overall, (1989-2012)

In our database, the total sum of these combined tech transfer metrics is 1,779 for the entire university from 1989 to 2012. (For more detail, see Section IV). However, in the time period following the creation of CSU Ventures, 2008-2012, the total count was 792, which is about 45 percent of total tech transfer activities in just five years or 21 percent of the total time period covered in this study, 1989-2012.

Figure 19 shows the time trend of the combined tech transfer metrics at CSU overall from 1989 to 2012. There was a remarkable increase of combined tech transfer metrics after 2007. However, the explanation for this increase is not only the foundation of the new tech transfer office (TTO), but it may also represent the cumulative effects from previous research as well as tech transfer activities. Nevertheless, it is important to test a structural change because of the enormous changes that occurred in 2007 and persist for all of the years thereafter.

The Chow test is the most commonly used technique for the analysis of structural changes, and is generally used for testing for a structural change of known timing. There are two different methods of Chow test for the structural change: one is a sum of squared error prediction and another is a dummy variable model. Most statistic software packages can generate both test methods. We want to test two different time periods, the pre-period 1989-2006, and the post-period of 2007-2012, and assume that CSU Ventures significantly affects the number of university invention disclosures, patent applications & published or granted patents, and startup companies during 2007 to 2012.

Using a Chow test based on the dummy variable method we adopt a log-log PDL model with White's robust standard errors, with the combined tech transfer metric research output per research FTE as a function of research expenditures per research FTE, following a PDL scheme and a two-year moving-window average of the value of research equipment acquisitions per research FTE, as in Section IV-A-4. Equation (27) shows the structural changes in the knowledge production function with group fixed effects and a polynomial distributed lag scheme of research expenditures,

$$\begin{aligned} y_{i,t} = \left(\alpha + u_i\right) + \varphi_0 D_t + \beta_o r_{i,t} + \varphi_1 r_{i,t} D_t + \beta_1 r_{i,t-1} + \varphi_2 r_{i,t-1} \cdot D_t + \cdots + \beta_k r_{i,t-k} \\ + \varphi_{k+1} x_{i,t-k} D_t + \beta_E eq_{i,t} + \varphi_E eq_{i,t} D_t + e_{i,t} \end{aligned} \tag{27}$$

where D is a dummy variable with a value of zero for years 1989 through 2006 and a value of one for years 2007 through 2012.

Running the regression, we assume a null hypothesis of no structural change between 2007 and 2012 and an alternative hypothesis that not all the dummy variable coefficients are zero. If the critical values of the F-distribution are smaller than the F-statistic values from the hypothesis test, within a 1 or 5 percent level of statistically significance, we can reject the null hypothesis. If otherwise, we must accept the null of no structural break in the data.

Table 13—The results of Chow test for the structural change after 2007 on the tech transfer metrics model at the department level, (1989-2012)

	Chow Test	
	Sum of squared errors of prediction	**Dummy variable of prediction**
F-statistics	8.4657	8.6402
p-value	0.0001	0.0001
Group*	33	33

Table 13 shows Chow test results for the structural changes after 2007 on the combined tech transfer metrics at the department level from 1989 to 2012. The result displays the values of F-statistics are greater than the critical value, so we can reject the null hypothesis at 1 percent level of statistical significance. This is strong evidence of a structural change in the relationship between the combined tech transfer activities and the independent variables measuring research inputs after 2007.

VI. Discussion

A. Knowledge Production Based on the Theory of Classical Production

In classical production theory, the principal goal of a firm, farmer, or other type of producer is to maximize the difference between the costs and the revenues from turning inputs into outputs, and the production function describes the quantitative relationship by which they turn inputs into outputs. The production function of a rational profit maximizing producer necessarily exhibits certain characteristics, such as non-negativity, weak essentiality, non-decreasing in input variables, and concavity. Several key concepts or measures used to describe production decision-making are marginal productivity, the scale of production, and elasticity. The knowledge production function is based on the same basic theories and assumptions. However, because knowledge is an intangible asset, which distinguishes it from the physical outputs of firms and farmers, not all of the typical production function characteristics and assumptions necessarily hold when describing the relationships between research inputs and outputs. Thus, the productivity of inputs with respect to research outputs—especially when measured by proxies such as publications or intellectual properties in both public and private research—varies due to the heterogeneous characteristics of the different types of research outputs and the different research environments across the colleges, departments, and research units of the university.

In section IV, the functions describing the "production" of publications and doctoral degrees exhibit decreasing returns to scale, at 0.9349 and 0.0458, respectively with respect to all inputs, past research expenditures and the value of research equipment. However, the functions describing "production" of the other two research outputs, those measured by the collaboration index and the combined tech transfer metric, exhibit increasing returns to scale, at 2.0239 and 1.4179, respectively (See model 4 for the effective labor PDL in the regression results of each research output model.) In light of these results, publications and doctoral degree awards might be understood to conform more closely to the typical assumptions of the theory of classical production, specifically the law of diminishing marginal productivity, than do the other two types of research outputs. The publication model's weighted average lag, which can be interpreted as the mean lag between research project inception and completion, is two or three years, which is shorter than the lag estimated for the other research outputs. In other words, production of a publication, as the predominate research output in the university, is faster than that of other research outputs, even without considering factors related to productivity.

Furthermore, with regard to lagged research expenditures in the effective labor PDL models, publications and combined tech transfer metrics have a concave production function, but the collaboration index has a convex production function, and doctoral degree award have a cubic production function. Thus, publications and tech transfer metrics have an estimated maximum point of impact among past research expenditures, but the collaboration index, on the other hand, has a minimum point; the doctoral degree award has both local minimum and local maximum points. One interpretations are that publications are the fundamental research output in the university while tech transfer represents are a newer channel of research output that, while it has a lot of potential for the future of research outcomes at the university, it is still in its early stages of implementations. In the long run, publications and tech transfer metrics are less elastic than is the collaboration index with respect to the past research expenditures. A possible interpretation is that, over the long run, conventional collaboration activities, such as collaborative R&D, informal conversations, consulting, co-supervising, and so on, are more sensitive to changes in research expenditures. This interpretation could be derived from the theories of classical production, in that publications and combined tech transfer metrics seem to have the attributes of “necessary goods” within the university.

B. Knowledge Production in a Single Institutional Data versus an Aggregation Data

According to Adams & Griliches (1998), in aggregated data research outputs exhibit constant returns to scale, but, they find, within individual universities, exhibit decreasing returns to scale, which they interpret to be because data errors are more important at the single institution level and because research spillovers exists only at the aggregate level. In our database, however, the regression results partially follow the critique of Adams and Griliches, as the results seemed to ignore the heterogeneous characteristics of the different research outputs, as in the increasing returns to scale for the collaboration index and combined tech transfer metrics.

Nevertheless, both advantages and disadvantages exist for the use of single institutional data. First, certain types of data are only available within the single institutional context, such as department level faculty extension budgets, collaborative contacts with industry, etc., which are rarely reported in aggregation datasets. Second, single institutional data makes it possible to investigate deeply within one institution or region, controlling for a host of exogenous characteristics by holding them relatively constant. These advantages therefore present the opportunity to unravel the university research system's generally unobservable and inherent characteristics, the inner workings of its so-called black box, and their implications for knowledge spillovers.

However, conclusions drawn from single institutional data may compromise findings' generality relative to other institutions or to aggregate data, because each institution has its own idiosyncratic conditioning characteristics, including levels of research expenditures, management skills, administrative policies, and so on. Thus, these factors may inherently bias findings based on single institutional data. In the optimistic view, we still hope to open a meaningful debate using data and empirical results from a single university, CSU, which in future analyses can be compared or combined with other universities' results. This makes it possible to discover the unobservable and heterogeneous research systems of different institutions, and eventually, their aggregation.

VII. Conclusion

This chapter has analyzed research production and technology transfer activities of the academic departments of a large research university (Colorado State University) and introduced a novel empirical technique for estimating knowledge production functions. First, we have utilized four different knowledge production functions as an empirical regression model and adopted a group fixed effect of panel negative binomial MLE with polynomial distributed lag (PDL) or Almon scheme of past research expenditures. The results show that PDL models have more statistically significant estimated slope coefficients, better model specification, and smaller AIC and SBIC values than *ad hoc* or Koyck scheme models, especially of slope coefficients within the mid-range lagged years. It is important to emphasize that the PDL scheme may unravel the unobserved lag effects found in previous studies. With regard to considering important input variables that have rarely appeared in previous studies of knowledge production, both full-time equivalent (FTE) of researchers as a labor variable and value of research equipment as a fixed input variable make it possible to better connect with neo-classical production theory, and these input variables increase the statistical significance of estimators as the best linear unbiased estimators (BLUE). These help overcome model misspecification errors, especially omitted relevant variables, affecting previous studies estimating knowledge production functions.

Further, an effective labor PDL model gives more intuitive results in this paper, having been adapted for investigating returns to scale of knowledge production, output elasticity, and the degree of co-production of outputs in a system of equations (SUR) model. The greatest advantage of this effective labor model is to control for the lag effects of the FTE variable, following the labor-augmenting or Harrod-neutral approach, with knowledge and labor entering multiplicatively. Human knowledge capital is generally unmeasurable in the production function system.

Second, merely from the summary statistics, publications as a knowledge dissemination channel seem to be a more productive and common research output across departments and research units than other research outputs. Indeed, most colleges and departments report the number of published journal articles as their major research output each year. In this regard, our empirical results have shown that in the estimated models publications have a more systematic relationship with research inputs, as well as a shorter lag length with respect to research expenditures. The mean lag between research project inception and the output of publications is two or three years shorter than for the output of other types: lags are, on average, 2.66 years for publications; 5.32 years for doctoral degree awards; 5.61 years for collaboration outputs; and 5.01 years for tech transfer outputs. Thus, the speed of output is fastest even without considering the quality of publications.

Publications are more closely linked with the assumptions of classical production theory, including the law of diminishing marginal productivity, and are less elastic with respect to past research expenditures than are the other research outputs. In contrast, the collaboration and tech transfer outputs exhibit increasing returns to scale with respect to all inputs and are more elastic with respect to past research expenditures than publications. In the models estimating doctoral degree award, impacts are less significant in all models, especially the restricted PDL model. Besides, the FTE and value of equipment variables are insignificant in all of the models. Hence, there does not seem to be a systematic relationship between research inputs and doctoral degrees as an output. Doctoral degree awards may be more related to the education mission of the university, but not seem to be directly related to research outcomes.

Third, we hypothesized co-production relationships among the different research outputs, expecting they would have a positive and significant correlation of residuals within departments or research units. This hypothesis was tested by SUR system of equations among the different types of research outputs. The results suggest that conventional collaboration activities may be substitutional or *ex ante* activities to the other research outputs, especially the tech transfer outputs. Furthermore, by the results of SUR models, estimations of the production of tech transfer outputs and collaboration outputs have inverse patterns of estimated slope coefficients—concave and convex over time, respectively—and the maximum point of impact of past research expenditures on the combined tech transfer metrics exactly matches the minimum point of impact of past research expenditures in the collaboration index.

Proceeding from what has been said above, it should be concluded that knowledge production and productivity in the university should rely on the heterogeneous characteristics of research outputs and research environments across the different departments and research units. Basic research outputs such as publications and doctoral degree awards are more closely connected with classical production theory, and publications and combined tech transfer metrics are the most primary research outputs in the departments and research units of the university. Furthermore, empirical model specifications, which should help to unravel the complex research system in the university, are also the central point of this chapter. Our proposed empirical models of the knowledge production function, a negative binomial MLE with both unrestricted and end-point restricted PDL models, and an effective labor PDL model, improve upon the statistical significance results of the relationship between research outputs and inputs, compared to previous studies. In general, further improvement upon the methods outlined may be fruitful particularly when combining multiple universities' data or utilizing aggregate data.

CHAPTER 2. SCIENTIFIC TEAMS AS QUASI-FIRMS: A NEW FRAMEWORK FOR UNDERSTANDING THE AGENCY OF KNOWLEDGE PRODUCTION WITHIN THE UNIVERSITY CONTEXT

I. Introduction

In the knowledge economy or information society, a knowledge-based view of economic activity is increasingly central to understanding production of new products and processes, for guiding the strategic management of the firm, and even for pioneering new market systems. This new knowledge-based paradigm can also lead to new conceptions of economic structures involving social networks. According to Spender (1996), organizations based on knowledge are best characterized as enduring alliances between "independent knowledge-creating entities". These independent entities—consisting of individuals or teams, together with tangible resources—generally have attributes of a semi-autonomous organic system, but its specific attributes depend on the type of organization. If the organization has a more hierarchical structure and culture, the degree of autonomy of the knowledge-creating entities within the organization may be low and their identity more machine-like. However, if the organization has a more network structure and culture, the independent knowledge creating entities that comprise the organization may have a high degree of autonomy and a self-regulating system. The organization based dynamic and network types of knowledge creation have positive advantages: pooling heterogeneous intangible resources or assets, increasing the speed of technological innovation, reducing externalities, and managing uncertainty or risk, and so on. While this new conceptualization is deeply rooted in the ideas of 1990s of its basis reaches to the years before World War II (Powell, 1990; Spender, 1996; Latour, 1993; Bush, 1945; Schumpeter, 1942).

In recent years, exploration of the inner logic of academic operations in the research university have suggested a firm-like quality of research teams (Etzkowitz, 2003), similar in many respects to the “independent knowledge creating entities” of Spender (1996). This conceptualization of research groups or scientific teams as quasi-firms operating within the context of the research university suggests a new paradigm for study of the organizational formation and evolution of the research university. According to Wuchty et al (2007), the traditional university ethos emphasized the role of individual genius in scientific discovery, but in recent developments, most academic research has shifted from an individual model to a teamwork model.

The study of university research teams is linked with the study of university knowledge production as well as university technology transfer. This approach views research project teams as the primary agents that contribute to new knowledge creation within the university, including potentially commercializable innovations, whether in the form of patented inventions or in the form of collaborative research projects conducted jointly with private sector industry partners. Project teams based in universities can have not only an interdisciplinary or inter-university composition but can also include team members from industry or from private entities. They can also extend across state and national borders. Regarding the formation of academic research teams, their size and composition is an important factor affecting the extent to which the university is fulfilling its outreach mission of economic and social development.

In U.S. research universities, the size of scientific teams increased by 50 percent between 1981 and 1999 (Adams, 2005). In recent times it has become much less common to find scientific articles that have one author, and usually there are several authors on a paper or project. In the case of Colorado State University (CSU), the number of authors per paper increased by 58 percent over the 24-year period between 1989 and 2012, while laboratory-based research papers were more prevalent than non-laboratory-based ones.

The most important reason for the organization of a research team is to generate efficient and synergic knowledge production that reduces the cost and time of knowledge production; in other words, they make it possible to produce a maximum amount or quality of research outputs with respect to the naturally limited inputs to the knowledge creation process, such as research expenditures, research equipment, highly specialized and talented personnel, etc. This impetus has led to experimentation with and the creation of new organizational forms within research universities, and, in this process, academic research teams have come to operate more and more like "quasi-firms," run like small businesses, optimizing their collective behavior albeit without being directly profit making (Etzkowitz, 2003).

The purpose of this chapter is to explore the agency of knowledge production. The chapter adopts the view of scientific research teams as "quasi-firms," arising as independent knowledge-creating entities within the university context. What are the key components for assembling a research project team, and how does understanding research processes as team-based help to reframe our conceptualization of university knowledge production and technology transfer activities? Of particular interest for this chapter is an examination of team assembly mechanisms and the impact of different aspects of team formation on eventual knowledge outputs. Exploring the situation inside research teams, social network analysis can be useful for understanding the structure of team member networks and their dynamics over time. How do existing network structures affect research team formation and performance? We can use ego-centric network analysis to gain insights on an individual principal investigator's (PI) professional network and relationships, while we can use socio-centric network analysis to concentrate on entire groups' patterns and the formations of teams within the network context. We expect that this kind of network analysis will allow us to look inside the more aggregate knowledge production function and structures of the research university. Although social network analysis can be helpful for understanding the inner workings of team assembly mechanisms and their team dynamics, empirical analysis of a sample of research teams, selected from the entire university, is used to explore what factors influence academic research team formation and performance.

The rest of this chapter consists of six sections. Section II reviews the previous studies of social network analysis and university collaborative research. This social network analysis covers not only economic models and topics, but also insights from the field of scientometrics which is concerned with the quantitative features and characteristics of science and scientific research. Section III shares summary statistics from an extensive database on Colorado State University of research team formation at the college and university level. Section IV analyzes the conceptual frameworks of university research teams and its formations. Section V explores two different empirical approaches to understand academic research teams, ego-centric social network analysis and regression analysis of a sample of research teams. Section VI provides empirical analysis of the impacts of outputs by research teams. This section evaluates the quality of research teams by different groups and hypotheses. Finally, section VII summarizes and concludes this chapter.

II. Literature Reviews

A. Social network analysis and research teams

Social networks and social network analysis have emerged and influenced many fields of endeavor within society and the economy. The scope of this is certainly not limited to Internet social media, but also includes co-authorship networks within academic literatures such as sociology, economics, computer science, statistical physics, mathematics, and so on. According to

Jackson (2008), social networks permeate our social and economic lives and play a central role in the transmission of information about everything from job opportunities to the trade of most goods and services. He also points out that social networks affect our well-being, making it critical to understand how social network structures affects behavior and which network structures are likely to emerge within a given part of society. This seeks to follow these crucial insights to analyze the formation and structure of academic research team.

The deeper motivations for exploring research teams emerging from within the structure of social networks arise from the seminal insights of Powell (1990) and Spender (1996). First, according to Powell, there are three distinct forms of economic organization: market, hierarchy, and network. The market formation has a contract normative basis, with a high degree of flexibility, low amount of commitment among the parties, and independent actor preferences or choices. The hierarchy formation has an employment relationship normative basis, with a low degree of flexibility, medium to high amount of commitment among parties, and dependent actor preferences. Finally, network has complementary strengths normative basis, with a medium degree of flexibility, medium to high amount of commitment among parties, and "inter-dependent" actor choices. Therefore, network forms of economic organization entail indefinite and sequential transitions, create trust and indebtedness. While transactions with uncertain outcomes and high resource costs are more likely to take place within hierarchical organized firms, relational contracting and collaboration, relying upon the network connections, can blur the firm's boundaries.

Second, from Spender, the theory of the firm takes on a new form when we compare the resource based view of the firm with a knowledge based view of the firm. The knowledge based view of the firm is more strategically significant, but it is also more complex than the resource based view. Both the knowledge based and the resource based views attempt to deconstruct the "Black Box" of the production function. However, according to the knowledge based view, tangible resources lie outside the firm, while intangible resources such as firm's specific knowledge are what give it a competitive advantage. Again, the organizations based on knowledge are viewed as enduring alliances between independent knowledge-creating entities. Moreover, the knowledge based view of the firm conforms to Latour's (1996) Actor-Network Theory (ANT) considers social entities, firms, governments, and other social institutions as identifiable actors in heterogeneous networks, not only individuals. The overall dynamic complex is called a "quasi-object". According to Spender, the firm, or the university, is a quasi-object with complex heterogeneous internal epistemological processes.

B. Academic research teams operate like quasi-firms

Although these previous studies of social networks focused on explaining the organization of economic transactions and the theory of the firm, they provide a useful approach for explaining the organization of academic research. The team perspective is also deeply linked with the newer mission of the university of outreach or engagement, given the entrepreneurial firm-like qualities of research teams (Etzkowitz, 2003). The formation of academic research teams is much better understood from a knowledge and social network based view rather than a hierarchical view of formations. Across multiple teams, over time, academic researchers create enduring alliances like a semi-autonomous organic system.

One important extension of the team based approach is to understand the collaboration between university and private entities, which is one of the important technology transfer activities in research universities. They significantly contribute to local economic development and commercial innovation through collaborative team based research projects. Private sponsorship of university research consistently averages 5 or 6 percent out of total university research expenditures (NSF, 2014[23]). Furthermore, the extent of collaboration by researchers in research universities in the U.S. or other advanced countries is not only inter-institutional, but can also be international and interdisciplinary.

The increased extent of external university collaboration may drive an increase in the size of research teams as well as more complex research funding arrangements in the university. Therefore, the university can be faced with such a dilemma as who is, in effect, the principal and who is the agent with respect to the management of research teams. Nevertheless, the size of teams —whether measured as the number of co-authors per paper, the number of co-inventors per patent, or the number of co-founders of startup companies —has been increasing gradually. According to Adams et al (2005), counting co-authors on U.S. scientific papers, research team size increased by 50 percent over the 19 year period, 1981-1999. They also found that foreign institutional collaborators on U.S. scientific papers increased significantly. Academic researchers in private universities, in departments whose scientists have earned prestigious awards, and in departments that have larger amounts of federal funding tend to participate in larger teams. In addition, they find that placement of former graduate students is a key determinant of inter-organizational collaborations, especially collaborations with firms and with foreign scientific institutions.

Lee et al (2015) point out that the increasing dominance of team science highlights the importance of understanding the effects of team composition on the creativity of research results. They unpack two facets of creativity in science: novelty and impact. They find that increasing team size has an inverted-U shaped relationship with novelty. The relationship between team size and

[23] Source NSF, Science and Engineering Indicators 2016: Total expenditures of U.S. colleges and universities on R&D in all fields totaled $67.3 billion in 2014. However, source of funds from business or private sector was 3.6 billion in 2014.

novelty is due to the greater field or task variety, consistent with an information processing perspective. Interestingly, they also find that team size has a continually increasing relationship with the likelihood of a high impact paper.

Previous studies of social network based (or embedded) research teams have focused on co-authorship networks as recorded in databases of published journal articles. Within this content, there are three different areas of research: first is a theoretical approach with mathematical models of network formation, dynamics over time, and network design for influencing behavior; second is an empirical and experimental approach using data that observes then patterns of real world networks in order to test conceptual social network theories; third is a methodological approach to the measurement and analysis of networks.

Some studies have adapted only theoretical or methodological approaches to study research teams, but most have utilized both theory and empirical analysis of network data. According to Barabasi et al (2002), dynamic and structural mechanisms govern the evolution and topology of the complex system of the social network of co-authors on scientific papers. They found three complementary approaches. First, empirical measurements make it possible to reveal the topological measures that characterize the network at a given moment, as well as the time evolution of these quantities. The results indicate that the network is scale-free, and that the network evolution is governed by the phenomenon of "preferential attachment", affecting both internal and external links. Second, a simple model captures the network's time evolution. Third, numerical simulations are used to uncover the behavior of characteristics that could not be predicted analytically because of the complexity of the system. In addition, Taramasco et al (2010) quantitatively explore the social and socio-semantic patterns of collaboration of academic teams.

They explain critical features of social networks of knowledge-based collaboration using team-centered approaches, hypergraphs, and *n*-adic interactions.

Guimera et al (2005) investigate the mechanisms by which creative teams self-assemble. They proposed three parameters that determine the structure of collaborative research networks: team size, the fraction of newcomers in each round of new production, and the tendency of incumbents within the network to repeat previous collaborations. Newman (2004) used bibliographic databases in biology, physics and mathematics, respectively, to construct networks in which the nodes are scientists, and two scientists are connected if they have coauthored a paper. He used these networks to answer questions about collaboration patterns, such as the numbers of papers authors write, how many people with whom they write, what the typical distance between scientists is through the network, and how these patterns vary between subjects and over time.

III. Trend of Academic Research Teams within Colorado State University

In terms of the aggregate level of research teams, the overall network of researchers who have collaborated at some point in time is distinguished from individual teams of researchers actively collaborating at this point in time. However, comparing these two perspectives helps to understand the behavior of both an individual team and the ever-evolving groups of research teams. In this section, we want to explore research teams both in the university overall and at the college levels. As shown in Chapter 1, the knowledge production by researchers at Colorado State University (CSU) are measured systematically in range of data, including published journal articles, degree awards (masters and doctoral), research contracts and grants, extension outreach activities, invention disclosures, patent applications and granted patents, and startup companies. In considering descriptions of research teams, publications and patent data have the advantages of data utilization. Moreover, publications are the main research output of university research and have a more systematic relationship with research inputs than other types of research output.[24]

Figure 20 shows the time trends of total number of papers and average number of authors per paper at the university overall level, from 1989 to 2012. Time is on the horizontal axis, and the total number of papers and the average number of authors per paper are on the vertical axes, left scale and right scale, respectively. The average number of authors per paper generally indicates the size of research teams in the university. The figure indicates that research team size has grown in proportion with the total number of papers, but their growth trajectories appear have begun diverging. By 2012, however, the divergence between these two indicators was the greatest: the average number of authors per paper is around five authors and the total number of articles for the year is about 2,000.

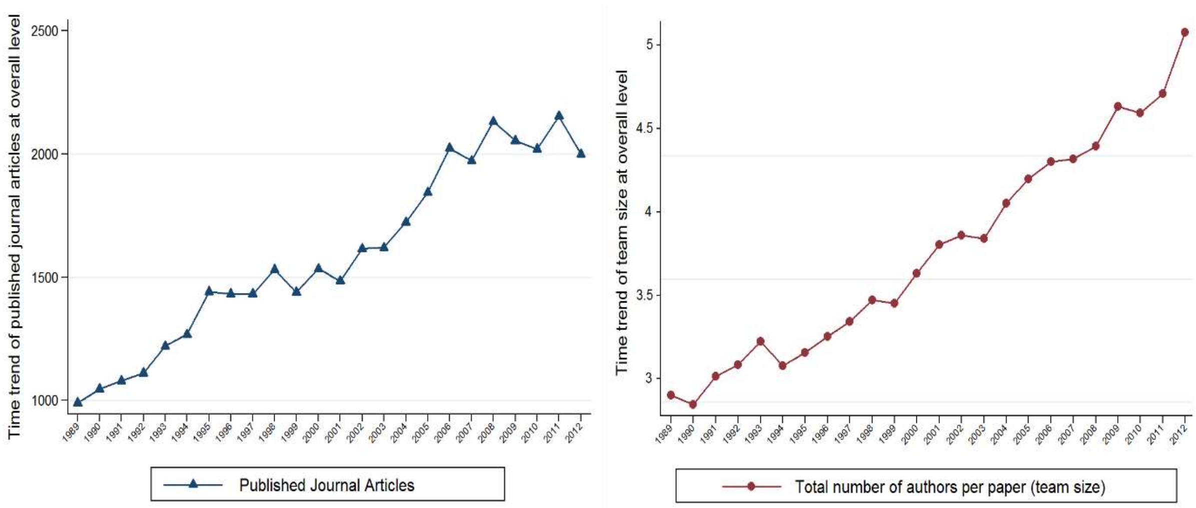


Figure 20—Time trends of journal articles and average co-author team sizes, at the university level, (1989-2012)

[24] It should be noted, however, that we modified the publication data such that if the total number of authors per article is over 30, we removed that article from the dataset. There were 753 journal articles with more than 30 authors out of 38,916 total articles. Most of these many-multi-authored papers were in the field of physics, astronomy, and astrophysics. The maximum number of authors found for a single article is 1,139. Such ‘outliers’ can cause statistical bias problems and distort the generality of the results of our analysis.

Table 14 reports summary statistics for the total number of papers, average number of authors per paper, and total number of citations per paper at the college level, by year, from 1989 to 2012. In terms of the total number of publications, College of Veterinary Medicine, the College of Natural Sciences, and the College of Engineering are the top three colleges at CSU and together have almost 67 percent of all publications at the university, as well as accounting for 73 percent of the total number of citations.

Table 14—Summary statistics of the total number of papers, average number of authors per paper, and total number of citations by college levels, (1989-2012)

	Total number of papers			Average number of authors per paper	Total number of citations		
	Mean	Sum	Share	Mean	Mean	Sum	Share
Agricultural Sciences	141.8	3,403	8.9%	4.1	2,317.3	55,614	7.3%
Business	14.6	351	0.9%	2.6	278.5	6,685	0.9%
Engineering	304.7	7,313	19.2%	3.9	5,209.3	125,023	16.5%
Health & Human Sciences	61.5	1,477	3.9%	3.5	757.8	18,188	2.4%
Liberal Arts	101.4	2,433	6.4%	1.8	444.8	10,674	1.4%
Natural Resources	142.4	3,417	9.0%	3.6	4,256.2	102,149	13.5%
Natural Sciences	399.7	9,592	25.1%	3.5	10,935.2	262,445	34.6%
Veterinary Medicine	352.9	8,470	22.2%	4.6	6,897.6	165,543	21.8%
Others	71.1	1,707	4.5%	3.6	545.2	13,085	1.7%
Total	1,590.1	38,163	100%	3.8	31,641.9	759,406	100%

Figure 21 represents the trend of the total number of articles (left scale) and average number of authors per article (right scale) at the college level, from 1989 to 2012. This figure helps us to begin to understand the relationship between the size of research teams and research productivity. Except for the College of Natural Sciences, growth in the average number of authors per article, which indicates the size of research teams, is roughly as fast as, or faster than, growth in the total number of articles. By contrast, the growth in the size of research teams in the College of Natural Sciences is slower than the growth in the total number of articles. The College of Liberal Arts has the smallest team size among other colleges, on average, 1.8 authors per article, but their article productivity is greater than the College of Business, and the College of Health and Human Sciences.

In the College of Veterinary Medicine and the College of Engineering, there seems to be a positive relationship between the size of research teams and the research productivity, because both indicators are growing at similar rates. However, in the College of Agricultural Sciences, the size of research teams decreased significantly from 2007 to 2008, but during the same period, the total number of articles published did not change significantly.

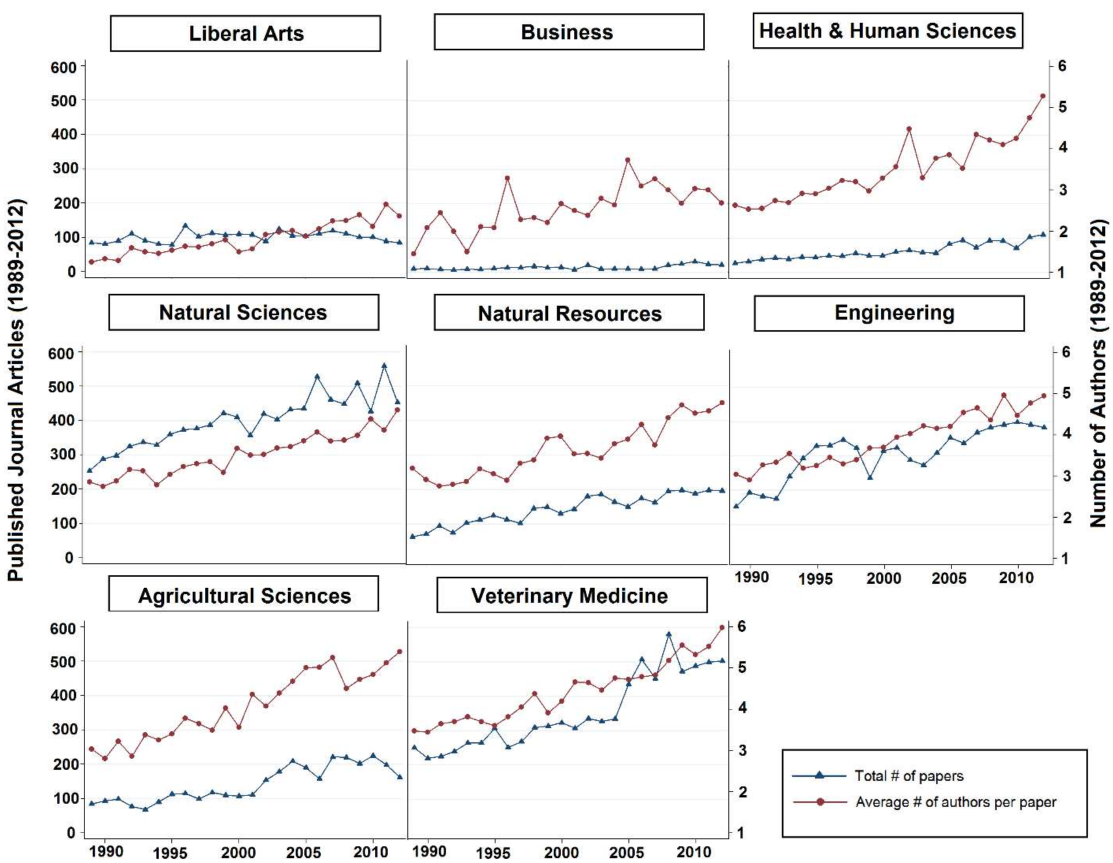


Figure 21—Trend of the total number of articles and average number of authors per article at the college levels, (1989-2012)

From the Figure 21, however, we are unable to confirm a positive relationship between the size of research teams and research productivity, but we are led to think that it may depend on the different research environments and other unobserved factors. Finally, in order to consider the relationship between team size and the quality of research, Figure 22 takes citations per article as a proxy of article quality into account. As indicated as Figure 22, the citations per article seems not to be related to the trend of team sizes across the different colleges.

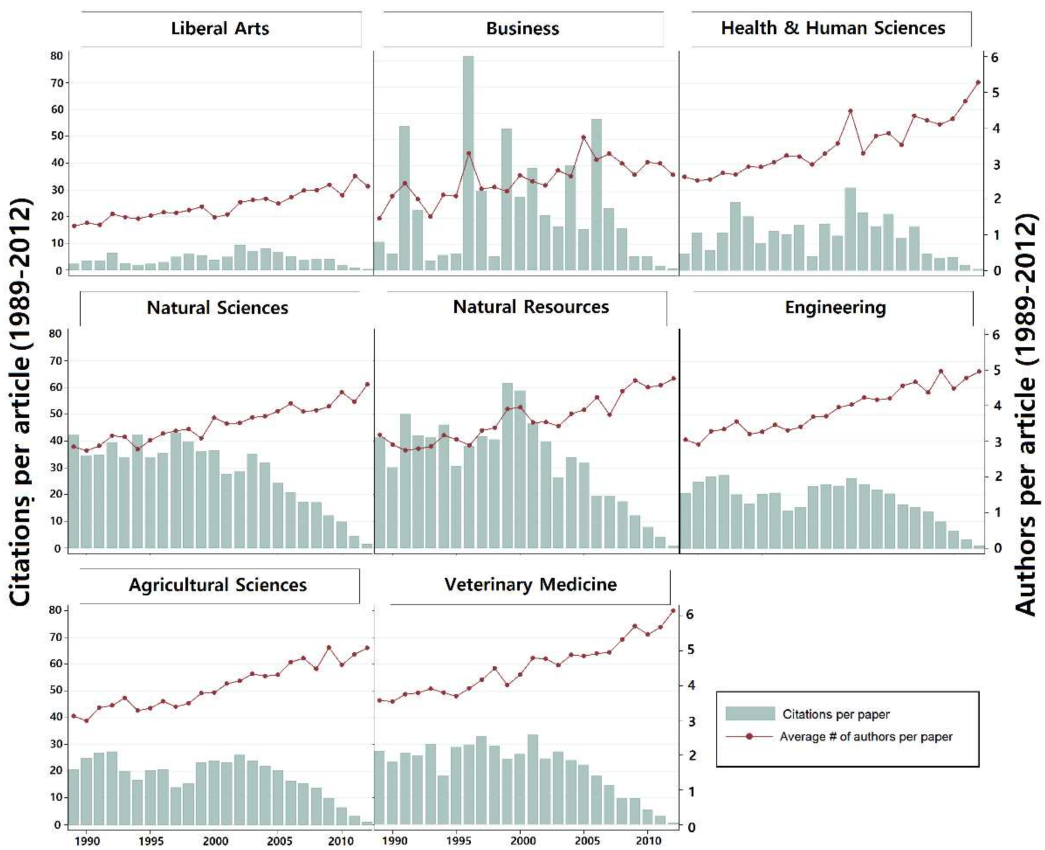


Figure 22—Trend of citations and authors per article at the college level: Citations display a natural right-hand-side truncation of the time series of up to ten years, (1989-2012)

For instant, the team sizes in the College of Natural Sciences are smaller than in the College of Engineering, but the magnitudes of citations are much greater. Similarly, the College of Business has high citation rates per article, but the team size is smaller than in the College of

Health and Human Sciences. Thus, we should consider the nature or characteristics of citations, including the fact that the distribution of citations is highly skewed across colleges and it has evolved over time. Thus, it is more appropriate to treat these in different ways, as discussed in a later section.

