Essays on Technological Change: Firm Organization, Problem Selection, and Diffusion

by

David P. Hall

Business Administration
Duke University

Date: ______________________________
Approved:

____
Victor Bennett, Advisor

____
Aaron Chatterji

____
Sharique Hasan

____
Ryan McDevitt

Dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy
in the Department of Business Administration
in the Graduate School of
Duke University

2020

ABSTRACT

Essays on Technological Change: Firm Organization, Problem Selection, and Diffusion

by

David P. Hall

Business Administration
Duke University

Date: ______________________________

Approved:

____
Victor Bennett, Advisor

____
Aaron Chatterji

____
Sharique Hasan

____
Ryan McDevitt

An abstract of a dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy
in the Department of Business Administration
in the Graduate School of
Duke University

2020

Abstract

This dissertation investigates three aspects of technological change. In the first essay, I build on prior research that suggests new ventures are often more innovative than established firms. I propose that one unexplored reason for this might be that new ventures work on different problems to begin with. I test this idea in the U.S. medical device industry and find evidence that established firms tend to select problems to which they can apply prior investments in complementary capabilities, while new ventures tend to undertake invention in less crowded areas. In the next chapter, I explore how artificial intelligence (AI) is being applied to medical devices, and which occupations are most affected. My analysis supports the notion that AI is different from prior digital technologies in its ability to perform non-routine tasks, but that concerns over its affect on high-skill labor might be misplaced. This provides some of the first systematic evidence of how AI is actually diffusing. Together, these two essays provide new insight into the direction of inventive activity in the medical device industry. Last, I look inside a firm to study how organizational factors influence communication patterns in an innovative financial services firm. This study complements the prior two by seeking to unpack the "black box" of a firm's innovation activities.

Contents

List of Tables

List of Figures

Acknowledgements

"Far better it is to dare mighty things, to win glorious triumphs, even though checkered by failure, than to take rank with those poor spirits who neither enjoy nor suffer much, because they live in the gray twilight that knows neither victory nor defeat."

– Theodore Roosevelt

This document is the culmination of an effort to dare something mighty. In the process, I've found the words of Theodore Roosevelt to be true: the joy of great accomplishment pales in comparison to the pain of setbacks, shortcomings, and failures along the way, as difficult as they may be in the moment. It is the difficulty of those moments that amplifies the joy of the ultimate victory. But just as every significant undertaking is certain to have setbacks and failures, it is as certain to have a supporting cast that elevates, enables, and encourages along the way. Here I express my sincere gratitude and appreciation for the people who have done this for me throughout this endeavor.

First, I want to thank those in the Fuqua Strategy community for the mentorship and friendship they have provided over the years. I want to thank Ramon Lecuona, who was an extraordinary example to me of hard work, determination, and rigor in research. In addition to providing excellent academic and storytelling training, Ramon provided life experiences in London, England, and Mexico City, Mexico, for which I am grateful. I also want to thank Victor Bennett, who selflessly provided endless hours of his time in personal mentorship. Victor taught me what it means to think straight. He also set an example of doing the hard intellectual work to go beyond finding the limitations of an idea to suggest avenues for improvement. It is not hyperbole to say that without his mentorship, I would not have accomplished what I have. Next, I express my thanks to Wes Cohen, Ronnie Chatterji, Ryan McDevitt, and Sharique Hassan for valuable guidance and feedback throughout the development of this dissertation. I also express my thanks to my fellow Fuqua Strategy PhD students who (almost) never missed an order from Enzo's pizza for our Thursday brownbag. This weekly ritual fed me in more ways than one. Lastly, I want to thank Colleen Cunningham, who was a friend and mentor in the early years of my PhD program when I was completely lost. I greatly appreciate her efforts to make me less lost and for providing encouragement along the way.

Although these and many other individuals entered the scene along the journey to provide support, there are a select few who were there from the beginning and who stayed to the end. My family deserves special credit for their constant patience, love, and support, in what likely appeared from the outside as a very uncertain and confusing process. Furthermore, credit goes to those who gave me the necessary foundation in life to qualify for this opportunity. To my mother, Melody Hall, who from a young age taught me that I can accomplish anything to which I apply myself, and who consistently and faithfully believed in me. And to my father, Peter Hall, who encouraged me to use my mind actively rather than as a passive receptacle of information, and who taught me the value of hard work through example and Saturday chores. From him I learned that in thoroughness is satisfaction.

Chapter 1

Introduction

A defining characteristic of the nineteenth and twentieth centuries was the unprecedented rate of technological advance. This growth in technology has raised many important questions for nations, industries, firms, and individuals about how to create and capture value from these innovations, how the innovations of others will affect themselves, and how to respond to the unexpected innovations introduced by others. This dissertation explores various aspects of this phenomenon. It includes how firms differ in their strategies for creating and capturing value in the U.S. medical device industry, how occupations are being affected by the innovations in artificial intelligence in the last decade, and how organizational factors influenced the flow of information in a firm generating innovative financial services in Mexico City, Mexico. Together, these essays provide insights into a variety of issues surrounding the creation and diffusion of new technology.

In chapter 2 I assemble a novel dataset to ask a simple question: which problems do firms undertake to solve, and why? One stream of innovation research emphasizes differences in the innovation activities of new ventures and established firms, typically portraying new ventures as more likely to introduce inventions of greater significance or novelty. However, rather than new ventures or established firms being "better" at innovation in some sense, some scholars propose they may tend to specialize in different steps of the innovation process. Building on these themes, this paper examines whether there are differences in which problems new ventures and established firms undertake to begin with, emphasizing the different roles they play in the direction of inventive activity. To study this I link the population of "medical activities" performed by healthcare providers to USPTO medical device patents. I find that new ventures are more likely than established firms to undertake invention for medical activities with fewer existing inventions. Additional analysis suggests that prior investment in complementary capabilities conditions which problems are undertaken by established firms. This paper contributes to the literature on the direction of inventive activity by proposing strategic reasons for why firms differ in what problems they undertake to solve.

Chapter 3 investigates a topic currently being debated in the public and academic spheres: how will advances in artificial intelligence (AI) affect the future of work, and which occupations will be most affected? To date, research on this topic has been largely speculative and has produced a wide range of estimates. We have little systematic evidence about the *actual* application of AI to the tasks of workers. This is an important distinction because there is often a significant gap between the potential for a new technology to diffuse and its actual course of diffusion, due to the incentives and inertia of individuals and firms. This paper provides some initial evidence of the actual diffusion of artificial intelligence in healthcare, and which occupations have been most affected. AI technology differs from prior digital technologies in its ability to perform non–routine tasks. This has been interpreted to imply that high–skill occupations are most at risk of being affected. The findings in this chapter support the notion that AI is capable of performing non–routine tasks, but do not suggest that the most high–skill occupations in healthcare have been affected. Rather, middle–skill occupations have been most susceptible to the application of AI technology to date. These findings align with prior research showing a hollowing out of middle–skill occupations, and suggest that while AI technology is different, it may follow a predictable course in its affect on labor.

The last chapter of this dissertation takes a turn toward the micro foundations of innovation to understand the organizational factors at work in the innovation process. The innovation process involves the integration and recombination of information across knowledge domains. One of the chief characteristics of firms, relative to markets, is their capacity to efficiently coordinate specialized knowledge. Prior research has shown that this coordination of knowledge (i.e. intra-firm communication) depends both on the work interdependencies and on the spatial distance between employees. This chapter expand this view of how communication patterns are shaped by taking focusing on the division of labor inside a firm. Employees are not stationary objects; rather, they move about the office as they accomplish their daily work tasks. Furthermore, the spatial movement of employees is likely to be greatly influenced by the location of their work interdependencies (the formal structure of the firm). Results from a 7-month field study support the idea that the frequency of face-to-face communication between two employees is influenced by their relative positions both spatially and organizationally, over and above the characteristics of the dyad itself.

Chapter 2

The Road Less Traveled: Problem selection in the medical device industry

2.1 Introduction

Differences in innovation activities between new ventures and established firms has been a central theme in innovation literature since Schumpeter (1942) suggested that innovation in established firms is fundamentally different than in new ventures. He argued that the innovation process in established firms was increasingly regimented and predictable. In contrast, by recombining resources into new products and new methods of production, new ventures tended to "undertake such new things ... [which] lie outside of the routine tasks which everybody understands" (Schumpeter, 1942). Subsequent work in this vein has investigated the extent to which small and large firms differ, in their propensity to perform R&D, in their inventive productivity, and in the qualitative differences of their inventions (Cohen, 2010). This literature has often portrayed new ventures as being "more innovative" than established firm in one sense or another. For example, established firms differ from new ventures in their incentives to take on incremental R&D projects (Cohen and Klepper, 1996), established firms are more rigid in their organizational structure and routines (Sørensen and Stuart, 2000), and established firms are more liable to suffer from entrenched cognitive biases of managers (Tripsas and Gavetti, 2000).

However, a different stream of research in the economics of innovation suggests that new ventures nor established firms have an overall advantage in innovation, but rather that they might specialize in different steps of the innovation process (Arora and Gambardella, 1995; Ziedonis, 2004). These studies suggest that new ventures sometimes specialize in upstream research while established firms tend to focus on downstream development, creating a division of labor in the innovation process. I build on the notion that new ventures and established firms might differ in what they do and ask whether there are differences in which problems they undertake to solve to begin with. In doing so I adopt a view of invention proposed by Rosenberg and Nelson (1994), who argue that it is useful to think about invention as a search problem in which inventors have some end-use (i.e. some "problem" to be solved) in mind. From this perspective, new ventures and established firms might play different roles in determining the direction of inventive activity in society by undertaking to solve different problems. Yet this question has never been examined empirically in the literature, to my knowledge.

One empirical challenge to answering this question is observing the population of problems a firm could potentially undertake to solve. In order to rigorously examine what firms do, the observed choices of firms must be compared to the complete set of alternatives *not* chosen by the firm. To address this challenge I construct a unique dataset of medical device patents. Medical devices are used by healthcare providers as they perform various medical activities. In this paper the term "medical activity" refers to any task performed in providing healthcare services to patients. For example, medical activities can be used for diagnosis (e.g. x-ray) or therapeutic (e.g. spinal fusion) purposes. Using a machine learning algorithm[1], I link USPTO medical device patents granted since 1976 to the population of medical activities healthcare providers can perform. This enables me to observe the set of medical activities or "problem areas" that are at risk of being worked on by medical device firms. I then hand collect data on a random sample of 260 patenting medical device firms for the years 2000-2017 and match these to the medical device patents linked

[1]The National Library of Medicine's Medical Text Indexer (MTI), see https://ii.nlm.nih.gov/MTI/

to medical activities.

I use the cumulative patents granted to *other firms* as a measure of the degree to which a medical activity has been explored. I find that firms are, on average, likely to undertake invention in problem areas which have been less explored by other firms. I also find that new ventures are *more* likely, relative to established firms, to undertake invention in less explored areas – "the road less traveled". Drawing on prior literature to explain this finding, I consider three theories for how firms select which problems they solve.

The first explanation is that prior investments in complementary capabilities change the cost-benefit calculus when deciding which problems to undertake to solve. Research in strategy (Helfat, 1997) and economics (Scott Morton, 1997) has found that prior investment in complementary capabilities increases the likelihood of undertaking invention in related technical domains because it makes these R&D projects relatively less costly. Additionally, prior investment in complementary capabilities can also influence which R&D projects are undertaken by increasing the ability of the firm to realize as well as appropriate the value of inventions in related areas (Teece, 1986; Mitchell, 1989).

A different explanation is that the decision of which problems to solve is largely a function of organizational rigidity (Leonard-Barton, 1992) and inertia in the structure and routines of the firm (Nelson and Winter, 1982b). The logic of this theory is that as firms age it becomes increasingly difficult for them to undertake innovation activities that require a different organizational structure or set of routines (Sørensen and Stuart, 2000). This difficulty can arise either from the firms inability to recognize new opportunities, or their failure to effectively respond to those opportunities. New technological opportunities therefore are more likely to be undertaken by new ventures, while established firms will tend to continue working in older technological areas in which they have previously worked.

Lastly, differences in which problems firms undertake to solve may be a consequence of variations in search costs between firms. Less explored areas might be characterized by higher costs of inventing (i.e. low technological opportunity). In these areas science can narrow the set of inventive alternatives to pursue and focuses the attention of the inventor, making the research process more efficient (Nelson, 1982). Firms closer to new scientific developments may have relatively lower costs of inventing in these areas, making them more likely to undertake invention.

I find support for the explanation that which problems established firms undertake to solve is conditioned by prior investments in complementary capabilities. Specifically, I find that for established firms, the extent to which the firm's technical capabilities (experience with a patent class) overlap with the capabilities required to invent in an area is significantly related to the probability of patenting in that area. Similarly, I find that for established firms, investments in marketing capabilities are significantly related to the probability of patenting in an area, conditional on patent stock. Furthermore, when firms have made prior investments in complementary capabilities they are more likely to undertake invention in more explored areas than when they have not.

Altogether, these findings provide some evidence that there are differences in which problems new ventures and established firms undertake to solve. This contributes to innovation research by taking a step back to consider why firms choose to work on some problems in preference to others. Related work that has proposed a division of innovative labor has emphasized environmental factors, such as markets for technology and strong intellectual property rights. In contrast, I show that differences in what firms do can arise from strategic considerations. These results provide empirical support for the idea that new ventures provide diversity in the types of problems that are pursued in society (Cohen and Klepper, 1992). New ventures are more likely to develop new solutions to less explored problems, rather than focusing on "the problem you know" (Eggers and Kaul, 2018). While for established firms, the benefits of being able to leverage complementary capabilities appears to outweigh the disadvantages of inventing in more technologically explored areas.