IV. Conceptual Frameworks of Academic Research Teams

In modern universities, research activities and collaboration are not only based on “disciplinary” principles but also on “inter-disciplinary” principles. In the same way, the boundaries between research activities are no longer limited to the university but also take in participants from outside the university. Furthermore, most research universities have been increasingly characterized by team-based research rather than to just individual authors or inventors.

Given these kinds of trends within the university environment, the formation of research teams and the evolving structure of collaborative networks is becoming more complicated and difficult to follow. Therefore, analysis of team-based research and team formation can help us to understand the agency of knowledge production within universities, as well as the agency of knowledge production outside of but affected by universities.

As we have seen from previous sections, exploring the formation of academic research teams is bound up with the mission of the university. Etzkowitz (2003) argues that firm-like and even commercially oriented qualities of academic research teams are strongly tied with the outreach mission of university, contributing to economic and social development. Table 15 provides a conceptual framework of academic research teams and their formation, motivated by and adapted from previous studies[25].

[25] Especially, Powell (1990), Parker (1992), Spender (1996), Etzkowits (2003), and Guimera et al (2005).

Table 15—A conceptual framework of different types of research teams

	Types of research teams		
	Functional teams	**Cross-functional teams**	**Research project teams**
Team size	• A few researchers.	• Relatively many researchers.	
Team formation	• Employment based on hierarchical structure.	• Inter-departmental • The network of the past collaborators. • Builds upon the work of a functional team.	• Temporary projects. • Self-directed • Extends or builds upon the work of cross-functional team.
Team purpose	• PI's functional expertise.	• A common goal among the team members.	• Special projects. • A common goal among the team members.
Team flexibility	• High degree of network density.	• Relatively lower degree of density than functional teams.	• Relatively lower degree of density than cross-functional teams.
Number of principals	• A single principal.	• Multiple principals.	• Multiple principals, but more affiliations, including private entities.
Research boundary	• Disciplinary. • Highly reliant on the PI's research area.	• Interdisciplinary. • Generally, bounded in the university.	• Can be varied by the purpose of a project. • Extends the boundary of cross-functional teams, but more flexible.
Funding system	• A single public R&D funding source.	• Multiple public R&D funding sources. • Single private R&D or several smaller private R&D funding sources.	• Relatively higher share of private R&D funding. •
Research outputs	• Journal articles. • Some traditional university-industry collaboration[a].	• Relatively more traditional university-industry collaboration[a], but some newer modes of technology transfer[b] • Journal articles.	• Relative higher in both types. • Journal articles.
Research output revenue	• May earn some revenue on fee-for-service basis for routine testing, conducting trials, etc.	• Relatively low probability of consulting revenues or licensing royalties.	• Relatively higher probability of consulting revenues or licensing royalties. Low probability of revenue from equity in startups.

Note: a) Traditional collaboration is consulting, conferences, collaborative research, co-supervising, and so on.
b) A newer mode of collaboration is patenting, licensing, and startup companies.

University research team formation may follow from and serve the overarching missions of university, but it depends on the heterogeneous characteristics of different research environments, the intrinsic interests of different fields, the extent of social networks, and the intrinsic research abilities of principal investigators, the numbers of potential outside participators, and so on.

As shown as Table 15, we can characterize three different types of research teams within the university: functional teams, cross-functional teams, and self-autonomous teams. First, functional teams are composed of organizational team members assembled from within hierarchical or vertical employment relationships, in formal units such as colleges, departments, and even individual professor's research labs. These functional teams perform employment-like functions, and each team member has relatively low degrees of participation in the creation of knowledge or research outputs, even though they may be subordinates of a principal investigator (PI) or the lead researcher for a particular laboratory. These teams have a relatively small number of team members and their research boundary tends to be limited to their own disciplinary field within the university. Research funding sources for these teams rely on institutional budgets, parts of research grants allocated for more routine research functions, and fee-for-service activities such as testing or conduction trials, etc.

Cross-functional research teams in the university consist of multidisciplinary or interdisciplinary groups of researchers. The cross-functional research teams are generally classified into the two broad types. One is the team of experts from different fields and affiliations. This type of research team can make decisions on their own management and expertise without outside consulting. Another is made up of combinations of different functional teams, a composite team to team organization. It is highly probable that these types of research teams have a high density of team members per research output and are supported by sponsorships of multiple research grants from both public and private entities.

Third, the self-governing research teams are effectively the "quasi-firm" postulated by Etzkowitz (2003) within the context of the research university. According to Parker (1992), "a self-directed team is an intact group … who are responsible for whole work process or segment.

Moreover, the team members work together to improve their operations, and plan and control their work". In other words, this type of research team, whether as a temporary project team or research project team, has come together to execute a specific research project or set of closely related research projects. These teams have a high degree of team flexibility and their boundaries of collaboration can be extended not only to be interdisciplinary and inter-university, but also to include industrial or private entities, as well as cross-national collaborators. The assembly of this type of research project teams depends on a variety of factors, such as the existing professional network of past collaborators or co-authors, which can span an entire discipline globally, incidental meetings at professional conferences, and so on. However, the most important distinction between cross-functional teams and autonomous research project teams are the voluntary or self-assembly mechanisms of forming the research teams (Newman, 2004; Guimera et al, 2005; Zhu et al, 2013).

V. Empirical Approaches to Analyzing Academic Research Teams

In this section, we explore how academic research teams are assembled and what kinds of factors are significantly associated with the team assembly mechanisms by using a highly detailed institutional dataset on CSU research inputs and outputs from 1989 to 2012. We utilize two different empirical approaches, including social network analysis, in particular the analysis of ego-centric networks, and OLS regression models of a selected sample of research teams and their research outputs, including published journal articles and patents.

The main purpose of these two empirical approaches is to explore different aspects of team assembly mechanisms, both from the perspective of an individual principal investigator (PI) and his research teams, which are denoted as "ego-centric research teams", and from the perspective of a random sample of research teams, as distinguished from a departmental or aggregate university-wide level of analysis of research teams.

First, the analysis of ego-centric research teams makes it possible to explore specific components internal to research teams and their assembly mechanism, especially the dynamics of research team formation and performance. However, an analysis of ego-centric research teams concentrates on the individual's network and relationships, as well as their intrinsic research abilities, so it is hard to generalize results to all researchers within the university.

Second, the analysis of a random sample of research teams has the advantages of analyzing a larger and more representative set of research teams and patterns of relationships between team characteristics and their performance, but it is more difficult to detect and follow the dynamics of research team formation for so many. Therefore, these two empirical approaches can be useful for understanding different aspects of academic research team formation and their behavior and research performance.

A. Social network analysis

In order to explore research team formation and assembly mechanisms, we employ ego-centric analysis on one particular individual's research network. Table 16 shows summary statistics of the research activities of an individual faculty member, Dr. Jorge J. Rocca, in the Department of Electrical and Computer Engineering at CSU, denoted "Rocca's team", from 1989 to 2012. We analyze the network of relationships evidenced in three different types of research outputs: (1) the network of co-authors on published journal articles, (2) the network of co-inventors on published patent documents, and (3) the network of co-founders in one startup company. Dr. Rocca has been very prolific, with 315 journal publications, and the average number of authors per article is 7.3 authors, which is quite large comparing with the average of 4.8 authors per article for the Department of Electrical and Computer Engineering, an average of 3.9 authors per article for the College of Engineering, and an average of 3.6 authors per article for CSU overall. The maximum total number of authors per a single article involving Dr. Rocca as a co-author is 31 authors, and the minimum is one, naturally being only Dr. Rocca himself.

1. The structure of ego-centric research teams

The boundaries of the ego-centric network, which is essentially the cumulative aggregation of multiple research teams formed over the years around specific research projects, can be defined by the members' affiliations.

Table 16—Summary statistics of the ego-centric research network of Dr. Jorge J. Rocca, of the Department of Electrical and Computer Engineering at CSU, (1989-2012)

	Research Outputs		
	Articles	Patents	Startup
Total number of each type of research output	315	6	1
Total number of individual as co-authors/co-inventors/co-founders	311	20	3
Mean individuals per output	7.3	4.2	1
Max individuals per output	31	7	1
Min individuals per output	1	1	1
Total number of individuals in the same dept.	154	20	3
Total number of individuals in other CSU depts	17	0	0
Total number of individuals outside of CSU	139	0	0
Total number of grant awards	217		
from public sector sponsors	173		
from private sector sponsors	44		
Total value of grant awards (million $)	51		
The share of value from public sector sponsors (%)	98		
The share of value from private sector sponsors (%)	2		

The number of co-authors in Dr. Rocca's network who are affiliated in the Department of Electrical and Engineering is 154; the number of co-authors who are affiliated with another CSU department, is 17, and finally the number of co-authors who are affiliated outside of CSU is 139. Because of data limitations, we cannot fully identify the affiliations of the outside members, but we can safely assume that many of these outside members likely are former graduate students and research staff (post-docs), who may now be faculty members at other institutions including overseas, as well as employees of private firms, and so on. Still, Rocca's network is highly concentrated on his own department, which is almost 50 percent. The cumulative value of research grants and contracts, from 1989 to 2012, is $51 million, and it is also highly concentrated on public sector (largely federal) sources, accounting for 98 percent.

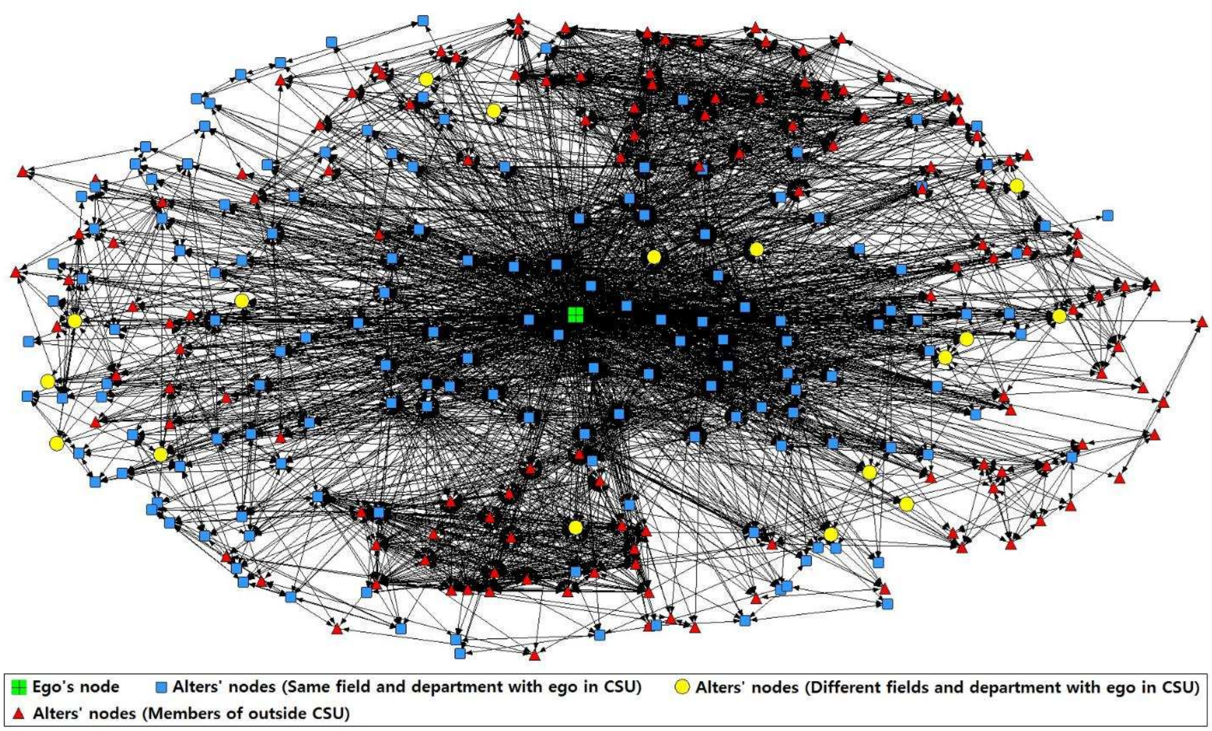


Figure 23—An ego centered co-authorship network, with all types of affiliations, (1989-2012)

The ego-centric network consists of nodes and ties. It describes the network around a single node, the ego, with a number of other nodes. In this case, Dr. Rocca is the ego node and his co-authors, co-inventors, and co-founders are the others, the alter nodes. Figure 23 represents the co-author network that has formed around Dr. Rocca over the time frame studied. The conceptual basis of this graphical analysis is managed in matrix form, which is called an "adjacency matrix[26]". This network diagram illustrates the data found in each row or column in a 311 x 311 matrix. The ego network analysis is processed here by UCINET 6.0[27]. According the Scott (2012), understanding graph theory in terms of an adjacency matrix significantly improves the sophistication of analysis.

In Figure 23, the large green box in the center is the ego's node, the small blue boxes are those alters in the same department with the ego, the yellow circles are the alters in other CSU departments, and the red triangles are the alters outside of CSU, who consist of a variety of affiliations, such as industry researchers, faculty members of other universities, former graduate students and research staff, and so on. This diagram shows a co-authored network of 310 other individuals anchored around the ego. The density, a measure of the overall degree of linkage among the nodes within the network, is 0.067, accounting for 6.7 percent of possible links, which is relatively low density. Network density can be calculated by equation (28)[28].

$$Density = \frac{l}{n(n-1)/2} \tag{28}$$

[26] When appropriating the adjacency matrix, "1" is denoted as the presence of a relation or tie and "0" is denoted absence of a relation with the ego.
[27] A software package for social network analysis. It is open access, https://sites.google.com/site/ucinetsoftware/home , and many scholars of social network analysis have adopted UCINET for research on social network theory (see Jackson, 2008; Scott, 1991; Borgatti et al, 2013).
[28] Social Network Analysis by John Scott (1992).

Where l is the number of links, n is the number of nodes, and $n(n-1)$ is the sum of degree. This measure can vary from 0 to 1. If the density is one, there are relationships between all nodes in the network, and if it is zero, there are no relationships (and thus, effectively, no network). Another measure of network connection is 'distance' which is the shortest path for a node to reach all other nodes, also called 'geodesic distance'. The average geodesic distance measures the average number of steps to reach all co-authors in the ego's network, and in the case of Dr. Rocca's co-author network is 1.933 steps. In the ego-centric network, there tends to be a shorter average geodesic distance than in a randomly generated 'reference' social network, because each co-author by definition already has a co-authorship relationship with the ego.

Now, we can decompose the ego's co-authorship network based upon different affiliations and measure the structural shape of each of the sub-networks, which indicates the extent to which the ego's contacts are connected to each other, including measures such as ego-net density, structural holes, and ego *betweenness* (Borgatti et al, 2012). Table 17 presents summary statistics of these decomposed ego-centric sub-networks by different affiliations. The sub-networks of co-authors in the same department and co-authors in other CSU departments have similar network densities, at 12 percent and 14 percent, respectively, but, the sub-network of co-authors outside CSU, perhaps not surprisingly, has a relatively lower density, at about 9 percent.

Table 17—Summary statistics of decomposed subsets of the ego-centric network by different affiliations

	Density	No. of Ties	Std. Dev	Avg. Degree	No. of authors	Effective size
Same department with ego	0.12	2777	0.32	17.92	154	136.08
Cross departments	0.14	44	0.35	2.44	17	14.56
Outside of CSU	0.09	1794	0.29	12.81	139	126.19
Overall network	0.07	6498	0.25	20.89	310	289.11

However, judging from the standard deviation, the sub-network of co-authors outside of CSU has a relatively smaller variance than the other sub-networks. Likewise, Figure 23 shows that there are two relatively tidy clusters of the sub-network, so it is highly probably that the network of outside co-authors are compactly anchored around each other. Thus, it seems reasonable to assume that the research teams associated with the network of outside co-authors are more likely to have formed a specific research project team.

In addition, in order to measure the structural shape of the ego-centric network, we adopt the concept of a structural hole, which is the lack of a tie between two alters within an ego network (Burt, 1995). The disconnected nodes or alters provide the ego's different points of view. Measuring the effective size is commonly used in a structural hole. The effective size can be measured as ego's degree, the number of co-authors in the network minus the average degree of the ego's alters within the network (Borgatti et al, 2012).

In Table 17, for example, the total number of co-authors represent the ego's degree, so 310 would be an ego's degree in the overall network and the average degree of ego's alters is 20.89. The effective size can be calculated by 310 minus 20.89, so the effective size of the overall network is 289.11. If no alter or node has ties with each other, the effective size is directly the number of authors, 310, and if all nodes are connected with each other, the effective size is one[29].

2. The dynamics of ego-centric research teams

One of the advantages of analyzing ego-centric research teams is the opportunity to explore the formations and sizes of discrete teams over time, as evidenced in each of the articles, patents, and startups associated in teams. Figure 24 displays the time of occurrence of patent applications

[29] From Table 3, the total number of authors is 311, included in Dr. Rocca, the ego. If all alters are had ties with all of the others, the average degree becomes 310, which equals the total number of co-authors without the ego. Thus, the effective size is 1, which can be calculated by 311 minus 310.

and article publications, as well as the timing of the startup company within the ego-centric network. It also indicates the different individual project-based team sizes, in terms of numbers of co-authors per article.

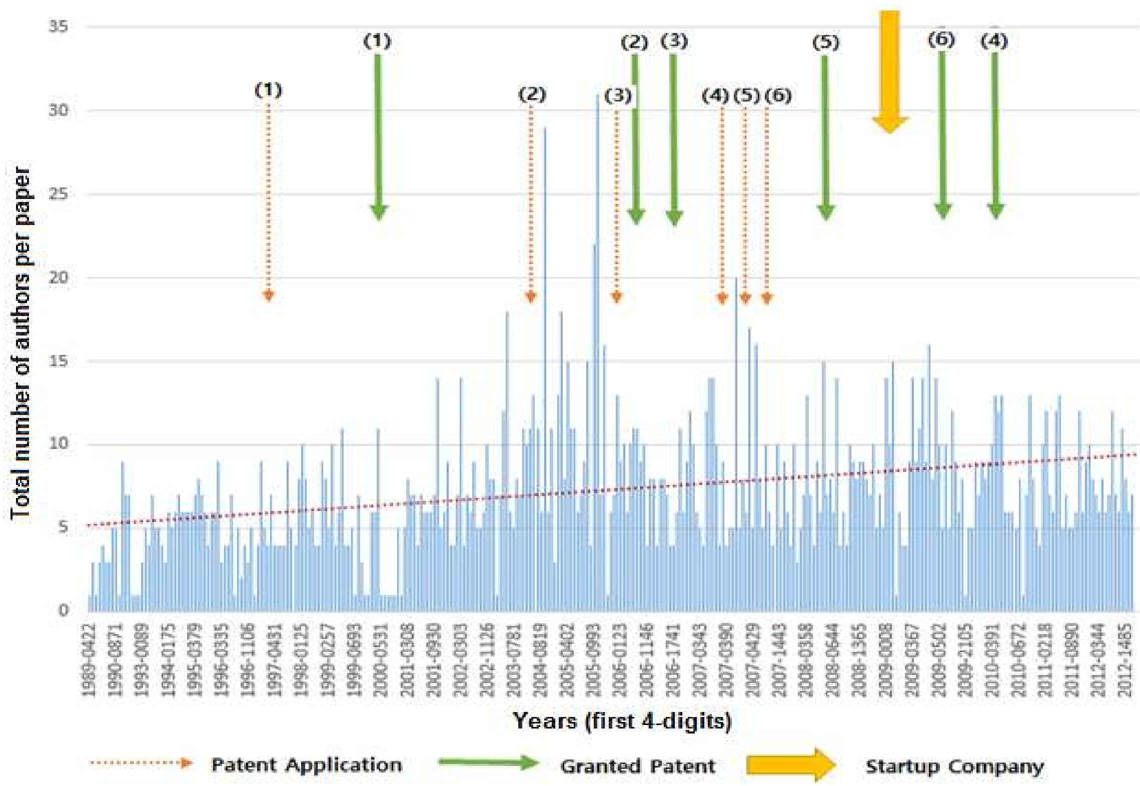


Figure 24—The dynamics of ego-centric research teams, (1989-2012): the time of occurrence of patent application & publication, and startup companies

On the horizontal axis, the first four digits represent years and the next four digits indicate the article ID number. The vertical axis is the number of authors per article which indicates the separate project team sizes. The linear trend-line in the team sizes per article is upward sloping over time. The most vigorous research activities of the project teams were between 2002 and 2007. In these time periods, Rocca's project-based research teams produced more than 100 research articles, accounting for 32 percent out of total article output, as well as submitting 5 patent applications and receiving 2 granted patents. At the same time, his research teams hit the largest team sizes, at 31 authors per article, in 2005. Eventually, in 2009, the ego and his co-founders established a startup company. Figure 25 decomposes the network by year and thereby represents the temporally distinct research teams that formed in the different years, and their relative sizes and densities, from 2002 to 2007. The distance between the ego and other alters is similar across the different years, but the team densities changed over time.

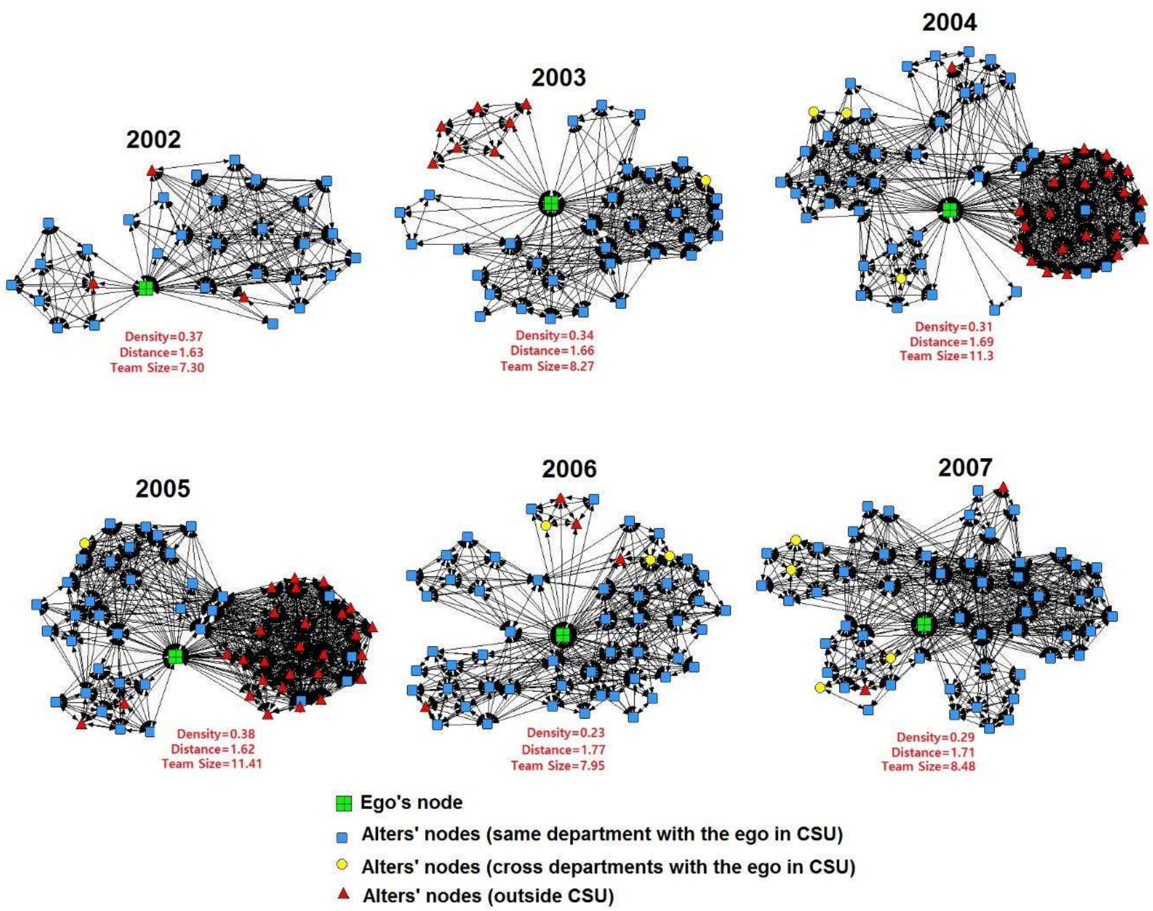


Figure 25—The dynamics of ego-centric research teams, (2002-2007): team formations and densities

It is reasonable from the figures that in 2004 and again in 2005 he participated with a large group of authors outside CSU on an individual publication which affect the size of the ego-centric network in those years significantly, as the pattern of the co-author teams for each of these publications is clustered tightly with high density. Thus, the ego's research boundaries and networks are clarified between those that are more predominantly inside and those that are outside the ego's own university. In addition, as shown in the previous section, the formation of a research team, especially as a project-based autonomous team, typically involves members with multiple affiliations per team (as evidenced by co-authors per article). In most cases we expect that all of the team members have come together to execute a special research project, and the team might include researchers affiliated with commercial or private sector entities.

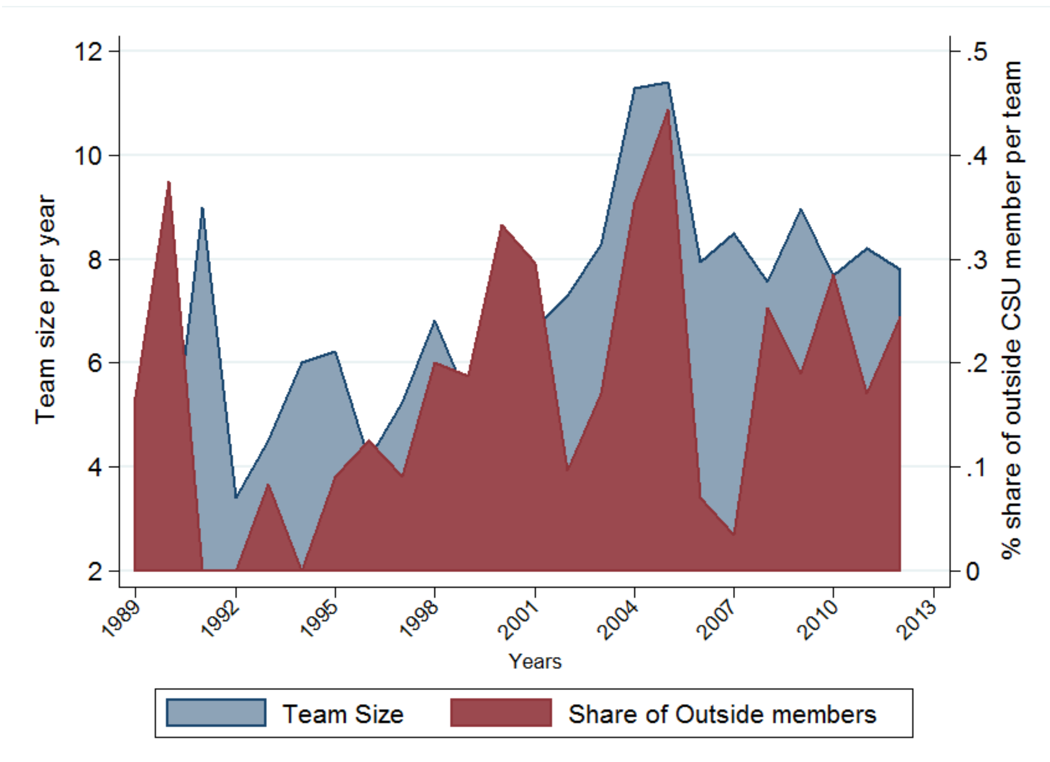


Figure 26—The dynamics of ego-centric research teams, (1989-2012): team size and share of non-CSU authors per article

For example, in Figure 25, in the network for year 2006, a cluster located on top consists of team members with all three types of affiliations, with two co-authors in the same department with the ego, one cross-departmental co-author, and two co-author outside CSU. So this network cluster is a good example of self-autonomous or a temporary research project team. Again, most of the team members in the self-directed team come together and work together based on their own decisions to improve their own research careers, and they jointly plan the operations and control the project.

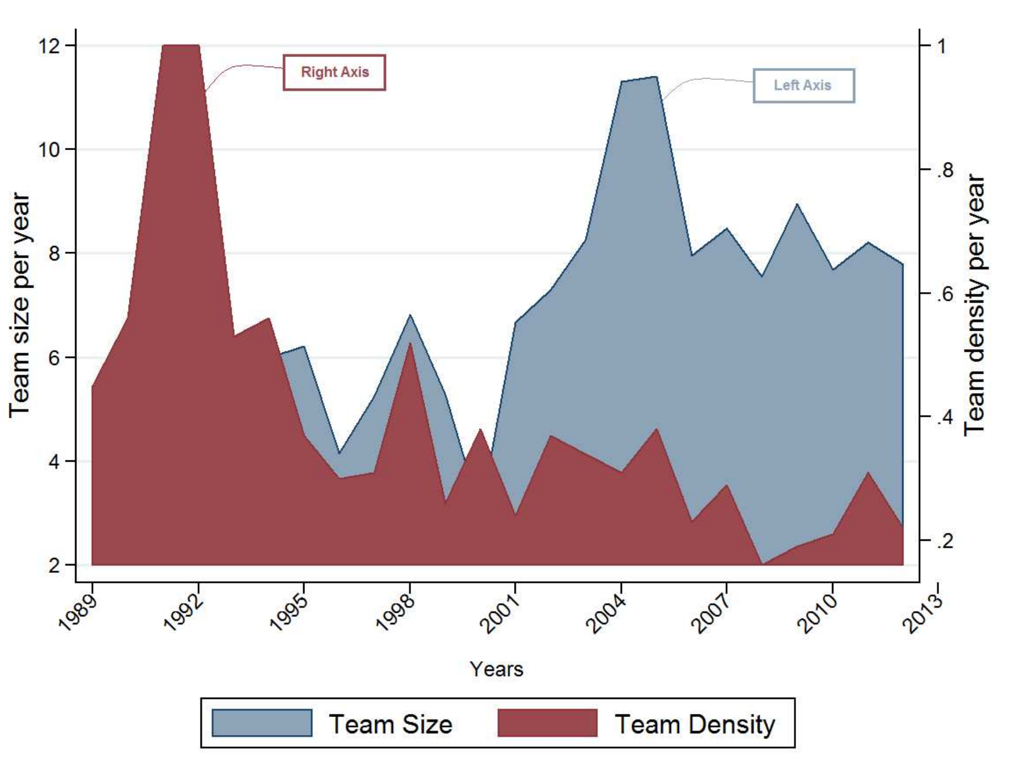


Figure 27—The dynamics of ego-centric research teams, (1989-2012): team size and density

Finally, the analysis of team dynamics not only considers the changes of team formations, but also the change between the size and density of teams over time. Figure 27 represents the dynamics of ego-centric research teams' sizes and densities over time. As indicated in the figure, team sizes tended to increase steadily over time, but conversely, the team density had a downward tendency.

From equation (1), team density is defined as the average strength of ties across all possible nodes, so the higher density rates refer to the higher degree of interconnection within the ego's network. Therefore, the ego's research activities were bounded up with closer relationships among all of the co-authors. Especially between 1991 and 1992, the team density was one, which means that all member knows each other and they have a perfect closed relationship. In other words, the team formation was more likely to be a hierarchy or a single large publication on which everyone participated together. Recently, over 5 years, the average team density is 0.218, in which lower density represents that the ego-centric teams have been characterized by more multiple smaller network-based or research project team formations.

B. Regression analysis

This section will continue to examine the assembly mechanism of research teams. The previous section, using ego-centric social network analysis, showed the components and the dynamic of research team formations within the context of an individual PI's research teams which, over time, built up that PI's network. However, it may be possible that the results are skewed due to field specific characteristics or the PI's inherent research environment, even though the analysis is still meaningful for exploring the formation of academic research teams. In this section, we select a random sample of research teams from the entire population of our database. From this sampling there are two different groups: one is 1,527 research teams drawn from co-authors on published journal articles between 1989 and 2012; the other is 123 research teams drawn from co-inventors on patent documents, from 1990 to 2011.

1. Hypotheses for academic research teams and its formations

The size of research teams plays a significant role in the collaboration activities across different fields and institutions. Relatively small sized teams tend to rely heavily on solitary authors or inventors, whereas relatively large sized teams consist of various affiliations and fields and tend to depend on the synergy effects of teamwork and an effective division of labor (Biffl and Gutjahr, 2001; Guimera et al, 2005; Wuchty et al, 2007). As shown in Table 15, the conceptual framework of research-team formations characterizes both cross-functional and self-autonomous research teams to have a larger number of team members than functional (hierarchical) teams. The team size also refers to the team flexibility and research boundaries. However, bigger is not always better, at least once a certain threshold is exceeded. As teams grow beyond a certain size the coordination costs increase and the marginal value of contributions of additional members with complementary skills decrease.

Therefore, we want to examine how the size of a team can vary and what factors affect the optimal team size. First, we assume that the number of participating research units and departments within the university, and the proportion of members in the team from outside the university, including the number from private entities, are crucial factors when it comes to the size of research teams. Although this assumption seems plausible, it requires further examination in the next section.

[**H1**] The extent to which the team members, who are from private companies and different research units and departments in the university, participate in the research project and the team plays a crucial role in determining the size of research teams and their assembly mechanism.

Furthermore, we want to examine the relationships between the size of the teams and their capability of utilizing R&D investments. From Table 15, it is clear that the funding system is one of the criteria for assessing the different types of team formation and their assembly mechanism, because the functional teams generally utilize the PI's grants or a single public R&D investment for their research budget. However, the larger teams are able to use a wider variety of funding sources and have higher chances of attracting both publically and privately sponsored R&D investment.

[**H2a**] The capability of attracting R&D investment and supporting research expenditures (financial components) is one of the most important criteria for assessing the types of team formations and their assembly mechanism.

[**H2b**] The magnitude and types of R&D investment and research expenditures have a positive relationship with the size of research teams.

[**H3**] The field variety takes on a much greater role in large-sized teams that are cross-functional than in relatively small-sized teams that are hierarchical or functional. Thus, we expect a positive relationship between the degree of field variety and the size of research teams.

In addition, we assume that there is a significant relationship between the field variety and the size of the research teams. However, signs of this relationship are still open to questioning. Although the importance of a multiple disciplinary perspective exists not only in cross-functional teams but also in functional teams, the field diversity and variety are more likely to have a positive relationship with large-sized research teams than small-sized teams.

2. University research teams generating published journal articles

We selected 1,527 sample research teams or journal publications by CSU-affiliated authors, covering 1989 to 2012. These sample teams were selected by several criteria: (1) the team must include at least one industry co-author, (2) in order to maintain consistent results, any teams with more than 30 authors per article were removed from the sample, and (3) the affiliations of all members of the team are identified. Table 18 represents the descriptive statistics of 1,527 sample research teams for the published journal articles from Thomson Reuters' Web of Science via CSU Library and Table 19 is its correlation matrix of all variables, from 1989 to 2012.

Table 18—Summary statistics of research teams generating published journal articles: 1,527 sample research teams from journal publications, (1989-2012)

Variables	Obs.	Mean	Std. Dev	Min	Max
	Summary Statistics				
1. Team size: Total number of authors per article	1527	5.49	3.01	2.00	27.00
2. The proportion of non-CSU affiliations per article (%)	1527	0.67	0.19	0.10	0.95
3. Number of participating private firms per article	1527	1.14	0.42	1.00	4.00
4. Number of CSU departments per article	1527	1.12	0.37	1.00	4.00
5. Number of WOS fields per article	1527	2.00	1.23	1.00	5.00
6. Field (dummy)	1527	0.13	0.34	0.00	1.00
7. Public R&D investment (million $) [a)]	1526	7.45	7.91	0.00	35.35
8. Private R&D investment (million $) [b)]	1526	0.99	1.30	0.00	5.37
9. R&D expenditure (million $) [c)]	1526	9.79	9.13	0.00	37.86
10. Value of R&D equipment stock per article (million $) [d)]	1527	1.24	1.41	0.00	7.15

Note: a) The value of public grant awards of PI or first author's department at each year, b) The value of private sponsorship grant awards of PI or first author's department at each year, c) Total R&D expenditure of PI or first author's department at each year, d) 3 years stock of the value of equipment at each year.

Table 19—Correlation matrix of all variables in Table 18: 1,527 sample research teams from journal publications, (1989-2012)

Variable	1	2	3	4	5	6	7	8	9	10
	Correlation									
1. Team size (count)	1.00									
2. Proportion of non-CSU members (%)	**0.47**	1.00								
3. Participating firms (count)	**0.18**	**0.11**	1.00							
4. CSU departments (count)	**0.06**	**-0.32**	0.00	1.00						
5. WOS fields (count)	**-0.04**	-0.01	0.00	**-0.05**	1.00					
6. Field variety (dummy)	**-0.34**	**-0.22**	**-0.07**	**-0.08**	**0.57**	1.00				
7. Public R&D investment (million $)	**0.21**	**0.23**	0.00	**-0.06**	0.00	**-0.10**	1.00			
8. Private R&D investment (million $)	**0.06**	**0.17**	**0.09**	-0.05	**-0.18**	**-0.10**	0.04	1.00		
9. R&D expenditure (million $)	**0.22**	**0.24**	0.00	**-0.07**	**-0.08**	**-0.14**	**0.92**	**0.19**	1.00	
10. R&D equipment (million $)	**0.18**	**0.24**	**0.06**	**-0.12**	-0.03	**-0.11**	**0.58**	**0.11**	**0.62**	1.00

Note: Bold numbers are the correlation with significance level for each entry at $p < 0.5$.

1) Dependent variables

a. Team size

One of the dependent variables in the assembly mechanism of research teams is the total number of authors per article. From the previous studies, there are various measures of the team size. One method is to handle the team size as the natural logarithm of total number of authors divided by total number of articles per year for the affiliated department or research unit, which represents the mean value of team size. This measurement is generally used in more aggregate levels of data, but not in team levels (Adams et al, 2005). Alternatively, we adopt the team size as the natural logarithm of total number of authors per article, as in Lee et al (2015). As shown in Table 18, the mean of team size is 5.49 authors per article, the minimum is 2 authors per article, and the maximum is 27 authors per article.

b. Relative proportion of non-CSU affiliations

The idea of this dependent variable comes from Adams et al (2005), and equation (29) shows their formulation. The P_i is a share of non-CSU authors per article *i*, and the entire term of equation (2) is called the natural logarithm of the relative proportion of non-CSU authors in team *i*.

$$\ln\left(\frac{p_i}{1-p_i}\right), \ where \quad i = 1, \cdots, 1527 \tag{29}$$

The main reasons for adopting the relative proportion of non-CSU affiliation in the research team as a dependent variable is to explore the structure and components of outside university members, and to examine what kinds of factors affect the participation of outside members. Again, the non-CSU authors generally consisted of research employees at private firms, faculty members from other universities, both of which may be former research staff (post-docs) and graduate students of CSU professors, or *vice versa*.

2) Independent variables

We adopt several different independent variables, which are each related with the different hypotheses. First of all, we assume that field diversity or variety is one of the key factors characterizing the assembly mechanism, and we want to examine the relationship between the field variety and team size. There are two different measures of field variety. One is the number of Web of Science (WoS) disciplinary field categories per article, and another is the dummy variable of field variety. The dummy variable is configured in binary terms. The first step is to calculate the ratio between the total number of WOS field categories and the total number of authors per article. If the ratio is over one, which means that the number of fields is greater than the number of authors, then we denote the article as having high field diversity, and assign a dummy variable value of 1. If the ratio is equal to or less than one, we denote the article as having low or moderate field diversity. Thus, the field dummy variable indicates that if the team has high field diversity, it is one, and if otherwise, it is zero.

Second, we also assume that R&D expenditure or investment is one of the important factors influencing team assembly. So, we use three different financial variables that serve effectively as proxies for the financial resources available to the team, total research expenditures, grant awards, and the value of research equipment of the home department of the lead CSU author. However, as shown as Table 19, these financial variables are highly correlated with each other. Thus, we will not use all of them at the same model. Again, the three financial variables indicate the size and relative research funding of PI or first author's department or organization.

3) Control variables

For testing the assembly mechanism of research teams, we use four control variables, which are in a generally accepted sense of the team components, (1) participating number of private firms per article (which is at least one, given the sample selection criteria), (2) participating number of the CSU departments per article, (3) proportion of non-CSU authors per article, and (4) afield category variable[30]. We want to examine the relationship between the number of private companies

[30] This one is distinguished from field variety dummy.

and the size of university-based research team. This relationship might indicate the degree of team flexibility and multiple principals in the team. By the same token, there also exists the significant relationship between the number of CSU departments and the team size. In the field category dummy variable, the Web of Science's subject research areas are comprised by 255 areas in science, social sciences, and arts & humanities. In our sample research teams' database, we have 174 WoS research categories and the 6 aggregate categories[31].

4) Regression results

Table 20 displays the result of OLS regression of the team assembly mechanism of published journal articles for the 1,527 sample research teams from 1989 to 2012. The regression results consist of 6 different team size models and the dependent variable is the natural logarithm of total number of authors per article, representing the size of the research team that produced the knowledge codified and published in that article. All models have same control variables, which are observed to strongly influence the results of team size. The six field categories of dummy variable are present in all of the regression models.

The logarithm of the number of participating private firms per article is positively related to the size of research team, at the 1 percent level of statistical significance across all six models. Moreover, the logarithm of the participated CSU departments per article also affects the team size, at 1 percent level. Although there exists a positive relationship between these two variables and the size of research teams, it does not directly reflect the total number of authors, but it might be considered to be related to the number of principals to which members of the team, as individual agents, are at least partly responding.

[31] Field 1: biology, chemistry, medical, agriculture, and life sciences (809 papers). Field 2: engineering, computer sciences, mathematics, and physics (422 papers). Field 3: environmental science and Geosciences (206 papers). Field 4: economics and business (13 papers). Field 5: social sciences and behavioral sciences (69 papers). Field 6: others (8 papers).

Thus, the involvement of members from multiple organizations, reflecting the involvement of multiple principals, indirectly affect the size of research teams. Besides, the factor of multiple principals in the research team represents a degree of the team flexibility, and it is more like a cross-functional research team, which is highly connected with the number of principals, such as collaborating between university teams and private firm's teams, and cross department teams.

Table 20—Regression results of team assembly mechanism from journal publications, dependent variable is team size measured by log (number of authors)

Variable	Models (Dependent variable: Team size)					
	[1]	[2]	[3]	[4]	[5]	[6]
Number of WOS research areas	-0.0237 (0.0228)	—	-0.0250 (0.0222)	—	-0.0311 (0.0233)	—
Field variety (dummy)	—	-0.5256*** (0.0279)	—	-0.5100*** (0.0259)	—	-0.5247*** (0.0281)
Value of R&D equipment	0.0380*** (0.0088)	0.0360*** (0.0081)	—	—	—	—
R&D expenditure	—	—	0.0598*** (0.0113)	0.0484*** (0.0103)	—	—
Public R&D investment	—	—	—	—	0.0427*** (0.0102)	0.0404*** (0.0098)
Private R&D investment	—	—	—	—	-0.0210** (0.0084)	-0.0142* (0.0078)
Participating firms	0.2735*** (0.0425)	0.2466*** (0.0408)	0.2841*** (0.0424)	0.2590*** (0.0405)	0.3010*** (0.0424)	0.2773*** (0.0402)
Participating Departments	0.4724*** (0.0499)	0.3734*** (0.0489)	0.4703*** (0.0469)	0.3720*** (0.0460)	0.4674*** (0.0514)	0.3674*** (0.0500)
Proportion of non-CSU authors	0.5565*** (0.0533)	0.4601*** (0.0484)	0.5710*** (0.0515)	0.4863*** (0.0469)	0.5589*** (0.0543)	0.4692*** (0.0499)
Field categories	YES	YES	YES	YES	YES	YES
Root MSE	0.4324	0.4006	0.4295	0.3974	0.4324	0.4005
Adjust R2	0.2287	0.3380	0.2517	0.3593	0.2378	0.3462
F-statistics	48.32	160.24	56.57	171.22	39.15	126.18
p-value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Observation	1,433	1,433	1,520	1,520	1,391	1,391

Note: The parentheses are White's robust standard error. *** 1%, ** 5%, and * 10% levels of statistically significant.

The logarithm of the proportion of non-CSU authors per article has a positive and statistically significant relationship, at the 1 percent level, with the size of research teams. Again, the non-CSU authors comprise outside team members, such as research employees of private companies and faculty members from other universities. The share of non-CSU authors is related with the size of research teams and the magnitude of estimated slope coefficients are higher than other variables. For example, in model 3, *ceteris paribus*, when there is a 1 percent increase in the proportion of non-CSU authors per article is associated with an increase the size of research teams by 0.57 percent. We have two different measures of the degree of field variety in the team. First, the logarithm of the number of WoS research fields per article is insignificant and the estimated slope coefficients are quite small across the model 1, 3 and 5. Furthermore, it has a negative relationship with the size of research teams.

Presumably, it might have some liner relationships with other independent variables, such as a collinearity problem, given that the values of the adjusted R-squared and F-statistics in model 1, 3, and 5, are smaller than in the other models, which do not include in the number of WoS research fields. Alternatively, the field variety dummy variable might be related to the size of research teams, at 1 percent statistically significant and the magnitude of estimated slope coefficients are much higher than the number of WoS field areas. However, both variables have a negative relationship with size of team. These indicate, for example in model 6, *ceteris paribus*, for every 1 percent increase in the field variety, the logarithm of the total number of authors per article will be decreased by 0.52 percent with 99 percent confidence. Thus, we expect that although it is still a negative impact on the size of teams, there still exist meaningful interpretations. First, literally, the field variety is not merely to predict the size of research teams, but it is to evaluate the quality of the research at the team level. Second, the field variety is possibly linked with the PI's subjective research interests or the purpose of team projects.

As shown as Table 19, the correlation matrix, the value of R&D equipment, R&D expenditures, and public & private contributions to grant and contract awards are highly correlated with each other, especially between public grants and contracts and R&D expenditures, and between the value of R&D equipment and R&D expenditures. Thus, in order to prevent a multicollinearity problem, we handle these financial variables within separate models. First, in model 1 and 2, the logarithm of the value of R&D equipment, which is a 3 year sum of investments in new equipment by the PI's or first author's department. This measure of the value of R&D equipment has a positive relationship with the size of research teams, at 1 percent statistical significance. This implies that the magnitudes of physical research capital is positively related to the size of research team. By the results, 1 percent increase in the value of R&D equipment of a department is associated with a larger size of research teams of 0.038 percent in model 1 and by 0.036 percent in model 2.

In model 3 and 4, we measure the overall magnitude of R&D conducted in the PI or first author's department using the logarithm of total R&D expenditures (Lee et al 2015). The results indicate that the overall magnitude of R&D is related to the size of research teams: the expenditures make it possible to hire more team members, such as research staff and GRAs, and to spend investment for the R&D equipment, see Table 18. In model 3, for every 1 percent increase in the R&D expenditures will make to increase the team size by 0.06 percent with 99 percent confidence.

Finally, we want to know how different types of R&D investments affect the size of research teams, so we attempted to run OLS regressions with public sector and private sector R&D investments. In models 5 and 6 of Table 20, the logarithm of public R&D investment has a positive impact on the size of research teams, at 1 percent level of significance, but the logarithm of private investment has a negative relationship with team size, at 5 and 10 percent in models 5 and 6, respectively. Interestingly, the results tell us that a 1 percent increase in the private R&D investment is associated with a smaller size of research team by 0.021 percent in model 5 and 0.014 percent in model 6.

Table 21—OLS regression results of team assembly mechanism (journal publications), dependent variable is relative proportion of non-CSU authors: log(proportion/(1-proportion))

Variable	Models (Dependent variable: Relative non-CSU author contribution)					
	[1]	[2]	[3]	[4]	[5]	[6]
WOS research areas	0.0565 (0.0379)	—	0.0763** (0.0368)	—	0.0572 (0.0384)	—
Field variety (dummy)	—	0.0108 (0.0615)	—	0.0916 (0.0564)	—	0.0470 (0.0624)
Value of R&D equipment	0.0899*** (0.0154)	0.0907*** (0.0155)	—	—	—	—
R&D expenditure	—	—	0.0919*** (0.0199)	0.0906*** (0.0196)	—	—
Public R&D investment	—	—	—	—	0.1292*** (0.0212)	0.1308*** (0.0214)
Private R&D investment	—	—	—	—	0.0508*** (0.0135)	0.0495*** (0.0135)
Team size	1.0090*** (0.0393)	1.0103*** (0.0446)	1.0287*** (0.0384)	1.0531*** (0.0436)	0.9969*** (0.0410)	1.0077*** (0.0468)
Participated firms	-0.0015 (0.0823)	-0.0024 (0.0823)	0.0345 (0.0819)	0.0317 (0.0821)	0.0543 (0.0835)	0.0529 (0.0836)
Field categories	YES	YES	YES	YES	YES	YES
Root MSE	0.7529	0.7535	0.7541	0.7546	0.7389	0.7393
Adjust R2	0.3633	0.3623	0.3683	0.3673	0.3933	0.3925
F-statistics	198.96	191.87	212.78	204.86	165.36	162.61
p-value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Observation	1,433	1,433	1,520	1,520	1,391	1,391

Note: The parentheses are White's robust standard error. *** 1%, ** 5%, and * 10% levels of statistically significant.

On the contrary, for every 1 percent increase in the logarithm of public R&D investment will make to increase the size of research teams by 0.043 percent in model 5 and 0.040 percent in model

6. Publicly funded science tends to be larger scale. Private grants and contracts, conversely, tend to be awarded to departments where research is conducted by smaller and more discrete team efforts.

As a final exercise with the random sample of research teams drawn from journal articles, Table 21 analyzes the proportional contribution of the various factors to the number of non-CSU authors on the research team. From equation (29), the dependent variable is the logarithm of the relative proportion of non-CSU authors per article. For the reason of taking this variable is how the non-CSU authors per article are assembled and what kinds of factor affect the assembly mechanism. The control variables are the logarithm of the total number of authors per article as team size and the logarithm of the number of participating private firms. In addition, the field category variable is included in all models. The structures and components are similar to Table 20, team assembly mechanism models.

Remarkably, the field variety and the number of participated private companies per article are not significant on the assembly mechanism of non-CSU authors, except model 3 with WoS research areas. Moreover, the logarithm of private R&D investment has a positive impact on the assembly non-CSU authors per article. Thus, to compare between Table 19 and 7, private R&D investment only affects the assembly mechanism in terms of non-CSU authors, but is not related to the total number of authors per article.

Team size is also found to be positively related to the relative proportion of non-CSU authors per article, with 1 percent level of statistical significance in all models. *Ceteris Paribus*, for every 1 percent increase in the size of research teams will make to increase the relative proportion of non-CSU authors per article by 1.053 percent in model 4.

3. University research teams producing patented inventions

As shown as in Chapter 1, the knowledge production function varies across the different research units as well as across the different types of knowledge dissemination channels, or research outputs.

Table 22—Summary statistics research teams producing patented inventions: 123 sample research teams (patents), (1989-2012)

Variables	Obs.	Mean	Std. Dev	Min	Max
	Summary Statistics				
1. Inventor teams	123	3.04	1.68	1.00	9.00
2. Proportion of Non-CSU inventors (%)	123	0.18	0.26	0.00	0.83
3. CSU departments (count)	123	1.24	0.56	1.00	3.00
4. Time lag (?)	123	2.21	2.05	0.00	12.00
5. Cited Patents (count)	123	19.63	70.05	0.00	643.00
6. Cited non-patents (count)	123	24.61	69.46	0.00	544.00
7. DWPI family members (count of documents)	122	6.22	9.17	1.00	68.00
8. DWPI family countries (count of countries)	122	66.11	51.36	1.00	125.00
9. Assignees (Count)	123	1.65	1.43	1.00	10.00
10. Count of Participating firms	123	0.32	0.53	0.00	2.00
11. Public R&D grants and contracts to lead department ($ millions)	123	7.12	6.60	0.26	35.35
12. Private R&D grants and contracts to lead department ($ millions)	123	0.84	1.05	0.00	5.37
13. R&D equipment acquisitions by lead department ($ millions)	123	1.57	1.24	0.00	6.10
14. R&D Expenditures by lead department ($ millions)	123	9.38	7.79	0.69	37.86

Note: a) The value of public grant awards of PI or first author's department at each year.
b) The value of private sponsorship grant awards of PI or first author's department at each year.
c) Total R&D expenditure of PI or first author's department at each year.
d) 3 years stock of the value of equipment at each year.

Following that line of reasoning, we assume that there may be significant disparity between the assembly of teams that tend to produce patented inventions and the assembly of research teams that only produce published journal articles. Table 22 provides summery statistics for a sample of research teams that created the inventions for the 123 patents assigned to CSU, from 1989 to 2012.

We have selected data on several variables, which might be related with the team assembly mechanism. Table 23 displays the correlation matrix of these variables. As we have seen in the previous section, the financial variables, such as R&D expenditures, the value of R&D equipment, and public and private R&D investments, are structurally analogous to each other, especially between public R&D investment and R&D expenditures where there is almost 90 percent correlation. Thus, we run these variables in different models.

Table 23—Correlation matrix of all variables: 123 research teams (from CSU patents), (1989-2012)

	Correlation													
Variables	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**
1. Inventors	1.00													
2. Non-CSU inventors	**0.59**	1.00												
3. CSU departments	**0.42**	**0.24**	1.00											
4. Time lag	-0.15	-0.05	0.03	1.00										
5. Cited Patents	-0.01	0.00	0.07	-0.01	1.00									
6. Cited non-patents	-0.01	-0.08	0.06	0.08	**0.85**	1.00								
7. Family members	0.01	0.03	**0.39**	**0.18**	**0.19**	**0.15**	1.00							
8. Family countries	0.10	0.12	0.08	**-0.25**	0.02	-0.01	**0.19**	1.00						
9. Assignees	**0.37**	**0.40**	**0.17**	**-0.16**	-0.02	-0.05	-0.02	**0.28**	1.00					
10. Participated firms	0.03	-0.06	0.08	0.08	-0.08	-0.08	0.02	-0.05	-0.07	1.00				
11. Public R&D	-0.06	-0.03	-0.16	**-0.19**	0.07	-0.08	-0.09	0.07	0.05	-0.09	1.00			
12. Private R&D	0.10	0.14	**0.22**	**-0.18**	-0.10	-0.11	-0.12	**0.24**	0.12	0.01	0.05	1.00		
13. R&D equipment	0.03	-0.10	-0.13	**-0.31**	-0.12	**-0.19**	**-0.18**	**0.26**	**0.19**	-0.03	**0.51**	0.12	1.00	
14. R&D Expenditures	-0.07	-0.02	-0.16	-0.14	-0.04	-0.14	-0.12	0.09	0.06	-0.11	**0.90**	0.13	**0.42**	1.00

Note: Bold numbers are the correlation with significance level for each entry at $p < 0.5$.

1) Dependent variable

The dependent variable in the assembly mechanism of patent research teams is the total number of inventors per patent, similar to our previous analysis of research teams on published journal articles. By comparison with the article team assembly mechanism, the mean value of patent team size is relatively small: it is 3.04 inventors per patent, but the mean value of article team size is

5.49 authors per article. Similarly, the maximum value of patent team size is 9 inventors per patent, but in the maximum authors per article is 27. (See Tables 17 and 21.)

2) Independent and control variables

Most of independent variables are identical with those used in the previous section to analyze formation of research teams on published journal articles, including public and private R&D grants and contracts made to the lead department, the value of R&D equipment acquisitions by the lead department, and R&D expenditures by the lead department, and the new variables here are the count of Derwent World Patents Index (DWPI) family members and countries[32] per patent. The control variables consist of (1) the number of participating CSU departments per patent, (2) the proportion of non-CSU inventors per patent, (3) the number of assignees[33] per patent, and (4) the number of private firms per patent.

3) Regression results

As shown as Table 24, we run six different models, analogous to the models in the previous section. The logarithm of the number of participating CSU departments is positively related to the size of co-inventor research teams on patents across all models with 1 percent level of significance. Similarly, the proportion of non-CSU co-inventors also has a positive relationship with the team size at 1 percent statistical significance. Interestingly, according to the regression results in both article and patent team assembly mechanism, the variable of participating CSU departments seems to be a significant factor for university research team formations.

However, the case of multiple departments in the CSU into one team is one of the different types of the cross-functional team formation: again, as shown as previous section IV, there are

---

[32] The DWPI family members and countries are a set of either patent applications or granted patents taken in multiple countries to protect a single invention by a common inventor

[33] The inventor(s) can make the assignment of the rights granted under the patent, as an assignor, to a third-person party or business, which is the assignee.

several different types of cross-functional team formation, such as combinations of different functional teams in the university, a composite team to team organization, (cross universities or university and private sector), a team supported by sponsorships of multiple research grants, and so on.

Table 24—Regression results of team assembly mechanism (patents), dependent variable is team size: log(inventors)

	Models (Dependent variable: Team size)					
Variable	[1]	[2]	[3]	[4]	[5]	[6]
DWPI family members	-0.0162 (0.0491)	—	-0.0190 (0.0453)	—	-0.0222 (0.0451)	—
DWPI family countries	—	0.0082 (0.0237)	—	0.0085 (0.0231)	—	0.0047 (0.0234)
Private R&D investment	0.0659 (0.0449)	0.0653 (0.0449)	—	—	—	—
Public R&D investment	-0.0117 (0.0352)	-0.0105 (0.0350)	—	—	—	—
R&D expenditure	—	—	0.0583 (0.0496)	0.0573 (0.0488)	—	—
Value of R&D equipment	—	—	—	—	0.0595* (0.0352)	0.0580 (0.0362)
CSU Departments	0.6244*** (0.1186)	0.6173*** (0.1165)	0.6277*** (0.1064)	0.6155*** (0.1038)	0.6388*** (0.1065)	0.6258*** (0.1022)
Proportion of non-CSU inventors	0.6375*** (0.1619)	0.6370*** (0.1619)	0.6142*** (0.1499)	0.6160*** (0.1496)	0.6355*** (0.1549)	0.6356*** (0.1542)
Count of assignees	0.1635* (0.0845)	0.1492* (0.0890)	0.1626** (0.0814)	0.1486* (0.0858)	0.1629* (0.0822)	0.1529* (0.0853)
Participated firms	0.0739 (0.0870)	0.0723 (0.0857)	0.0624 (0.0870)	0.0605 (0.0859)	0.0727 (0.0844)	0.0706 (0.0836)
Field categories	YES	YES	YES	YES	YES	YES
Root MSE	0.4650	0.4649	0.4635	0.4635	0.4612	0.4615
Adjust R2	0.2615	26.1800	0.2716	0.2717	0.2682	0.2672
F-statistics	8.0900	8.1500	10.5900	10.6900	11.0400	11.3100
p-value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Observation	117	117	122	122	120	120

Note: The parentheses are White's robust standard error. *** 1%, ** 5%, and * 10% levels of statistically significant.

Although, we need to consider more types of the cross-functional team formation, this result seems to show one possibility of the importance of cross-functional teams for the size of research team. However, it does not mean that the variable is also an important factor for research impact, such as measured by citations in the next section. The logarithm of the count of assignees per patent affects the positive influence to the size of research teams of patents. For example, every 1 percent increase in the number of assignees per patent corresponds to an increase the size of research teams by 0.163 percent, at 5 percent level of statistical significance in model 3: the other models have a positive correlation, but 10 percent level of significance. However, the variable of the number of participating private companies per patent is insignificant across all models. From the point of view of the heterogeneous research environments across the different knowledge outputs, the major purpose of patenting the inventors or teams' idea and invention is to protect their knowledge. In other words, it is deeply aligned with the intellectual property rights (IPRs). Thus, there is a contrast between article and patent team assembly mechanisms.

Another possibility of the results is the different perspectives between university and business inventors. Basically, the university is a non-profit organization and most often pursues a public domain mechanism even though the research team formations in modern universities are like small businesses (Etzkowitz, 2003). However, the private company is a profit maximizer and more often pursues a IP based mechanism, so if the purpose of university patenting is for protecting knowledge, the degree of participation of private firms is quiet low. As shown as Table 22, the mean value of participating private firms is 0.32 per patent.

In counterpoint to the article team assembly mechanism in Table 20, most of the financial variables are insignificant across all models in Table 24, except the value of R&D equipment in model 5. It is quite probable that financial factors might not be important for the assembly mechanism of research teams of the patents, because patents are imperfect indicator of the number of new inventions and are applied for at an intermediate stage in the process of transforming research input into benefits from knowledge output (Pakes and Griliches, 1989).

Finally, DWPI patent family in both family members and countries per patent is insignificant and the family members per patent has a negative impact on the size of research teams, but not the count of family countries. Again, the DWPI family members and countries are a set of either patent application or publications taken in multiple countries to protect a single invention, so it seems not to seriously affect the size of research, but we expect that it affects to the research impact like the number of citation per patent as the quality of research, which is further examined in the next section.

VI. Impacts of Team Research

The impacts of the team based organization of research influences not only the quality of research, but also plays a role in the economic and social benefits with respect to knowledge spillovers generated. Moreover, engagement in a team setting is beneficial for the researchers, in terms of developing new skills, improving their methodologies, and extending their professional networks. In general, the number of citations per article or patent is one of the measurements or assessments of research quality and impact[34].

However, the impact of research outputs, such as citation rates of published journal articles and patents, tends to have highly skew distributions in the most academic fields and technologies, (Scherer and Harhoff, 2000). Moreover, when collecting data on journal articles via Web of Science or other sources, there exists a time lag problem in citation counts. Therefore, regression analysis is not likely the best measurement of testing the impact of research with the number of

[34] In some cases, journal impact factors and top 1 %, high impact of paper are to measure the quality of research.

citation counts directly. In addition, this chapter will examine the difference between team based research impact and departmental or aggregate level of research impact.

A. Impact of published journal articles as research output

In order to measure the impact of research on the published journal articles, we attempt to classify binary or categorical groups. Table 25 provides summary statistics of times cited of published journal articles for different groups: 1) single private firm per article versus multiple private firms per article, 2) single CSU department per article versus multiple CSU departments per article, 3) three different team sizes, such as large sized teams (more than 9 authors per article), mid-sized teams (between 4 and 9 authors per article), and small sized teams (less than 4 authors per article), and 4) four different disciplinary field groupings.

Table 25—Summary Statistics of times cited of published journal articles by different groups, (1989-2012)

	Summary Statistics (Times cited)				
Variables (different groups)	**Obs**	**Mean**	**Std. Dev.**	**Min**	**Max**
Single participated firm	1352	17.41	63.30	0.00	1507.00
Multiple participated firms	175	12.13	18.72	0.00	95.00
Single CSU department	1366	16.22	61.24	0.00	1507.00
Multiple CSU departments	161	21.78	47.10	0.00	499.00
Large sized teams	139	25.78	65.07	0.00	499.00
Mid-sized teams	993	17.23	67.49	0.00	1507.00
Small sized teams	395	12.57	30.08	0.00	381.00
Field dummy for Biology/medical/life sciences	809	18.71	75.70	0.00	1507.00
Field dummy for Engineering/Computer/Math/Physics	422	16.88	40.24	0.00	499.00
Field dummy for Environmental/Geosciences	206	12.01	22.04	0.00	215.00
Field dummy for Social Sciences & Others	90	10.28	24.89	0.00	163.00

1. The impact of participated private companies per article

Figure 28 represents the cumulative density function (CDF) of citations received by published journal articles in binary groups of those with and without participating firms, and the two distributions seem to have quite similar patterns of cumulative density.

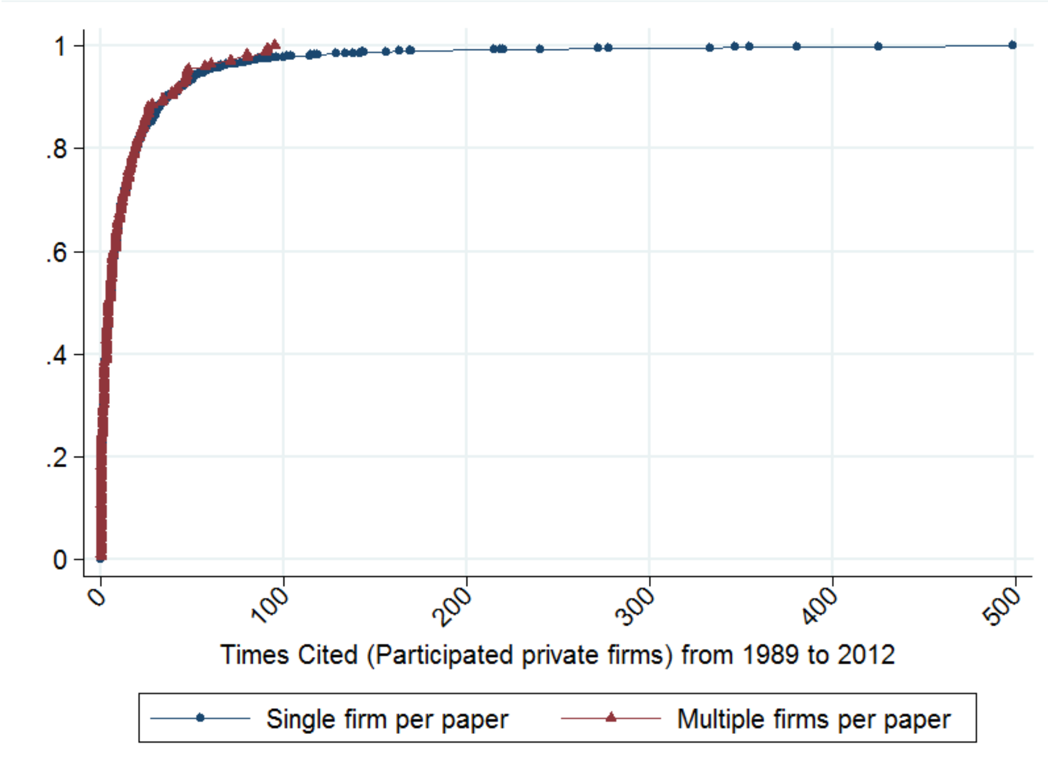


Figure 28—Cumulative Density Function (CDF) of times cited at participating private firms: 1527 sample research teams (journal publication), (1989 to 2012)

Thus, we cannot say which group has higher impact. The group with a single firm has relatively higher frequency than the group with multiple firms. Moreover, in Table 25, the mean value of times cited in the group with a single firm per article is 17.41 which is higher than the mean for the group with multiple private firms per article group, at 12.13 per article. Although the non-parametric test results are insignificant, the group of articles associated with a single firm per article appears to have higher research impact than the group of articles associated with multiple firms per article.

2. The impact of the participation of CSU departments per article

Figure 29, the group of multiple CSU departments per article has a higher research impact than single CSU department. Table 25, the mean value of the multiple departments is 21.78 times cited per article, but the mean value of single department is 16.22 per article. Therefore, the group of multiple CSU departments per article has a higher research impact than the group of single department per article.

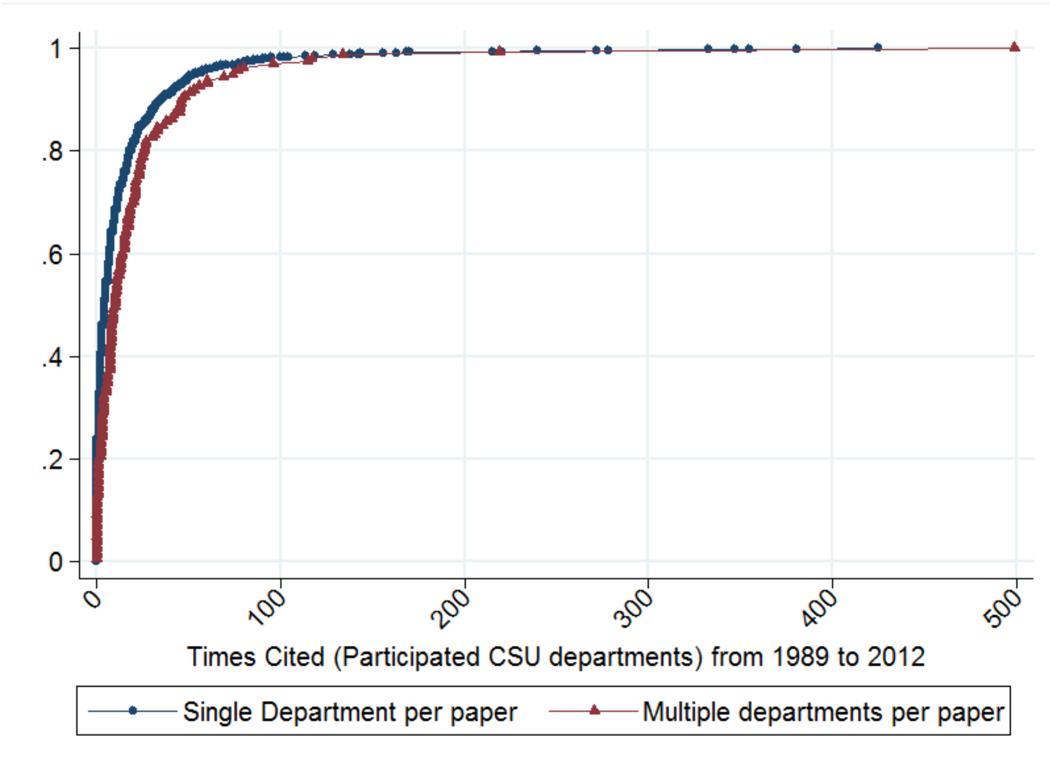


Figure 29—Cumulative Density Function (CDF) of times cited at participated CSU departments: 1,527 sample research teams (journal publication), (1989-2012)

3. The impact of team size per article

The role of research team size is still controversial in the study of team assembly mechanisms and its research impact. However, most previous studies have shown that the size of teams has a positive effect on the quality of research and assembly mechanisms, (Guimera et al, 2005; Adams et al, 2005; Wuchty et al, 2007; Zhu et al, 2013; Uzzi et al, 2013; Lee et al, 2015). According to Guimera et al (2005), the main purpose of team assembly mechanisms is for incorporating individual members with diverse ideas, skills and resources. Although the size of research teams is a major carrier of the assembly mechanisms in the previous section, it did not tell us the impact of team sizes. Thus, we want to evaluate the quality of team based research and how the different team sizes affect the number of citations per article.