2.2 Related Literature

A vast body of literature across strategy and economics examines the patenting behavior of firms. Three streams of work in this tradition are relevant to the central question of this paper. The first of these has focused on the innovative differences between small and large firms. Part of this literature has sought to understand how, and why, innovation varies by firm size. A different part has emphasized the strategic implications of these differences. Second, a related literature has suggested what is often referred to as a division of innovative labor. This work departs from the first by thinking about how different types of firms cooperate in the innovation process. Lastly, a stream of research has emphasized the importance of understanding the economic and social forces influencing the direction of inventive activity. These streams are related and overlapping to some extent. I distinguish them here based on their distinct contributions and the opportunities they provide for further research.

One stream of work, largely growing out of the work of Joseph Schumpeter, has examined the relationship between firm size and innovative activity. This has typically been examined empirically as the differences between small, "entrepreneurial", firms and large established firms. Building on the suggestion that large firms might have an advantage in R&D (Schumpeter, 1942), this literature has sought to understand whether R&D rises more than proportionately with firm size. The empirical consensus in this literature is that while large firms conduct more R&D than small firms, they do not conduct *disproportionately* more R&D. This has generally been interpreted as support for that idea that size does not confer an advantage in performing R&D (Cohen, 2010). Using data at a more fine grained level than the firm, Henderson and Cockburn (1993) study whether there are scale and scope effects in drug development. They examine R&D in therapeutic classes and find no scale effects at the firm level, but do find scale and scope effects within the firm across therapeutic classes. Although they do not directly tie their findings to an R&D advantage in firm size, their paper highlights the point that disaggregating firm level data into a more fine-grained level of analysis can provide insights into the innovation activities of firms.

A feature of this stream of research is that it focuses on the inputs (R&D) to innovation as opposed to the output. A related stream of literature spanning economics and strategy places greater emphasis on the differences between innovative output of firms. A number of early studies found that smaller firms tended to produce a disproportionate share of inventions for their size (Scherer, 1965; Acs and Audretsch, 1988; Associates, 1982). For example, Lerner (2006) found that smaller firms in the financial services industry accounted for a disproportionate share of innovations. This research has also found that small firms are often more likely to introduce more significant or "radical" inventions, while larger firms tend to introduce more incremental inventions (Henderson, 1993; Mansfield, 1981; Roessner, 1981). This work is closely related to work in strategy that examines the failure of established firms to commercialize or adopt radically new technology (Henderson, 1993; Henderson and Clark, 1990; Christensen, 1997) due to various reasons such as organizational inertia (Sørensen and Stuart, 2000), cognitive biases (Tripsas and Gavetti, 2000), and incentives to allocate resources based on ties to larger markets (Christensen, 1997). However, while there is a great deal of research showing why and when established firms fail to introduce radically new technology, there is surprisingly little empirical work showing that new ventures, as opposed to small firms more generally, tend to undertake more radical invention.

Overall, this stream of literature has focused on identify whether and when large or small firms have an "advantage" in innovative activity. A separate stream of work has instead suggested that firms might specialize in different aspects of the innovation process (Arora and Gambardella, 1995; Arora et al., 2016). This stream of literature draws on the fundamental concept proposed by Adam Smith that there are gains from specialization and trade. In terms of innovative activity, firms might realize better outcomes by specializing in different aspects of the innovation process, relying on the market as a means of organizing the innovation process rather than completely internalizing the process. In this view, small or large firms aren't "better" at innovation, rather, they play distinct roles. For example, Hall and Ziedonis (2001) found that stronger property rights in the semiconductor industry related to increased entry of firms that specialized in the design of chips and performed no manufacturing, suggesting that well-functioning markets for technology can facilitate specialization and trade in the innovation process.

Lastly, a related stream of research has sought to understand the determinants of the direction of inventive activity. This stream of research has proposed numerous explanations of what shapes the path of scientific technological progress. At the level of individuals and teams, collocation has been found to be an important determinant of which projects are undertaken (Catalini, 2017), as well as the cost of accessing knowledge (Teodoridis, 2017). At the firm level, internal organization of the R&D has been found to be an important factor. For example, Argyres and Silverman (2004) found that centralized R&D is associated with drawing on inventions of other organizations and pursue a broader set of technologies than decentralized firms. Broader environmental factors can also influence the direction of inventive activity, such as the government funding (Corredoira et al., 2018), and the market for ideas (Chatterji and Fabrizio, 2016). Although this body of work contributes greatly to our understanding of the direction of inventive activity, this literature has largely been silent on how new ventures and established firms differ with respect to the direction of inventive activity. One exception to this is Cohen and Klepper (1992), who suggest there might be a trade-off between firm size and diversity in what technology is pursued, though this proposition has never been examined empirically. This is surprising given that differences between new ventures (and small firms more generally) and established firms has been a central theme in much of the innovation literature, as discussed above.

To address this gap I suggest that new ventures and established firms might undertake invention for different problems to begin with. In doing so, I bridge the literature on the direction of inventive activity and the literature on firm size and innovation. While I build on the notion that small and large firms differ in their inventive activities, I explicitly focus on new ventures rather than small firms more generally. Furthermore, I focus my analysis on the end use of the inventions, as opposed to the characteristics of the technology itself, and ask whether new and established firms differ in their propensity to invent in areas that have been more or less explored by other firms.

2.3 Empirical Context

I explore these ideas within the context of the U.S. medical device industry. This setting as many advantages. Prior research has found that patents are used extensively for medical device inventions (Cohen et al., 2000; Arora et al., 2016), making this an appropriate means for observing inventive activity by firms. Additionally, studying a single industry allows me to be specific in my operationalization of a "problem" and make reasonable comparisons between problems. Furthermore, by focusing on a single industry I fix other unobserved, industry level determinants of inventive activity (e.g. appropriability).

A challenge to addressing which problems firms undertake to solve is that one must observe the entire population of problems that are *at risk* of being chosen. This includes both the problems they choose to work on and the problems they choose not to work on. This is important for at least two reasons. First, having the entire risk set alleviates concern that a firm selected a problem that was not observed. Second, having the entire risk set allows one to compare what was selected to the complete set of alternatives. To address this challenge I leverage the Unified Medical Language System (UMLS), a database of medical terminology established in 1986 and maintained by the United States National Library of Medicine for use in medical informatics (see for example Bodenreider 2004; Wu et al. 2012; Carrell et al. 2017). The UMLS brings together over 1 million distinct concepts from 214 incorporated and maintained medical dictionaries across 25 languages, providing a comprehensive and detailed set of medical terminology. In addition to concept definitions and synonyms, this database incorporates taxonomies and relationships between concepts, making it a useful tool in natural language processing tasks, such as text classification.

I focus my analysis on concepts that are defined as therapeutic or diagnostic activities in the database. Thus, my analysis accounts for the set of diagnostic and therapeutic activities for which firms can develop medical devices. There are two advantages of using the activity as the unit of analysis. First, a medical activity indicates the intended use of a medical device, rather than the characteristics of the underlying technology such as patent classification codes. This builds on the idea suggested by Rosenberg and Nelson (1994) that inventors have an end-use in mind as they search for solutions. Second, a medical activity reflects the problem the medical device is solving. For example, spinal fusion fixes weak or deformed spines in patients, and x-ray allows patients to be diagnosed non-invasively. This enables me to observe the set of medical activities or "problem areas" that are at risk of being worked on by medical device firms.

Healthcare providers such as medical doctors, nurses, and medical technicians, perform various activities to provide health services to patients. Medical device firms develop technology to facilitate the performance of such procedures. For example, one such therapeutic activity is an atherectomy, a minimally invasive surgical technique to remove plaque buildup within the walls of an artery. Various types of devices have been developed by firms to aid physicians performing this procedure (see Appendix A for example patents).

One solution uses a laser to ablate the plaque buildup in the artery (Dippel et al., 2015). Another solution uses a blade that scrapes the inner wall of the artery (Topol et al., 1993), and yet another solution uses a grinding wheel that breaks down the buildup as it rotates within the artery (Shih et al., 2016). Patents on these and other medical devices are an indication that a firm has committed resources to finding and improving solutions for the needs of physicians and their patients.

2.3.1 Data Construction

Constructing my sample involved multiple stages. In the first stage I link the population of medical activities (available from the UMLS) to patents granted by the USPTO using a machine learning algorithm developed by the National Library of Medicine. Since 2002 this algorithm has been improved upon to facilitate medical publication indexing for databases such as MEDLINE. This algorithm has many advantages. First, it combines traditional feature engineering approaches such as bag-of-words with more sophisticated approaches such as n-grams, noun phrases, and related publications (Yepes et al., 2015). Second, various types of learning models within the algorithm have been tested and evaluated against each other to significantly improve the performance of the text classification (Yepes et al., 2015)[2].

I obtained medical patents granted by the USPTO between January 1976 and November 2017 from PatentsView.org, a platform supported by the USPTO that provides bulk downloads of granted U.S. patents. I restrict my sample to medically related patents based on Cooperative Patent Classification (CPC) codes. A detailed list of which CPC codes were included can be found in Appendix B. This results in 544,047 distinct patents. I then apply the algorithm to the title and abstract text. Of the 544,047 medically related patents,

[2]Additional details on the algorithm as well as a web interface for accessing it can be found at https://ii.nlm.nih.gov/MTI/.

165,303 (30.4%) were linked to diagnostic or therapeutic medical activities[3]. An inspection of the unmatched patents indicates they are for devices not associated with a particular medical procedure (e.g. hospital bed software, patent number 5787528), or they are patents on drugs (e.g. US patent number 8906940).

In the second stage I hired nine medical residents and three fourth year medical students to rate the medical activities drawn from the UMLS on various dimensions. The purpose of this was twofold. First, a manual inspection of the medical activities revealed that they varied in their level of generality. For example, the UMLS labels "Diagnostic Tests, Routine" as a medical activity, despite the fact that this concept encompasses many specific medical activities. To ensure a suitable level of analysis I asked respondents to indicate whether the activity was too vague or general to provide meaningful responses. These activities were excluded from the analysis. See the Appendix D for examples of medical activities included in the analysis. Second, additional data important for the analysis, such as the frequency with which an activity is performed, were not available through other means.

The questions were developed by first consulting with nurses and other medical professionals (i.e. asking whether question make sense) and the word choice (i.e. whether the question will be interpreted the way I intend). See Appendix C for the full list of questions. A pilot using 200 activities was conducted, after which I spoke with the pilot raters to identify sources of confusion or ambiguity. Next, nine medical residents and three fourth year medical students were hired to rate the medical activities on various dimensions. The list of medical activities were broken up into blocks of 200, and each block activities was coded by two separate raters. Responses were aggregated based on the level of familiarity of the respondent with the activity and their seniority. I kept the responses of the respondents

[3]My analysis links medical device patents to the activities, or procedures, for which they are used. My analysis does not investigate patents directly granted for medical procedures. See Appendix E for a discussion on the legal history of patenting medical procedures.

that were most familiar with the activity. Ties were broken based on the seniority of the respondent (year in residency) or using an average if respondents were the same seniority.

In the last stage I selected a random sample of 300 assignees from the matched patents between 2000-2017. I exclude independent inventors, universities, patent holding companies, and assignees for which there was not a definitive match based on name, location, and/or inventor names in the case of startups. This results in a random sample of 260 medical device firms[4]. I then hand collect data for each firm for each year between 2000-2017. Because my sample includes firms of all types, from newly founded firms with a single founder to global conglomerates over one hundred years of age, I use multiple sources to collect information. Sources include BvD Orbis, Crunchbase, company websites, LinkedIn profiles of companies and founders, and other private company databases. My final dataset is comprised of 950 unique medical activities and 260 firms for the years 2000-2017. Slightly over 50% of the firms in my sample are classified as a startup at some point between 2000-2017. Founding teams often include a founder with a Ph.D. (30%), M.D. (36%), or prior experience founding a firm (28%).

2.3.2 Variables

The dependent variable in my main analysis is a binary indicator that is equal to one if a firm *applies* for a patent related to a medical activity in a given year and the patent is subsequently granted. I use the application date because this indicates an action over which the firm has control, as opposed to the grant date. While I am only able to observe granted patents, studies have estimated that USPTO patent grant rates are between 80-97% (Quillen Jr and Webster, 2005). This is because a patent application is similar to a negotiation process (i.e. concerning what the inventor can and cannot claim), rather than

[4] Although my current analysis is limited to 260 firms due to the time-intensive nature of hand collecting these data, I also run my main analysis on a separate sample of 160 firms I collected for a separate project and find qualitatively identical results.

an application for a job that will be awarded to only one applicant. I use a binary outcome because I am primarily interested in whether or not a firm is working on a problem, rather than the "intensity" of those efforts.

To measure the extent to which a medical activity is technologically explored I calculate the cumulative patents related to each medical activity every year since 1976 *not granted to the focal firm* (*Cum activity patents*$_{-f}$). This measure includes all medical device patents from the USPTO that were linked to the activity, not only those granted to firms in my sample.

Medical device firms seeking regulatory approval in the United States take approximately three to seven years to receive approval (Van Norman, 2016). Additionally, the average time from first institutional funding to acquisition for medical device firms is between five to eight years (Norris et al., 2014). Therefore, I define new ventures as firms that are 5 years of age and younger (*Startup*). Firms appear in the sample the year they are founded, at age 0, so a surviving firm will be classified as a new venture (=1) for six years, after which it transitions to incumbency (=0). Data for founding date was hand collected from Crunchbase, LinkedIn, and company websites.