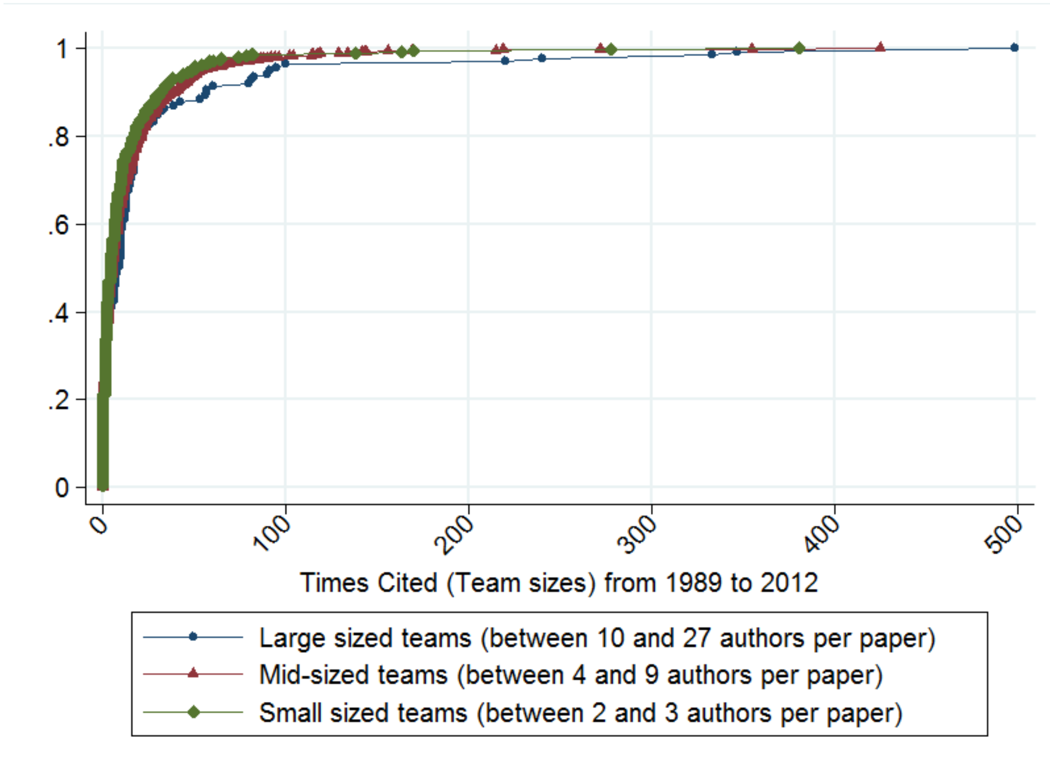


Figure 30—Cumulative Density Function (CDF) of times cited based on team sizes: 1,527 sample research teams (journal publication), (1989-2012)

Table 25 shows that the mean value of the large sized research teams' citation counts is 25.78 per article, which is a relatively higher impact than the mean values of mid-sized and small sized teams, 17.23 and 12.57, respectively. Figure 30 displays the CDFs of different team sizes. The large sized research teams have the highest research impact on the number of citations per article and conversely the small-sized research teams have the lowest impact. According to the results of Lee et al (2015), the team size has a continually increasing relation with the likelihood of a high-impact article, which is analogues in our results of team sizes.

In addition, as shown as Figure 22 in Section III, the number of citations per article seems not to relate the team size. However, there is a significant disparity between the Figure 22 and 12, because it is a different measurement of the team size. In aggregate or department levels, the team size can be calculated by the average number of authors per article[35], but in team levels, it can be measured by the total number of authors per article.

4. The impact of field categories per article

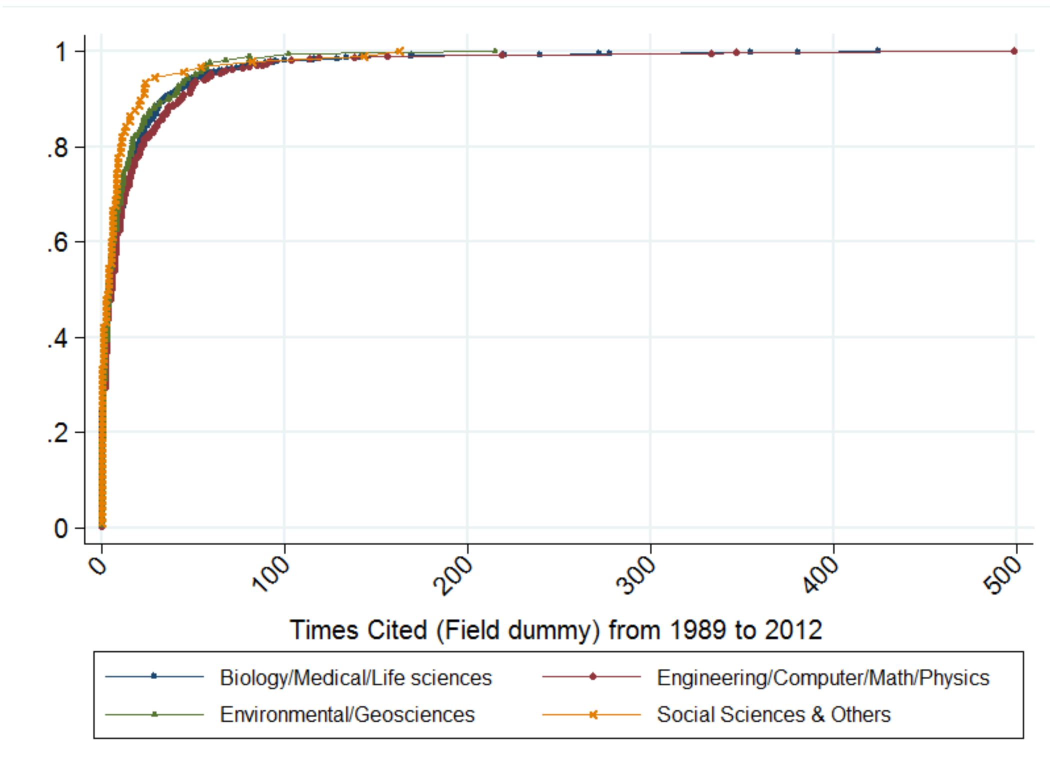


Figure 31—Cumulative Density Function (CDF) of times cited for four different field groupings: 1,527 sample research teams (journal publication), (1989-2012)

[35] The total number of authors are divided by the total number of papers at *t* time and *i* department or college.

In the regression analysis in the previous section, the field category variable was absorbed in the OLS models in order to maintain model stability[36], so we could not provide field specific differences. However, in this section, non-parametric analysis allows us to consider field effects. In Figure 31, the CDF of journal article citations the group of engineering, computer science, mathematics and physics have a greater probability than other fields of high citation counts per article in the range exceeding 100 times cited.

However, from Table 25, the mean value of the group of biology, medical, and life sciences, which includes the agriculture and veterinary sciences, is 18.71 per article, exceeding the mean value of the group of engineering, computer science, mathematics and physics at 16.88 per article. The magnitudes of research impact as measured by citation counts do vary across the different fields or disciplines.

B. Research impact of the patents

Given previous results and interpretations, it appears that the different research outputs of published journal articles and patents have different research impacts, likely associated with the heterogeneous research environments across the university. Table 26 provides summary statistics of citing patents by different groups: 1) corporate co-assignment of patents versus patents assigned only to the university, 2) single CSU department versus multiple CSU departments per patent, 3) single assignee versus multiple assignees per patent, 4) single cited reference: patents versus multiple cited references: patents ,5) single DWPI family versus multiple DWPI families, included in family members and countries, 6) different team sizes, and 7) time lags between patent application and publication.

---

[36] The absorbing of one categorical factor is designed for datasets with many groups. However, in some cases, field dummy variables rather than a category variable make it possible to harm the degrees of freedom in each model.

Table 26—Summary Statistics of citing patents by different groups, (1989-2012)

Variables (different groups)	Summary Statistics (citing patents) Obs	Mean	Std. Dev.	Min	Max
No corporate co-assignment per patent	88	6.966	12.871	0	71
One or more corporate co-assignment per patent	35	4.171	7.106	0	34
Single CSU department per patent	102	4.990	8.809	0	46
Multiple CSU departments per patent	21	11.905	19.565	0	71
Single cited reference: patents	28	2.893	9.678	0	51
Multiple cited references: patents	95	7.137	11.943	0	71
Single DWPI family member per patent	18	5.555	10.579	0	45
Multiple DWPI family members per patent	105	6.276	11.773	0	71
Single DWPI family country per patent	38	4.131	7.669	0	45
Multiple DWPI family countries per patent	85	7.082	12.873	0	71
Large sized teams (more than 5 inventors)	14	11.857	15.913	0	51
Mid-sized teams (between 5 and 3 inventors)	52	6.096	11.749	0	71
Small sized teams (two or single inventor)	57	4.842	9.856	0	46
Time lag btw application and publication (< 1 year)	57	4.526	9.657	0	51
Time lag btw application and publication (2 or 3 years)	42	9.190	14.797	0	71
Time lag btw application and publication (> 3 years)	24	4.792	8.193	0	34

1. The impact of the corporate co-assignments per patent

The impact of collaborations between university and industry is complex and controversial, especially privately sponsored research. In the CSU database, the total number of industry sponsored grant awards and contracts from 1989 to 2012 is 4,387, and the cumulative value of private awards over this period is $268.7 million. In 2012, the value was $21.84 million. The average annual growth rate over 24 years is 12.1 percent. According to Wright et al (2014), corporate-sponsored research appears valuable for university innovation with private sponsored university inventions appearing to be more accessible and useful, having higher patent citation rates than inventions sponsored only by grants from the government or non-profit organizations.

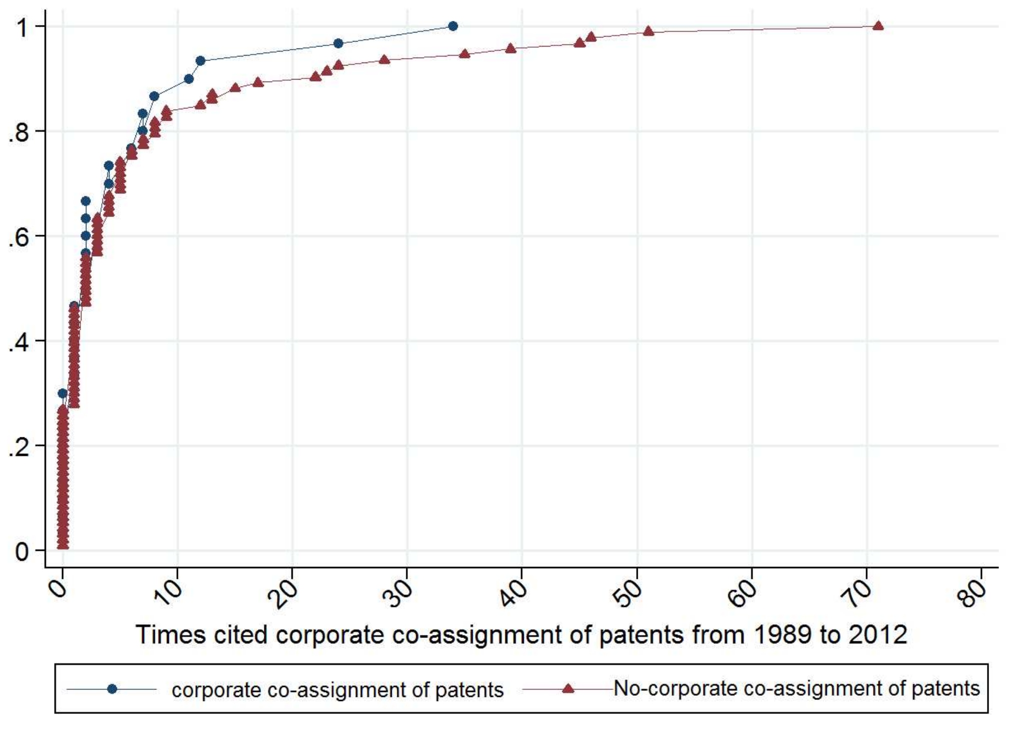


Figure 32—Cumulative Density Function (CDF) of citations to CSU patents with corporate co-assignment of patents, (1989-2012)

However, our empirical evidence appears contrary to the results in Wright et al (2014). Table 26 shows that the mean value of citing patents to the group of CSU patents with participating private firms is 4.171, but the mean value of citing patents to the other group of patents, with no firm participation, is 6.966 per patent. In the CSU data, thus, it appears that patented outputs from privately sponsored or industry collaborating research teams has less impact than the patented outputs from publicly sponsored research teams. In Figure 32, the CDF of citing patents indicates that the impact of patents without private firm sponsorship is higher than those with private sponsorship.

Although there are numerous reasons for the different results regarding the role of corporations in university research and invention, it seems reasonable to consider initially the size of corporate R&D investments as the main factor explaining the differences. This factor is intimately related with the academic rankings and reputations of universities. For example, Wright et al (2014) report that the oil company British Petroleum (BP) announced in 2007 that the company invested $500 million to fund a decade of alternative energy research at the University of California, Berkeley. The size of just that one (major) private R&D funding award was greater than the total sum of private R&D investment over 24 years at CSU. In the CSU data, 46 percent of privately sponsored patents were from smaller local Colorado companies, and only 6 percent were sponsored by relatively large companies like Chevron Corporation.

2. The impact of the participation of researchers from multiple CSU departments

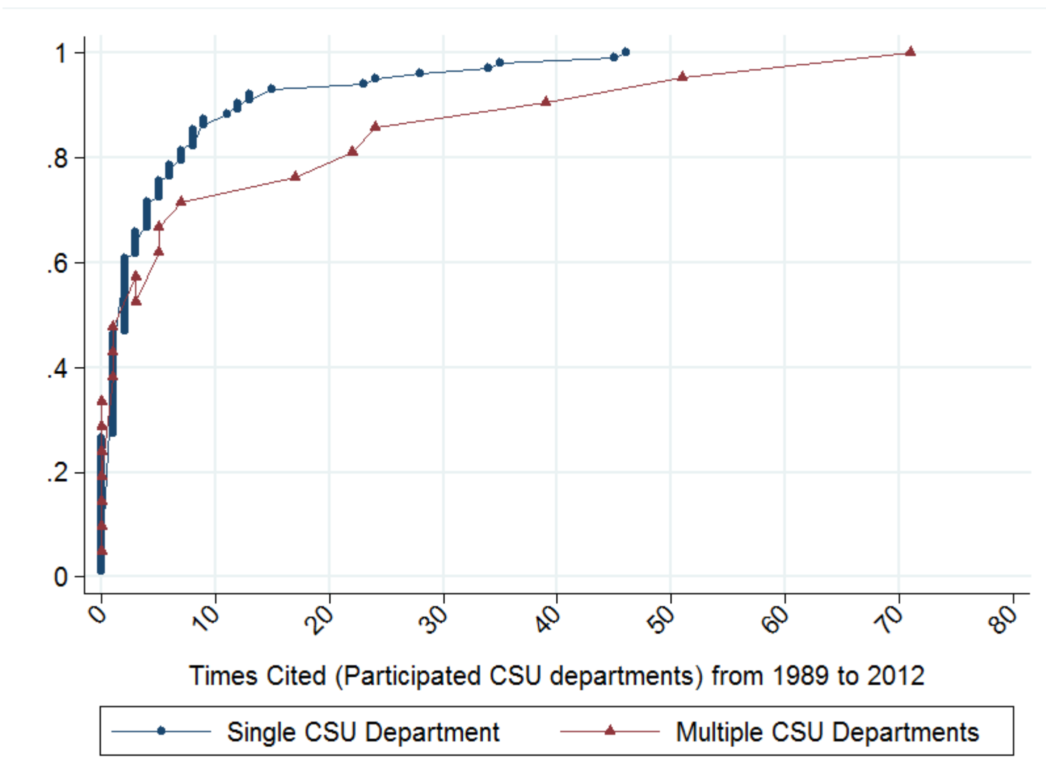


Figure 33—Cumulative Density Function (CDF) of patent citations by the number of participating CSU departments, (1989-2012)

As shown as Table 26 and Figure 33, the group of patents with inventors from multiple CSU departments has higher citation rates than the group of patents by inventors from a single department. This result is an obvious parallel to the research teams of published journal articles. The mean value of citations to patents from multiple departments, at 11.905, is almost twice that of patents from single departments, 4.990. Thus, research teams from multiple departments have better performance and impact than single department research teams.

3. The impact of the cited references per patent

Citation mapping tracks a patent's cited and citing references to discover technological relationships. A "backward" citation is defined as a citation made by a patent to patents that were issued earlier.

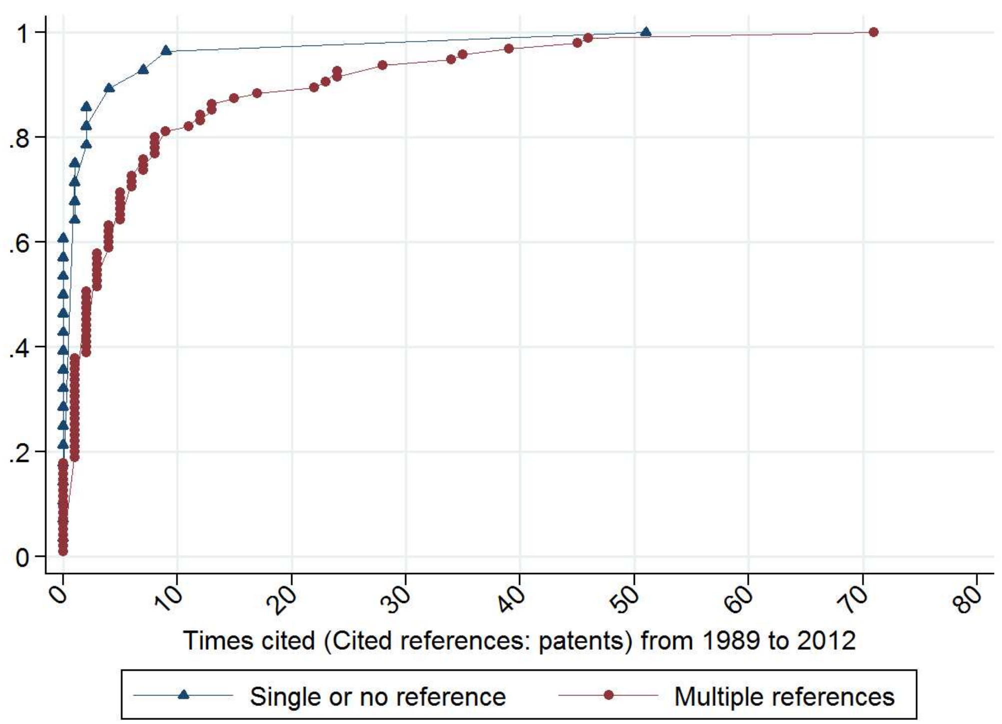


Figure 34—Cumulative Density Function (CDF) of patent citations received, based on the number of patent citations made, (1989 to 2012)

Accordingly, a "forward" citation is a citation made to a patent by more recently issued citing patents. Here, we examine how the number of cited patents, or backward citations, related to the number of citing patents, or forward citation. The sample is divided into binary groups: 1) those

CSU patents that make just a single or no backward citations to earlier patents, and 2) those CSU patents that make multiple backward citations to earlier patents. Table 26 shows that the mean value of forward citations received by the group of patents with multiple backward citations is higher than the other group, at 7.14 and 2.89, respectively. Thus, the number of cited patents is positively related to the number of citing patents.

4. The impact of the DWPI patent family per patent

A patent family is the set of either patent application or publication taken in multiple countries to protect a single invention. The World Intellectual Property Organization (WIPO), currently has 188 member states, and the Derwent World Patent Index (DWPI) maintained by Thompson Reuters covers more than 30.33 million inventions represented in over 64.8 million patent documents worldwide.[37] The aim of this section is to examine how the DWPI patent family size and structure is related to the impact of the research teams patented research outputs, as measured by forward patent citations.

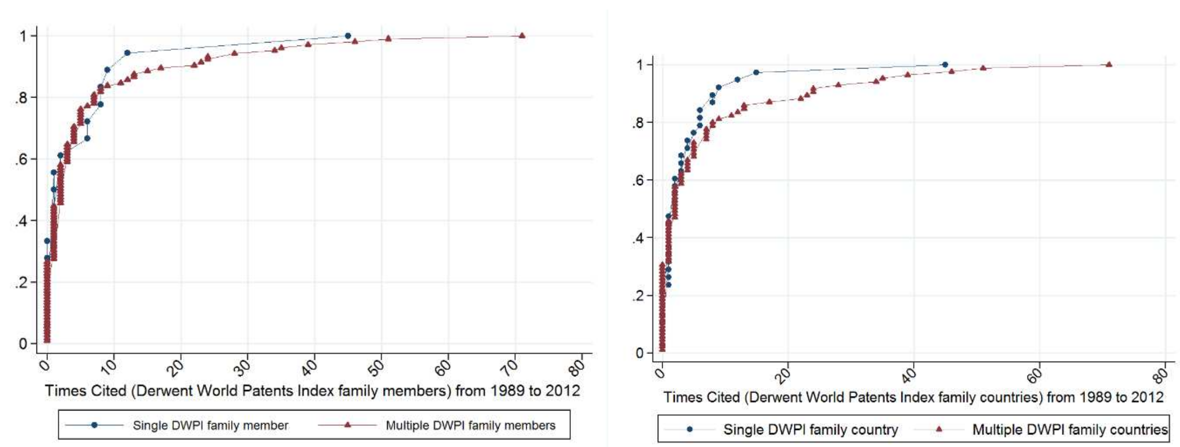


Figure 35—Cumulative Density Function (CDF) of patent citations based on the number of DWPI family members and the number of countries covered by DWPI families, (1989-2012)

[37] Source: Thomson Reuters Innovation (http://info.thomsoninnovation.com/)

In Table 26, the mean number of citing patents to the group of CSU patents with just a single DWPI family member is 5.55 and the mean citations to CSU patents protected in just one country is 4.131. These values are slightly smaller than the group of CSU patents with multiple DWPI family members and countries. It seems reasonable to assume that patents with multiple DWPI family members registered in multiple countries can have more chances to be cited. Moreover, it is likely also to increase self-citations within the members of the patent family. But most fundamentally, it is likely that patents on inventions with greater importance are likely both to be protected more widely, with more patent family members, and at the same time to be cited more frequently by subsequent patents. Figure 35 shows the CDFs of two types of DWPI patent family, members and countries per patent, by binary groups.

With regard to the results of patent team assembly mechanisms from previous section, the research impact indicated by DWPI patent family members or countries per patent is insignificant and has ambiguous results. However, we expect that the binary independent distribution might vary across different groupings than those chosen here. For example, the binary groups can be organized between above average and below average number of DWPI patent family members or countries per patent, rather than single versus multiple family members of countries per patent.

5. The impact of team size on patent citations

Figure 36 shows that the group of CSU patents made by large sized research teams, defined here as more than 5 inventors per patent, has the highest impact as measured by the number of citing patents. Similarly, the group of patents by mid-sized research teams has higher impacts in terms of citing patents than the group of patents by small sized teams. In Table 26, the mean value of citing patents per patent in the group by large sized teams is 11.85 per patent, which is more than two times greater than in the group by small sized research teams, at 4.84 per patent. Thus, it appears to support the conclusion that team size affects the impact of an invention as measured by the number of citing patents positively, and the result is an obvious parallel to the number of citations per article found in the previous section.

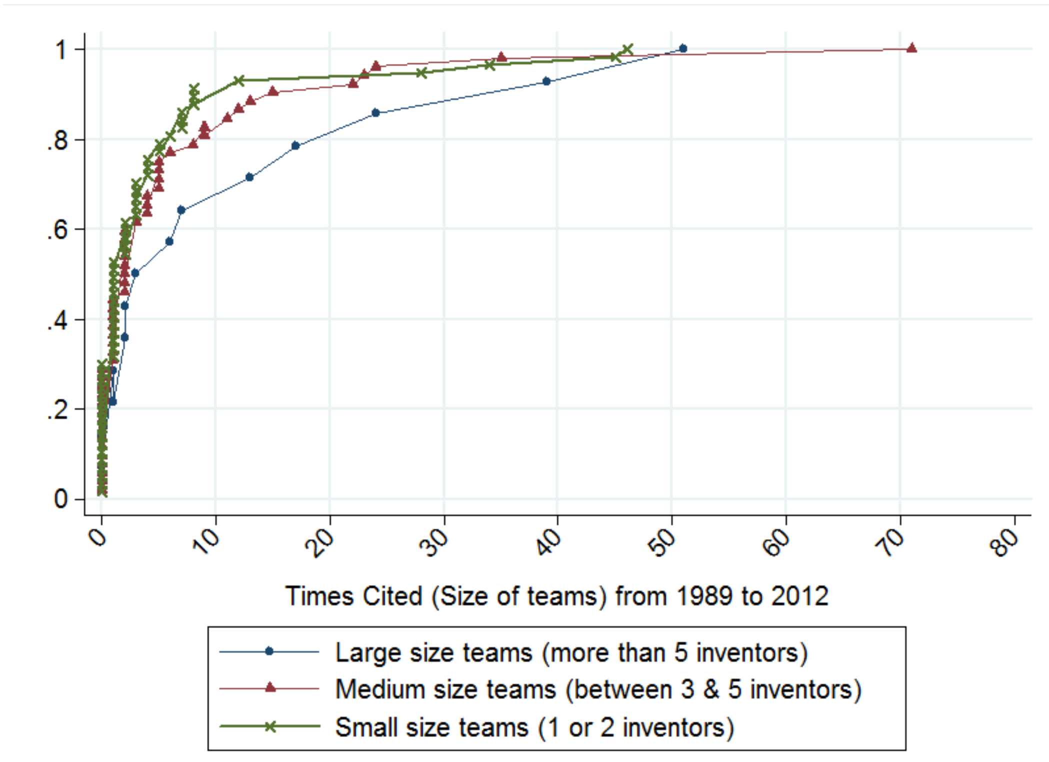


Figure 36—Cumulative Density Function (CDF) of times cited at the team size: the citing patents by three sample groups, (1989-2012)

6. The impact of time lag between application and publication per patent

According to Heher (2007), the of benchmarking time lags in the technology transfer process is typically 6-10 years elapsing from invention disclosure to significant income from a license. Moreover, the general time lag between patent application and publication of a granted patent is 2-3 years in most of the major patent offices. Thus, we examine how the length of time lags affect the quality of patents, as measured by the count of citing patents per patent. From Table 26, the highest mean value of citing patents is 9.19 for patents in the group with 2 or 3 years of time lag between application and publication. The groups of patents with a lag of one year or less and with more than 3 years have similar results, with means of 4.53 and 4.79 citations per patent, respectively. Figure 37 displays the CDFs of patent citations per patents based on the time lags between patent application and publication by three different sample groups, and shows the group with 2-3 year time lags distinguished from the other two groups.

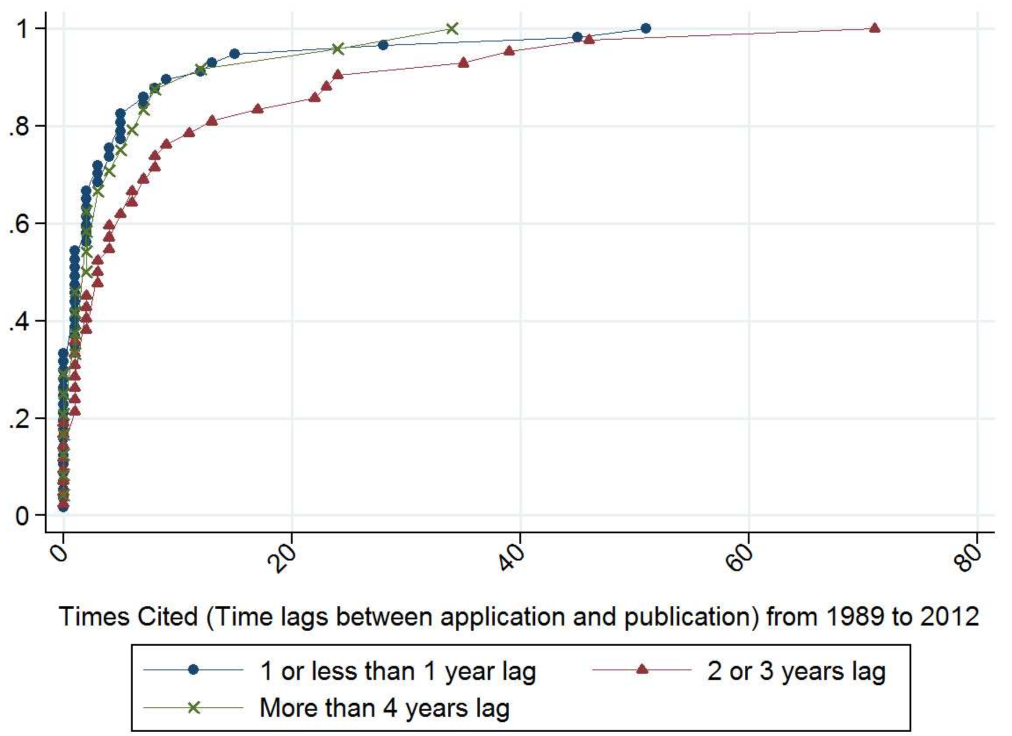


Figure 37—Cumulative Density Function (CDF) of patent citations based on the time lag between application and publication, (1989-2012)

VII. Conclusion

This chapter had attempted to explore the agency of knowledge production, viewing scientific research teams as quasi-firms arising as independent knowledge-creating entities within the university context. To sum up, the following methods were used: (1) summary statistics of the aggregate level of research teams and their time trends; (2) a conceptual framework of research teams; (3) analysis of ego-centric social networks of research teams (4) OLS regression of the sample research teams selected from entire groups, and (5) analysis of individual factors on impact of team research as measured by forward citations. Proceeding from what has been said above, it should be concluded that the size of research teams has a positive impact on university knowledge production and technology-transfer activities, and the team assembly mechanisms are significantly associated with key factors such as the interdisciplinary structure of teams and the participation of outside members. In the empirical approaches attempted, we explore two different views of academic-research teams. First, the findings from the egocentric social network illustrate how the existing networks and their structures can affect team formations and performances. Particularly, the participation of members from outside increases the size of the egocentric research team, and the growth patterns of the percent share from outside are an obvious parallel to the patterns of team size. Moreover, the pattern of outside members' networks clustered around each other, defining the ego's external research networks outside the university. As a result, the dynamics of the egocentric research teams were that the growth rate of team size came at the cost of reduced team density. The downward-sloping trend in density tells us that the egocentric research teams have been transformed from closer collaborations to more network-based relationships over time, with team size having an upward tendency over time.

Even though it makes it possible to explore the inside of research teams and their dynamics, the egocentric team analysis might be skewed, due to specific field characteristics and the research environments or idiosyncratic features of the ego that was selected for the case study. So, as an alternative approach, we selected a sample of 1,527 research teams associated with published research articles and a sample of 123 research teams associated with patents. The findings of this analysis indicate that the number of participating private companies has a positive and significant effect on the research article team's assembly but an insignificant effect on the patent team's assembly. As shown in Chapter 1, we can expect different results across the different types of knowledge outputs.

Other findings from the regression results show that the number of CSU departments, which represents cross-functional and interdisciplinary team formation, is statistically significant in influencing the team's assembly mechanism and resulting team size for teams producing both articles and patents as outputs. Thus, it seems reasonable to conclude that cross-functional team formation is more effective and common in the university context. Similarly, the proportion of outside members, which largely consists of former graduate students, faculty members from other institutions, and collaborators at private firms, also had a positive and significant impact on the team's assembly mechanism for both articles and patents. In addition, as hypothesized in Section IV, research project team formations are also influenced by private sector participations, such as involving members of private entities or attracting privately sponsored R&D investment.

Finally, when it comes to the quality or impact of research teams' knowledge production, the number of forward citations are evaluated. The cumulative density functions (CDFs) of forward citations to published articles tell us that research teams with members from multiple departments have a higher research impact than research teams from a single department. By the same token, larger-sized teams have higher impact than smaller-sized teams, as well as field variety. In the teams that produced patent, those with multiple backward citations per patent had a higher impact in terms of forward citations than those with a single or no backward citations per patent. This indicates that teams building on already existing inventions may be a significant factor for the research teams' impacts on the economic benefits with respect to knowledge spillovers, which follows from previous studies.

CHAPTER 3. UNIVERSITY KNOWLEDGE SPILLOVERS TO THE AGRICULTURAL ECONOMY: THE IMPACT OF AGRICULTURAL RESEARCH AT COLORADO STATE UNIVERSITY ON THE COLORADO ECONOMY, AND BEYOND

I. Introduction

Technology or knowledge spillovers are one of the most important sources of externality benefits in our society and economy. Defined as the spreading of ideas not mediated by market transaction, knowledge spillovers are an important issue for economic development, underlying many of our commonly held assumptions about commercial innovation processes. Moreover, historically, knowledge was understood by most economists to have largely public-good attributes, but more recently, the characteristics of knowledge have come to be viewed as more various, with different degrees of appropriability, due to differences in tacitness, embeddedness, or legal excludability because of well-developed intellectual property rights. In industry, knowledge is one of the most important assets today. Firms protect their knowledge or technology by various types of intellectual property and contractual mechanisms, including patenting and licensing. Therefore, questions of knowledge spillovers in industrial or commercial innovation have been analyzed extensively by many researchers (Jaffe, 1989; Mansfield, 1991 & 1995; Henderson et al, 1998; Jensen and Thursby, 2001; Adams, 2002; Cohen et al, 2002; Shane, 2002; Thursby and Thursby, 2011) and are important to both industry leaders and policymakers.

In high-tech industries, such as biotechnology and semi-conductors, spillovers of technology among similar firms are quickly transmitted and improved upon by rivals' follow-on inventions and innovation. Moreover, in technologically advanced countries, such as the United States, Japan, United Kingdom, Germany, France, and Israel, there is an ongoing exchange of people and ideas among private firms, universities, and research institutes located in close proximity to one another

(Krugman and Wells, 2013). Positive externalities can increase the incentives for cooperative research and development (R&D) between universities and industries, and are a fundamental reason that governments support the costs of R&D.

Most research universities in the United States are independent non-profit or state-affiliated knowledge organizations. They perform not only an educational function, but they also create and disseminate new knowledge through their other core functions as well. The roles and missions of universities have been shaped by a long history of national government policy changes, such as the Morrill Land-Grant Act of 1862, the Hatch Act of 1887, the Smith-Lever Act of 1914, and the Bayh-Dole Act of 1980. Following these formative policies, universities in the United States have generally come to embrace three missions: an educational mission, a research mission, and an outreach mission.

These different missions have spurred the emergence of different types of knowledge dissemination channels used in universities, such as the public domain, tacit dissemination through close collaboration, patenting/licensing of inventions and technical knowledge, and venture creation. These encompass not only traditional modes of university knowledge dissemination, such as publications, conference presentations, collaborative research with industry partners, consulting, co-supervising of internships, and so on, but also newer modes of university knowledge dissemination, such as university invention disclosures, patenting and licensing of new technologies, and the startup of new tech ventures. In agriculture, the Land Grant universities have long focused on agricultural research and the commercial dissemination of university innovations for regional and state economic growth, as well as for national and international development.

The main purpose of this chapter is to examine the geospatial pattern of knowledge spillovers from agricultural research at Colorado State University (CSU) and thus, by implication, the geospatial pattern of their commercial economic impacts, within Colorado's agriculturally-related sectors or industries, but also nationally and globally. In particular, this chapter seeks to identify how different channels of knowledge dissemination from the university differ in the type and location of impact. What is the relationship between geographic distance (proximity) and the types of university knowledge dissemination mechanisms used? Can we differentiate between knowledge dissemination channels specialized in disseminating 'sticky' (tacit) versus 'slippery' (codified) knowledge? We expect the more tacit or 'sticky' knowledge to stay within the local or regional economy and the latter more codified or 'slippery' types more readily to spill over nationally or even globally. Moreover, the use of different knowledge dissemination channels is likely to be varied across different technological categories.

The rest of this chapter consists of four sections. Section II reviews and discusses previous studies of knowledge spillovers and geographic proximity. Section III provides information on the structure of the agricultural economy and the industry value chain within Colorado. Section IV shows an empirical analysis of university knowledge spillovers across different dissemination channels. Finally, Section V summarizes the main conclusions.

II. Literature Review

A. Technology change and interaction between university and industry

The extent of new knowledge diffusion from universities and public research labs is important for private R&D activities and economic growth. Many studies have addressed and evaluated the effects of university knowledge spillovers with respect to commercial invention and innovation. Mansfield (1991 & 1995), Henderson, Jaffe, & Trajtenberg (1998), and Adams and Griliches (1998) study trends in university research and their effects on industrial innovation in the United States, finding that a range of new products and processes are based on academic research, and that some industries--such as electrical equipment, instruments, chemicals, drugs, mining, and petroleum--would not have developed in the way they did without the influence of academic research. University knowledge production activities have the potential both to directly affect commercial innovation and indirectly affect economic growth and development.

Universities, as a group of knowledge creating entities, are a fundamental source of highly skilled human capital to industry, and labor-augmenting technological change is a major factor of endogenous economic growth. According to Romer (1990), and Grossman and Helpman (1991), the role of knowledge capital, together with the stock of human capital, determines the rate of aggregate economic growth. Moreover, the distinguishing feature of knowledge is that it is a non-rival, but potentially excludable good. As mentioned before, however, the range of university knowledge is various, characterized by not only public good attributes (non-rivalry and non-excludability) but also some degree of private good attributes (rivalry and excludability). Agrawal and Henderson (2002) study the impacts of university knowledge spillovers focusing on MIT. They explore the magnitude, direction, and impact of patenting activities. According to this study, university R&D activities, especially faculty members' research, have a variety of knowledge dissemination channels, such as publications, conferences, consulting, informal conversations, collaborative research, patents/licenses, co-supervising, and so on. As such, different firms use different channels to access knowledge coming from the university. Similarly, Cohen et al (2002) and Laursen and Salter (2004) study the influence and role of university and government laboratory research on industrial R&D. These studies suggest that university research has a considerable impact on industry's new R&D projects and activities, across the different knowledge channels.

B. Geographic proximity and localization

Active interaction between university researchers and industry is important for dissemination of some types of knowledge. Because of this, the geographic location of impacts is arguably influenced by the characteristics of the knowledge being disseminated. Jaffe (1989) studies the importance of geographically mediated spillovers from universities and research labs to commercial innovation at the state level. He provides evidence that corporate patenting activities depend on university research, indicating that the university research system can increase local commercial innovations, but these effects are clearer within specific technical areas, such as drugs, chemicals, and electronics than in total. Jaffe et al (1993) compare the geographic location of citing patents with the patents they cite. The proportion of citations that geographically matched the originating patents indicate that that inventors are more likely to cite patents from the same country, state, and even the same metro area (MSA) and thus that the geographic location of knowledge spillovers is localized.

According to Adams (2002), the localization of university knowledge spillovers is greater than that of industrial spillovers. He finds the degree of localizations depends on the nearby stocks of R&D, but the degree of localization decreases with the size of firms. This study suggests that the results on localized university spillovers reflects the dissemination of normal science and the industry-university cooperative movement, implying that geographic localization occurs with the public good attributes of academic research. Similarly, Ponds et al (2009) find that university spillovers can be localized by geographically bounded mechanisms, but university-industry collaborations are not limited to the regional scale. Their findings show that university research impacts local innovation and inventions not only due simply to geographic proximity, but also due to collaboration networks.

Buenstorf and Schacht (2013) study the geographic information about patent licensing activities of the German Max Planck Society. They point out that proximity does not always effectively lead to better commercialization outcomes. Sometimes, proximity causes a negative association with foreign licensees. Moreover, there is little evidence to support that local licensees have superior information about the quality of academic inventions. Hong and Su (2013) study Chinese patent data for collaborations between university and industry. This study indicates that geographic distance is an obstructive factor in achieving university-industry collaborations in some cases, such as a central government system, and the problem of lock-in, indicating a lack of openness and flexibility. Nevertheless, a university's research and knowledge dissemination activities still play a significant role in providing knowledge inputs to industry innovation and inventions within its region, at least in particular fields and technologies (Anselin et al, 1997 and 2000).

Evaluating both the public and the private benefits of university knowledge spillovers relies considerably on the different types of knowledge dissemination. According to Jaffe (1989), when the channel is published journal articles, then geographic proximity is unimportant, but when the channel is informal interaction, then geographic locations is important in capturing the benefits of spillovers. According to Lester (2004), university contributions to regional commercial innovation processes can be achieved in various ways. Many universities are searching to develop their discoveries and findings by patenting and licensing to local companies, yet the most important contribution of the university may be through education and informal interactions as a public service to local communities and businesses.

III. The Regional Agricultural Economy and Value Chains

The state economy of Colorado had long depended on agriculture and innovation as drivers of economic growth and development. In 2011, the supply of agricultural inputs by Colorado agribusinesses contributed $2 billion, crop and livestock sales contributed more than $8 billion, and commodity marketing, processing, and food/beverage manufacturing contributed $15 billion to the state economy (Graff et al, 2013). Also in the one year of 2011, in agriculturally related fields[38], CSU researchers received just over $5 million in grants and contracts awarded from businesses in or closely related to agriculture, co-authored 65 scientific articles with industry partners, made 22 invention disclosures, submitted 11 patent applications, and founded one new startup company, again all involving technologies related to agriculture.

Colorado has a diverse agriculture and food sector, and several subsectors play important roles in the state economy. Colorado State University, as the Land Grant university in Colorado and a world leader in agricultural sciences, has had economic impact on these subsectors. Colorado is a major producer of beef and dairy, at both the farm level and in processing and manufacturing. Colorado State University has leading programs in veterinary medicine and animal science with emphases on large animal and bovine. Colorado is the 5th largest producer of potatoes in U.S., and CSU's Potato Breeding Selection Program has developed more than 60 percent of the potato varieties that are planted in Colorado. Colorado is a major wheat producing state and is home to the largest wheat milling company in the U.S., and the Colorado Wheat Breeding and Genetics program at CSU has improved more than 30 percent of wheat varieties grown in Colorado (CSU Ventures' Annual Report, 2012). Colorado maintains a good reputation for organic and natural

[38] In that year, the range of research fields were included in animal health, dairy (organic milk), pest control, crop varieties, soil fertilizer, ground water & irrigation, and food processing & packing.

foods. And Colorado hosts not only the two top brewing companies in the nation, but is also known for the high quality of local brewing firms[39].

Table 27—The numbers of actively innovating private companies within Colorado's agricultural and food value chain in 2014, congregated into technological categories, with cumulative Web of Science (WoS) publications and U.S. patents in each category, (1990-2013)

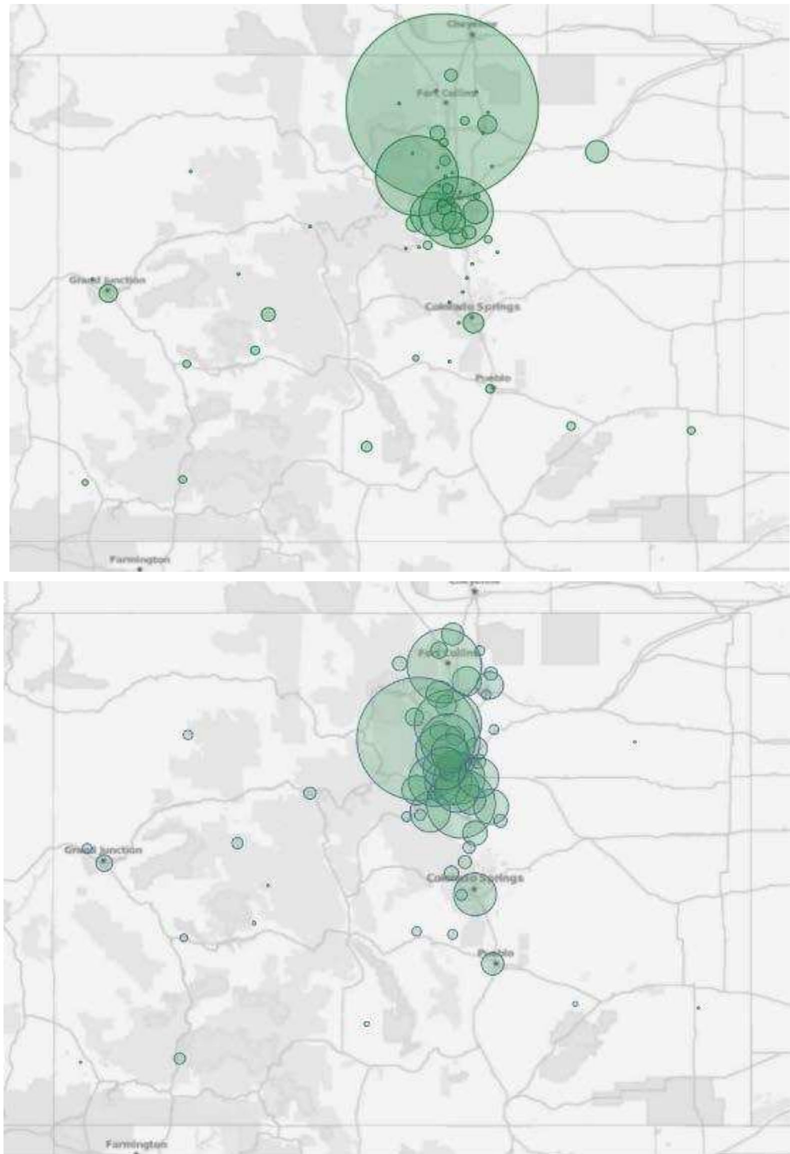
Technology	No. of company	No. of WoS publications	No. of U.S. patents
1. Water technology, infrastructure, analytics, and management	93	488	94
2. Soil fertility and pest control	23	21	68
3. Plant genetics and new crop varieties	20	30	527
4. Animal health, nutrition, and herd management	49	339	329
5. Agricultural information systems	22	3	97
6. Sensors, testing, and analytics for product quality and biosafety	32	119	99
7. Bio-energy & fuel	25	7	250
8. Commodity processing and food manufacturing	38	50	514
9. Dairy production and dairy product manufacturing	12	14	88
10. Beer, wine, & spirits production and marketing	67	26	37
11. Natural, organic, and local foods and marketing	33	7	36
12. "Fast & Fresh" food service	13	0	0
13. Other emergent subsectors	33	5	74
Total	460	1109	2,213

Source: Graff, Berklund, and Rennels, 2014.

According to Graff et al (2014), an innovation cluster in the agricultural and food industries appears to be forming in the Colorado Front Range. Innovation clusters can be defined as the geographically proximate sets of interconnected companies and associated institutions in particular fields and technologies. The structure of the value chain of Colorado's agriculture and food industries is closely associated with the emerging innovation cluster. The agricultural value chain in Colorado includes 550 innovators, of which 460 are private-sector companies and 90 are public-sector (academic, nonprofit, and government) organizations, and the innovating organizations are

[39] Source: Colorado Office of Economic Development & International Trade. www.advancecolorado.com

categorized according to a dozen areas corresponding to different subsections of the value chain. Table 27 displays these areas of agricultural innovation going on in Colorado, and Figure 38 shows the geographic landscape of Colorado's agricultural innovation.



A. Scientific journal publications in agricultural and food related fields, by city of author's institutional affiliation

B. U.S. patents on agricultural and food related techniologies filled and granted by city of inventor's residence

Source: Graff, Berklund, and Rennels, 2014.

Figure 38—A landscape analysis mapping two agricultural R&D outputs: scientific journal publications and U.S. patents filled and granted, (1990-2013)

The two most active areas in Colorado in terms of scientific research are water and animal agriculture (Table 27). Water related companies in Colorado are highly involved in scientific research papers and engage in projects for water storage, transmission, and irrigation infrastructure. Many of these companies consist of consulting and analytics firms or civil engineering firms. Moreover, the companies associated with animal health and nutrition—such as beef, dairy cattle, horses and sheep—participate actively in R&D for improving the quality of animal health and nutrition and herd management. Furthermore, as shown in Figure 38, the geographical locations of Colorado authors in the agricultural value chain are highly concentrated in Northern Colorado, between Denver, the main urban center, and Fort Collins, where CSU is located.

In addition, over the 5 years, 2008-2012, the College of Veterinary Medicine and Biomedical Sciences at CSU was top ranked in terms of privately sponsored grant awards and contracts, at an average of $5.8 million per year, and industry co-authored journal articles comprised an average of 75.6 articles per year. Moreover, the College of Engineering was second ranked over that time period, and the Department of Civil and Environment Engineering, which is the center of CSU's water related research, received private sector grants averaging $1.2 million per year, and published an average of 12.8 articles per year with industry co-authors. From this evidence, we might expect that the roles and missions of land-grant universities are crucially important for the local commercial innovation and inventions in the agricultural value chain, as well as the importance of geographical proximity.

IV. Empirical Study of University Knowledge Spillovers in Agriculture

We seek to take into consideration the full range of potential knowledge dissemination channels, as indicated by such measures as academic journal publications, industry co-authorship on journal publications, private sponsorship of research grants and contracts, patent applications and granted patents, and startup companies, all from data collected at the level of the different research units of Colorado State University (CSU), from 1989 to 2012. These various measures make it possible to analyze the extent to which the different types of knowledge dissemination channels work. Four general types of knowledge channels are introduced in Chapter 1, including (1) the public domain channel, (2) the collaboration channel, (3) the intellectual property rights and licensing contracts channel, and (4) the venture creation channel.

Table 28—Summary of data used to measure four different types of agriculture and food-related knowledge dissemination channels

Ag. Related CSU Knowledge Channels	Years	Sum
1. Public Domain Channel		
Published journal articles	2008-2010	1,202
Citations of these published journal articles	2008-2012	6,883
2. Collaboration Channel		
Privately sponsored grant awards and contracts	1989-2012	543
Total amount of the private grant awards (million $)	1989-2012	17.58
Industry co-authored journal articles	1989-2012	290
3. Patenting/Licensing Channel		
Patent applications and granted patents	1990-2013	76
Citations of these patents (Excluding self-cites)	1990-2013	1,868
4. Venture Creation Channel		
CSU affiliated startups	1989-2012	11

In this section, we examine CSU's knowledge dissemination and its impact on commercial innovation within the state economy, specifically in sectors related to agricultural industry, in terms of the four different types of knowledge dissemination channels. Moreover, this chapter attempts to determine the locations of private industry innovators impacted by or associated with CSU agricultural research in both agricultural and related sectors, to confirm the relationship between geographic distance and the types of CSU knowledge transfer mechanisms employed. In agriculture and food-related research activity, the four different types of knowledge channels are measured using the citations a sample of CSU journal publications, privately sponsored grants awards and industry co-authorship on articles, citations of CSU patent applications and granted patents, and CSU startups.

Table 28 provides summary statistics of the data used to measure activity in four different types of knowledge dissemination channels for CSU's agriculture and food-related research activities. First, within the context of agriculture (using Web of Science keywords, including agriculture, agronomy, entomology, food science, horticulture, plant sciences, soil science, veterinary science, and water) we select a target sample of 1,202 journal publications by CSU authors from the Web of Science database, published from 2008 to 2010. Using the Web of Science forward citations reporting tool, we find these 1,202 CSU articles have been cited 6,883 times, collecting from 2008 to 2012. We then analyze the location of the authors and other characteristics of these citing papers to understand the geographic footprint and the nature of spillovers via the public domain channel.

Second, from 1989 to 2012, 543 grant and contract awards were received by CSU from private sector sponsors to conduct agriculturally related research. The total amount of these awards was $38 million, and came from 169 private companies engaged in some aspect of agriculture or food related business. CSU researchers have collaborated with and co-authored 290 agriculture and food-related journal publications with authors from 194 private companies from 1989 to 2012. We then analyze the locations and characteristics of these companies that awarded grants and contracts to CSU and the companies that co-authored with CSU to understand the geographic footprint and the nature of spillovers via the research collaboration channel.

Third, CSU inventors had 76 agriculture and food-related patent applications and granted patents from 1990 to 2013. By 2015, all of these patent applications and grants had received 1,868 forward citations from other patents, owned by 206 companies. We then analyze the location of the inventors and the assignee firms of these citing patents to understand the geographic footprint and the nature of spillovers that occur via the intellectual property licensing channel.

Finally, 11 startup companies were created from 1989 to 2012 from research in CSU's departments and research units around technologies that can be associated with or applied in the agriculture or food industries. We analyze the locations and characteristics of these startup companies to understand the geographic footprint and the nature of spillovers via the venture creation channel.

Across these four different types of knowledge dissemination channels, this section seeks to determine the locations of private industry innovators in agricultural sectors that utilize the university's knowledge outputs, to ascertain the relationship between geographic distance and the type of channel employed. We expect the different channels to specialize in disseminating sticky or slippery (tacit or codified) forms of knowledge in agriculturally related sectors, with some staying within Colorado and others more readily spilling over nationally or even globally.

A. Geographic footprint of university knowledge spillovers

1. Public domain mechanism of knowledge dissemination

Journal publications are the major research output of universities, and knowledge primarily disseminated through the public domain channel has the strongest public good attributes, defined as being both non-excludable and non-rivalrous. Once published (and if not otherwise protected, such as by a patent), it is not possible to exclude anyone from accessing this knowledge or to prevent simultaneous use or access, which means full freedom of use and open access. Again, according to Jaffe (1989), when the mechanism of dissemination is primarily journal publications, then geographic location is generally unimportant for recipients to access the knowledge transfers.

Journal publications are the channel most likely to be used to transfer "slippery" information, that which is relatively codifiable and transmissible at lower transaction costs, so that there is no boundary of the knowledge spillovers. These spillovers are effectively worldwide, and they have the highest speed of transmission.

Table 29—The location of authors of the 6,883 articles that cite a randomly sampled target set of 1,202 agriculturally-related journal articles by CSU authors in several different agricultural research fields and technologies

Research Field/Technology	No. of citing articles, by authors in Colorado	No. of citing articles, by authors in US but outside of Colorado	No. of citing articles by authors outside the US	Average distance from CSU (miles) for those in U.S. only
1. Water tech & management	133 (9.8)	444 (32.7)	782 (57.5)	950.46
2. Crop genetics, soil fertility, & pest control	209 (8.3)	861 (34.4)	1,433 (57.3)	1,032.24
3. Animal health & nutrition	202 (10.3)	680 (34.6)	1,086 (55.2)	976.01
4. IT and data systems in food & agriculture	13 (8.9)	53 (36.3)	80 (54.8)	1,071.91
5. Bioenergy	34 (10.8)	106 (33.5)	176 (55.7)	962.49
6. Food & beverage processing & manufacturing	63 (10.7)	195 (32.9)	333 (56.3)	930.47
Total	654 (9.5)	2,339 (34.0)	3,890 (56.5)	988.79

Note: Parentheses are percent share of total citing papers.
See the names and addresses of company in Appendix

Table 29 shows a target sample of articles citing a randomly selected subset of CSU's agriculturally-related journal publications, within seven Web of Science's categories chosen for disciplines encompassing the whole agricultural value chain. In Table 29, of the total of 6,883 citing articles, 3,890, or 56.5 percent, are by authors in other countries. Of the total, only 654 articles are by Colorado authors, accounting for just 9.5 percent. Within the United States, the average distance from CSU—which is located in Fort Collins, Colorado—to the location of citing papers' authors is 989 miles. While this is a somewhat simplistic measure, it provides a basic sense of the geographic footprint of spillovers via CSU's agriculturally-related journal publications.

Among the various sub-fields or technology categories, the farthest average distance within the US is 1,072 miles from CSU, in the research field of IT and data systems for food and agriculture.

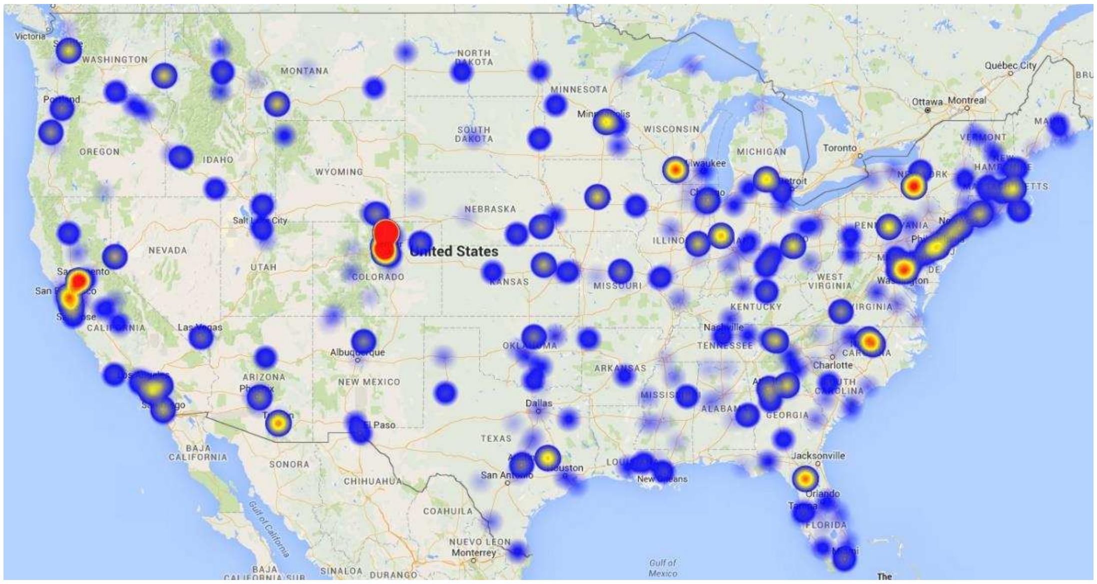


Figure 39—Heat map of the geographic locations in the U.S. of authors of papers citing CSU's agriculturally-related journal publications

Figure 39 displays the mapping of the geographic footprint in the United States of the location of authors of articles citing CSU's journal publications. Significantly, there are relatively high densities of citing authors located in northern Colorado, northern California (the San Francisco Bay area), southern Texas, Wisconsin, Florida, North Carolina, New York, and the Boston area. These geographic locations are all associated with the location of land-grant universities in the United States, other universities with a similar profile of research in agricultural sciences and veterinary health. Similarly, Figure 40 displays the geographic locations of citing papers based on authors' affiliations within Colorado. Citations of CSU's agriculture-related journal publications are highly concentrated in three locations within the state, each associated with a major research institution: Fort Collins (CSU), Boulder (University of Colorado), and Golden (Colorado School of Mines and the National Renewable Energy Laboratory). Moreover, among the other minor locations are those associated with the experimental farms of the Colorado State Agricultural Experimental Station (SAES).

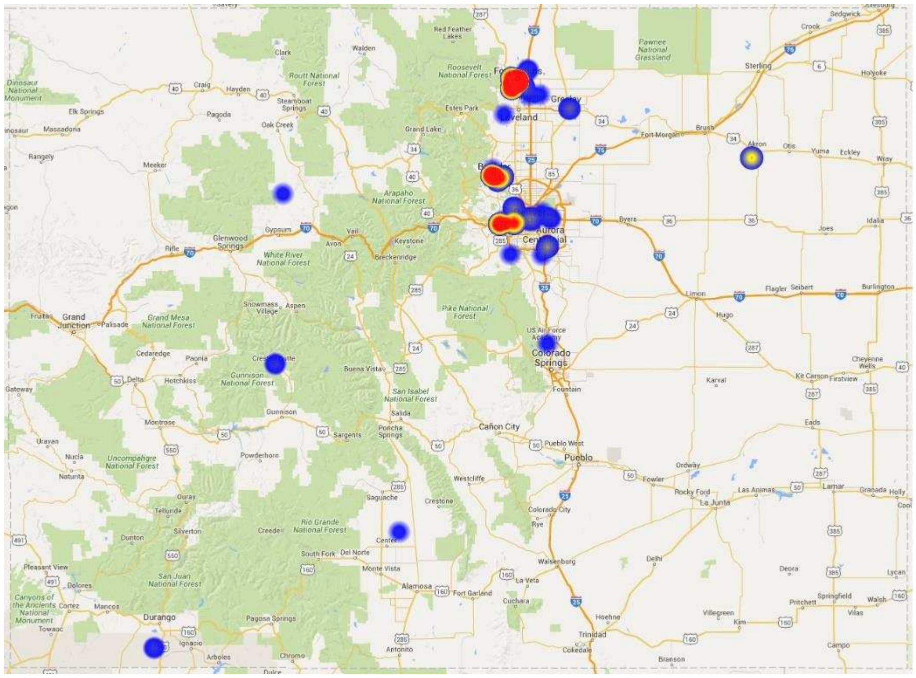


Figure 40—Heat map of the geographic locations in Colorado of authors of papers citing CSU's agriculturally-related journal publications

Therefore, knowledge spillovers of CSU's agriculture-related journal publications via the public domain mechanism have impact locally, nationally, and globally. As previous studies have pointed out, the knowledge channel of journal publications does not depend upon geographic proximity for realizing the social and economic benefits of knowledge transfers.

2. Collaboration mechanism of knowledge dissemination

The collaboration mechanism of knowledge commercialization is characterized by close interaction between university and industry researchers, such as in research collaboration or outreach activities. It is generally well suited for conveying tacit or sticky knowledge. Faculty members in the university work with colleagues in the private sector in a number of ways, including consulting, conference presentations, informal consultations, collaborative research projects, co-supervising of interns, and such. However, these collaboration activities are harder to detect and to systematically measure in terms of their magnitude, size, and scope. So we proposed three proxy variables in Chapter 1, including (1) the number of industry co-authored journal publications, (2) privately sponsored grant awards, and (3) departmental level cooperative extension budgets. Two of these measures contain information about the geographic location of the collaborating industry partner--the industry co-authored journal publications and the privately sponsored grant awards. Thus, in this section, we utilize these two to assess the geographic scope of impact of collaboration as a mechanism for CSU's knowledge dissemination within agriculture.

1) Grants and contracts awarded from private sector sponsors

One of the proxy measures of university traditional collaboration activities is the announced award of research funding from external sources via grants and contracts. The total value of privately sponsored grants and contracts in agriculturally-related sectors from 1989 to 2012 is \$17.6 million, in awards from 169 private companies. Table 30 presents summary data on private sponsors of grants and contracts across the different agriculturally related sub-fields and technologies. The average distance from CSU to the locations of private sponsors in U.S. is 663 miles. Among the sub-fields or technology categories, the nearest average distance is 279 miles, in the field of water technology and management.

The 13 companies associated with water technology located in Colorado account for 76.5 percent of all of CSU's private sector collaboration on water–related research and projects. These in-state companies have invested \$1.43 million in their collaborations with CSU, which is almost 73 percent out of total private R&D investment in water-related research and projects at CSU. Moreover, there is only one foreign company, KOWACO-Korean Water Resources Corporation, and significantly, this company is strongly associated with the Department of Civil Engineering at

CSU[40]. Similarly, almost 50 percent of companies associated with IT and data systems in food & agriculture, in bioenergy, and in food and beverage manufacturing technologies are located in Colorado. However, only 15 percent of the collaborating companies in animal health technology are located in Colorado.

Table 30—The location of privately sponsors of contract and grant awards to CSU in several different agricultural research fields and technologies

Research Field/Technology	No. of firms in Colorado	No. of firms in the US but outside of Colorado	No. of firms outside the US	Average distance from CSU (miles in U.S. only)	Amount of awards (million $)
1. Water tech & management	13 (76.5)	3 (17.6)	1 (5.9)	278.87	1.97
2. Crop genetics, soil fertility, & pest control	19 (38.8)	23 (46.9)	7 (14.3)	575.51	2.96
3. Animal health & nutrition	5 (14.7)	25 (73.5)	4 (11.8)	1051.27	6.24
4. IT and data systems in food & agriculture	9 (42.9)	10 (47.6)	2 (9.5)	727.26	1.28
5. Bioenergy	5 (50.0)	4 (40.0)	1 (10.0)	620.33	0.95
6. Food & beverage processing & manufacturing	16 (42.1)	19 (50.0)	3 (7.9)	570.32	4.19
Total	67 (39.6)	84 (49.7)	18 (10.7)	662.90	17.58

Note: Parentheses are percent share of total firms.
See the names and addresses of company in Appendix.

Figure 41 display the geographic footprint across the U.S. of private firms that have awarded grants and contracts for agriculturally related research at CSU. Although sponsoring companies are distributed widely throughout the various states in the U.S., many of them are in the northern Front Range of Colorado. Figure 42 shows the geographic footprint within Colorado, and the locations of companies are relatively compacted in the Front Range of Colorado, particularly in northern Colorado near CSU and around the Denver Technology Center (DTC) in south Denver.

[40] The Department of Civil Engineering at CSU and KOWACO have a sisterhood relationship, by which many faculty members in the Department of Civil Engineering at CSU have participated in national water projects in South Korea via KOWACO.

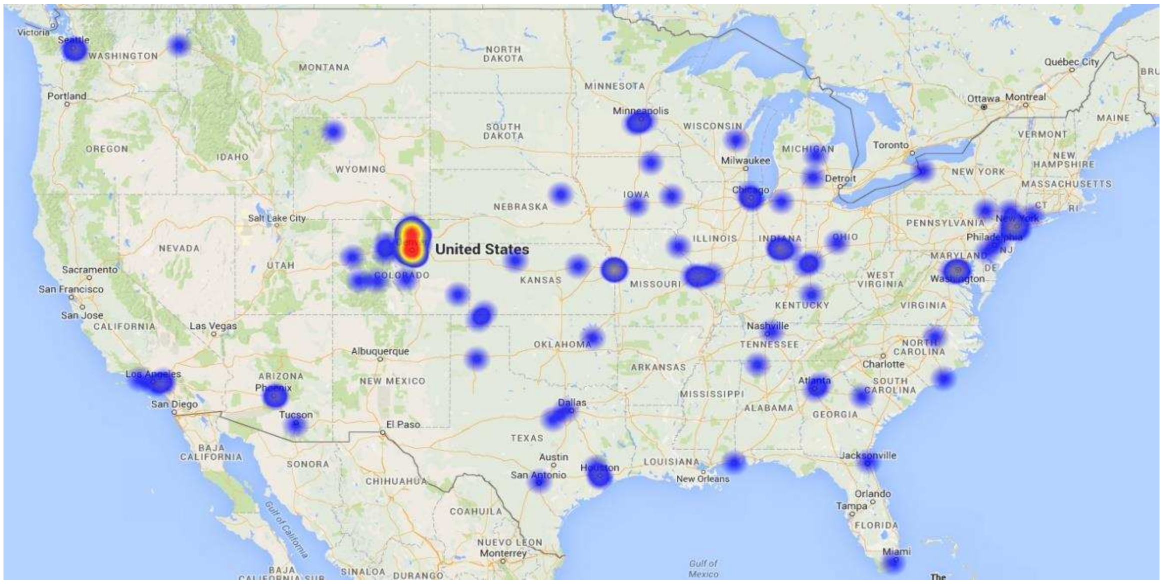


Figure 41—Heat map of the geographic footprint of locations in the U.S. of private companies that have sponsored research contract and grant awards for agriculture and food related research at CSU

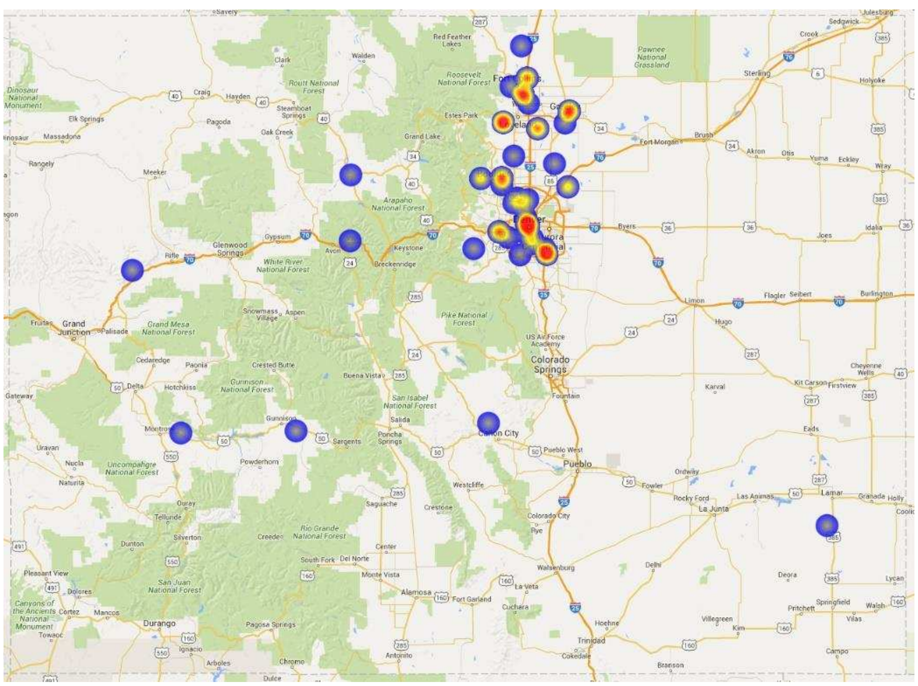


Figure 42—Heat map of the geographic locations, within Colorado, of private companies that have sponsored contract grant awards for agricultural and food related research at CSU

2) Industry co-authorship on academic journal articles

Co-authored journal publication is another more direct indicator of the university collaboration activities with industry. We find a total number of 290 papers in agriculturally related fields by CSU authors written with co-authors at 194 companies. While this particular knowledge output (the article being published) is being placed in the public domain, we assume that it is indicative of associated tacit knowledge being generated that is transferred through the mechanism of interpersonal contact and collaboration. As such, we assume that it may have greater reliance upon geographic proximity. Major characteristics of this channel is that it is significantly associated with CSU authors' social and professional research networks, and their capacity of maintaining collaboration activities with these private sector colleagues.