I include a number of controls in my analysis. First, to control for unobserved differences between medical activities I include activity fixed effects. For example, some medical activities might present greater opportunity for invention than others. To the extent these differences are stable over time, including activity fixed effects will account for these differences. Similarly, I include firm fixed effects in my analysis to control for unobserved characteristics of firms that might influence their propensity to patent. Because this propensity might change over time, and might be directly related to *firm size* (employee count) or *patent stock*, I also include time varying controls of these firm characteristics. I also include year fixed effects to account for differences in the technological environment over time. Finally, I standardize all variables in my models for ease of interpretation. Ta-

Table 2.1: Summary statistics of key variables

Firm-Year-Activity Obs.	N	Mean	StDev	Min	Max
Patent (0/1)	3,374,400	.0015	.039	0	1
Cum activity patents$_{-f}$	3,374,400	109	373	0	5,580
Startup	3,374,400	.181	.385	0	1
Patent stock	3,374,400	33.4	174.8	0	2,960
Employees	3,374,400	13,550	52,170	0	472,500

Table 2.2: Correlations between key variables

Panel A: Full sample	(1)	(2)	(3)	(4)	(5)
(1) Patent (0/1)	1.000				
(2) Cum activity patents$_{-f}$	0.080	1.000			
(3) Startup	-0.013	-0.014	1.000		
(4) Patent stock	0.065	0.009	-0.090	1.000	
(5) Employees	0.047	-0.001	-0.112	0.459	1.000

bles 2.1 and 2.2 provide summary statistics for the key variables and their correlations, respectively.

2.3.3 Data Description

Table 2.3 provides descriptive statistics on the medical activities in my data. The majority of these activities are performed in a hospital (77%) by a medical doctor (68%), and the sample is largely composed of therapeutic activities (70%). Interestingly, most medical activities have at least one patent (92%). Although the cumulative number of patents per medical activity is highly skewed: about 50% of the total patents on medical activities are associated with only 3% the activities (about 30 activities). See Figure 1 for a plot of the cumulative patents by medical activity. Examples of highly patented diagnostic activities include CT scanning and ultrasound. Highly patented therapeutic activities include procedures such as renal dialysis and spinal fusion.

Table 2.4 provides additional descriptive statistics on the nature of these medical ac-

Table 2.3: Medical activity descriptive statistics

Activity Summary	N	Mean	StDev	Min	Max
Diagnostic	950	.300	.456	0	1
Therapeutic	950	.700	.456	0	1
Performed by M.D.	950	.671	.470	0	1
Performed in hospital	950	.778	.415	0	1

Table 2.4: Medical activities by quartiles of cumulative patents

Quartile Summary	0-25%	26-50%	51-75%	76-100%
Cum activity patents$_{-f}$	0.584	7.75	29.38	397.6
RVU (reimbursment)	16.04	14.42	16.33	9.33
Frequency (weekly)	308.0	461.0	495.29	854.85
Share Diagnostic	0.323	0.334	0.288	0.259

tivities across the distribution of total patents. I break down the reimbursement rate, the frequency with which it is performed, and the share that are diagnostic activities, by quartile of cumulative patents. Activities in the top quartile of patents are performed more frequently and have a lower reimbursement rate on average, than those in the bottom quartile.

New ventures and established firms vary dramatically in terms of the number of medical activities for which they invent. The average firm in my sample patents on 58 medical activities, though this is driven largely by the established firms: new ventures patent, on average, on 5 medical activities, while the mean for established firms is 59. The maximum number of activities any single firm in my sample develops patents for is 149 (16% of the total medical activities).

2.4 Main Results

This section presents my main results. My main specification employs a linear probability model. There are two major advantages to using a linear probability model in this setting. The first is the ease of interpreting the coefficient estimates, which can be interpreted as an increase/decrease in the probability of the outcome event occurring. The second advantage is the ability to cluster on multiple groups using the `reghdfe` package in STATA. When including time fixed effects, the coefficients can be interpreted as the marginal increase/decrease in probability of an outcome at a given point in time. To analyze which problems startups and incumbents undertake to solve, I estimate the the probability that firm f will develop a patent targeting activity i in year t, conditional on the cumulative patents for a medical activity. The main specification is

$$\begin{aligned} Patent_{fit} \;=\; & \alpha + \beta_1 \big(Cum\ activity\ patents_{-f}\big)_{it} + \beta_2 \big(Startup\big)_{ft} + \\ & \beta_3 \big(Cum\ activity\ patents_{-f}\ x\ Startup\big)_{fit} + \Gamma_t + \Phi_i + \varepsilon_{fit} \end{aligned}$$

where Γ is a vector of controls for patent stock and firm size, and Φ is a vector of fixed effects for the year, activity, and firm. The coefficient on interaction between the cumulative patents for an activity and the indicator for a startup is the primary coefficient of interest. Because the data are at the firm-activity-year unit of analysis, the errors are not independent (e.g. correlated within firm and within activity). I account for this by using clustered standard errors clustered at both the level of the firm and the level of the activity. This is implemented using the `reghdfe` package in STATA, which allows for multi-way clustering[5].

[5] The `reghdfe` package is a generalization of the areg command that allows for specifying multiple absorbed fixed effects and clustering on multiple groups. Details can be found at http://scorreia.com/software/reghdfe/.

Table 2.5: Differences between new ventures and established firms

	(1)	(2)	(3)	(4)	(5)	(6)
	LPM	LPM	LPM	LPM	OLS	ZINB
DV =	Patent (0/1)	Patent (0/1)	Patent (0/1)	Patent (0/1)	Patent count	Patent count
Cum activity patents$_{-f}$	-0.00341***	-0.00341***	-0.00352***	-0.00352***	-0.02763**	-0.11903***
	(0.00100)	(0.00100)	(0.00102)	(0.00102)	(0.01310)	(0.02149)
Startup		-0.00082***	-0.00090***	0.00012	-0.00263***	-0.09779
		(0.00029)	(0.00031)	(0.00016)	(0.00094)	(0.08505)
Startup × Cum activity patents$_{-f}$			-0.00314***	-0.00322***	-0.01831**	-0.10405***
			(0.00092)	(0.00095)	(0.00737)	(0.01849)
Constant	0.00165***	0.00178***	0.00178***	0.00144***	0.00504***	-0.73390***
	(0.00018)	(0.00019)	(0.00019)	(0.00015)	(0.00096)	(0.18721)
Control for patent stock	Yes	Yes	Yes	Yes	Yes	Yes
Control for firm size	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Activity FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	No	No	No	Yes	No	No
Obs.	3,374,400	3,374,400	3,374,400	3,374,400	3,374,400	3,374,400
F	14.39	11.99	9.36	19.76	90.47	
R^2(Adj.)	0.018	0.018	0.019	0.028	0.017	
Sample	Full	Full	Full	Full	Full	Full

Standard errors in parentheses. Standard errors clustered at the firm and procedure levels.

* p<0.10, ** p<0.05, *** p<0.01

Table 2.5 presents my main results. Column 1 estimates the probability a firm will patent on an activity in a year conditional on the cumulative patents on that activity by other firms. The coefficient on *Cum activity patents*$_{-f}$ implies that a one standard deviation increase in the cumulative patents (about 370 patents) is related to a decrease in the probability of patenting of about .34 percentage points, about double the baseline probability. This result indicates that on average firms are less likely to apply for patents for a medical activity as the number of patents for that activity increases. Column 2 indicates that, unsurprisingly, new ventures are less likely than established firms to patent on a given activity in a year. In column 3 and 4 I interact *Startup* with *Cum activity patents*$_{-f}$. The coefficient is negative and significant, indicating that new ventures are less likely than established firms to patent in areas with higher levels of patenting by other firms.

One feature of my data is that the dependent variable has a large proportion of zeros. Unsurprisingly, most firms do not develop patents for most medical procedures in a given year. For example, in my sample the maximum number of procedures a single firm develops patents for is 149 (16% of the total number of procedures) and the average firm develops patents for 58 activities (6%). To address this issue I use a zero-inflated negative binomial (ZINB) model. The intuition for the ZINB model is that a negative outcome (0) may be common but for different reasons across latent groups. The ZINB model generates two separate models to take into account the possibility of two separate data generating processes. In the first stage, the ZINB uses logistic regression to estimate the likelihood that the outcome is always zero. In the second stage it uses a negative binomial model to estimate the outcome for those observations that are not certain zeros. Because the ZINB model uses a negative binomial model, I use the count of patents rather than a binary indicator as the dependent variable. Models comparing OLS and ZINB to my main results are reported in columns 5 and 6 of Table 2.5. Overall these results suggest that, on average, new ventures are more likely than established firms to take “the road less traveled”.

2.5 Additional Analysis

This section explores why new ventures are more likely than established firms to undertake invention in less explored areas. I draw on existing theories in strategy and economics to propose different explanations for this result. Each theory has its own distinct mechanisms from which I draw predictions. The theoretical development in this paper is not intended to develop a new "grand theory" or resolve conflicting world-views about the innovation process. Rather, the intended theoretical contribution of this paper is in the question it asks and in its investigation of which mechanisms most likely explain the observed empirical patterns[6] (Makadok et al., 2018). Thus, I draw on existing theories to address a question different than what has been previously asked in the innovation literature, and to develop tests for plausible explanations.

2.5.1 Complementary Capabilities

A fundamental tenet of strategic management is that firms draw competitive advantage from their accumulated resources and capabilities (Peteraf, 1993; Dierickx and Cool, 1989). While firm resources and capabilities are an important source of competitive advantage, new ventures and established firms differ dramatically on this dimension. New ventures are limited in their access to resources, and by definition have not had time to develop capabilities[7] (Dierickx and Cool, 1989). Prior literature has emphasized two important capabilities: technical capabilities and marketing capabilities.

[6] As Makadok et al. (2018) point out, "Without a research question, there is nothing to theorize about ... One obvious way to make a contribution to theory is to change the research question, either by asking a new question, modifying an existing question in some way, or applying an existing theory to address a different question."

[7] Related literature in entrepreneurship has looked at how new ventures overcome these liabilities. For example, the resource mobilization literature studies how new ventures gain access to resources necessary for survival (Clough et al., 2019), and the literature on bricolage examines how new ventures "create something out of nothing" (Baker and Nelson, 2005). However, my focus is not on how entrepreneurs overcome these liabilities, but rather the implications for their behavior relative to established firms

First, prior investment in technical capabilities lowers the cost of undertaking future invention that requires similar capabilities. This comes from experience with the technology (Lieberman, 1989), and economies of scope in R&D (Penrose, 1959). For example, (Helfat, 1997) finds that U.S. oil companies with experience in refining R&D were more likely to undertake R&D in synthetic fuels, which required the same technical knowledge as R&D in refining. Similarly, (Scott Morton, 1997) finds that in the pharmaceutical industry established firms were more likely to enter markets in which they had experience with similar manufacturing capabilities.

Second, prior investment in marketing capabilities increases the value and appropriability of inventions in related areas. Mitchell (1989) finds in medical imaging sub-markets that if an incumbent has specialized complementary assets, it will enter these sub-markets sooner than if it does not, presumably to preempt competitors and realize the value of their specialized assets. Likewise, Teece (1986) describes a case in which incumbents were better positioned to appropriate the value of an invention by an entrant due to their complementary capabilities. Although EMI invented the technology for the CT scanner, "Two competitors, GE and Technicare, already possessed the complementary capabilities that the scanner required, and they were also technologically capable. In addition, both were experienced marketers of medical equipment, and had reputations for quality, reliability and service. GE and Technicare were thus able to commit their RD resources to developing a competitive scanner, borrowing ideas from EMI's scanner, which they undoubtedly had access to through cooperative hospitals, and improving on it where they could while they rushed to market" Teece (1986).

Capabilities: Data and Analysis

I measure the degree to which the technical capabilities of a firm are a good fit for what is required to invent for a medical activity (*Tech overlap*). First I take the count of distinct patent classes previously associated with a medical activity. Next I find which of those patent classes a focal firm has prior experience with, and divide this by the distinct number of patent classes of the activity. For example, if X-Ray has 20 patent classes, and firm ABC has previously patented on 5 of them, the technological overlap for firm ABC with respect to X-Ray will be 5/20 = 0.25. On average, there are 169 patent classes per medical activity[8], though there is significant variation (standard deviation of 426). The advantage of this measure is that it reflects the degree of fit between the technical capabilities of the firm and a particular medical activity. This is conceptually distinct from the experience of the firm with respect to a technical capability, which might be measured, for example, using the count of prior patents on an activity. Thus, in my analysis I distinguish between the technical *experience* of the firm and the technical *fit* of the firm, and conceptualize capability as a technical fit between the firm and the medical activity.

For my measure of marketing capabilities, I access a relatively new source of data available from the Center for Medicare and Medicaid Services (CMS). Starting in 2013, CMS began collecting information on payments made by medical device firms to physicians. This database is referred to as the Open Payments Database[9]. These data are reported at the transaction level and include such information as the reason for the payment (gift, food and lodging, speaking fee, etc.), the specialty of the physician receiving the payment, and the location of the physician. I match my set of medical activities to the set of physician specialties and calculate the stock (Dierickx and Cool, 1989) of payments made by a firm to physicians *within a specialty* (*Marketing stock*). Because physicians are often instrumental in providing knowledge for inventions, I exclude transactions explicitly made to physicians for consulting, royalties or licenses. Thus, this measure largely reflects payments made to physicians for speaking at conferences, teaching continuing education courses, and other

[8]I use the USPTO Cooperative Patent Classification (CPC) system, and specify a patent class as the "Subgroup ID". This is the most fine-grained level of classification, with about 250,000 entries.

[9]www.cms.gov/openpayments/

Table 2.6: Capabilities by firm type (firm-year obs.)

New ventures	N	Mean	StDev	Min	Max
Tech overlap	644	0.0006	0.0026	0	0.0337
Marketing stock	112	277	2,326	0	23,510
Established firms					
Tech overlap	2,996	0.0135	0.0383	0	0.3392
Marketing stock	957	51,941	262,536	0	4,110,212

"gifts" given during visits by sales representatives.

Measuring marketing capabilities in this way is advantageous for several reasons. First, this allows my to measure the level of marketing investment by a firm with respect to a medical specialty. For example, I observe the extent to which a focal firm has made payments to cardiothoracic surgeons relative to endocrinologists. Second, this facilitates within firm analysis by providing variation of the marketing capabilities for specific medical activities. Third, since the majority of my firms are privately held, obtaining traditional financial metrics isn't feasible. Thus, this data source gives me finer grained marketing data, and on a set of firms for which it is difficult to find financial data. The limitation of this measure is that payments to physicians may also indicate a relationship between the firm and the physician, which may be correlated with other factors influencing the firms propensity to patent, such as consulting on technical matters. Table 2.6 provides the correlations between the various types of payments made by firms to physicians.

Table 2.6 reports *Tech overlap* and *Marketing stock* broken down by new ventures and established firms. As expected, the average *Marketing stock* for a new venture is $277, compared to $51,941 for established firms. Similarly, new ventures have lower levels of technical capability than established firms.

Because I am interested in how established firms might be conditioned by prior investments in capabilities, and because new ventures have not made prior investments, I restrict the following analysis on the role of capabilities to established firms. Table 2.7 reports the

Table 2.7: Prior investments in capabilities

	(1)	(2)	(3)	(4)
DV =	Patent (0/1)	Patent (0/1)	Patent (0/1)	Patent (0/1)
Firm experience$_{t-1}$	0.00432***	0.00421***	0.00242***	0.00237***
	(0.00111)	(0.00108)	(0.00072)	(0.0007)
Tech overlap		0.00414***		0.00215***
		(0.00071)		(0.00038)
Marketing stock			0.00044**	0.00037**
			(0.00022)	(0.00018)
Constant	0.00164***	0.00169***	0.00196***	0.00222***
	(0.00008)	(0.00007)	(0.000509)	(0.00032)
Control for patent stock	Yes	Yes	Yes	Yes
Control for firm size	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Activity FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Obs.	2,764,500	2,764,500	718,679	718,679
F	14.69	21.75	15.45	16.98
R^2(Adj.)	0.083	0.089	0.085	0.089
Sample	Incumbents	Incumbents	Incumbents 2013-2017	Incumbents 2013-2017

Standard errors in parentheses. Standard errors clustered at the firm and procedure levels.

* p<0.10, ** p<0.05, *** p<0.01

empirical tests for the idea that prior investments in capabilities influence which problems established firms undertake to solve.

As a baseline I include the firm's prior patenting experience on a focal medical activity (Column 1). Column 2 reports estimates of the probability a firm will patent on an activity in a year conditional on the technical capabilities the firm has to invent for a focal medical activity. Because firm capabilities are measured with respect to problems, I can include firm fixed effects to account for unobserved characteristics of the firm that might be related to a propensity to patent for an activity. The coefficient on *Tech overlap* indicates that a one standard deviation increase in the technical fit between a firm and the medical activity increases the probability of patenting by .41 percentage points. Given the baseline of 0.15, this is a meaningful increase. In column 3 I estimate the probability a firm will patent on an activity in a year conditional on the firms' investment in marketing to a given medical specialty. A $200,000 increase in marketing stock (one standard deviation) is associated with an increased probability of patenting of 0.04 percentage points, roughly a 33% increase in the probability of patenting.

The above results indicate that prior investments in capabilities is related to the probability established firms will undertake invention for a medical activity. This tendency to leverage prior investments in capabilities suggests that established firms might be more likely to undertake invention in more explored areas when they have these capabilities than when they do not, because the benefits of leveraging these capabilities outweighs the downside of being in more highly patented areas. I test this idea by interacting *Tech overlap* and *Marketing stock* with *Cum activity patents*$_{-f}$ to compare the probability a firm will undertake invention for a medical activity with many patents when they have related capabilities, compared to when they do not. Table 2.8 reports the results.

Table 2.8: Capabilities and invention in more explored areas

	(1)	(2)	(3)	(4)	(5)	(6)
DV =	Patent (0/1)	Patent (0/1)	Patent (0/1)	Patent (0/1)	Patent (0/1)	Patent (0/1)
Tech overlap	0.00498***	0.00497***	0.00745***			
	(0.000521)	(0.000809)	(0.000786)			
Cum activity patents$_{-f}$		-0.00473***	-0.00723***		-0.0206***	-0.107***
		(0.00134)	(0.00214)		(0.00445)	(0.0385)
Tech overlap × Cum activity patents$_{-f}$			0.0135***			
			(0.00119)			
Marketing stock				0.000447*	0.000446*	0.000807
				(0.000239)	(0.000238)	(0.000666)
Marketing stock × Cum activity patents$_{-f}$						0.00214**
						(0.00106)
Constant	0.00187***	0.00189***	0.00194***	0.00251***	0.00492***	0.0201***
	(0.0000238)	(0.0000415)	(0.0000468)	(0.000318)	(0.000555)	(0.00479)
Control for patent stock	Yes	Yes	Yes	Yes	Yes	Yes
Control for firm size	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Activity FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	2,764,500	2,764,500	2,764,500	890,150	890,150	890,150
F	38.229	17.039	32.147	13.731	17.519	8.50
R^2(Adj.)	0.039	0.040	0.101	0.022	0.028	0.026
Sample	Incumbents	Incumbents	Incumbents	Incumbents 2013-2017	Incumbents 2013-2017	Incumbents 2013-2017

Standard errors in parentheses. Standard errors clustered at the firm and medical activity levels.

* p<0.10, ** p<0.05, *** p<0.01

On average, established firms are less likely to undertake invention for medical activities with many existing patents by other firms. However, when established firms have made prior investments in capabilities they are *more likely* than when they have not to undertake invention for activities for which many other firms have patents. These results are suggestive that while firms generally seek to differentiate from other firms, prior investments in complementary capabilities incentivize firms to change which types of problems they undertake to solve.

2.5.2 Organizational Inertia

A different literature suggests that differences in the inventive activity between new and established firms might arise because of organizational factors. Prior research has argued that the organizational practices of established firms become increasingly inert as firms age (Leonard-Barton, 1992), making them less able to adapt to changes in the environment. As Sørensen and Stuart (2000) put it, "because they are less flexible, older firms may be less likely to incorporate the technological advances of other firms into their activity, effectively ceding the development of newer and potentially more influential areas of technology to start-up organization". Different mechanisms contributing to inertia in organizations have been proposed. Some mechanisms suggest established firms are inept at recognizing new opportunities, while other mechanisms suggest established firms may recognize new opportunities but lack the ability to respond. For example, over time organizational structure can come to mirror the architecture of a product, making it difficult for established firms to respond to drastic changes in product architecture (Henderson and Clark, 1990). Tripsas and Gavetti (2000) propose that rather than organizational features, it may be a cognitive bias against new business models that make established firms inept at recognizing new opportunities. These inertial forces bias established firms away from undertaking new activities. Furthermore, employees of established firms working on the technological frontier might develop new technology that the firm is unwilling or unable to exploit, providing an ideal situation for the employees to found a new firm to exploit the technology (Klepper and Sleeper, 2005; Agarwal et al., 2004).

Taking this theoretical framework seriously implies that new ventures and established firms will undertake to solve different problems, but for different reasons than prior investments in complementary capabilities. Specifically, this theory predicts a difference based on the inability of established firms to be flexible enough to explore new areas. Newer medical activities will by nature be less explored than older ones, which might account for the increased propensity of new ventures to patent in less explored areas. Taken to its logical extreme, this theory predicts two empirical patterns. First, experience patenting in an area should be associated with a higher probability of patenting in the future in the same are, and this relationship should be stronger for older firms than younger firms. Second, established firms should be less likely than new ventures to undertake invention in newer areas.

Inertia: Data and Analysis

To explore this empirically I calculate two variables. First, I calculate the prior experience a firm has patenting on a focal medical activity ($\textit{Firm experience}_{t-1}$). This is simply the stock of patents a firm has been granted for a medical activity. The average firm experience for a single medical activity is 33 patents, with a standard deviation of 174. Second, I calculate a proxy for the age of the medical activity by taking the number of years since the first patent on the activity. This is not a perfect measure of when the medical activity first came into existence, but it does measure the time since medical device firms began developing inventions for it.The average age of a medical activity is 22 years, with a standard deviation of 11.

First, I ask whether experience inventing for a medical activity is more strongly related to the probability of future patenting for established firms than new ventures. The notion is that if established firms are more inert, they will be likely to continue to work on what they have worked on in the past. The first column of Table 2.9 reports results. In column 1, I interact the firm's prior experience patenting on a medical activity ($\textit{Firm experience}_{t-1}$) with an indicator for whether the firm is a new venture (*Startup*). Results suggest that prior experience patenting on an activity is associated with an increased probability of patenting for that activity, and that this is stronger for new ventures than it is for established firms. This is opposite what we should expect if inertia is constraining what established firms undertake.

Next I ask whether the difference between new ventures and established firms is driven by the newness of the medical activity. If less explored medical activities are simply newer, and new ventures are more likely to undertake newer medical activities, this would explain the observed relationship. In the second column I add the age of the medical activity to the baseline results from my main specification. On average, the older an activity is the higher the probability a firm will apply and be granted a patent for it. Column 3 indicates that new ventures are indeed more likely to undertake invention for newer medical activities, but that this does not account for the main result that new ventures are also more likely to patent in less explored areas.

2.5.3 Technological Opportunity

Finally, a different body of literature, drawing on the economics of innovation, emphasizes the role of science as a means to lower the cost of invention. One reason some areas might be less explored is that in these areas it is inherently more difficult and costly to develop inventions. The precise definition of this concept, often referred to as "techno-

Table 2.9: Age of medical activity

	(1)	(2)	(3)
DV =	Patent (0/1)	Patent (0/1)	Patent (0/1)
Firm experience$_{t-1}$	0.00434***		
	(0.00111)		
Startup	0.000267**	0.000116	0.0000108
	(0.000135)	(0.000164)	(0.000158)
Startup $\times$ Firm experience$_{t-1}$	0.0952***		
	(0.0197)		
Cum activity patents$_{-f}$		-0.00354***	-0.00355***
		(0.00102)	(0.00103)
Startup $\times$ Cum activity patents$_{-f}$		-0.00322***	-0.00311***
		(0.000953)	(0.000926)
Activity age		0.000381***	0.000338**
		(0.000137)	(0.000137)
Startup $\times$ Activity age			-0.000410***
			(0.000148)
Constant	0.00126***	0.00144***	0.00145***
	(0.000117)	(0.000104)	(0.000103)
Control for patent stock	Yes	Yes	Yes
Control for firm size	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Activity FE	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes
Obs.	3,374,400	3,374,400	3,374,400
F	14.118	20.442	18.258
R^2(Adj.)	0.079	0.028	0.028
Sample	Full	Full	Full

Standard errors in parentheses. Standard errors clustered at the firm and procedure levels.
* p<0.10, ** p<0.05, *** p<0.01

logical opportunity", differs slightly among researchers. I adopt the definition by Jaffe (1986) of technological opportunity as "exogenous variations in the cost and difficulty of innovating in different technological areas". Scholars have proposed that greater scientific understanding in an area can facilitate greater opportunities for invention and decrease the cost of inventing in an area. Specifically, science can act as "a powerful heuristic guiding the search process associated with technological change" (Cohen, 2010). As Nelson (1982) argued, science narrows the set of inventive alternatives to pursue and focuses the attention of the inventor, making the research process more efficient [10].

This implies that firms incorporating science in their search processes will have a relatively lower cost of inventing in areas with low technological opportunity (i.e. less explored areas). Studies have shown that scientists often have a preference for self-employment and are willing to sacrifice financial rewards for scientific positions in more entrepreneurial ventures (Elfenbein et al., 2010; Stern, 2004). Additionally, recent work has shown a decline in the tendency of incumbents to conduct internal basic scientific research in recent decades (Arora et al., 2018). One explanation for the main result in this paper is that new ventures incorporating science into their inventions are driving the difference between new ventures and established firms. If this is the case, we would expect to see a difference in what problems new ventures and established firms undertake to solve primarily in cases where firms draw on scientific knowledge.

[10] One example of how science can increase the inventive opportunity for a medical device is the drug-eluding stent. Prior to drug-eluding stents, bare-metal stents were used to treat restenosis (the re-narrowing of a coronary artery). There are naturally a limited number of technical changes that can be made to a bare-metal stent, such as the type of metal used, the strut thickness, or the linkages between struts. However, the drug-eluding stent was invented by incorporating scientific understanding of cell reproduction to limit the high risk of restenosis with which bare-metal stents were associated. Thereafter, there were additional features of the stent with which inventors could experiment, such as the type and quantity of the drug or polymers used. In this case, science guided the search for a stent that could prevent restenosis, and was an important means of broadening "the inherent characteristics of the technology" (Jaffe, 1986) for stents.

Technological Opportunity: Data and Analysis

To explore this possibility I introduce two additional variables. First, I hand collect information on founders from LinkedIn, Crunchbase, and company websites. I code the origin of the new ventures based on the most recent positions held by the founders. For example, if one founder listed Medtronic as an employer immediately prior to listing their founding position on LinkedIn, I code the new venture as an incumbent spawned firm. Likewise, if the founder held an academic position prior to founding the firm, the new venture is coded as a university spawned firm. Second, I use the count of citations by a patent to non-patent literature, based on data made available by Marx and Fuegi (2019)[11]. This proxies as a measure of the extent to which a patent uses scientific knowledge.