Table 31—The location of industry co-authors on journal publications with CSU authors in several different agricultural research fields and technologies

Research Field/Technology	No. of firms in Colorado	No. of firms in the US but outside of Colorado	No. of firms outside US	Average distance from CSU (miles in U.S. only)	No. of co-authored papers
1. Water tech & management	21 (60.0)	12 (34.3)	2 (5.7)	482.85	54
2. Crop genetics, soil fertility, & pest control	9 (21.4)	25 (59.5)	8 (19.0)	854.26	55
3. Animal health & nutrition	6 (16.2)	23 (62.2)	8 (21.6)	1094.93	62
4. IT and data systems in food & agriculture	12 (48.0)	13 (52.0)	0 (0.0)	636.48	31
5. Bioenergy	6 (42.9)	7 (50.0)	1 (7.1)	758.92	19
6. Food & beverage processing & manufacturing	10 (24.4)	24 (58.5)	7 (17.1)	622.07	69
Total	64 (33.0)	104 (53.6)	26 (13.4)	749.92	290

Note: Parentheses are percent share of total firms.
See the names and addresses of company in Appendix.

The average distance in Table 31 is 750 miles, which is slightly longer than the average distance of 663 miles, in Table 30 in the previous section, to companies that sponsored grants and contracts, an alternative measure of the collaboration channel. By the same token, Figure 43 displays the geographic footprint of co-authors from private companies, and it is virtually identical with that of Figure 41 of private sector co-authoring companies, with a relatively high density in Colorado. However, there is some contrast between Figure 42 and 44, with the locations of private sector co-authors more highly concentrated near CSU. This may be due to the fact that the kind of relationship involved in sponsoring may require more active interaction than the kind of relationship involved in co-authoring.

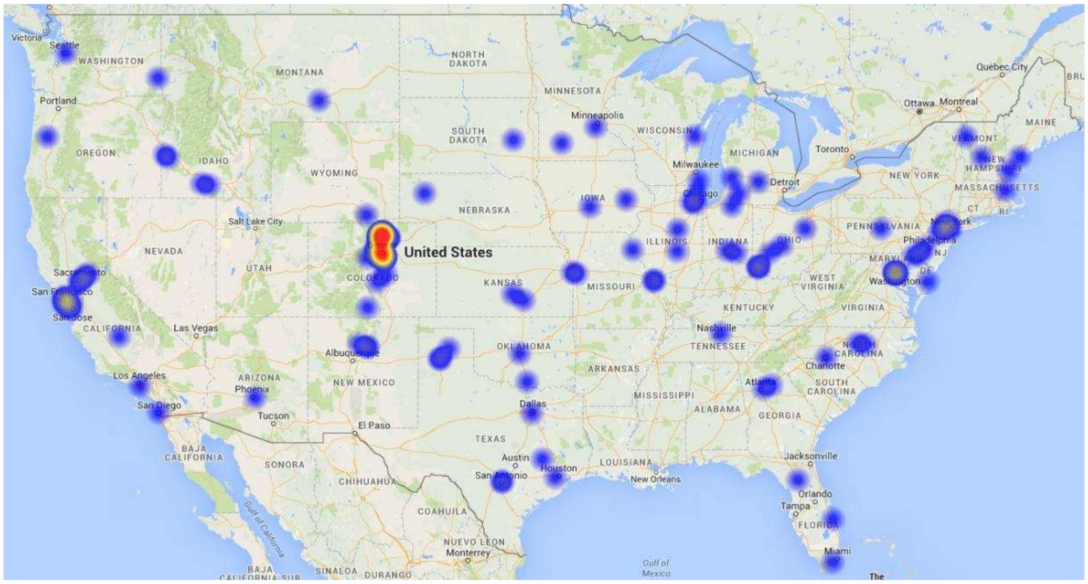


Figure 43—Heat map of the geographic location of companies in the U.S. that have co-authored journal publications with CSU researchers in agriculturally related fields

Privately sponsored grants and contracts and industry co-authored journal publications as proxy measures of the traditional collaboration mechanism of knowledge dissemination that are more likely to involve transfer of sticky or tacit forms of knowledge, which has relatively higher transaction costs and capacity requirements. It is possible to exclude others from accessing (at least some key aspects of) this knowledge, by virtue of its intrinsic stickiness, merely by not including them in the collaborative relationship. Therefore, geographic proximity is relatively more important. Within these two proxy measures of traditional collaboration, we find the geographic location of industry co-authors on journal articles have a longer average distance from CSU, perhaps because it involves more codified forms of knowledge than the other measure of grant and contract awards from private sponsors.

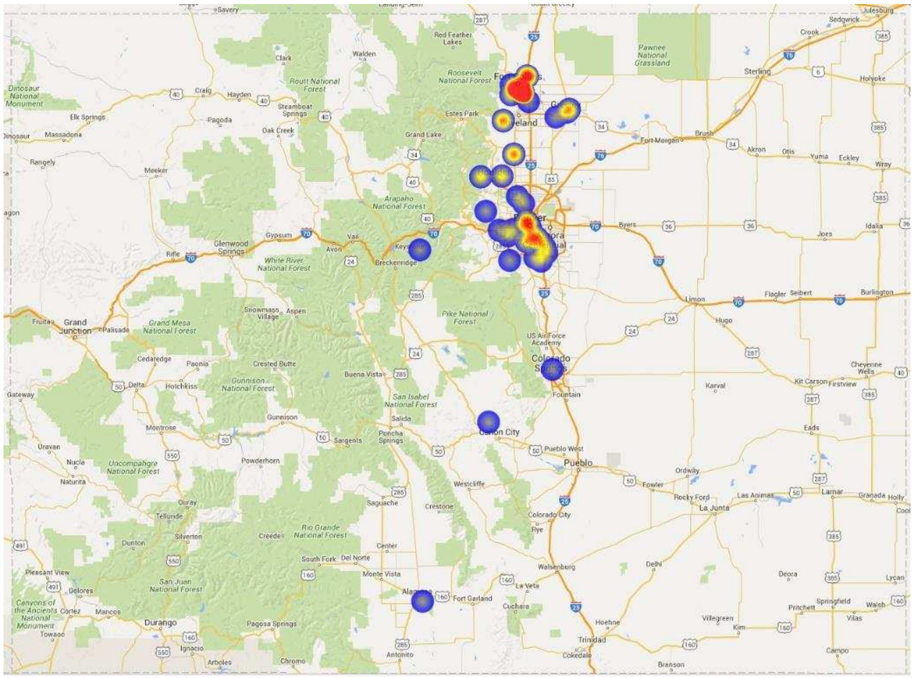


Figure 44—Heat map of the geographic location of companies in Colorado that have co-authored journal publications with CSU researchers in agriculturally related fields

3. Patenting/licensing mechanism of knowledge dissemination

The patenting and licensing mechanism for knowledge commercialization, which is characterized by the utilization of the intellectual property rights (IPRs) and licensing contracts to control access, is best suited when a certain degree of excludability is required to create sufficient incentives for follow-on investments in otherwise non-excludable and non-rivalrous knowledge outputs. To better understand the geographic impact of CSU's patented knowledge, we utilize citation mapping. In general, citation mapping of patents consists of connecting cited and citing references (backward and forward, respectively), which allows us to track the relationships between existing and new technologies, as well as the impacts of the existing technologies on the emergence of new technologies in the field. CSU inventors filed for protection on 76 inventions in agricultural and food related technologies between 1989 and 2013. Our analysis found 1,868 newer patents had made citations to these CSU patents and patent applications. Table 32 summarizes data on these citing patent documents across the different research fields and technologies related to agriculture.

Table 32—Patent documents citing CSU's portfolio of patent applications and granted patents across the different agricultural technologies

Research Field/Technology	No. of citing patents by inventors in Colorado	No. of citing patents by inventors in the U.S. but outside of Colorado	No. of citing patents by inventors outside of the U.S.	Average distance from CSU (miles) for those in the U.S. only
1. Water tech & management	0 (0.0)	25 (71.4)	10 (28.6)	1278.18
2. Crop genetics, soil fertility, & pest control	11 (8.0)	103 (75.2)	23 (16.8)	1037.57
3. Animal health & nutrition	187 (29.3)	329 (51.5)	123 (19.2)	888.06
4. IT and data systems in food & agriculture	25 (4.7)	466 (87.3)	43 (8.1)	994.82
5. Bioenergy	6 (2.0)	233 (77.9)	60 (20.1)	1303.70
6. Food & beverage processing & manufacturing	16 (7.1)	145 (64.7)	63 (28.1)	886.25
Total	245 (13.1)	1301 (69.6)	322 (17.2)	1001.87

Note: Parentheses are percent share of total citing patents.
See the inventors' affiliated companies in Appendix.

As shown as Table 32, the average distance from CSU to the locations of inventors of citing patents is 1002 miles, which is similar to the average distance seen for citing papers, in the public domain knowledge dissemination channel, and longer than the average distance to private sector research sponsors grants and co-authors, in the collaboration knowledge dissemination channel. Thus, by virtue of the slippery nature of information codified within patents, geographic proximity is unimportant for spillovers, at least to the extent of other inventors learning about the new technology disclosed by the patent documents. Although it is still possible to exclude others from accessing and using the technology described in the patents in those jurisdictions where the patents are issued. We note that the share of foreign inventors citing CSU's patents, at 17 percent, is relatively much smaller than the share of foreign authors citing CSU's journal publications, at 57 percent.

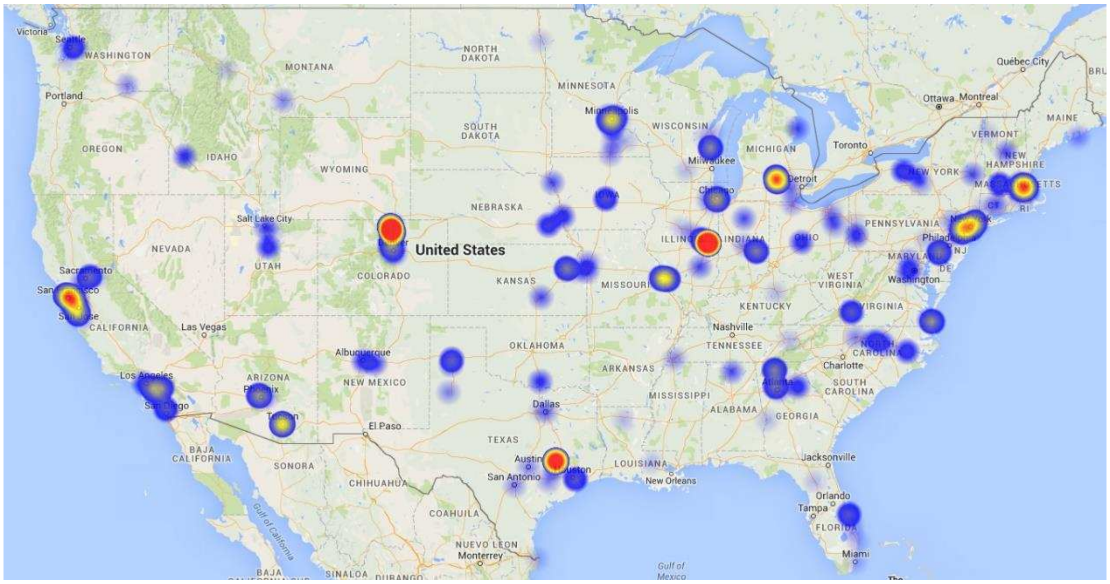


Figure 45—Heat map of the geographic location in the U.S. of inventors on patents that cite CSU's agriculturally related patent applications and granted patents

Figure 45 displays the geographic locations of inventors on the patents that are citing CSU's agricultural technology patents. This map demonstrates the commercial spillovers of CSU's agriculture-related research throughout the U.S. The distribution of the geographic footprint is similar to that of authors citing CSU's agricultural journal publications in Figure 39, but not all of the geographic locations are the same. It seems reasonable to assume that the distinction between citing patents and citing papers can be explained by not only by the different magnitudes of citation rates, but also the different characteristics of the respective knowledge dissemination channels. Both channels specialize in disseminating relatively slippery forms of knowledge, but the patents have an invoked exclusion by virtue of IPRs.

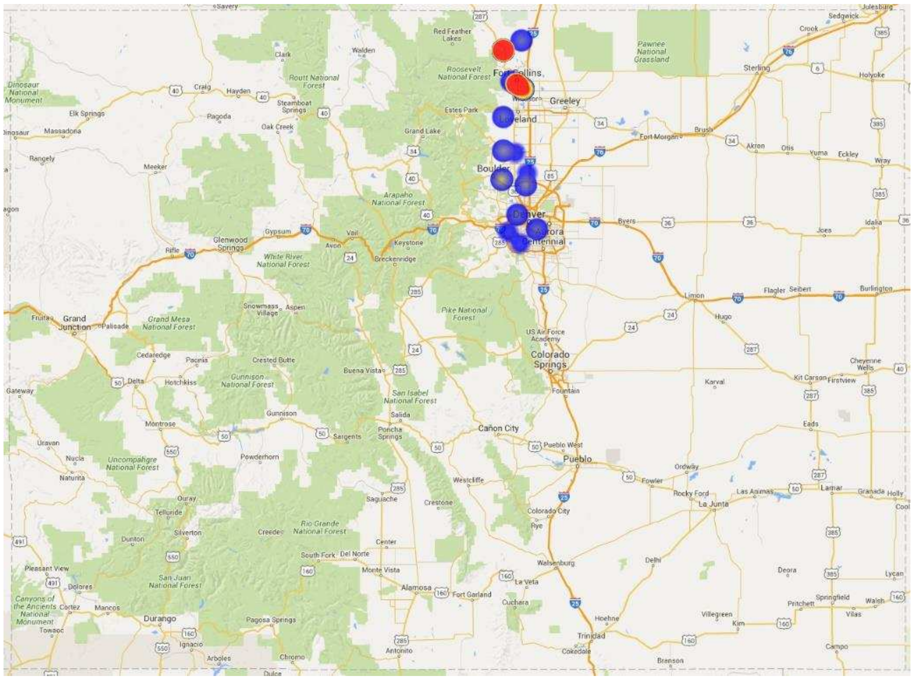


Figure 46—Heat map of the geographic location within Colorado of inventors on patents that cite CSU's agriculturally related patent applications and granted patents

Similarly, Figure 46 shows the location of inventors within Colorado that are citing CSU's agriculturally related patents. Of these, 91 percent are concentrated near CSU and elsewhere in the northern Front Range, and we observe that many of these areas are strongly associated with the presence of CSU's startup companies. Therefore, spillovers via the patenting and licensing mechanism appears to have a wide impact on commercial innovation in agriculturally-related sectors. While a patented invention is a codified or slippery form of knowledge, it is possible to exclude others from accessing and making use of it, which is the key distinction from research results only disseminated via journal publications through the public domain mechanism.

4. Venture creation mechanism of knowledge dissemination

The venture creation mechanism of knowledge commercialization is perhaps best suited for raising private capital for the further development of knowledge, treating that knowledge most like a private good, which has intrinsic stickiness or context dependence, making it possible to exclude others from accessing this knowledge. Both patenting/licensing and venture creation are newer mechanisms of knowledge disseminations, distinguished from some of the more traditional industry collaboration activities.

Table 33 summarizes data about the ten startup companies created by CSU across a range of technologies related to agriculture and food. Nine of the ten startup companies are located in Colorado, mostly near the CSU campus, while one company recently moved from Fort Collins, CO, to Navasota, TX.

Table 33—CSU's startup companies across the different agricultural technologies

Technology	Startups in Colorado	Startups in the US but outside of Colorado	Startups outside the US	Average distance from CSU (miles in U.S. only)
1. Water tech & management	1 (100.0)	0 (0.0)	0 (0.0)	5.00
2. Crop genetics, soil fertility, & pest control	2 (100.0)	0 (0.0)	0 (0.0)	27.00
3. Animal health & nutrition	3 (75.0)	1 (25.0)	0 (0.0)	257.75
4. Bioenergy	2 (100.0)	0 (0.0)	0 (0.0)	12.50
5. Food & beverage processing & manufacturing	2 (100.0)	0 (0.0)	0 (0.0)	33.50
Total	10 (90.9)	1 (9.1)	0 (0.0)	107.45

Note: Parentheses are percent share of total startup companies.
See the names and addresses of company in Appendix.

Figure 47 illustrates the geographic locations of these startup companies, almost entirely limited to Colorado, and Figure 48 shows that the startups are mostly located near CSU in the northern Front Range. Thus, the venture creation mechanism appears to be highly reliant upon geographic proximity. Involving quite sticky knowledge and context dependence, it has much higher transaction costs and capacity requirements than the other knowledge dissemination channels we have considered.

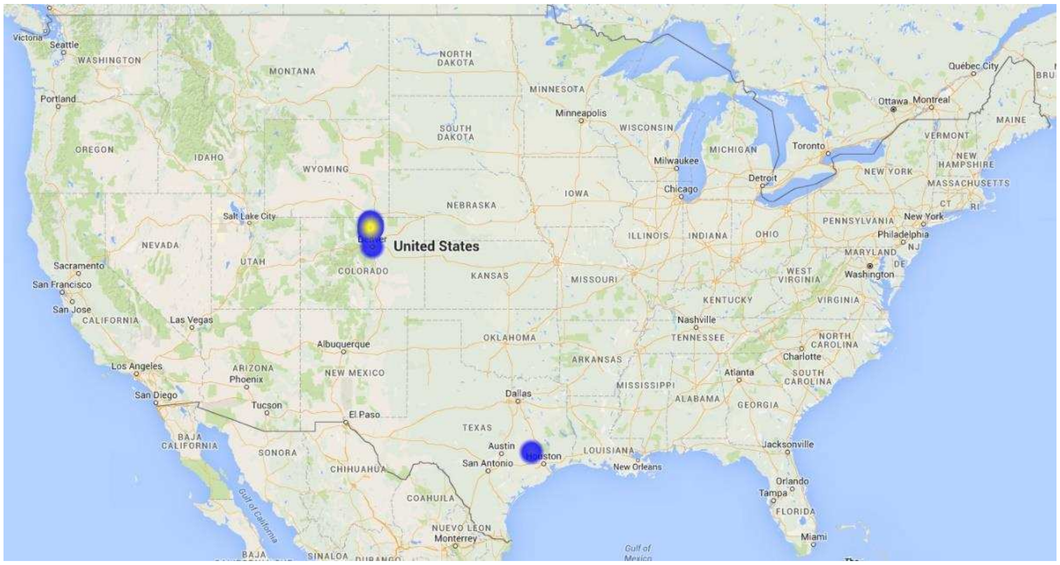


Figure 47—Heat map of the geographic locations in the U.S. of CSU's agricultural and food related startup companies.

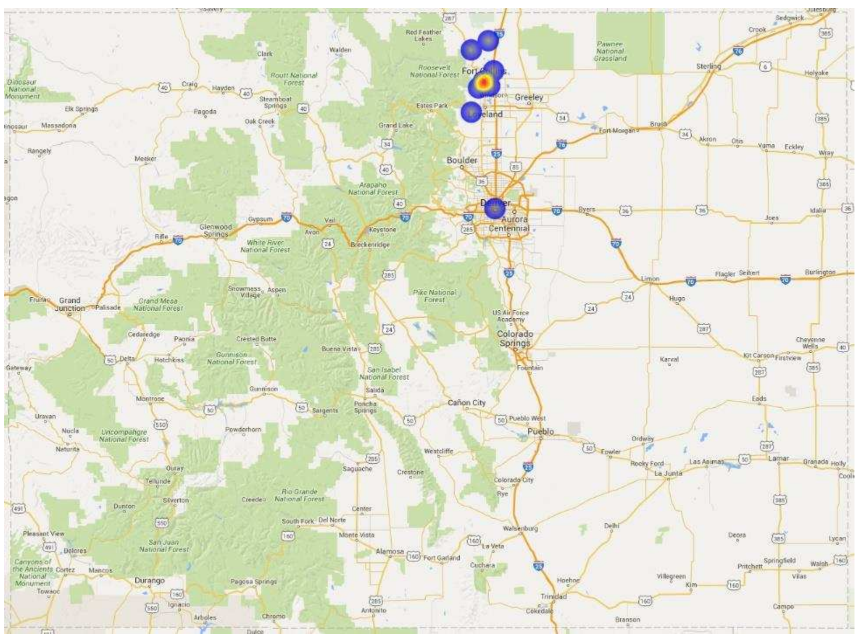


Figure 48—Heat map of the geographic locations within Colorado of CSU's agricultural and food related startup companies.

B. Non-parametric tests of university knowledge dissemination channels

In this section, we attempt to examine the different degrees of knowledge "stickiness" across CSU's dissemination channels by estimating the one-sample Kolmogorov-Smirnov non-parametric test, which is based on Kolmogorov (1933), Smirnov (1933), and Conover (1999). This test involves measuring a random sample from some unknown distribution for testing the null hypothesis, which specifies some distribution function, $F^*(x)$ as the cumulative distribution function.

A random sample $X_1, X_2, \cdots, X_n$ is drawn from some population, such as the geographic distances of university knowledge spillovers, and it is compared with the true distribution function of the random sample, $F^*(x)$. In the test, we hypothesize that the true distribution is a normal distribution. In order to compare the random sample with $F^*(x)$, the empirical distribution function of the random sample is defined by Definition (1)[41]

Definition (1) Let $X_1, X_2, \cdots, X_n$ be a random sample. The empirical distribution function, $S(x)$, is a function of *x*, which equals to the fraction of $X_i s$ that are less than or equal to *x* for each *x*, $-\infty < x < \infty$.

The equation (30) represents the empirical distribution function and where, $I_{[-\infty,x]}\left(X_i\right)$ is the indicator function, which equals to 1 if $X_i \leq x$ and equals to zero if otherwise.

$$S(x) = \frac{1}{n}\sum_{i=1}^{n} I_{[-\infty,x]}\left(X_i\right) \tag{30}$$

So, this test definition can compare the empirical distribution function, $S(x)$, with the theoretical distribution function, $F^*(x)$, to see if there is good agreement.

[41] W.J, Conover (1999)

The geographic distances of university knowledge spillovers data consist of a random sample, $X_1, X_2, \cdots, X_n$, of size *n* associated with some unknown distribution function, denoted by $F(x)$.

Hypothesis (1): Let $F^*(x)$ be a completely specified theoretical distribution function, such as a normal distribution: Two-sided test

$$\begin{cases} H_0 : F^*(x) = F(x) \ \forall x \ \text{from} -\infty \text{ to } \infty \\ H_1 : F^*(x) \neq F(x) \ \text{for at least one value of } x \end{cases}$$

Suppose that the Kolmogorov-Smirnov statistic, *D*, be the greatest vertical distance between the empirical distribution function, $S(x)$, and the theoretical distribution function $F^*(x)$, which is given in equation (31) below.

$$D = \sup_x \left| F^*(x) - S(x) \right| \tag{31}$$

Where the "*D*" equals the supremum, over all *x*, of the absolute value of the difference, $F^*(x) - S(x)$.

Thus, for testing the "stickiness" or the characteristic behaviors of the tacit versus codified forms of university knowledge, the one-sample Kolmogorov-Smirnov non-parametric tests (K-S tests) and its CDFs can be used to compare empirical distribution function and the normal distribution function across the different types of university knowledge dissemination channels.

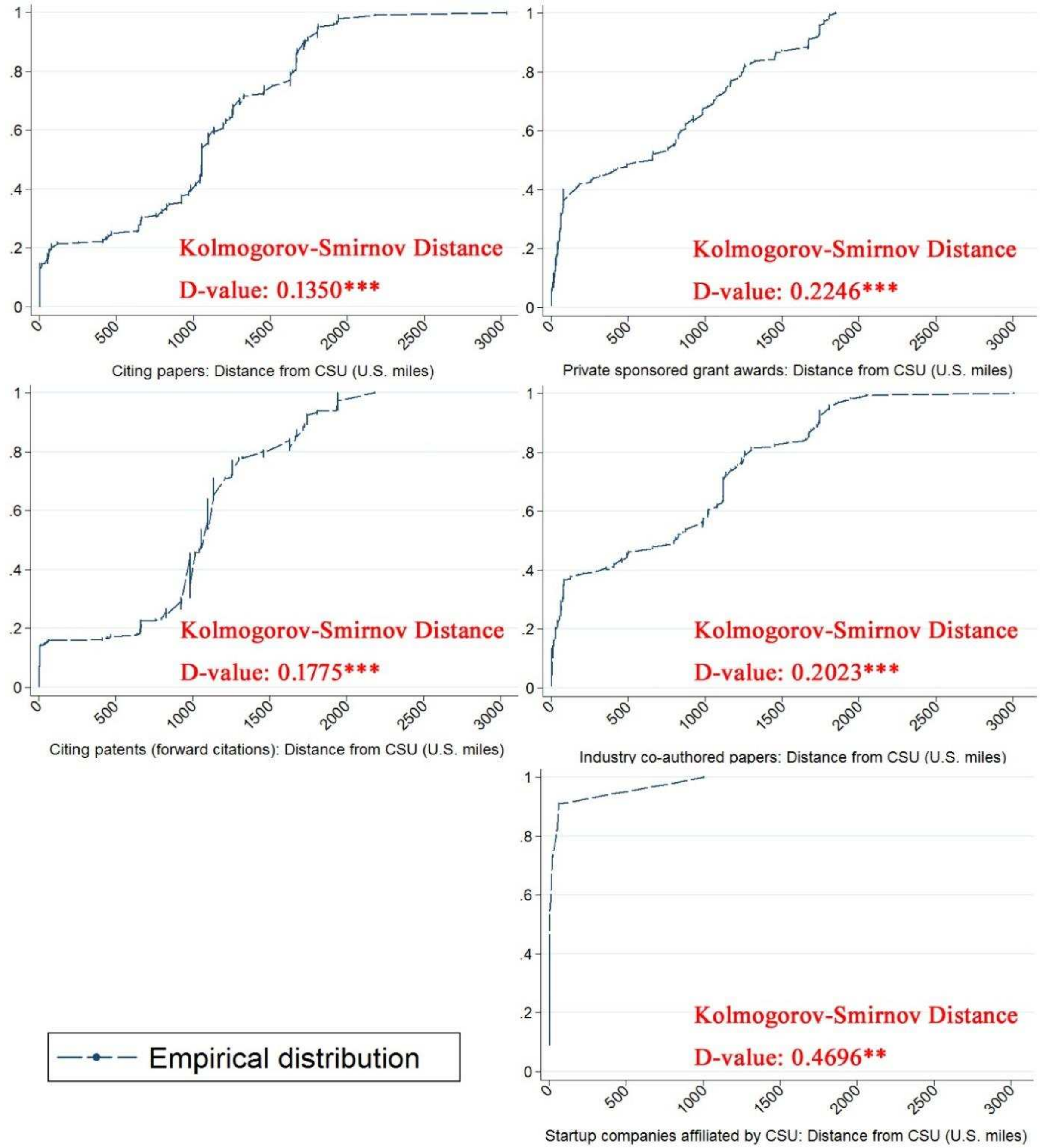
Table 34 shows summary statistics of the geographic distances from the location of CSU to the recipients of spillovers, across the different knowledge dissemination channels. The table considers only distances to those in the United States; it excludes foreign distances. As shown in previous sections, the mean distance between CSU and the location of authors on citing papers and the location of inventors of citing patents in the United States are the farthest and are similar to each other. Thus, these channels seem to involve more slippery forms of knowledge, the spillovers of which are less likely to be geographically bounded. Figure 49 displays the results of the K-S nonparametric tests, which consist of D-values and their statistical probability values, and the CDFs of the geographic distances from the location of CSU across the different knowledge dissemination channels in the United States.

Table 34—Summary statistics: Geographic distances from the location of CSU (U.S. only) across the different knowledge dissemination channels (All technology)

	Summary Statistics (U.S. miles)				
Knowledge Channels	**Obs.**	**Mean**	**Std. Dev**	**Min**	**Max**
Citing papers	2,989	988.79	637.99	5	3,037
Privately sponsored grant awards	148	662.90	623.33	5	1,852
Industry co-authored papers	167	749.92	676.79	5	3,010
Citing patents	1,546	1,010.27	550.24	5	2,183
Startups	11	107.45	298.67	5	1,006

The one-sample nonparametric test results show that the K-S tests reject the null hypothesis at a 1% level of statistical significance, which means the empirical distributions cannot converge to a normal distribution. In other words, at the given level of significance, $\alpha = 0.01$, the K-S statistical value exceeds the critical value of the quantiles of the K-S test statistic. However, it is still meaningful for testing the mechanisms of university knowledge dissemination channels by assessing the CDFs and K-S's D-value.

Figure 49, of the citing papers, has the smallest D-value of all, which means the distribution of citing papers is more likely to converge to a normal distribution than the others, even though it is insignificant. Thus, the channel is not likely to be localized and, rather, therefore likely involves a slippery form of knowledge. The CDF plot seems not to be skewed within 300 miles but rather to extend over several thousand miles. Similarly, the CDF plot of citing patents is less likely to be within Colorado, but it is skewed around 1,000 miles, with the maximum distance between CSU and a citing patents being 2,183 miles, which is shorter than the maximum for citing papers.



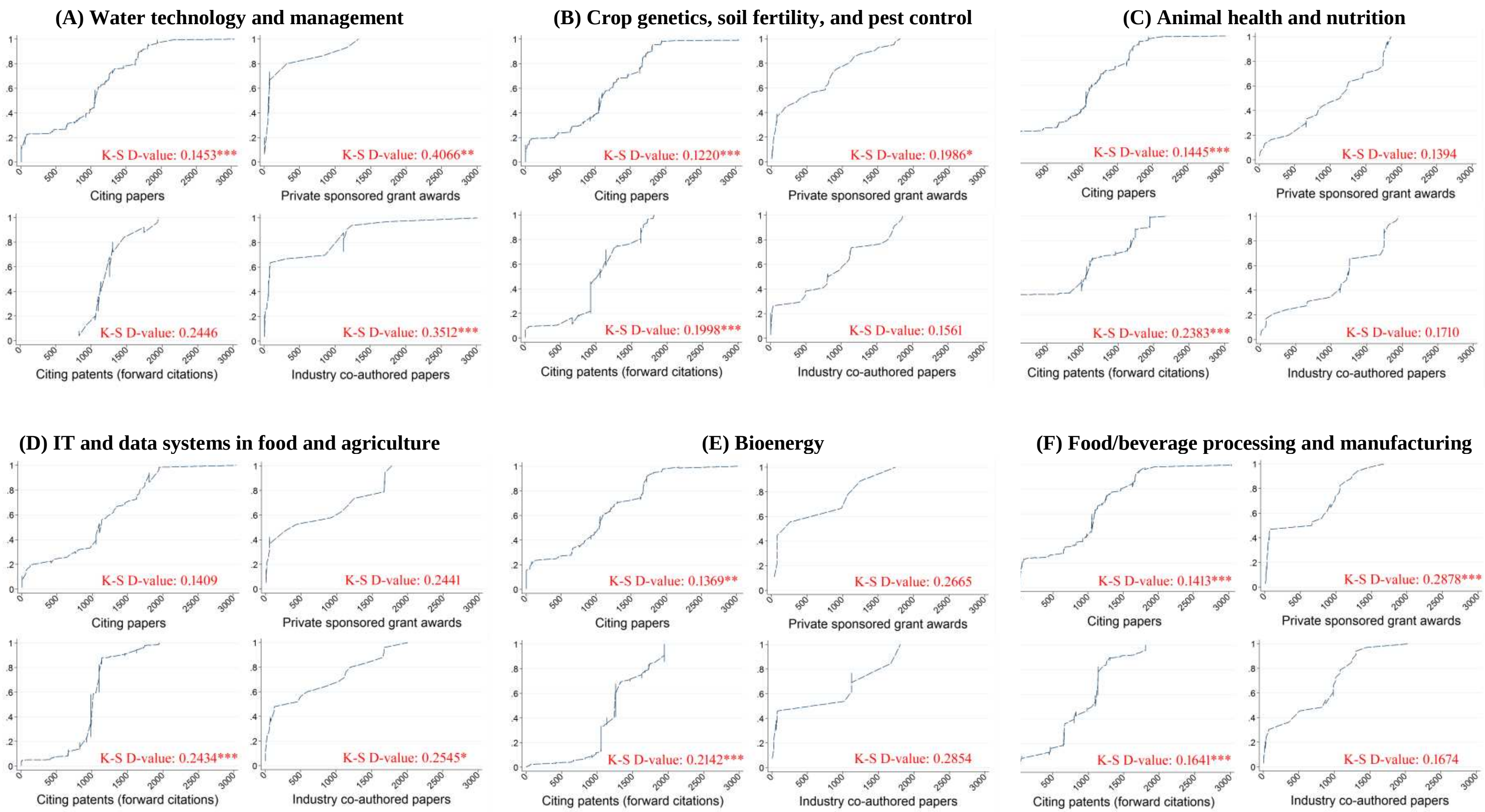
Note: *** at 1%, ** at 5%, and * 10% level of statistical significance.

Figure 49—The CDF plots and one-sample Kolmogorov-Smirnov non-parametric tests: Geographic distances from the location of CSU (U.S. only) across the different knowledge dissemination channels: All technology.

It is highly probable that the inherent characteristics of both knowledge dissemination channels are non-rivalrous, but the patents involve a certain degree of excludability by utilizing intellectual property rights and contracts, so the citing patents are likely to involve a less slippery form of knowledge than citing papers.

Of particular interest is the collaboration mechanism of knowledge dissemination. The privately sponsored grant awards and industry coauthored papers are relatively localized, with almost 40 percent of them within 200 or 300 miles. These channels apparently involve a more sticky form of knowledge, and geographic proximity is more important for these collaboration activities. Nevertheless, they have different scopes of geographic locations. In Table 34, the maximum distance between CSU and industry coauthored articles is much longer than the maximum distance for privately sponsored grant awards—3,010 miles and 1,852 miles, respectively—because again of the heterogeneous features between the channels. It should be pointed out that the inherent characteristics of industry coauthored journal articles involve both public domains and collaboration mechanisms of knowledge dissemination, so it may be less likely to involve sticky forms of knowledge. In addition, the D-value of industry coauthored journal articles is smaller than that of the privately sponsored grant awards.