Table 2.10 first presents the main result from Table 2.5 as a baseline. The second column breaks out *Startup* into three groups, university spawns, incumbent spawns, and new ventures that were not spawned by either (which I refer to as "de novo" ventures). If the main result is being driven by university spawned ventures then we should see a different in the interaction coefficients of these three type of ventures.

The interaction coefficients for university spawns and incumbent spawns are different at the ten percent level (p=0.051). University spawns appear to be slightly more likely than incumbent spawns to undertake invention in areas with fewer inventions. Additionally, the interaction coefficient for university spawns is different from that of de novo ventures, significant at the five percent level (p=0.016). These results suggest science might be related to undertaking invention for less explored medical activities. I investigate this further by asking whether new ventures of different origins are more or less likely than incumbent firms to use science in their inventions. Additionally, I ask whether less explored areas, which might be characterized by lower technological opportunity, are more likely to in-

[11]Data available for download at http://mattmarx.com/

Table 2.10: Breaking out new ventures by origin

	(1) Patent	(2) Patent
Cum activity patents$_{-f}$	-0.00352***	-0.00357***
	(0.00103)	(0.00102)
Startup	0.00012	
	(0.00016)	
Startup × Cum activity patents$_{-f}$	-0.00322***	
	(0.00095)	
De novo		0.000039
		(0.00016)
De novo × Cum activity patents$_{-f}$		-0.00319***
		(0.00089)
University Spawn		0.00020
		(0.00023)
University Spawn × Cum activity patents$_{-f}$		-0.00447***
		(0.00119)
Incumbent Spawn		0.00038
		(0.00024)
Incumbent Spawn × Cum activity patents$_{-f}$		-0.00221
		(0.00138)
Constant	0.00144***	0.00143***
	(0.00010)	(0.000148)
Control for patent stock	Yes	Yes
Control for firm size	Yes	Yes
Year FE	Yes	Yes
Activity FE	Yes	Yes
Firm FE	Yes	Yes
Obs.	3,374,400	3,374,400
F	19.76	11.79
R^2(Adj.)	0.028	0.028
Sample	Full	Full

Standard errors in parentheses. Standard errors clustered at the firm and procedure levels.
* p<0.10, ** p<0.05, *** p<0.01

Table 2.11: Use of science by new ventures

Dv =	(1) NPL cites	(2) NPL cites	(3) NPL cites	(4) NPL cites
Startup	4.149			7.177
	(5.265)			(7.911)
University spawn		-0.717		
		(1.893)		
Incumbent spawn		-1.363		
		(1.542)		
Denovo firm		8.783		
		(9.396)		
Cum activity patents$_{-f}$			-0.299	-0.313
			(0.237)	(0.245)
Startup $\times$ Cum activity patents$_{-f}$				-1.215
				(1.135)
Constant	4.243***	4.247***	5.534***	5.520***
	(0.349)	(0.349)	(1.091)	(1.079)
Year FE	Yes	Yes	Yes	Yes
Activity FE	Yes	Yes	Yes	Yes
Obs.	14,024	14,024	14,024	14,024
F	0.621	0.745	1.582	0.809
R^2(Adj.)	0.190	0.192	0.190	0.192
Sample	Full	Full	Full	Full

Standard errors in parentheses. Standard errors clustered at the firm and procedure levels.

* $p<0.10$, ** $p<0.05$, *** $p<0.01$

corporate science than more explored areas. Results reported in Table 2.11 do not provide support for these ideas.

2.5.4 Discussion of Theories and Results

Overall, the preceding empirical analysis is most consistent with the explanation that differences in which problems new ventures and established firms undertake arise from differences in complementary capabilities. Established firms are more likely to undertake invention for problems when they have related technical and marketing capabilities because prior investments in these capabilities changes the cost-benefit calculus. In general, established firms are likely to pursue invention for problems that have been less explored. However, it appears that the presence of complementary capabilities modifies the incentives of a firm to pursue invention in more explored areas. When a firm has made prior investments it more likely to pursue invention in more explored areas than when it is not. In contrast, new ventures have not made prior investments in complementary capabilities. Thus on average, established firms are more likely to undertake invention in more explored areas because they seek to leverage their complementary capabilities, despite the fact that these areas may have less "white space".

Two additional analyses produce mixed results. One reason some areas might be less explored is that they are relatively newer. Organizational theory predicts that new ventures will be more likely than established firms to pursue invention in these newer areas. I find support for this idea, but this does not account for the empirical observation that new ventures are more likely to undertake invention in less explored areas. Furthermore, prior experience appears to be more strongly related to what problems are pursued by new ventures than established firms. But theories of inertia suggest as firms age, experience should be more influential. Another reason some areas might be less explored is that they are relatively more costly to invent in. I consider whether the use of science as a means to lower the cost of invention might explain the observed difference between new and established firms. New ventures spawned by universities are likely to be closer to scientific discovery, lowering their search costs for inventing in these areas. While university spawned ventures are more likely than their non-university spawned counterparts to pursue invention in less explored areas, they do not appear to be any more likely to use scientific knowledge in their inventions. This finding is in line with recent research showing that established firms are just as likely as ever to use science in their inventions, even though they are less likely to perform it internally (Arora et al., 2018).

2.5.5 Demand Heterogeneity

A concern is that firms' decisions to undertake invention is endogenous to market demand (Schmookler, 1962). Higher demand for medical activities is expected to incentivize firms to invest in capabilities, and is also expected to increase the probability a firm will invest the time and resources to invent. Furthermore, these incentives could be underlying the observed interaction between firm capabilities and more explored areas, and the differences between new ventures and established firms. I take a first step to address this concern by matching medical activities to the 2018 Medicaid and Medicare reimbursement rates[12] for the subset that are reimbursable (e.g. there is no reimbursement code for taking temperature or checking lungs with a stethoscope). I then multiply this reimbursement rate by the estimated frequency of the medical activity provided by medical residents (Appendix C). This serves as a proxy for the market size of the medical activity.

Two additional steps to address this concern are in progress. First, the reimbursement rate in my data does not vary over time because I currently have only one year matched. I am in the process of collecting additional years' rates to account for changes in the reimbursement rate over time. As such, including fixed effects for the medical activity absorbs the coefficient for the reimbursement rate and market size. Second, implementing an instrumental variable can lessen concerns about the endogenous relationship between the number of patents for an activity and the market incentive. One potential instrumental variable is the affiliations between the rate setting committee and medical specialties. Recent work in healthcare economics has found that a stronger affiliation between the rate setting committee and a medical specialty results in a higher reimbursement rate for procedures in that specialty (Chan and Dickstein, 2019). If firms are unlikely to make patenting decisions based on the affiliations of the committee, this might serve as a good instrument

[12]Specifically, I match them to Current Procedural Terminology (CPT) codes. Each CPT code has an associated Relative Value Unit (RVU), which is used for billing Medicaid and Medicare.

for reimbursement rates.

Using the matched 2018 reimbursement rates I perform two analyses. The first analysis investigates whether heterogeneity in demand is partially explaining the propensity of new ventures to undertake invention in less explored areas. This would be problematic if new ventures, for example, are more likely than established firms to undertake invention for small markets that might not "move the needle" for larger firms (Christensen, 1997). In Table 2.12, column 2, I include an interaction between *RVU* and *Startup* to account for this possibility. The main interaction between *Startup* and *Cum activity patents*$_{-f}$ is stable when accounting for variation in demand. Results also suggest that in fact there may be differences between the propensity of new ventures and established firms to invent in areas with higher demand. Specifically, some evidence suggests new ventures may be more likely than established firms to undertake invention for activities with higher reimbursement rates. This would make sense if high entry barriers in the medical device industry justify entry only when there is a large incentive.

Table 2.12: Demand incentives for patenting

	(1)	(2)	(3)
DV =	Patent	Patent	Patent
Cum activity patents$_{-f}$	-0.00352***	-0.00344***	-0.00343***
	(0.00102)	(0.00106)	(0.00106)
Startup	0.000116	0.000266	0.000266
	(0.000164)	(0.000223)	(0.000223)
Startup × Cum activity patents$_{-f}$	-0.00322***	-0.00279***	-0.00280***
	(0.000953)	(0.000995)	(0.000988)
Startup × RVU		0.000349*	
		(0.000191)	
Startup × Market size			-0.000171
			(0.000264)
Constant	0.00144***	0.00190***	0.00190***
	(0.000103)	(0.000111)	(0.000111)
Control for patent stock	Yes	Yes	Yes
Control for firm size	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Activity FE	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes
Obs.	3,374,400	1,594,848	1,594,848
F	19.756	10.705	9.069
R^2(Adj.)	0.028	0.029	0.029
Sample	Full	Reimbursed	Reimbursed

Standard errors in parentheses. Standard errors clustered at the firm and medical activity levels.
* p<0.10, ** p<0.05, *** p<0.01

Table 2.13: Demand incentives and capability development

	(1)	(2)	(3)	(4)	(5)	(6)
DV =	Patent	Patent	Patent	Patent	Patent	Patent
Tech overlap	0.00744***	0.00850***	0.00850***			
	(0.000786)	(0.00110)	(0.00110)			
Cum activity patents$_{-f}$	-0.00554***	-0.00527***	-0.00549***	-0.0183***	-0.0199***	-0.0191***
	(0.00176)	(0.00195)	(0.00186)	(0.00391)	(0.00454)	(0.00423)
Tech overlap $\times$ Cum activity patents$_{-f}$	0.0136***	0.0136***	0.0136***			
	(0.00119)	(0.00187)	(0.00187)			
RVU $\times$ Cum activity patents$_{-f}$		0.000741			-0.00240	
		(0.00138)			(0.00439)	
Market size $\times$ Cum activity patents$_{-f}$			-0.000260			0.000657
			(0.000259)			(0.000457)
Marketing stock				0.000331*	0.000206	0.000205
				(0.000196)	(0.000174)	(0.000174)
Marketing stock $\times$ Cum activity patents$_{-f}$				0.000731**	0.00101***	0.00101***
				(0.000282)	(0.000238)	(0.000238)
Constant	0.00166***	0.00228***	0.00225***	0.00420***	0.00545***	0.00551***
	(0.0000872)	(0.000191)	(0.000129)	(0.000506)	(0.000821)	(0.000780)
Control for patent stock	Yes	Yes	Yes	Yes	Yes	Yes
Control for firm size	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Activity FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	3,374,400	1,594,848	1,594,848	988,000	466,960	466,960
F	32.175	15.733	16.685	14.262	15.65	16.13
R^2(Adj.)	0.095	0.086	0.086	0.027	0.028	0.028
Sample	Full	Reimbursed	Reimbursed	Reimbursed 2013-2017	Reimbursed 2013-2017	Reimbursed 2013-2017

Standard errors in parentheses. Standard errors clustered at the firm and medical activity levels.

* p<0.10, ** p<0.05, *** p<0.01

Second, I seek to understand whether the observed interaction between technical and marketing capabilities and the number of inventing firms is due to heterogeneous demand. For example, higher reimbursement rates cardiovascular procedures will incentivize firms to develop capabilities and also make them more likely to patent on devices used in these procedures. To account for this possibility I include an interaction between the reimbursement rate (*RVU*) and *Cum activity patents*$_{-f}$. I do the same using the estimated market size in place of the reimbursement rate. Importantly, these results are also restricted to medical activities for which there is a financial incentive, assuring that the results are not determined by differences between reimbursable and non-reimbursable activities. Results are reported in Table 2.13. The main interactions between capabilities and the number of inventing firms is stable, suggesting that the propensity of firms to patent in more explored areas when they have related capabilities is not primarily due to market incentives.

2.6 Conclusion

Do new ventures and established firms differ in the problems they undertake to solve? If so, why? This paper investigates these questions in the U.S. medical device industry. I find that, on average, firms are more likely to undertake invention in areas with fewer patents by other firms. However, this tendency is stronger for new ventures than established firms: new ventures are significantly more likely than established firms to undertake invention in less explored areas. Additional analysis suggests that this is due to prior investments in capabilities by established firms, which change the cost-benefit calculus of undertaking invention in areas with a greater number of existing patents.

This paper has limitations. First, resources and capabilities are not randomly allocated to firms. This introduces selection bias into my estimates because firms are selecting into which capabilities to develop, and consequently opens the door for various alternative explanations. I seek to address one of these explanations – that heterogeneous market demand and might partially explain my findings. While I do find some evidence that this influences which problems firms work on, these factors do not seem to explain my results. Still, these results are not causally identified and should be interpreted with caution. Second, I do not directly observe the problems firms pursue, only those for which a patent was eventually granted (i.e. successful outcomes). This implies that my results may be biased if a firm 1) doesn't apply for a patent on a problem it is working on, or 2) isn't granted a patent for a problem it is working on. This concern is reduced partially by two facts . First, patents are a very important means of IP protection in the medical device industry, more so than any other industry (Arora et al., 2016), so it is unlikely patents will not be applied for. Second, the vast majority of patents which are applied for are granted (Quillen Jr and Webster, 2005), with estimates around 80-97%. Lastly, my study focuses on a single industry. This is helpful because it allows me to be precise about how I measure a "problem" and to obtain the risk-set of problems firms can work on. The trade-off is that these results may or may not generalize to other settings. Additional work needs to be done in additional settings to verify the robustness of these results.

My research contributes to several streams of innovation and strategy research. First, I contribute to the literature on the direction of inventive activity (Nelson et al., 1962). By assembling a unique dataset of the population of problems that firms can work on I am uniquely positioned to look upstream in the inventive process to understand differences in which problems are worked on, rather than taking the problem as given. To my knowledge, this has not previously been done. This paper also contributes by better demonstrating the sources of inventive diversity in society. Cohen and Klepper (1992) propose a model in which many small firms (as opposed to a few large firms) in an industry provide greater diversity in the ideas pursued, but that this comes at the expense of the amount of effort put into any given idea. My paper offers empirical support for this theory; new ventures do provide diversity in what is pursued in society. Indeed, one reason new ventures might often portrayed as introducing more novel innovations is that they tend to work on different problems altogether. This is particularly important in a setting such as medical devices, where solutions to customer problems have a direct bearing on longevity and quality of life.

This paper also offers a different perspective on the observed differences in inventive activity between new and established firms. While scholars have proposed that a "division of innovative labor" is possible under certain conditions, empirical research on this has been limited to the biotechnology (Arora and Gambardella, 1995), semiconductor (Hall and Ziedonis, 2001), and automobile (Lee and Berente, 2012) industries. My study adds to this by demonstrating a division in the inventive activity between new and established firms in the U.S. medical device industry. Additionally, existing research on the division of innovative labor emphasizes the role of strong intellectual property rights and markets for technology. In contrast, my study proposes that such a division might arise from strategic considerations.

Finally, strategy research has long been interested in how resources and capabilities shape strategic behavior of firms (Klepper and Simons, 2000; Bayus and Agarwal, 2007). I add to this literature by offering another important strategic choice to consider – which problems firms undertake to solve. This question is a central component of value creation, and therefore of strategic importance to the firm (Felin and Zenger, 2015). My paper uses a novel dataset to show that firm capabilities shape which problems firms undertake to solve, and by implication, where they create value. For established firms, it appears that the benefits of being able to leverage complementary capabilities outweigh the disadvantages of inventing in more competitive areas.

In conclusion, I explore differences in inventive activity between new ventures and established firms. I do so by considering inventions in terms of the problems they solve, or their end-use. This paper suggests that in the U.S. medical device industry, strategic factors, specifically prior investments in capabilities, influence which problems firms undertake to solve.

Chapter 3

Predicting Prediction: Evidence of the diffusion of artificial intelligence in healthcare

3.1 Introduction

Artificial intelligence (AI) technology is predicted to affect almost every aspect of the economy, including healthcare (He et al., 2019), financial services (Fuster et al., 2018), manufacturing (Wuest et al., 2016), legal services (Alarie et al., 2016), and defense. This has raised questions about how the diffusion of AI will affect labor markets, the nature of occupations, and who the ultimate winners and losers will be (Brynjolfsson and McAfee, 2011). Current research investigating the consequences of AI on labor ranges from "business as usual" to near-apocalyptic. For example, Manyika et al. (2013) estimate that increases in automated knowledge work could equate to productive output of between 75–90 million knowledge workers in advanced economies (the total U.S. labor force is around 165 million). Frey and Osborne (2017) estimate that 47% of jobs in the United States could be replaced by computerization technology. A less aggressive estimate was suggested in a study by the OECD, which found only 9% of jobs are at risk of replacement (Arntz et al., 2016). (Mann and Puttmann, 2018) project that although manufacturing jobs will decline, this will be made up for by the increase in service sector jobs. And Brynjolfsson et al. (2018) suggest that all occupations will be impacted to a degree, but none completely.

While these estimates are helpful, they are speculative in nature. No work to-date has measured the actual diffusion of AI in the most recent decade. This is understandable for at least two reasons. First, although AI has been around since the 1950's, the recent resurgence of interest in the field is still relatively new. Data collection efforts have begun, but are still in their infancy (McElheran, 2019). Second, because AI lives in computer code, it is difficult to observe. We do not see AI show up in administrative data like imports and exports, and survey's might be an insufficient means to measure adoption because of the ambiguity around AI's definition. However, it is important to find ways to measure actual adoption because the gap between what could happen and what does actually happen can be large. New technology does not automatically and inevitably diffuse, let alone at a uniform rate across industries, firms, and occupations. Adoption is influenced by factors such as individual and firm incentives, complementary technologies, and inertia. As Dian Baily put it at a recent conference on the impact of AI, "technological progress, or lack thereof, is shaped by people making choices." (Fisher, 2016).

The purpose of this paper is to provide empirical evidence of the actual diffusion of AI and investigate the implications for workers. In doing so I make an empirical contribution to the current research on the future of work by operationalizing and validating a measure of the diffusion of AI. Leveraging a comprehensive database of medical terminology, I assemble a dataset of the population of tasks healthcare providers can perform (e.g. MRI, Spinal Fusion, etc.). I then use a machine learning algorithm developed by the National Library of Medicine to link USPTO medical device patents to these tasks. Next, to observe the application of AI to medical devices and the tasks for which they are used, I construct a dictionary of specific AI algorithms (e.g. support vector machine, artificial neural network) as well as general AI terminology (e.g. learning algorithm, unsupervised learning) and search for these terms in the title, abstract, and claims of USPTO medical device patents. This allows me to observe inventions that explicitly incorporate AI as a component of the invention. Linking these medical device patents to the medical task being performed with the device allows me to observe the application of AI to tasks and the types of workers performing them.

My empirical analysis first validates my measure of the diffusion of AI. Similar to (Autor et al., 2003) – who made simple observations about what computers do - I acknowledge that the technology underlying AI is fundamentally prediction technology (Agrawal et al., 2018). Consistent with this, my measure shows tasks requiring higher levels of prediction are more likely to have AI applied than those requiring less prediction. Additionally, I find that as the performance of AI algorithms improves, the rate of AI application increases, and that this is more pronounced for tasks requiring high levels of prediction. After verifying the measure, I look at the differences in the application of AI to tasks performed by different occupations in healthcare. I find that on average, no occupation is more likely than others to have AI applied to its tasks. However, high-prediction tasks performed by physicians are less likely to have AI applied to them than high-prediction tasks of other occupations. Conversely, high-prediction tasks performed by technicians are more likely to have AI applied to them than high-prediction tasks of other occupations. These findings are interesting in light of the recent concern about the potential for AI to substitute for the tasks of physicians. On one hand, AI is shown to impact more complex tasks (such as non-routine tasks requiring tacit knowledge) than those impacted by previous digital technologies (which performed repetitive, codifiable tasks). This aligns with the concerns about the potential of AI to impact high-skill jobs. On the other hand, application of AI has been lowest for the highest skills jobs, and most pronounced for middle-skill occupations.

3.2 Background

3.2.1 The Technology of Artificial Intelligence

To ground the discussion in concrete ideas, I first discuss what I mean by artificial intelligence (AI) in this paper. This is important because the term has been used to mean different things to various people in different contexts. It is also important to recognize that AI technology includes hardware, or computing power, algorithms, and data. It is the integration of all three that make AI possible. However, this paper focuses primarily on the algorithms as the defining feature of AI. Because the field of artificial intelligence is broad, below I provide a very brief overview of two algorithmic approaches to AI, contrast historical uses of the term with what is commonly meant by its use today, and explain how I will use the term in this paper.

The field of AI encompasses disciplines such as computer science, psychology, and philosophy. In the broadest sense, there have been two major approaches to AI: symbolic systems and statistical or machine learning. Each has many sub-fields, such as expert systems and logic systems in the former, and evolutionary computation and Bayesian learning in the latter. Historically, symbolic systems were the champion of the AI landscape, having been given the nickname "Good Old-Fashioned Artificial Intelligence", or GOFAI for short (Haugeland, 1985). An example of this approach is the machine built by I.B.M. named Deep Blue – the machine that beat chess world champion Garry Kasparov in 1997.

Expert systems require creators to embed knowledge explicitly in the system, in the form of facts, rules, and relationships. Consequently these AI systems only "know" what they are "told". Along this vein, Autor et al. (2003) made this general observation about computing almost two decades ago, six years after the victory of Deep Blue

> Tasks demanding flexibility, creativity, generalized problem-solving, and com-

> plex communications–what we call non-routine cognitive tasks–do not (yet) lend themselves to computerization (Bresnahan, 1999). At present, the need for explicit programmed instructions appears a binding constraint.

In contrast to GOFAI, statistical learning methods do not necessarily require the creator to explicitly tell the AI what it "knows". Instead, these approaches learn from observation and develop their own "beliefs" and rules for decisions. The quality (i.e. accuracy) of those beliefs and decision rules then depends on the data the algorithm observes, and the type and architecture of the algorithm. Currently the use of the term AI in public discourse most often refers to these statistical learning approaches. This is largely due to significant advances in this branch of AI over the recent decade, as will be discussed below.

One way to interpret the basic functionality of this approach (statistical learning) to AI is that of prediction. In an NBER working paper, Agrawal et al. (2018) suggest that AI technology (i.e. statistical learning) is fundamentally a prediction technology. Prediction can be defined as taking available data and transforming it to provide an estimate about unavailable data. The unavailable data need not be in the future, individuals often confront unavailable data on current scenarios, in which case we might refer to this technology by another name, such as classification. One important feature of this is that prediction and classification tasks often require tacit knowledge, or knowledge that is difficult or impossible to explicitly state. For example, the ability to recognize a familiar face goes beyond the ability to write down that individuals eye color, hair color, and other features. The limited explainability of the task is partially what makes it complex. These are the types of tasks for which AI is designed. From this point on in this paper, when using the term AI I will be referring to statistical learning, with the connotation that it primarily functions as a prediction technology.

3.2.2 Technological Change and Labor

Concerns that technology will reduce the demand for labor have existed at the latest since the early 19^{th} century. On the face of it, this is understandable because technology is often invented with the stated objective of automating tasks. Yet we have seen unprecedented advances in technology in the last two centuries, without an accompanying decline in the aggregate need for individuals to perform work. Rather, some jobs have been created while others have disappeared, and the skills required to perform most jobs have evolved. Autor (2015) points out that the reasons for the long run sustainability of the need for work is that in addition to technology substituting for tasks, it also complements other tasks, and raises productivity such that rising earnings generate demand for new types of labor. However, Autor (2015) also acknowledges and provides evidence of the polarizing effect technology can have on labor markets . For example, a substantial body of work has found that new technology has historically complemented high skill workers, raising their productivity and increasing demand for these these workers (Autor et al., 2003; Berman et al., 1994; Autor et al., 1998; Machin and Van Reenen, 1998; Berman et al., 1998; Gera et al., 2001; Levy and Murnane, 1996; Doms et al., 1997; Bresnahan et al., 2002).

From this perspective, the recent "automation anxiety" (Akst, 2013) over white-collar jobs being replaced by AI appears to be unjustified. There is substantial evidence that, historically, technological change has benefited high-skill workers more than it has hurt them. However, the future does not necessarily look like the past. As described above, AI technology differs in important ways from prior technologies. The routine tasks performed by computers in the 1990's consisted mainly of mathematical and other logical operations (hence the name, which means "one who calculates"). Today, the tasks being performed by AI were previously thought to include too much judgement, tacit knowledge, or contextual understanding to ask of a machine. The performance of AI algorithms has achieved superhuman performance in areas such as strategy games (Silver et al., 2017; Brown and Sandholm, 2018), image recognition (He et al., 2015), and medical diagnosis (Abbott, Abbott). It is not clear at the moment whether AI performing these more complex tasks will continue to complement high-skill labor, and if so how. From this perspective, the anxiety over white-collar jobs being replaced appears reasonable.

Because of the uncertainty about how it is going to play out, the question of how AI will affect labor has received much attention recently. The general approach to studying this question has been to create a score measuring the suitability or likelihood that an occupation will be impacted by AI. For example, Webb (2019) develops a score indicating the "exposure" of an occupation to AI, Frey and Osborne (2017) calculate the probability of computerization for occupations, and Felten et al. (2018) develop a measure of the suitability of machine learning for specific tasks, and make predictions about the impact of AI on the occupation based on the tasks that are involved. One advantage of this approach is that it is general enough to be applied to any occupation in the economy. This approach is also attractive because it directly links capabilities of the technology to the skills required to perform a set of tasks. However there is little systematic evidence at the moment about how AI technology has actually diffused in the most recent decade, and how that relates to occupations at different skill levels.

3.3 Setting and Data

I explore this question within the context of U.S. healthcare. This setting has many advantages. First, it is particularly relevant to the question this paper seeks to address, as concerns about the impact of AI on the future of certain physician specialties has been a recent topic of debate (Liew, 2018; Davenport and Dreyer, 2018; Langlotz, 2019; Choy et al., 2018). Another advantage of healthcare is the availability of a codified database of *tasks* performed by workers. This is different from many of the current and historical approaches to studying the affect of AI on labor, which use *skills* necessary to carry out a job as the unit of analysis. By contrast, this paper uses specific tasks (i.e. medical procedures) performed by a worker, and link these tasks to the inventions (i.e. medical devices) used to perform the task. This can be an important distinction because it provides insight into the way the technology is being applied. For example, AI is skilled at identifying objects in images, but that skill applies to a variety of tasks and occupations. Rather than observing what AI is capable of, I observe the tasks to which AI has been applied. Lastly, prior research has found that patents are used extensively for medical device inventions (Cohen et al., 2000; Arora et al., 2016), making this an appropriate means for observing the application of AI to inventions.