Finally, the start-up companies affiliated with CSU appear to involve the stickiest form of knowledge, and their D-value is much larger than other channels at 0.4696. The start-up companies are highly localized, within 10 or 20 miles of the university, because most of the founders and employees are CSU faculty members, research staff, or GRAs, as well as local, private entities. In addition, in the early stage start-up companies need support from their original university (e.g., for the utilization of university facilities and equipment), which is an important part of the venture creation knowledge transfer mechanism.



Note: *** at 1%, ** at 5%, and * 10% level of statistical significance.

Figure 50—The CDF plots and one-sample Kolmogorov-Smirnov non-parametric tests: Geographic distances from the location of CSU (U.S. only) across the different knowledge dissemination channels: 6 different technologies

Again, according to Jaffe (1989), “the effect of geographically mediated spillovers from university to local commercial innovation comes through more clearly within technical areas than it does in the total across areas”, as well as (Anselin et al, 1997 and 2000). Figure 50 represents the results of the K-S tests and the CDFs of geographic distances from the location of CSU across the four different knowledge dissemination channels in the United States, congregated into six different technological categories in agriculture.

The one-sample nonparametric test results across the different technological categories are distinguished from the results of aggregate levels in Figure 49. First, the CDF patterns and D-values of citing papers seem not to change significantly across the different technologies, but in statistical result, the technological category of IT and data systems in food and agriculture is to converge to a normal distribution by accepting null hypothesis. Thus, CSU’s published journal articles as the public domain mechanism of knowledge, and its citing papers are more likely to have a slippery form of knowledge across the different technologies.

However, the stickiness of other types of knowledge dissemination channel can be varied across the different technological categories. Figure 50, in looking at the private sponsored grant awards, it is localized to within 200 or 300 miles, accounting for almost 70-40 percent, across the water, crop genetics, IT & Data systems, bioenergy, and food processing. By contrast, the animal health and nutrition technology seems to involve a slippery form of knowledge and it is localized to within 200 miles, accounting for less than 20 percent. Similarly, in the industry co-authored papers, water, IT & data systems, bioenergy, and food processing technologies are highly skewed within Colorado, accounting for 60-40 percent, but crop genetics, and animal health technologies do not be localized to within Colorado area.

In the geographic location of citing patents, excluding self-cites, the extent to which CSU's knowledge spillovers are geographically localized is considered across the different technological categories. As shown as Figure 50, the citing patents in technological categories of crop genetics, IT& data systems, bioenergy, and food processing can be seen as analogous to the results of aggregate levels in Figure 49, such as the long tail to the left, but not in water and animal health technologies. In water technology, the locations of inventor who cite CSU's patents are highly concentrated on between 800 and 1000 miles from the location of CSU. However, in animal health, it is highly localized to within Colorado area, accounting for 40 percent.

Therefore, the sticky versus slippery forms of knowledge dissemination channel are more likely to rely on the technological categories rather than aggregate level of technology. In other words, the evaluation of the benefits for university knowledge spillovers to commercial innovation should be considered not only across the different knowledge dissemination channels, but also across the different technological categories.

V. Conclusions

Despite debate about the geographic proximity of university-industry collaborations, university knowledge spillovers do appear to generate localized impacts on regional commercial innovation that are likely to beneficial both to private industry and the public. Particularly in agriculture, the land-grant universities play a significant role in agricultural research and innovation in the agribusiness sectors. However, rather than thinking that "one size fits all," the spillover benefits from university knowledge generation should be considered by the different types of dissemination channels that are utilized. In this chapter, we have focused on Colorado State University (CSU)'s knowledge spillovers within agriculturally related fields and technologies.

We have examined the various mechanisms of university knowledge spillovers and the geographic scope of impact on the agricultural economy associated with each. We find evidence that academic knowledge spillovers are geographically bounded, but they are not strictly limited to the regional scale. Crucially, the impact of university spillovers on agriculturally-related industries depends upon which type of knowledge dissemination channel or transfer mechanism is utilized by university researchers. Broadly speaking we evaluate four types of channels—including the public domain or publication mechanism, the industry collaboration and extension mechanism, the technology patenting/licensing mechanism, and the venture creation mechanism—each of which are variously adapted to transmitting different degrees of sticky (tacit) versus slippery (codified) knowledge.

Our findings show that in both aggregate level of technology and six different technological categories, the spillover impacts of journal publications, through the public domain mechanism of knowledge dissemination, are rarely localized within Colorado; rather, the geographic scope of these impacts are national and even global. Thus, geographic proximity is not a question with this channel. However, the extent to which the spillover impacts of patented knowledge is localized within Colorado is open to question because it is possible to control permissions for use, but at the same time it is impossible to limit everyone's awareness and use of it, particularly in foreign jurisdictions where patents are not taken out by the university. Therefore, the degree of localization of university knowledge spillovers when using the patent and licensing mechanism might depend on the different types of technology involved as well as the intellectual property and contract strategy pursued. Thus, university journal publications and patents are most appropriate for dissemination of more slippery forms of knowledge, but commercial innovation impacts can be localized nearer to the location of a university when using patents and licensing than when using journal publications alone.

However, the collaboration mechanism of knowledge dissemination, such as indicated by industry coauthorship on journal articles and private sponsorship of grants and contracts, which are more rivalrous by virtue of the more tacit qualities of knowledge being disseminated and because of the higher transaction costs, requires closer interaction and greater geographic proximity, which usually prevents global dissemination. Thus, we observe geographic proximity is significantly important for these channels. However, there are even distinctions within these. For example, we find industry coauthorship on articles to be less likely to be localized than privately sponsored grant awards. Nevertheless, the stickiness of these channels might depend also on the different technological categories. As mention as above, the geographic proximity is important only in aggregate level of technology, but it can be varied across the different technological categories, especially the slippery form of knowledge in animal health and nutrition health technology. Finally, university start-ups are highly geographically bounded near universities because in the early stages start-up companies need support from their host university.

One conclusion we can draw from this study is that both public and private benefits of university knowledge transfers rely on the different types of knowledge dissemination channels and the intrinsic characteristics of the different types of technology being disseminated. Moreover, our unique data set of Colorado State University's research activities and a full range of potential knowledge dissemination channels make it possible to examine the extent to which the different types of knowledge dissemination channels work within the context of agricultural fields and technologies. Most previous studies of university-industry collaboration activities have used aggregate data.

Despite these interesting preliminary findings, there are at least two major shortcomings in our approach. First, which we hope to address in further studies, we will attempt to build relevant regression models for measuring university knowledge spillovers, via mechanisms of knowledge dissemination for both sticky and slippery types of knowledge. Second, although both advantages and disadvantages exist for the use of single institutional data, it may compromise findings' generality relative to other institutions or to more aggregate economy-wide data because each institution has its own idiosyncratic conditioning characteristics, including levels of research expenditures, management skills, administrative policies, and so on. Thus, we hope to add more institutional data and compare the results.

OVERALL CONCLUSION

This study analyzes the economic impacts of university knowledge production and spillover activities, especially on innovations in agriculturally-related industries. The three chapters provide insights into the dynamics of university knowledge production and dissemination from three interrelated perspectives: (1) the university knowledge production function, (2) the organization and agency of research production within research teams, and (3) the geographical scope of economic impacts from university knowledge dissemination. The overall finding of this study shows how university research and its knowledge spillovers drive and create benefits to society and the economy, both publicly and privately. As the roles and missions of the university have changed, the composition of the university research system has also been transformed, affecting the trend of its impact on industrial innovation directly, and on economic growth indirectly. In agriculture, especially, the land grant universities play a significant role in supporting and stimulating commercial technological innovation within the state's agricultural value chain, and beyond.

To sum up, chapter 1 has analyzed Colorado State University's research production, disaggregated across the different colleges, departments, and research units, range of knowledge dissemination channels. This chapter introduced and utilized a uniquely detailed dataset together with a novel empirical technique for estimating three different knowledge production function models, including unrestricted and restricted negative binomial polynomial distributed lag (PDL) models and an effective labor log-log PDL model. The empirical test results indicated that the PDL models have advantages over the ad hoc lag scheme common in the literature, in terms of: (1) better goodness of fit according to two information criteria (AIC & SBIC), and (2) more meaningful explanations of unknown lag effects.

The effective labor PDL model provides some preliminary economic intuition regarding the "returns to scale" of knowledge production. The results showed that published journal articles and doctoral degree awards, outputs disseminated via the public domain channel, exhibit decreasing returns to scale, which is closely linked with the law of diminishing marginal productivity, whereas outputs disseminated via the other two channels, of collaboration and tech transfer, exhibit increasing returns to scale. The regression results also indicate that the mean gestation lag between research expenditures inception and completion of journal articles is 2.5 years, which is two or three years shorter than the gestation on collaborative outputs (4.8 years) and tech transfer outputs (4.7 years). In the SUR models, journal articles and tech transfer outputs have positive relationships with past research expenditures, whereas collaborative outputs have a negative relationship with past expenditures. Perhaps, researchers turn to collaboration activities when their research funds from public sources dry up. And, the collaboration index and the combined tech transfer metrics have an inverse relationship.

Chapter 2 has explored a new framework for understanding the agency of knowledge production within the university context. First, findings from an exploration of an ego-centric social network provide insights and intuition about research team participation dynamics over the course of the career of an individual university researcher. It finds that the share of co-authors from outside CSU contributes the size of ego-centric research teams, with the time trend between the share of outside team members and research team size both grow over the time. Conversely, the trends of the teams' percent share of co-authors from the ego's home department and team size have opposite patterns over time with downward and upward tendencies, respectively. Moreover, the downward trend of the percent share of team members from the ego's home department interpreted as "inverse team flexibility" tells us that the ego-centric research teams have been transformed from closer collegial or hieratical forms to broader network or project-based relationships.

Second, regression results from a representative sample of research teams shows that participation of authors from other CSU departments, other institutions, and private firms all contribute to the size of research teams. Interestingly, field variety seems to affect the size of research teams negatively, but it does affect the inclusion of non-CSU authors positively. R&D expenditures generally have a positive influence on team size. However, private R&D money has a small but negative relationship with team size, but it affects the inclusion of non-CSU authors positively. It is possible that private R&D investments and field variety affect the size of research team indirectly, rather than directly.

Finally, research quality is analyzed by comparing CDFs of various indicators. These indicate that both in teams generating journal articles and in teams generating patents, the involvement of team members from multiple CSU departments improves the impact of their research. There is no significant difference between teams that include members from just a single private firm and those that include members from multiple firms. Larger teams are more likely to have higher impacts than small teams in generating both articles and patents.

Chapter 3 has examined the geospatial pattern of Colorado State University (CSU) knowledge spillovers and their commercial economic impact, especially within Colorado's agriculturally-related sectors. The geographical scope of impacts and benefits realized are determined by the different types of knowledge dissemination channels suitable for different inherent characteristics of (sticky versus slippery) knowledge. Our findings show that the spillover impacts of journal publications are rarely localized within Colorado; rather, the geographic scope of these impacts is national and even global. For this slippery form of knowledge geographic proximity is unimportant.

However, the extent to which the spillover impacts of patented knowledge is localized within Colorado is open to question because it is possible to control permissions for use, but at the same time it is impossible to limit everyone's awareness and use of it, particularly in foreign jurisdictions where patents are not taken out by the university. By contrast the knowledge disseminated via the collaboration mechanism, such as indicated by industry coauthorship on journal articles and private sponsorship of grants and contracts, is more sticky; it is more rivalrous and excludable by virtue of the more tacit qualities of knowledge being disseminated. The location of private sector collaborators is closer to CSU, but can still be national in scope. Looking at the venture creation mechanism, also involving sticky knowledge, we see university start-ups are highly geographically bounded and located near the university, likely because early stage start-up companies need support from their faculty founder and the host institution. Because of the higher transaction costs, these channels require closer interaction and greater geographic proximity, which usually prevents global dissemination. Thus, the geographic proximity is significantly important for the collaboration and venture creation channels.

Despite these interesting preliminary findings, one major issue in this dissertation is its reliance upon a single institution's data and perspective. There are both advantages and disadvantages in the use of single institutional data. It may compromise findings' generality relative to other institutions or to more aggregate economy-wide data because each institution has its own idiosyncratic conditioning characteristics, including levels of research expenditures, management skills, administrative policies, and so on. Thus, we hope to add more institutional data and compare the results. Nevertheless, we expect the results of this study to be of value in future economic studies of university knowledge production and impact on commercial innovation, as well as of practical value to CSU administration as well as state and federal policymakers.

REFERENCES

Adams, J.D. "Comparative Localization of Academic and Industry Spillovers." *Journal of Economic Geography*, 2002, Vol. 2, pp. 253-278.

Adams, J.D., and J.R. Clemmons. "The role of search in university productivity: inside, outside, and interdisciplinary dimensions.", *Industrial and Corporate Change*, 2011, Vol. 20, No.1, pp. 215-251.

Adams, J.D., and Z. Griliches. "Research Productivity in a System of Universities." *Annales d'Economie et de Statistique*, 1998, Vol. 49/50, pp.127-162.

Adams, J.D., G.C. Black, J.R. Clemmons, and P.E Stephen. "Scientific Teams and Institutional Collaborations: Evidence from U.S. Universities, 1981–1999." *Research Policy*, 2005, Vol. 34, pp.259-285.

Agrawal, A. and R. Henderson. "Putting Patents in Context: Exploring Knowledge Transfer from MIT." *Management Science*, January 2002, Vol. 48, No.1, pp. 44-60.

Almon, S. "The Distributed Lag Between Capital Appropriations and Expenditures." *Econometrica*, 1965, Vol. 33, No. 1, pp. 178-196.

Ancona, D.G., and D.F. Caldwell. "Bridging the Boundary: External Activity and Performance in Organizational Teams." *Administrative Science Quarterly*, 1992, Vol. 37, No. 4, pp. 634-665.

Anselin, L., and A. Varga. "Entrepreneurship, Geographic Spillovers and University Research." *ESRC Centre for Business Research, University Cambridge Working Paper*, 1997, No. 59.

Anselin, L., A. Varga, and Z. Acs. "Geographical Spillovers and University Research: A Spatial Econometric Perspective" *Growth and Change*, 2000, Vol. 31, pp. 501-515.

Arellano, M., and B. Honore. "Panel Data Models: Some Recent Developments." *Handbook of Econometrics, Amsterdam, North Holland*, 2001, Vol. 5, pp. 3229-3296.

Arora, A., P.A. David, and A. Gambardella. "Reputation and Competence in Publicly Funded Science: Estimating the Effects on Research Group Productivity." *Annales d'Economie et de Statistique*, 1998, Vol. 49/50, pp.163-198.

Artz, K.W., P.M. Norman, D.E. Hatfield, and L.B. Cardinal." A Longitudinal Study of the Impact of R&D, Patents, and Product Innovation on Firm Performance." *Journal of Innovation Management*, 2010, Vol. 27, pp. 725-740.

Aube, C., V. Rousseau, and S. Trembley. "Team Size and Quality of Group Experience: The more the Merrier?" *Group Dynamics: Theory, Research, and Practice*, 2011, Vol. 15, No. 4, pp. 357-375.

Audretsch, D.B., and E.E. Lehmann. "Mansfield's Missing Link: The Impact of Knowledge Spillovers on Firm Growth." *Journal of Technology Transfer*, 2005, Vol. 30, pp. 207-210.

Barabasi, A.L., H. Jeong, Z. Neda, E. Ravasz, A. Schubert, and T. Vicsek. “Evolution of the Social Network of Scientific Collaborations.” *Physica A,* 2002, Vol. 311, pp.590-614.

Belenzon, S., and M. Schankerman. “Spreading the Word: Geography, Policy, and university Knowledge Diffusion.” LSE STICERD Research Paper, 2010, No. EI 50.

Bettencourt, L.M.A, J. Lobo, and D. Strumsky. "Invention in the city: Increasing returns to patenting as a scaling function of metropolitan size." *Research Policy,* 2007, Vol. 36, pp. 107-120.

Biancani, S., and D.A. McFarland. “Social Networks Research in Higher Education.” *Springer Science Press*, 2013, Higher Education: Handbook of Theory and Research, Ch.4, pp.151-215.

Biffl, S., and W. Gutjahr. “Influence of Team Size and Defect Detection Technique on Inspection Effectiveness.” *Software Metrics Symposium, METRICS, IEEE,* 2001, pp.63-75.

Borgatti, S.P., M.G. Everett, and J.C. Johnson. “Analyzing Social Networks”, *Sage Publication Press,* 2013.

Boschma, R.A. “Proximity and Innovation: A Critical Assessment.” *Regional Studies,* 2005, Vol. 39, No. 1, pp.61-74.

Buenstorf, G., and A. Schacht. "We need to talk – or do we? Geographic distance and the commercialization of technologies from public research." *Research Policy*, 2013, Vol. 42, pp. 465-480.

Bulut, H., and G. Moschini. “U.S. Universities’ Net Return from Patenting and Licensing: A Quantile Regression Analysis.” *Economics of Innovation and New Technology*, 2009, Vol.18, No.2, pp. 123-137.

Burt, R.S. “Structure Holes: The Social Structure of Competition.” *Harvard University Press,* 1995.

Bush, V. “Science, the Endless Frontier: A report to the President by Vannevar Bush, Director of the Office of Scientific Research and Development.” *University of Michigan Library,* 1945.

Chamberlain, G. “Analysis of Covariance with Qualitative Data.” *Review of Economic Studies*, 1980, Vol. 47, pp. 225-238.

Cohen, W.M. “Patents and Appropriation: Concerns and Evidence.” *Journal of Technology Transfer*, 2005, Vol. 30, pp. 57–71.

Cohen, W.W., R.R. Nelson and J.P. Walsh. “Links and Impacts: The Influence of Public Research on Industrial R&D.” *Management Science*, January 2002, Vol. 48, No.1, pp. 1-23.

Conover, W.J. "Practical Nonparametric Statistics." *Wiley,* 1999, 3rd edition.

Crane, D. “Social Structure in a Group of Scientists: A Test of the "Invisible College" Hypothesis.” *American Sociological Review*, 1969, Vol. 34, No. 3, pp. 335-352.

Crespi, G.A., and A. Geuna. “An Empirical Study of Scientific Production: A Cross Country Analysis, 1981-2002.” *Research Policy,* Vol. 37, 2008, pp.565-579.

Etzkowitz, H., A. Webster, C. Gebhardt, and B.R.C. Terra. “The Future of the University and the University of the Future: Evolution of Ivory Tower to Entrepreneurial Paradigm.” *Research Policy*, 2000, Vol. 29, pp. 313-330.

Etzkowitz, H., and L. Leydesdorff. “The Dynamics of Innovation: from National Systems and "Mode 2" to a Triple Helix of University–Industry–Government Relations.” *Research Policy*, 2000, Vol. 29, pp. 109-123.

Etzkowitz, H. “Research Groups as ‘Quasi-firms’: the Invention of the Entrepreneurial University.” *Research Policy*, 2003, Vol.32, pp.109-121.

Feller, I., and M. Feldman. “The Commercialization of Academic Patents: Black Boxes, Pipelines, and Rubik’s Cubes.” *Journal of Technology Transfer*, 2010, Vol. 35, pp. 597-616.

Graff, G.D., A. Heiman and D. Zilberman. “University Research and Offices of Technology Transfer.” *California Management Review*, Fall 2002, Vol. 45, No.1, pp. 88-115.

Graff, G.D., R. Mortenson, R. Goldbach, et al. “The Value Chain of Colorado Agriculture.” *Colorado State university*, February 2013.

Graff, G.D., A. Berklund, and K. Rennels. “The Emergence of an Innovation Cluster in the Agricultural Value Chain along Colorado's Front Range.", *Colorado State university*, November 2014, pp. 1-87.

Graff, G.D., G.C. Rausser, and A.A. Small. “Agricultural Biotechnology’s Complementary Intellectual Assets." *Review of Economics and Statistics*, May 2003, Vol.85, No.2, pp. 349-363.

Greene, W.H. “Econometric Analysis.” *Prentice Hall*, 2011, Seventh Edition.

Griliches, Z. “Issues in Assessing the Contribution of R&D to Productivity Growth”. *Bell Journal of Economics*, 1979, Vol.10, pp. 92-116.

---------------. “Productivity, R&D, and the Data Constraint.” *American Economic Review*, 1994, Vol.84, pp. 1-23.

Grossman, G.M., E. Helpman. “Innovation and Growth in the Global Economy.” *MIT Press*, 1991.

Guimera, R., B. Uzzi, J. Spiro, and L.A. Numes Amaral. “Team Assembly Mechanisms Determine Collaboration Network Structure and Team Performance.” *Science, New Series*, 2005, Vol. 308, No.5722, pp.697-702.

Gujarati, D.N. “Basic Econometrics.” *McGraw Hill*, 2009, Fifth edition.

Hall, B.H. “Exploring the Patent Explosion.” *Journal of Technology Transfer*, 2005, Vol. 30, pp. 35–48.

Harhoff, D., F. Narin, F.M. Scherer., and K. Vopel. “Citation Frequency and the Value of Patented Inventions.” *Review of Economics and Statistics*, 1999, Vol. 81, No. 3, pp. 511-515.

Hausman, J.B., H. Hall, and Z. Griliches. “Econometric Models for Count Data and with Application to the Patents-R&D Relationship.” *Econometrica*, 1984, Vol.52, No.4, pp.909-938.

Heckman, J. “Statistical Models for Discrete Panel Data. Structural Analysis of Discrete Data with Econometric Applications.” MIT Press: Cambridge, 1981.

Henderson, R., A.B. Jaffe, and M. Trajtenberg. “Universities as a Source of Commercial technology: A Detailed Analysis of University Patenting, 1965-1988.” *Review of Economics and Statistics*, February 1998, Vol. 80, No. 1, pp. 119-127.

Henderson, R., and I. Cockburn. "Scale, Scope, and Spillovers: The Determinants of Research Productivity in drug discovery.", *RAND Journal of Economics*, 1996, Vol. 27, No. 1, pp. 32-59.

Heher, A.D. “Benchmarking of Technology Transfer Offices and What it means for Developing Countries.” *Intellectual property management in health and agricultural innovation: a handbook of best practices*, 2007, Vol. 1 and 2, pp. 207-228.

Herrigel, G.B. “Power and the Redefinition of Industrial Districts: the Case of Baden-Württemberg.” *The Embedded Firm: On the Socioeconomics of Industrial Networks. London: Routledge*, 1993, pp.227-251.

Hong, W., and Y. Su. “The Effect of Institutional Proximity in Non-local University–Industry Collaborations: An Analysis Based on Chinese Patent Data.” *Research Policy*, 2013, Vol. 42, pp. 454-464.

Huffman, W.E., and R.E. Just. “Setting Efficient Incentives for Agricultural Research: Lessons from Principal-Agent Theory.” *American Journal of Agricultural Economics*, 2000, Vol. 82, No. 4, pp.828-841.

Jackson, M.O. “Social and Economic Networks.” *Princeton University Press*, 2008.

Jaffe, A.B. “Real Effects of Academic Research.” *American Economic Review*, December 1989, Vol. 79, No. 5, pp. 957-970.

Jaffe, A.B., M. Trajtenberg, and R. Henderson. “Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations.” *Quarterly Journal of Economics*, 1993, Vol. 108, No. 3, pp. 577-598.

Jaffe, A.B., and M. Trajtenberg. “Flows of Knowledge from Universities and Federal Laboratories: Modeling the Flow of Patent Citations over Time and across Institutional and Geographic Boundaries.” *National Academy of Sciences*, 1996, Vol. 93, No. 23, pp. 12671-12677.

Jensen, R., and M.C. Thursby.. “Proofs and Prototypes for Sale: The Licensing of the University Inventions.” *American Economic Review*, March 2001, Vol. 91, No. 1, pp. 240-259.

Johnes, J. and G. Johnes. “Research Funding and Performance in UK University Departments of Economics: a Frontier Analysis.” *Economics of Education Review*, Vol. 14, pp. 301-314.

Jones, B.F., S. Wuchty, and B. Uzzi. "Multi-University Research Teams: Shifting Impact, Geography, and Stratification in Science." *Science*, 2008, Vol. 322, No. 5905, pp. 1259-1262.

Katz, J.S., and B.R. Martin. "What is Research Collaboration?" *Research Policy*, 1997, Vol. 26, pp.1-18.

Kolmogorov, A.N. "Sulla Determinazione Empirica di Una Legge di Distribuzione." *Giornale dell' Istituto Italiano degli Attuar*, 1933, Vol. 4, pp. 83-91

Krugman, P., and R. Wells. "Microeconomics." *Worth Publishers*, 3rd edition.

Laband, D.N., and R.D. Tollison, "Intellectual Collaboration." *Journal of Political Economy*, 2000, Vol. 108, No. 3, pp. 632-662.

Latour, B. "We have Never been Modern." *Harvard University Press*, 1993.

-------------. "Aramis or the Love of Technology." *Harvard University Press*, 1996.

Laursen, L. and A. Salter. "Searching High and Low: What Types of Firms Use Universities as a Source of Innovation?" *Research Policy*, 2003, Vol. 33, pp. 1201-1215.

Lazega, E., M. Jourda, L. Mounier, and R. Stofer. "Catching up with Big Fish in the Big Pond? Multi-level Network Analysis through Liked Design." *Social Network*, 2008, Vol. 30, pp. 159-176.

Lester, R.K. "Universities, Innovation, and the Competitiveness of Local Economics" *MIT Industrial Performance Center Working Paper*, 2005, No. 05-010.

Lee, Y., J.P. Walsh, and J. Wang. "Creativity in Scientific Teams: Unpacking Novelty and Impact." *Research Policy*, 2015, Vol. 44, pp.684-697.

Lee, Y.S. "Technology Transfer and the Research University: a Search for the Boundaries of University—Industry Collaboration." *Research Policy*, 1996, Vol. 25, pp. 843-863.

Liu, S. "Spillovers from Universities: Evidence from the Land-Grant Program." *Journal of Urban Economics*, 2015, Vol. 87, pp. 25-41.

Lissoni, F., P. Lierena, M. McKelvey, and B. Sanditov. "Academic Patenting in Europe: New Evidence from the KEINS Database." *Research Evaluation*, 2008, Vol. 17, No. 2, pp. 1-16.

Mansfield, E. "Academic Research and Industrial Innovation." *Research Policy*, 1991, Vol. 20, pp. 1-12.

---------------. "Academic Research Underlying Industrial Innovations: Sources, Characteristics, and Financing, " *Review of Economics and Statistics*, February 1995, Vol. 77, No.1, pp. 55-65.

Min, I, and P. Choi. "STATA Advanced Panel Data Analysis." *Jiphil Media*, 2012, First Edition.

Min, I, and P. Choi. "STATA Time Series Data Analysis." *Jiphil Media*, 2014, First Edition.

Moody, J. "The Structure of a Social Science Collaboration Network: Disciplinary Cohesion from 1963 to 1999." *American Sociological Review*, 2004, Vol. 69, No. 2, pp. 213-238.

Mowery, D.C., and B.N. Sampat. “University Patents and Patent Policy Debates in the USA, 1925-1980.” *Industrial and Corporate Change*, 2001, Vol. 10, No. 3, pp. 781-814.

--. “The Bayh-Dole Act of 1980 and University–Industry Technology Transfer: A Model for Other OECD Governments?”, *Journal of Technology Transfer*, 2005, Vol. 30, pp. 115–127.

Narin, F., K. Hamilton, and D. Olivastro. “The increasing Linkage between US Technology and Public Science.” *Research Policy*, 1997, Vol. 26, pp. 317-330.

Newman, M.E.J. “Clustering and Preferential Attachment in Growing Networks.” *Physical Review E*, 2001, Vol. 64, Issue. 2, pp.1-13.

--------------------. “Coauthorship Networks and Patterns of Scientific Collaboration.” *PNAS*, 2004, Vol. 101, Suppl. 2, pp.5200-5205.

O’Shea, R.P., T.J. Allen, A. Chevalier, and F. Roche. “Entrepreneurial orientation, technology transfer and spinoff performance of U.S. universities.”, *Research Policy*, 2005, Vol. 34, pp. 994-1009.

Pagano, M., and M. J. Hartley. “On Fitting Distributed Lag Models Subject to Polynomial Restrictions,” *Journal of Econometrics*, 1981, Vol. 16, pp 171-198.

Pakes, A., and Z. Griliches. “Patents and R&D at the Firm Level: A First Report.” *Economics Letter*, 1980, Vol.5, pp. 377-381.

----------------------------------.”Patents and R&D at the Firm Level: A First Look.” NBER Book Title: R & D, Patents, and Productivity, *University of Chicago Press*, 1984, Ch. 3, pp. 55-72.

Pardey, P.G. “The Agricultural Knowledge Production Function: An Empirical Look.” *Review of Economics and Statistics*, 1989, Vol. 71, No. 3, pp.453-461.

Parker, G.M. “Cross-functional Teams: Working with Allies, Enemies, and other Strangers.” *Wiley press*, 1992.

Payne, A., and A. Siow. "Does Federal Research Funding Increase University Research Output?" *Advances in Economic Analysis and Policy*, 2003, Vol. 3, Iss 1, Art. 1.

Paytas, J., R. Gradeck, and L. Andrews. “Universities and the Development of Industry Clusters.” *Carnegie Mellon Center for Economic Development Administration, U.S. Department of Commerce*, 2004.

Ponds, R., F. van Oort, and K. Frenken. “Innovation, Spillovers and University-Industry Collaboration: an Extended Knowledge Production Function Approach.” *Journal of Economic Geography*, 2010, Vol. 10, pp. 231-255.

Powell, W.W. “Neither Market nor Hierarchy: Network Forms of Organization.” *Research in Organizational Behavior*, 1990, Vol, 12, pp.295-336, ISBN: 1-55938-029-2.

Rasmussen, E., S. Mosey, and M. Wright. “The Influence of University Departments on the Evolution of Entrepreneurial Competencies in Spin-off Ventures.” *Research Policy*, 2015, Vol. 43, pp. 92-106.

Rasmussen, E., O. Moen, and M. Gulbrandsen. “Initiatives to promote commercialization of university knowledge.” *Technovation*, 2006, Vol. 26, pp. 518-533.

Rodriguez, M.A., and A. Pepe. “On the relationship between the structural and socio-academic communities of a coauthorship network.” *Journal of Informetrics*, 2008, Vol. 2, pp. 195-201.

Romer, P.M. “Increasing Returns and Long-Run Growth.”, *Journal of Political Economy*, 1986, Vol. 94, No. 5, pp. 1002-1037.

----------------. “Endogenous Technological Change.” *Journal of Political Economy*, 1990, Vol. 98, No. 5, pp. S71-S102.

Ruef, M. “A Structural Event Approach to the Analysis of Group Composition.” *Social Networks*, 2002, Vol. 24, pp. 135-160.

Samuelson, P.A. “The Pure Theory of Public Expenditure.” *Review of Economics and Statistics*, 1954, Vol.36, No.4, pp.387-389.

Scherer, F.M., and D. Harhoff. “Technology policy for a world of skew-distributed outcomes.” *Research Policy*, 2000, Vol. 29, pp. 599-566.

Schumpeter, J.A. “Capitalism, Socialism, and Democracy.” *Harper and Brothers Press*, 1942.

Scott, J. “Social Network Analysis.” *Sage Publication Press*, 2012, 3rd edition.

Shane, S. "Selling University Technology: Patterns from MIT." *Management Science*, 2002, Vol. 48, No. 1, pp. 122-137.

Smangs, M. “The Nature of the Business Group: A Social Network Perspective.” *Organization*, 2006, Vol. 13, No. 6, pp. 889-909.

Smirnov, N.V. “Estimate of Deviation between Empirical Distribution Functions in Two Independent Samples.” *Bulletin Moscow University*, 1933, Vo. 2, pp.3-16.

Solow, R.M. “A Contribution to the Theory of Economic Growth.”, *Quarterly Journal of Economics*, 1956, Vol. 70, No. 1, pp. 65-94.

Spender, J.C. “Making Knowledge the Basis of a Dynamic Theory of the Firm.” *Strategic Management Journal*, 1996, Vol. 17, pp.45-62.

Stokols, D., K.L. Hall, B.K. Taylor, and R.P. Moser. “The Science of Team Science: Overview of the Field and Introduction to the Supplement.” *American Journal of Preventive Medicine*, 2008, Vol. 35, No. 2S, pp. S77-S89.

Sunding, D., and D. Zilberman. “The Agricultural Innovation Process: Research and Technology Adoption in a Changing Agricultural Sector.” *Handbook of Agricultural Economics*, 2001, Vol. 1, Chapter 4, pp. 207-261.

Taramasco, C., J. Cointet, and Camille Roth. “Academic Team Formation as Evolving Hypergraphs.” *Scientometrics*, 2010, Vol. 85, pp.721-740.

Thursby, J.G., A.W.Fuller, and M.C. Thursby. "US faculty patenting: Inside and outside the university." *Research Policy*, 2009, Vol.38, pp. 14-25.

Thursby, J.G. and M.C. Thursby. “University-Industry Linkages in Nanotechnology and Biotechnology: Evidence on Collaborative Patterns for New Methods of Inventing.” *Journal of Technology Transfer*, 2011, Vol. 36, pp. 605-623.

---. “University Licensing and the Bayh-Dole Act.”, *Science*, 2003, Vol. 301, pp. 1052.

Tohidi, H., and M.J. Tarokh. “Productivity Outcomes of Teamwork as an Effect of Information Technology and Team Size.” *International Journal of Production Economics*, 2006. Vol. 103, pp.610-615.

Uzzi, B., S. Mukherjee, M. Stringer, and B. Jones. “Atypical Combinations and Scientific Impact.” Science, 2013, Vol. 342, No. 25, pp.468-472.

Uzzi, B., J. Spiro. “Collaboration and Creativity: The Small World Problem.” *American Journal of Sociology*, 2005, Vol. 111, No. 2, pp. 447-504.

Wagner, C.S., and L. Leydesdorff. “Network Structure, Self-organization, and the Growth of International Collaboration in Science.” *Research Policy*, 2005, Vol. 34, pp. 1608-1618.

Whalley, A, and J. Hicks. “Spending Wisely? How Resources Affect Knowledge Production in Universities.” Economic Inquiry, 2014, Vol. 52, No. 1, pp. 35-55.

Wooldridge, J.M. “Econometric Analysis of Cross Section and Panel Data.” *MIT Press*, 2001, First Edition.

Wuchty, S, B.F. Jones, and B. Uzzi. “The Increasing Dominance of Teams in Production of Knowledge.” *Science*, May 2007, Vol. 316, pp.1036-1039.

Zhu, M., Y. Huang, and N.S. Contractor. “Motivations for Self-assembling into Project Teams.” *Social Networks*, 2013, Vol. 35, pp.251-264.