3.3.1 Data Construction

The data were constructed following the process outlined in Chapter 1 of this dissertation, with additional steps unique to this chapter described below. The set of tasks used in this paper comes from the Unified Medical Language System (UMLS), a database of medical terminology established in 1986 and maintained by the United States National Library of Medicine for use in medical informatics (see for example Bodenreider 2004; Wu et al. 2012; Carrell et al. 2017). This database includes a comprehensive list of the tasks performed by medical professionals. For example, it includes complex medical procedures such "Pulmonary Transplantation", as well as simple tasks such as "Auscultation", the act of listening to a patients lungs using a stethoscope. These tasks are then linked using a machine learning algorithm to USPTO medical device patents. The machine learning algorithm was developed by the National Library of Medicine to facilitate indexing medical publications. Using this algorithm has many advantages. First, the algorithm uses a combination of feature engineering approaches, such as bag-of-words, with more sophisticated approaches such as n-grams, noun phrases, and related publications (Yepes et al., 2015). Second, the creators have tested and evaluated various types of learning models in the algorithm to improve the performance of the text classification (Yepes et al., 2015)[1]. Additional data on the medical tasks was collected by having medical residents rate each task on various dimensions. See Chapter 1 of this dissertation for detail on the process.

USPTO patents from January 2000 to November 2017 were downloaded from PatentsView.org, a platform supported by the USPTO. My sample is restricted to medically related patents based on Cooperative Patent Classification (CPC) codes. A detailed list of the included CPC codes can be found in Appendix B. This results in 380,225 unique medical device patents. Applying the algorithm to the title and abstract text resulted in 141,377 (37%) patents being linked to diagnostic or therapeutic medical tasks. Inspecting the unmatched patents revealed they are for devices not associated with a particular medical procedure (e.g. hospital bed software, patent number 5787528), or they are patents on drugs (e.g. US patent number 8906940). My final dataset contains a balanced panel of 21,168 task-year observations from 2000 to 2017.

Last, I code each patent as being an AI patent or not based on the patent text and the technological classes of the patent. Because there is no standard agreement about what constitutes AI technology, discretion in selecting which terms and patent classes to include is unavoidable. I selected the terms based on two criteria. First, as discussed in my definition of AI outlined above, I consider AI as primarily statistical learning models. Therefore, I include general terms such as "machine learning" and "supervised learning", etc. Second, from a technical perspective, the use of the term AI today is commonly associated with the use of artificial neural networks. Therefore I also include search terms for names of specific algorithms, such as "convolutional neural network" and "recurrent neural network".

[1]See https://ii.nlm.nih.gov/MTI/ for additional details.

Likewise, patent classes were selected based on their reference to statistical learning and artificial neural networks (e.g. G06N 3/08 COMPUTER SYSTEMS BASED ON SPECIFIC COMPUTATIONAL MODELS – Learning methods). See Appendix F for the full list of search terms and patent classes included. Appendix G provides the abstracts of two AI patents for illustration.

3.3.2 Variables

Table 3.1 provides a description of the key variables in the analysis and their summary statistics. My outcome variable of interest is $\textit{AI Patents}_{ij}$. This measures the application of AI to medical task j in year i. I include an indicator for whether the medical task is diagnostic, and use this to validate my measure of AI, as described below. Additionally, indicators for the type of medical professional are included. Last, I include two characteristics of the medical tasks that might be associated with the application of AI. More frequently performed tasks provide an incentive for firms to invest in the application of AI because of its low marginal cost once developed. Additionally, AI algorithms require high volumes of data on which to train. More frequently performed tasks are more likely to have high volumes of data, which facilitate the application of AI. More general tasks – those that are performed in a greater number of hospital units – may also be likely to provide incentives for firms to invest in AI technology, because it allows the firm to spread its fixed investment in the technology over a greater number of applications and/or units.

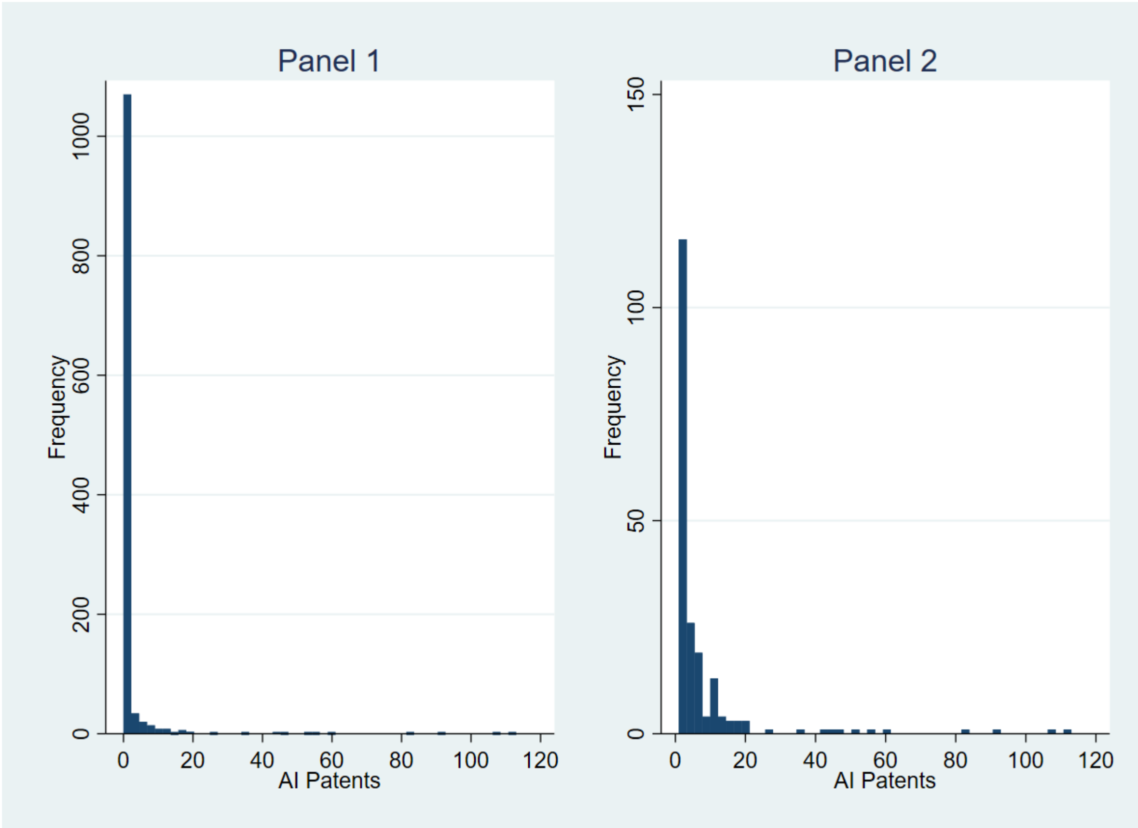
3.3.3 Descriptive Statistics

Unsurprisingly, the application of AI to medical tasks is a relatively infrequent event. Figure 3.1 shows the distribution of AI patents across medical tasks is highly skewed; only 17% of the tasks have at least one AI patent, and of those that have at least one, 70% have

Table 3.1: Variable Descriptions

Variable	**Description**	**N**	**Min**	**Max**	**Mean**	**SD**
Dependent Variables						
AI Patents$_{ij}$	The count of patents granted that incorporate AI technology	21,168	0	25	0.074	0.604
Independent Variables						
Diagnostic$_j$	Binary indicator for diagnostic tasks.	21,168	0	1	0.230	0.421
Medical doctor$_j$	Binary indicator for whether the task is performed by medical doctors.	21,168	0	1	0.178	0.382
Nurse$_j$	Binary indicator for whether the task is performed by nurses.	21,168	0	1	0.581	0.493
Technician$_j$	Binary indicator for whether the task is performed by technicians.	21,168	0	1	0.258	0.438
Controls						
Frequency$_j$	The approximate number of times a task is performed in a hospital in a week.	21,168	0	62,222	760	3,762
Generality$_j$	The approximate number of hospital units in which the task is likely to be performed.	21,168	0	100	5.136	10.565

Figure 3.1: Distribution of AI Patents



five or fewer. Panel 1 of Figure 1 shows the complete distribution, and Panel 2 excludes tasks with zero patents.

Physicians perform the largest share of the tasks, but these tasks are on average much less frequent than the tasks performed by nurses and technicians. As we would expect, the tasks of physicians also tend to be more specialized (i.e. performed in fewer units of a hospital) than those performed by nurses. However, the tasks performed by technicians are as specialized as those of physicians. The share of tasks performed by physicians that are diagnostic is about the same as for nurses, but lower than the share for technicians. Table 3.2 provides these descriptive statistics and Table 3.3 provides correlations between variables.

Table 3.2: Occupation Summaries

Occupation	Share of total tasks	Mean frequency	Mean generality	Share diagnostic
Medical doctor	0.58	581	4.20	0.21
Nurse	0.10	1280	12.1	0.18
Technician	0.26	745	4.17	0.30
Other	0.06	1662	7.16	0.20

Table 3.3: Correlations Between Variables

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1) AI Patents	1.000						
(2) Diagnostic	0.120	1.000					
(3) Medical doctor	-0.060	-0.047	1.000				
(4) Nurse	-0.002	-0.041	-0.383	1.000			
(5) Technician	0.071	0.092	-0.695	-0.193	1.000		
(6) Frequency	0.029	0.078	-0.056	0.045	-0.002	1.000	
(7) General	0.000	0.104	-0.105	0.215	-0.054	0.341	1.000

3.4 Empirical Analysis

One objective of this paper is to provide empirical evidence of the actual, not speculative, diffusion of AI. Therefore, the empirical analysis that follows is organized into two sections. First, I validate my measure of AI diffusion based on two facts of AI technology. 1) As described above, AI is primarily a prediction technology, and 2) the performance of AI improved dramatically around 2010. Second, I explore differences in the application of AI to tasks of healthcare providers. In all analyses, results are interpreted as correlations; the overarching goal is to observe a phenomenon that is characteristically difficult to observe, and describe what is there.

3.4.1 AI Measure Validation

As discussed above, much of AI in the present day falls within the sub-field of machine, or statistical, learning. This technology is fundamentally about the ability to take existing data and provide some insight into unavailable data. This is exactly what healthcare providers do when they diagnose a patient: a patient enters a physicians office with certain data available – the symptoms – but neither the patient nor the physician initially knows the underlying cause of the symptoms. The task of the physician to "predict" the underlying cause of the symptoms. To the extent that my measure of the diffusion of AI accurately captures the technology I have in mind, there should be a strong relationship between *AI Patents* and *Diagnostic* tasks. In other words, diagnostic tasks should be significantly more likely than therapeutic and preventative tasks to have AI patents. Results of this test are reported in Table 3.4.

Table 3.4: Difference between the application of ML to diagnostic and therapeutic tasks

	OLS		Logit		ZINB	
	AI Patents$_{ij}$		I(AI Patents)$_{ij}$		AI Patents$_{ij}$	
	(1)	(2)	(3)	(4)	(5)	(6)
Diagnostic$_j$	0.1726***	0.1645***	3.8789***	3.8015***	1.8590***	1.6235***
	(0.0590)	(0.0559)	(0.7476)	(0.7187)	(0.2553)	(0.2332)
Constant	0.0344***	0.0395***	.0223***	.0083***	-1.6197***	-2.9531***
	(0.0087)	(0.0097)	(.0028)	(.0022)	(0.2827)	(0.3102)
Control for frequency	No	Yes	No	Yes	No	Yes
Control for generality	No	Yes	No	Yes	No	Yes
Year FE	No	Yes	No	Yes	No	Yes
Occupation FE	No	Yes	No	Yes	No	Yes
Obs.	21,168	21,168	21,168	21,150	21,168	21,168
F / Wald Chi^2	8.620	3.530	49.46	323.48	53.01	519.67
R^2(Adj.)	0.014	0.029	0.048	0.105		

Notes: Standard errors in parentheses. Standard errors clustered on medical task.

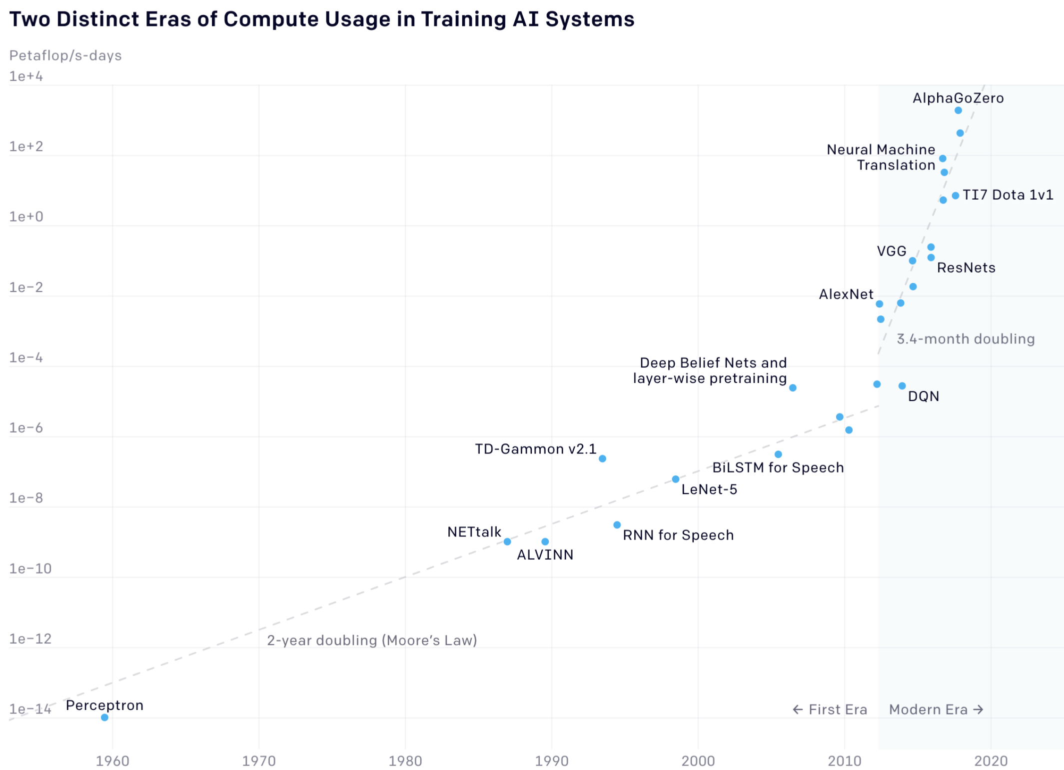
Odds ratios shown for the Logit models.