APPENDIX

Table A-1—Major companies active in water technology, infrastructure, analytics, and management innovation

Company	Country	State	Privately sponsored grant awards	WoS articles	Citing Patents
3DGeo Dev Inc	USA	CA	$0	1	0
AECOM	USA	CO	$10,504	3	0
AGRO ENGN INC	USA	CO	$0	1	0
Analyt Technol Inc	USA	CO	$0	2	0
Anderson Associates LLC	USA	CO	$18,265	1	0
Andritz Inc.	USA	PA	$0	0	1
Aqua Engn Inc	USA	CO	$0	1	0
AQUA TERRA CONSULTANTS	USA	CA	$0	1	0
Balance Hydrol Inc	USA	CA	$0	1	0
Bay Bioscience Kabushiki Kaisha	Japan	Foreign	$0	0	1
Bioconversion Technologies Limited	USA	GA	$0	0	3
BIO-Logic Environmental	USA	CO	$2,856	0	0
Black Veatch	USA	CO	$0	1	0
Brown & Gay Engineers Inc	USA	TX	$0	1	0
Brown and Caldwell	USA	CO	$40,000	0	0
Camp Dresser McKee	USA	CO	$11,276	1	0
Cbec Ecoengn	USA	CA	$0	1	0
CDM Smith	USA	CO	$0	3	0
Centre National De La Recherche Scientifique	Frence	Foreign	$0	0	1
CH2M Hill	USA	CO	$83,351	13	0
Chingoo Research Partnership	USA	CA	$0	0	2
Cleareso Llc	USA	FL	$0	0	2
Clearwater Solut	USA	CO	$0	1	0
Coastal Waters Biotechnology Group Llc	USA	CA	$0	0	2
Elk River Management Corporation	USA	AL	$1,463	0	0
Endress & Hauser	Switzerland	Foreign	$0	0	2
FRICO-Farmers Reservoir and Irrigation	USA	CO	$57,689	0	0
GeoTrans Inc	USA	CO	$114,549	1	0
Geowatersheds Sci	USA	AK	$0	1	0
Groundwater Technol Team Chevron Energy Technol CO.	USA	CA	$0	1	0
GST Growth LLC	USA	DE	$0	0	1
Hach Co	USA	CO	$0	1	0
Halliburton Company	USA	TX	$53,488	0	0

Heinz Ploechinger	UK	Foreign	$0	0	2
Hemisphere Gps Llc	Canada	Foreign	$0	0	1
HydroQual Inc	USA	NJ	$0	2	0
Hydrosphere Resource Consultants	USA	CO	$0	1	0
IER Niono	Mali	Foreign	$0	1	0
Institute Of Geological & Nuclear Sciences Ltd.	New Zealand	Foreign	$0	0	1
Invitrogen Corporation	USA	CA	$0	0	2
KOWACO-Korean Water Resources Corp.	S. Korea	Foreign	$473,737	0	0
Life Technologies Corporation	USA	CA	$0	0	2
Lumiere Diagnostics Inc	USA	CO	$0	0	0
McLaughlin Water Engineering LTD	USA	CO	$360,400	1	0
MWH Amer Inc	USA	CO	$37,895	2	0
NW Hydraul Consultants	USA	CA	$0	1	0
Plato Industries Ltd	USA	TX	$0	0	2
Porzak Browning & Bushong LLP	USA	CO	$0	1	0
Regenesis Manageinent Grp	USA	CO	$46,140	1	0
Riken	Japan	Foreign	$0	0	2
Soldier Canyon Filter Plant	USA	CO	$5,000	0	0
Stanley Consultants Inc.	USA	IA	$0	0	4
Streamline Automation Llc	USA	AL	$0	0	3
Stuart Products Inc.	USA	TX	$7,064	0	0
Swiss Fed Inst Technol Inst Hydromech & Water Resources Management	Switzerland	Foreign	$0	1	0
Tech Serv Ctr Bur Reclamat Sedimentat & River Hydraul Grp	USA	CO	$0	1	0
Tetra Tech Surface Water Grp	USA	CO	$0	1	0
TransAlgae Ltd	Israel	Foreign	$0	0	1
TRC Hydro Geo Consultants	USA	CO	$0	1	0
Vision For You Llc	USA	NJ	$0	0	1
WATER ENGN & TECH INC	USA	CO	$0	1	0
Western Water Consultants	USA	WY	$0	1	0
Wright Water Engn Inc	USA	CO	$0	1	0
Yellow Springs Instrument Co Inc	USA	OH	$0	1	0

Table A-2—Major companies active in soil fertility or pest control innovation

Company	Country	State	Privately sponsored grant awards	WoS articles	Citing Patents
ACM-Texas LLC	USA	CO	$102,084	0	0
Agriliance LLC	USA	CO	$0	1	0
American Cyanamid Company	USA	NJ	$101,500	0	0
Amino Chem Co Ltd	Japan	Foreign	$0	1	0
Applied Chemical Magnesias Corp.	USA	CO	$16,412	0	0
BASF Corporation	Germany	Foreign	$102,000	0	6
Brotica Inc	USA	CO	$0	1	0
Cargill Inc	USA	CO	$0	1	0
Chisso Corporation	Japan	Foreign	$18,647	0	0
CIBA-GEIGY Chemical Company	Switzerland	Foreign	$120,000	0	0
Decagon Devices Inc	USA	WA	$0	1	0
Dow Chemical Company	USA	MI	$10,000	0	0
E.I. Dupont Company	USA	DE	$88,190	0	0
EMBRAPA Pecuaria Sul	Brazil	Foreign	$0	1	0
Fluid Fertilizer	USA	KS	$4,250	0	0
Fmc Corporation	USA	PA	$0	0	3
Genetics & Ivf Institute	USA	VA	$0	0	4
Hauser Company	USA	CO	$22,000	0	0
Ice Nine Environm Consulting	USA	VT	$0	1	0
Isca Technologies	USA	CA	$0	0	2
Isoprime Ltd Cheadle Hulme	UK	Foreign	$0	1	0
Laboratoires Expanscience	France	Foreign	$0	0	3
Lenat Consulting Serv	USA	NC	$0	3	0
LGI Ellsworth	USA	IA	$0	1	0
Nutrinsic Corporation	USA	CO	$0	0	1
Proviron Holdings	Belgium	Foreign	$0	0	2
Roche Colorado Corp	USA	CO	$0	2	0
Roses Inc.	USA	OK	$31,515	0	0
Sekisui Chemical Company Ltd.	Japan	Foreign	$120,000	0	0
Seroctin Research and Technology Inc	USA	UT	$0	0	1
Showa Denko K.K.	Japan	Foreign	$37,760	0	0
Small Businesses Sustainabil Coordinat Acquisit	USA	VA	$0	1	0
Soil Enhancement Technologies	USA	CO	$35,235	0	0
Syngenta Crop Protect Inc	USA	FL	$0	1	2
Taensa Inc.	USA	CT	$12,180	0	0
TERRAGEN DIVERS INC	Canada	Foreign	$0	1	0
Thin Air Nitrogen Solutions	USA	CO	$0	0	0
TIGR	USA	MD	$0	1	0

U.S. Borax Inc.	USA	CO	$12,000	0	0
UGG Ltd	Canada	Foreign	$0	1	0
United Phosphorous Inc	USA	CA	$0	1	0
Valent BioSciences Corporation	USA	IA	$17,211	0	0
Western Mobile Northern Inc	USA	CO	$5,750	0	0
Whitmire Micro-Gen Research Laboratories Inc.	USA	MO	$0	0	3
Yulex Corp	USA	AZ	$39,854	1	0

Table A-3—Major companies active in animal health and bio-medical science innovation

Company	Country	State	Privately sponsored grant awards	WoS articles	Citing Patents
Advanced Integrative Medicine	USA	CO	$0	1	0
Advanced Tissue Sciences Inc	USA	CA	$0	0	1
Agropecuaris CFM Ltd	Brazil	Foreign	$0	2	0
Agroscope Changins Wadenswil ACW	Switzerland	Foreign	$0	1	0
Alltech, Inc.	USA	KY	$204,524	0	0
Alpharma Animal Health	USA	CO	$200,253	1	0
American Network of Lipolysis	USA	CO	$20,736	0	0
Antel Biosyst Inc	USA	MI	$0	2	0
Anthrogenesis Corporation	USA	NJ	$0	0	15
Aperion Biologics, Inc.	USA	TX	$0	0	2
Applera Corporation	USA	CT	$0	0	7
Applied Biosystems Inc.	USA	MA	$0	0	10
Balchem Corporation	USA	NY	$42,154	0	0
Bayer Materialscience Llc	USA	KS	$0	1	3
Beckman Coulter, Inc.	USA	CO	$0	0	2
Becton Dickinson And Company	USA	NJ	$0	0	6
Biassex Pty Ltd	Australia	Foreign	$0	0	1
Boehringer Ingelheim Animal Health, Inc.	Germany	Foreign	$13,970	0	0
Boston Scientific Scimed, Inc.	USA	MA	$0	1	5
Bovine Reproduction Specialists	USA	CO	$0	1	0
Brigham And Women's Hospital	USA	MA	$0	0	1
Btg International Limited	UK	Foreign	$0	0	20
Cancer League of Colorado, Inc.	USA	CO	$59,780	0	0
Cedus, Inc.	USA	CO	$0	0	0
Cellx Inc.	USA	MD	$0	0	1
CENT Vet Insitute	Netherlands	Foreign	$0	1	0
Centaur Equine Products, Inc.	USA	PA	$4,900	0	0
Codexis, Inc.	USA	CA	$0	0	8

Coley Pharmaceutical Gmbh	Germany	Foreign	$0	0	2
Colgate Palmolive Company	USA	NY	$0	0	1
Corlife GbR	Germany	Foreign	$0	0	3
CSL Limited	Australia	Foreign	$0	0	2
Ctr Med Agr & Vet Entomol,	USA	FL	$0	1	0
Dakocytomation Colorado, Inc.	USA	CO	$0	0	6
Dfb Pharmaceuticals, Inc.	USA	TX	$0	0	11
Doskosil Co Inc	USA	KS	$0	1	0
Dr Joyce Donkersgoed Vet Serv Inc	Canada	Foreign	$0	1	0
Dymatize Enterprises, Llc	USA	TX	$0	0	2
Elanco Animal Health	USA	IN	$652,259	14	0
Embl Heidelberg	Germany	Foreign	$0	0	3
Encelle, Inc.	USA	NC	$0	0	4
Energy Enzymes, Inc.	USA	KS	$0	0	1
FMC Corporation	USA	PA	$57,848	0	0
Fort Dodge Animal Health (old Cyanamid)	USA	KS	$233,275	0	0
Fraunhofer-Gesellschaft Zur Forderung Der Angewandten Forschung E.V.	Germany	Foreign	$0	0	1
Gardens Alive	USA	IN	$42,956	0	0
Geissler Technologies, Llc	USA	MN	$0	0	1
Gerigene Medical Corp	USA	WI	$0	0	4
Global Animal Products, Inc.	USA	TX	$35,878	0	0
Hancock Jaffe Laboratories	USA	CA	$0	0	4
Heska Corp	USA	CO	$1,012,096	20	0
Hill Top Pharmatest Inc	USA	OH	$0	1	0
Hms Veterinary Development Inc	USA	CA	$0	1	0
Hoechst-Roussel	USA	MO	$86,249	0	0
Hoffmann La Roche Inc	USA	NJ	$0	1	0
Icyt Mission Technology, Inc.	USA	IL	$0	0	1
IDEXX Laboratories, Inc.	USA	CO	$104,100	0	0
Inguran, Llc	USA	TX	$0	0	198
Intervet Inc	USA	DE	$102,942	3	0
IVY Laboratories, Inc.	USA	KS	$76,189	0	0
J. Craig Venter Institute	USA	CA	$0	0	2
Juvaris Biotherapeutics, Inc.	USA	CA	$0	0	2
LifeCell Corp,	USA	NJ	$0	1	0
Lifenet Health	USA	VA	$0	0	8
Lilly Research Laboratories	USA	IN	$36,250	0	0
Masterrind Gmbh	Germany	Foreign	$0	0	7
McMahon Bioconsulting & Stockhausen Chem	Germany	Foreign	$6,500	0	0
Medarex, Inc.	USA	NJ	$0	0	1
Merial, Ltd.	USA	GA	$11,000	0	0

Moormans	USA	IL	$28,500	0	0
Mwi Veterinary Supply Co.	USA	CO	$0	0	7
Novartis Nutrition Ag	Switzerland	Foreign	$0	0	2
Novus International, Inc.	USA	MO	$87,594	0	0
NPC-Nutrition Physiology Company	USA	OK	$61,194	0	0
Nusirt Sciences, Inc.	USA	TN	$0	0	1
Nutri-Turf, Inc.	USA	FL	$12,703	0	0
Optibrand	USA	CO	$29,237	0	0
Organogenesis Inc.	USA	MA	$0	0	7
Peptide Technology Ltd	Australia	Foreign	$0	1	0
Perkin Elmer Corp, Appl Biosyst Div,	USA	CA	$0	1	0
Pfizer Animal Health	USA	MO	$1,197,386	2	0
Pharmacia & Upjohn Anim Hlth	USA	MI	$0	1	0
PIC N Amer	USA	TN	$0	2	0
Prairie Aquatech	USA	SD	$0	0	2
Qlt Phototherapeutics, Inc.	USA	CO	$0	0	4
Quali Tech, Inc.	USA	MN	$48,808	0	0
Roche Vitamins Inc	USA	NJ	$451,209	1	0
Sakura Properties, Llc	USA	UT	$0	0	2
Schering-Plough Corporation	USA	NJ	$22,999	1	0
Select Breeders Serv Inc	USA	MD	$0	2	0
Silliker, Inc	USA	IL	$15,000	0	0
SmithKline Beecham Animal Health	USA	FL	$75,747	0	0
Source Integration, Inc.	USA	CA	$0	0	1
Swing Aerobics Licensing, Inc.	USA	TX	$0	0	3
Syntex	Mexico	Foreign	$43,296	0	0
Tei Biosciences, Inc.	USA	MA	$0	0	1
Tissue Engineering, Inc.	USA	NY	$0	0	9
UCD Sch Agr Food Sci & Vet Med	Ireland	Foreign	$0	1	0
Upjohn Company	USA	MI	$14,300	0	0
Vet Adm Med Ctr	USA	VT	$0	1	0
Vet Canc Specialists	USA	CO	$0	2	0
Vet Infect Dis Org	Canada	Foreign	$0	1	0
Vet Labs Agcy Lasswade	Scotland	Foreign	$0	1	0
VET LIFE	USA	NC	$13,583	0	0
Wolfgang Jochle Associates Inc	USA	NJ	$0	1	0
XY Inc	USA	CO	$0	5	298
Zapata Haynie Corporation	USA	MS	$13,600	0	0
Zero Discharge Pty Ltd.	Australia	Foreign	$0	0	1
Zymogenet Inc	USA	WA	$0	1	0

Table A-4—Major companies active in agricultural information system and business management innovation

Company	Country	State	Privately sponsored grant awards	WoS articles	Citing Patents
Advantage Int,	USA	WI	$0	1	0
AEI Econ Consultants	USA	CO	$0	1	0
AESOP Enterprises Ltd	USA	DC	$0	1	0
American Express Travel Related Services Company, Inc.	USA	NY	$0	0	14
Appl Geosolut LLC	USA	NH	$0	1	0
Ascendant Partners Inc	USA	CO	$0	1	0
At&T Intellectual Property I, L.P	USA	TX	$0	0	5
Battelle Memorial Institute	USA	OH	$0	0	1
Bbi International, Inc.	USA	ND	$0	0	2
Bioprocessh20 Llc	USA	RI	$0	0	1
Broadcom Corporation	USA	CO	$0	0	2
Camas Technologies, Inc.	USA	CO	$57,608	0	0
Care of Smith S, Regenesis Management Grp,	USA	CO	$0	1	0
Catalyst International, Inc.	USA	PA	$6,439	0	0
Chemonics	USA	DC	$579,130	0	0
Crosscart, Inc.	USA	CA	$0	0	19
Dfb Technology Holdings, Llc	USA	TX	$0	0	8
DM International, Inc.	USA	NY	$18,900	0	0
E. I. Du Pont De Nemours And Company	USA	DE	$0	0	11
Employers Reinsurance Co	USA	CO	$0	1	0
Epia:Kk	Japan	Foreign	$0	0	1
Evers Associates,	USA	MD	$0	1	0
Iii Holdings 1, Llc	USA	DE	$0	0	2
Invention Science Fund I, Llc.	USA	WA	$0	0	1
Lead Core Fund, L.L.C.	USA	NY	$0	0	3
Lee French Agr Res Inc,	USA	MN	$0	1	0
Optibrand Ltd., Llc	USA	CO	$0	0	2
Overseas Projects Corp.	Canada	Foreign	$42,963	0	0
Pragma Corp.	USA	VA	$4,700	0	0
Pursell Industries, Inc.	USA	CA	$900	0	0
Res Management Syst Inc	USA	CO	$0	1	0
Research Management Systems, Inc.	USA	CO	$79,905	0	0
Ronco Consulting Corporation	USA	DC	$103,448	0	0
Sci Applicat Int Corp	USA	CO	$0	2	0
Scientific Methods, Inc.	USA	IN	$23,539	0	0
Southeast Colo. Enterprise Develop. Inc.	USA	CO	$9,000	0	0
Tenera Technology, LLC	USA	CO	$14,500	0	1

Terra Global Capital	USA	CA	$0	1	0
Winterlab Limited	USA	NY	$0	0	1
Xatra Fund Mx, Llc	USA	DE	$0	0	13

Table A-5—Major companies active in sensors, testing, analytics, and equipment for product quality and biosafety innovation

Company	Country	State	Privately sponsored grant awards	WoS articles	Citing Patents
Abelbeck Partners, Ltd.	USA	CO	$0	0	1
Abs Corporation	USA	NE	$0	0	3
Access Security Protection, Llc	USA	TX	$0	0	1
AgDia Inc,	USA	IN	$0	1	0
Agr Informat Technol,	USA	SD	$0	1	0
AgriHouse Inc	USA	CO	$0	2	0
Amc Italy S R L	Italy	Foreign	$0	0	1
Amerasia International Technology, Inc.	USA	NJ	$0	0	2
Analytics Dev Corp	USA	CO	$0	1	0
Arryx Inc	USA	IL	$0	0	16
Bella Technologies, Llc	USA	TX	$0	0	2
Bellsouth Intellectual Property Corporation	USA	GA	$0	0	2
Best Analyt Serv,	USA	DC	$0	1	0
BH&H Engn Inc,	USA	CO	$0	1	0
Bioscience Consultants	USA	FL	$0	0	3
Biotronics, Inc.	USA	PA	$0	0	5
Body Surface Translations, Inc.	USA	GA	$0	0	1
Camas Technol Inc	USA	CO	$0	2	0
Canon	Japan	Foreign	$0	0	3
Celleration, Inc.	USA	MN	$0	0	4
COMP ASSISTED DEV INC,	USA	CO	$0	2	0
COMP SCI CORP,	USA	GA	$0	1	0
Cryolife, Inc.	USA	GA	$0	0	21
C-Scan, L.L.P.	USA	AZ	$0	0	1
Goldfinch Solutions, Llc	USA	NJ	$0	0	5
Google Inc.	USA	CA	$0	0	2
Hewlett Packard	USA	CO	$2,722	0	0
Hid Global Corporation	USA	CO	$0	0	2
HydroBio Adv Remote Sensing,	USA	NM	$0	2	0
IEH Labs & Consulting Grp	USA	CO	$0	1	0
International Business Machines (IBM) Corporation	USA	NY	$0	0	3

Kaiser - Hill Company, L.L.C.	USA	CO	$11,760	0	0
Local Independent Administrative Institution Aomori Prefectural Industrial Technology Research Center	Japan	Foreign	$0	0	1
Lumidigm, Inc.	USA	NM	$0	0	2
MANTECH ENVIRONM TECHNOL INC,	USA	OR	$0	2	0
Marctec, Llc	Australia	Foreign	$0	0	1
Mc10, Inc.	USA	MA	$0	0	2
Menon & Associates, Inc.	USA	CA	$0	0	2
Microbix Biosystems Inc.	Canada	Foreign	$0	0	1
Microsoft Corporation	USA	WA	$0	0	4
Nickel Brand Software Inc	USA	NM	$0	0	4
Non Typical, Inc.	USA	WI	$0	0	1
Non Typical, Inc.	USA	WI	$0	0	2
Optibrand Ltc. LLC	USA	CO	$29,237	0	0
Phillips Petr Co	USA	TX	$0	1	0
Phoenix Research Laboratories	USA	CA	$0	0	1
PRAXAIR,Inc.	USA	CO	$29,900	0	0
Renergon International AG	Switzerland	Foreign	$0	0	2
Smart Machine Vision, Inc.	USA	WA	$16,380	0	0
Smartin Technologies, Llc	USA	TX	$0	0	2
Smith & Nephew, Inc.	UK	Foreign	$0	0	5
Sony Corporation	USA	NY	$0	0	2
St. Jude Medical, Inc.	USA	AK	$0	0	5
Systems Research & Applications Corp	USA	VA	$91,576	0	0
United Power	USA	CO	$17,000	0	0
Vibralung, Inc	USA	AZ	$32,549	0	0
Y-TEX Corporation	USA	WY	$6,522	0	0

Table A-6—Major companies active in bioenergy innovation

Company	Country	State	Privately sponsored grant awards	WoS articles	Citing Patents
Abengoa Bioenergy New Technologies, Inc.	USA	MO	$0	0	5
Algae Systems, LLC	USA	NV	$0	0	1
Algal Scientific Corporation	USA	MI	$0	0	1
Algenol Biofuels Switzerland GmbH	Switzerland	Foreign	$0	0	1
Alliance For Sustainable Energy Llc	USA	CO	$0	0	1
Amoco Production Company	USA	IL	$15,000	0	0
Anadarko Petroleum Corporation	USA	TX	$78,216	0	0
Aurora Biofuels, Inc.	USA	CA	$0	0	25
Bio Finance Consultants	USA	ID	$0	1	0
Biolite Llc	USA	NY	$0	0	1
Biota N Amer Program,	USA	NC	$0	1	0
Blue Sun Biodiesel	USA	CO	$75,600	0	0
Boulder Innovative Technologies, Inc.	USA	CO	$10,536	0	0
Carbon Sink, Inc.	USA	IL	$0	0	5
Chevron Energy Technol Co,	USA	CA	$0	3	0
Cobalt Technologies, Inc.	USA	CA	$0	0	7
Community Power Corporation	USA	CO	$15,495	0	0
Coulter Corp,	USA	FL	$0	1	0
Edenspace Systems Corporation	USA	KS	$0	0	2
Fairfield & Woods PC,	USA	CO	$0	1	0
GENENTECH INC,	USA	CA	$0	1	0
Gevo	USA	CO	$21,433	0	0
Global Research Technologies, Llc	USA	AZ	$0	0	3
GlobeImmune Pharmaceut	USA	CO	$0	1	0
Grad Sch Biotechnol	S. Korea	Foreign	$0	1	0
Hitachi, Ltd.	Japan	Foreign	$0	0	1
Honeywell International Inc.	USA	NJ	$0	0	25
ICF Int	USA	DC	$0	1	0
Intelligent Decis Technol,	USA	CO	$0	1	0
Interlink Biotechnol LLC,	USA	CA	$0	1	0
Iogen Energy Corp	Canada	Foreign	$0	0	14
Joule Unlimited Technologies, Inc.	USA	MA	$0	0	3
Kilimanjaro Energy, Inc.	USA	WI	$0	0	10
Korea Institute Of Energy Research	S. Korea	Foreign	$0	0	1
Mascoma Corporation	USA	MA	$0	0	2
Matheson Gas Prod,	USA	CO	$0	1	0
Neatech, Llc	USA	CO	$0	0	2
Novozymes North America, Inc	USA	NC	$71,979	0	6
Photofuel Sas	Frence	Foreign	$0	0	2

Photon8, Inc.	USA	TX	$0	0	1
Phycal Llc	USA	OH	$0	0	1
Phytodetectors, Inc.	USA	CO	$0	0	0
Point Source Power, Inc.	USA	CA	$0	0	1
Pond Biofuels Inc.	Canada	Foreign	$0	0	3
Qteros, Inc.	USA	MA	$0	0	4
Ra Energy Corporation	USA	MI	$0	0	3
Senesco, Inc.	USA	NJ	$0	0	11
Solar Energy Res Inst	USA	CO	$0	3	0
Solix Biofuels	USA	CO	$0	2	1
Sunsource Industries	USA	IL	$0	0	2
Syngenta	Switzerland	Foreign	$236,968	0	0
Targeted Growth, Inc.	USA	WA	$29,575	0	0
Univerve Ltd.	Israel	Foreign	$0	0	1
V35A Enterprises, Llc	USA	WI	$0	0	2
Williams Production RMT Company	USA	CO	$88,102	0	0
Xyleco, Inc.	USA	MA	$0	0	13

Table A-7—Major companies active in agricultural commodity processing and management innovation

Company	Country	State	Privately sponsored grant awards	WoS articles	Citing Patents
Agridigit Inc	Canada	Foreign	$0	0	1
Attec Danmark A/S	Danmark	Foreign	$0	0	1
Australian Meat & Live-Stock Corporation	Australia	Foreign	$51,998	0	0
Beijing Poultry Breeding Company	China	Foreign	$158,100	0	0
Bemis Company, Inc.	USA	WI	$88,018	0	0
Cargill Meat Solution	USA	KS	$0	2	0
CATIE-Centro Agronomico Tropical de Inve	Costa Rica	Foreign	$4,355	0	0
Certified Angus Beef LLC	USA	OH	$0	3	0
Clougherty Packing Co	USA	CA	$0	1	0
Conagra Red Meat Co	USA	CO	$0	1	0
ContiBeef LLC	USA	CO	$152,345	5	0
CSIRO, Livestock Ind	Australia	Foreign	$0	1	0
Denver Buffalo Company	USA	CO	$187,994	0	0
Diamond Bar Land & Livestock	USA	CA	$38,099	0	0
F Menard	Canada	Foreign	$0	2	0
Farm to Table	USA	CO	$9,000	0	0
Farmland Industries	USA	MO	$14,000	0	0
Farnam Companies, Inc.	USA	AZ	$4,500	0	0

Five Rivers Ranch Cattle Feeding LLC	USA	CO	$840,697	2	0
Gerber Agri Inc	USA	CO	$0	1	0
Grandin Livestock Handling Syst Inc	USA	CO	$0	1	0
JBS Five Rivers Cattle Feeding	USA	CO	$58,552	4	8
John Morrell and Company	USA	OH	$17,944	0	0
Keen Ingredients, Inc,	USA	CO	$0	0	4
Leachman Cattle Company	USA	CO	$14,210	1	0
MaGiix	Germany	Foreign	$16,655	0	0
Magness Land & Cattle Company	USA	CO	$10,195	0	0
Meyer Natural Angus, LLC	USA	CO	$27,866	0	0
Micro Beef Technologies, Ltd	USA	TX	$0	0	16
Monfort, Inc.	USA	CO	$13,243	4	0
Paper Pak, Inc.	USA	CA	$5,796	0	0
PIC USA, Inc	USA	TN	$63,472	0	0
Saueressig Gmbh + Co.	Germany	Foreign	$0	0	1
Smithfeld Beef Grp	USA	WI	$0	2	0
Swift & Company	USA	CO	$78,450	0	3
Tyson Fresh Meats, Inc.	USA	SD	$0	0	3

Table A-8—Major companies active in plant genetics, and crop variety innovation

Company	Country	State	Privately sponsored grant awards	WoS articles	Citing Patents
ADM Alliance Nutrition Inc	USA	IL	$0	1	0
Agrigenetics, Inc.	USA	IN	$5,880	0	0
AlgEternal Technologies, LLC	USA	TX	$0	0	2
Archer Daniels Midland Co	USA	IL	$0	1	2
AsRes Ltd	New Zealand	Foreign	$0	1	0
Ayers Associates,	USA	CO	$0	1	0
Bartlett Grain Pty. Ltd.	Australia	Foreign	$0	0	1
Bartlett Tree Experts,	USA	NC	$0	1	0
Bayer Cropscience Lp	USA	NC	$0	0	2
Biocare Gesellschaft Fuer Biologische Schutzmittel Mbh	Germany	Foreign	$0	0	3
Blue Valley Ranch	USA	CO	$51,816	0	0
Booth Creek Ski Holdings, Inc.	USA	CO	$43,678	0	0
Btf Pty, Ltd.	Australia	Foreign	$0	0	1
Busch Agr Resources LLC,	USA	CO	$0	1	0
Cactus Feeders Inc	USA	TX	$0	1	0
Cargill, Inc.	USA	CO	$28,500	0	2

Cattlemans Choice Loomix, LLC	USA	CO	$160	0	0
Coca-Cola Company	USA	GA	$219,893	0	0
Crenshaw & Douget Turfgrass, Inc.	USA	TX	$4,000	0	0
DAS-Dow AgroSciences, LLC	USA	IN	$55,092	0	0
Diamond V Mills, Inc	USA	IA	$33,463	1	0
Dow AgroSci LLC,	USA	IN	$0	1	0
DuCoa	USA	IL	$8,491	0	0
DuPont Stine Haskell Res Ctr,	USA	DE	$0	1	0
E.G. & G. Rocky Flats, Inc.	USA	CO	$8,320	0	0
Ecoduna Ag	Germany	Foreign	$0	0	1
GAIA Management International, Inc.	USA	TX	$72,654	0	0
Gunnison Hay Products, LLC	USA	CO	$11,172	0	0
Heliae Development, Llc	USA	AZ	$0	0	1
Hill's Pet Nutrition, Inc.	USA	KS	$0	0	20
Horton Feedlot	USA	CO	$6,600	0	0
Jacklin Seed Company	USA	ID	$17,443	0	0
L. Johnson Farms, LLC	USA	KS	$11,892	0	0
Landcare Res NZ Ltd, Palmerston North,	New Zealand	Foreign	$0	4	0
Machinery Developments Ltd	UK	Foreign	$0	0	1
MIN AD Inc	USA	TX	$0	1	0
Multigrain Int LLC,	USA	CO	$0	1	0
Neptune & Co,	USA	NM	$0	1	0
Newleaf Symbiotics, Inc.	USA	MO	$0	0	1
Noble Fdn Inc,	USA	OK	$0	3	0
Northrup King Lawn/Garden Corp	USA	MN	$9,629	0	0
Nutribasics Company	USA	NY	$57,506	0	0
Oxion, Inc.	USA	KS	$13,743	0	0
PACE Turf,	USA	CA	$0	1	0
Penford Food Ingredients Company	USA	CO	$7,976	0	0
Pioneer Hi-Bred International, Inc.	USA	CO	$56,717	5	15
Planalytics Inc,	USA	PA	$0	1	0
PLANT TRADEMARK & COPYRIGHT OFF,	USA	CA	$0	2	0
Premium Genetics (Uk) Ltd.	Ireland	Foreign	$0	0	5
ptimal Ag Consulting Inc	USA	CO	$0	1	0
Purina Mills, Inc.	USA	CO	$12,500	0	0
Ralston-Purina Company	USA	MO	$45,986	0	0
Rhodia Acetow Ag	Germany	Foreign	$13,558	0	2
RINGGER FEED INC, NUTR,	USA	IL	$0	1	0
SEEDEX INC,	USA	CO	$0	1	0
Seminis Vegetable Seeds,	USA	ID	$0	1	0
Society for Range Management	USA	CO	$16,229	0	0
Tagawa Greenhouse Enterprises, Llc	USA	CO	$849,700	0	3

The Samuel Roberts Noble Foundation, Inc.	USA	OK	$0	0	3
United Feeds	UK	Foreign	$3,969	0	1
Vit-E-Men Company, Inc.	USA	NE	$25,356	0	0
W.C. Bradley Co.	USA	GA	$0	0	1
ZedX Inc,	USA	PA	$0	1	0

Table A-9—Major companies active in dairy and food product innovation

Company	Country	State	Privately sponsored grant awards	WoS articles	Citing Patents
Abbott Laboratories	USA	IL	$0	0	3
ACR, LLC	USA	CO	$49,900	0	0
Agrifood Solut Int	USA	TX	$0	1	0
American Diabetes Association, Inc.	USA	CO	$79,167	0	0
Angel Yeast Co., Ltd.	China	Foreign	$0	0	2
Aurora Organic Dairy	USA	CO	$281,856	0	0
Bar-S Foods Co.	USA	AZ	$18,000	0	0
Bil-Mar Foods	USA	MI	$0	1	0
Cooking Kids Inc	USA	NM	$0	1	0
Csb-System Software-Entwicklung & Unternehmensberatung Ag	Germany	Foreign	$0	0	3
CSR LTD, INGHAM,QLD 4850,	Australia	Foreign	$0	1	0
DairyNZ Ltd, Hamilton,	New Zealand	Foreign	$0	1	0
Energy Enzymes Inc	USA	MO	$0	0	4
Food Friends Inc.	USA	CO	$30,007	0	0
Food Safety Net Services, Ltd.	USA	TX	$129,041	4	0
Foodbrands Amer	USA	OK	$0	1	0
Hershey Company	USA	PA	$0	0	1
Hormel Foods, Llc	USA	MN	$0	0	1
Hort & Food Res Inst	New Zealand	Foreign	$0	1	0
IFT Food Microbiol Div	USA	CO	$0	1	0
Immucell Corp	USA	ME	$0	1	0
Johnson Res	USA	ID	$0	1	0
Keen Ingredients	USA	CO	$0	0	0
Kemin Industries, Inc.	USA	IA	$17,388	0	0
King Soopers	USA	CO	$16,500	0	0
Kroger Co	USA	GA	$19,350	1	0
Leprino Foods	USA	CO	$88,192	0	0
Mars Incorporated	USA	WA	$0	0	7
McDonalds Corp	USA	IL	$0	2	0

Minerva Ltd	Brazil	Foreign	$0	1	0
Monsanto	USA	MO	$295,988	4	27
Mountain View Harvest Coop/Gerards Baker	USA	CO	$161,015	0	0
MPC-Medical Packaging Corporation	USA	CA	$11,760	0	0
Mrp Group, Inc.	USA	GA	$0	0	1
Nat Way Inc,	USA	CO	$0	1	0
Nederlandse Organisatie Voor Toegepast-Natuurwetenschappelijk Onderzoek Tno	Netherland	Foreign	$0	0	1
Nestec S.A.	Switzerland	Foreign	$0	0	4
Procter & Gamble Co	USA	OH	$0	2	0
Purac, Incorporated	USA	KS	$29,609	2	0
Raps Gmbh & Co. Kg	Germany	Foreign	$0	0	2
Ruakura Res Ctr, AgRes	New Zealand	Foreign	$0	1	0
Select Sires Incorporated	USA	OH	$11,963	1	0
SoyPLUS/West Central	USA	GA	$7,572	0	0
SSI Food Serv Inc	USA	ID	$0	1	0
Sweetwater Energy, Inc.	USA	NY	$0	0	6
The Hershey Company	USA	PA	$0	0	1
Voogd Consulting Inc	USA	IL	$0	2	0
WESTERN SUGAR CO,	USA	NE	$0	1	0
Zinpro Corporation	USA	MN	$834,076	1	0

Table A-10—Major companies active in beer and wine production innovation

Company	Country	State	Privately sponsored grant awards	WoS articles	Citing Patents
Adolph Coors Company	USA	CO	$43,500	0	0
ANHEUSER BUSCH INC,	USA	MO	$0	3	0
Brewer Environmental Industries	USA	HI	$48,742	1	0
New Belgium Brewing Co	USA	CO	$0	1	0