* p<0.10, ** p<0.05, *** p<0.01

Three types of statistical models are reported, each first without any controls, and then including controls. Across the difference specifications, the coefficient on *Diagnostic* is fairly stable when including controls. Ordinary regression estimates are provided in models 1 and 2. On average, diagnostic tasks have about .17 more AI patents than therapeutic and preventative tasks. Given the mean task has .07 AI patents, this seems to be an economically significant difference. Models 3 and 4 use a binary dependent variable equal to one if the task is granted an AI patent in a given year. The odds ratios reported indicate that diagnostic tasks are roughly 3.8 times more likely than therapeutic and preventative tasks to have AI applied. Last, because the dependent variable *AI Patents* has many zeros, I use a zero-inflated negative binomial model to check robustness. Results are qualitatively identical to the other models.

A second test of the measure corresponds to the performance of AI technology over time. Around 2009-2011, there were significant advances in the performance of AI, as well as an increasing number of tools (e.g. machine learning libraries) that lowered the cost of development and adoption. This increase in the performance of AI algorithms came primarily from artificial neural networks with many layers, which subsequently came to be known as "deep learning". In addition to the anecdotal evidence of these performance improvements, such as Geoffrey Hinton's team winning the ImageNet competition in 2012 by a 10.8% margin of victory, a 41% improvement on the next best algorithm, there is some systematic evidence that AI improved dramatically around this time period. One approach to documenting the performance of AI algorithms was undertaken by OpenAI, a not for profit research organization based in San Fransisco, CA. Figure 1 shows the amount of compute to train various AI algorithms over time, which they suggest is an appropriate proxy for the performance of these algorithms, because increases in compute are typically associated with better performance Hestness et al. (2017) (see https://openai.com/blog/ai-and-compute/ for the original report.). According to OpenAI, the post-2012 period can be

Figure 3.2: Compute Used to Train AI Models

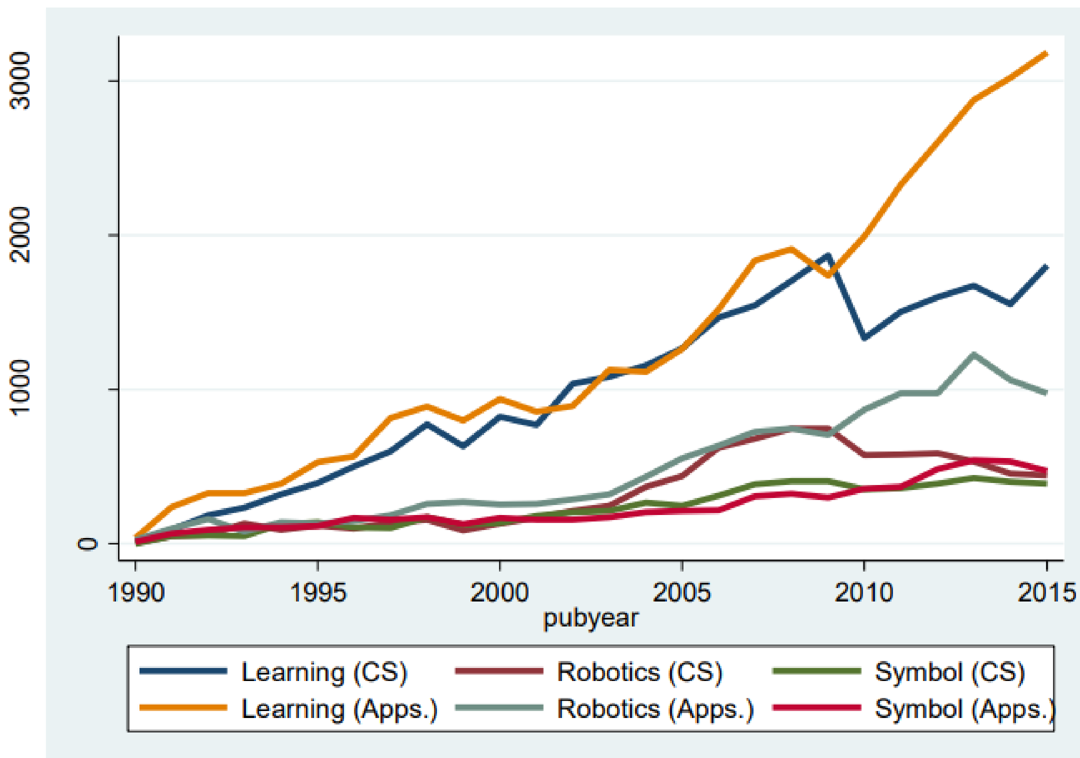


Source: https://openai.com/blog/ai-and-compute/

defined as a different era of AI. This is because "since 2012, the amount of compute used in the largest AI training runs has been increasing exponentially with a 3.4-month doubling time." Figure 3.2 illustrates these two distinct era's.

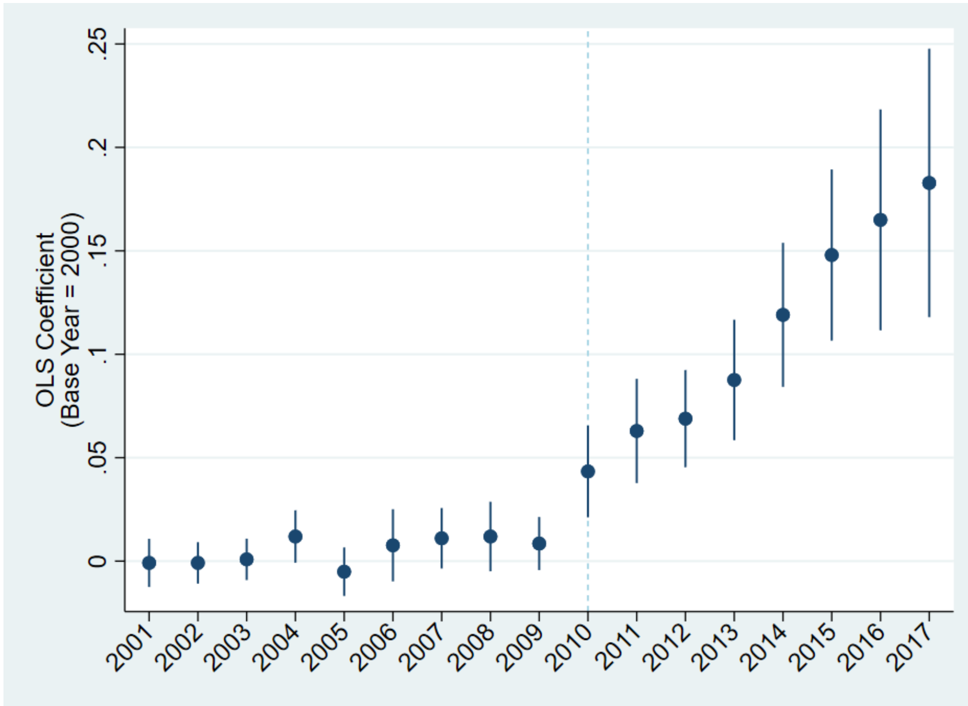
Another systematic analysis comes from Cockburn et al. (2018), who find a rise in application-oriented machine learning research after 2009, as well as a rise in AI patents around the same time. Figure 3.3, taken from Cockburn et al. (2018), plots the count of publications in scientific journals versus application journals, by different areas of AI. Application oriented machine learning publications grow at a faster rate after 2009 than all other fields, and in particular, faster than scientific machine learning publications, with which it had previously kept pace.

Figure 3.3: Publications in Computer Science versus Application Journals, by AI Field



Source: Cockburn et al. (2018)

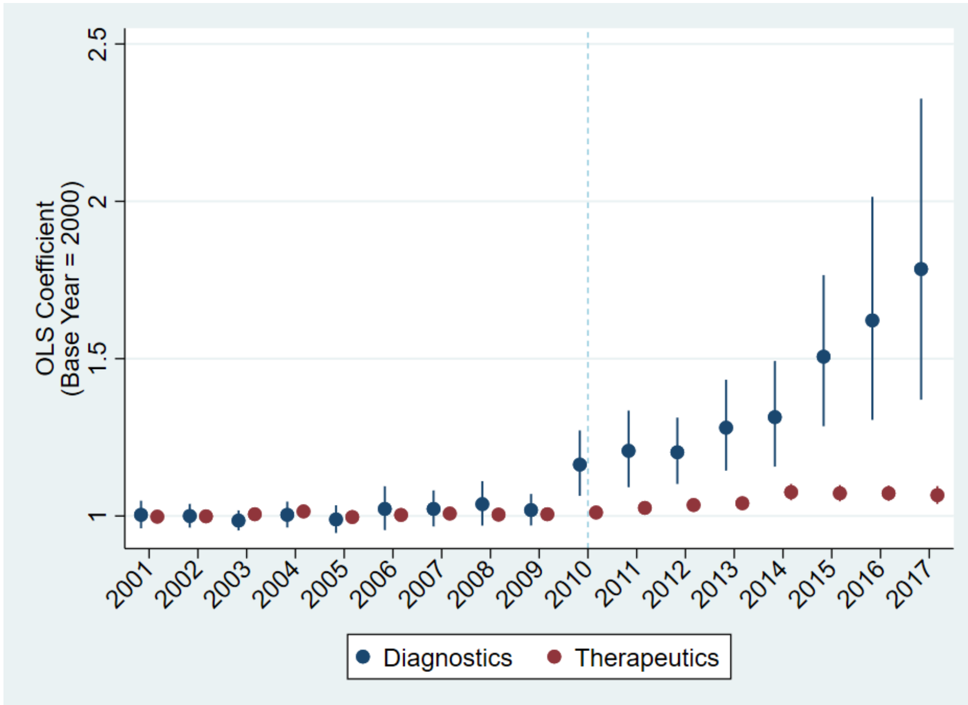
Figure 3.4: OLS coefficients of AI patent count on year



The foregoing evidence suggests something changed – either improvements to the algorithms or a reduction in cost to deliver them, or both – around 2009-2011. I compare my data to the trends provided by others and find a consistent pattern. Figure 3.4 plots OLS coefficients of each year on my dependent variable, *AI Patents*, with 2000 as the base year. From 2001 to 2009 there is no change in the level of AI patenting, after which there is a consistent increase the AI patenting each year.

As a final set of checks, I combine the prior two facts that AI is primarily prediction, and that the performance of AI dramatically improved around 2010. Overall, I find that the advances in AI after 2010 were associated with higher rates of application to diagnostic tasks, relative to therapeutic and preventative tasks. Figure 3.5 builds on Figure 3.4 by breaking out diagnostic tasks from therapeutic and preventative tasks. Figure 3.6 shows the share of total tasks that have at least one AI patent granted in each year, by task type.

Figure 3.5: OLS coefficients of AI patent count on year, by activity type



Lastly, Table 3.5 provides the results of a difference in difference model, supporting the notion that AI impacted diagnostics to a greater extent after 2010. Combined, these results lend validity to my measure of AI diffusion.

3.4.2 Occupation Analysis

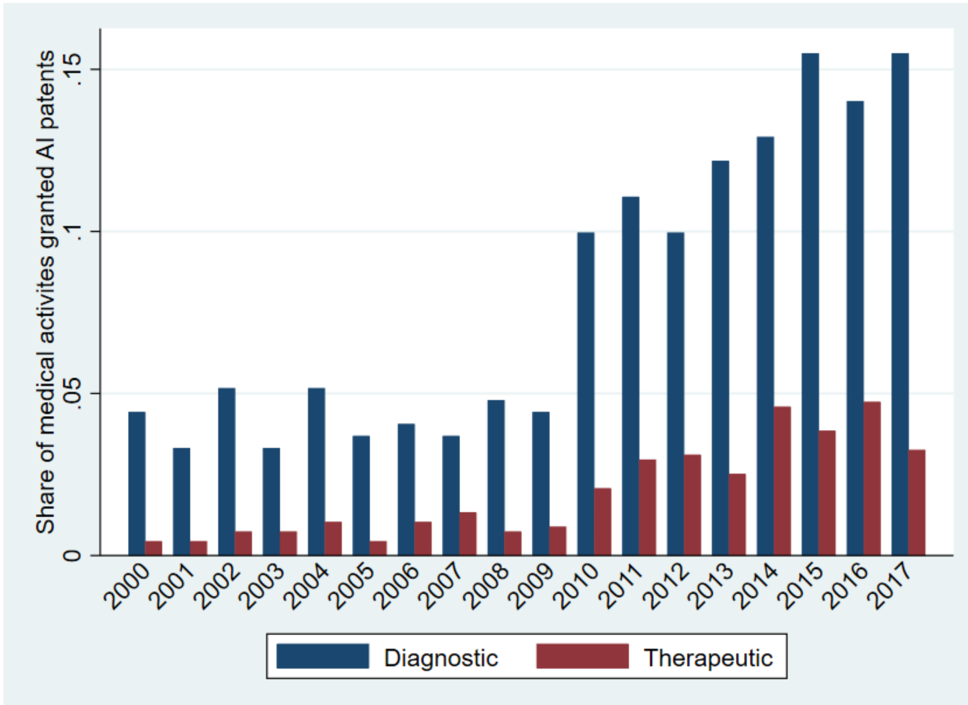
How is the diffusion of AI affecting occupations in health care? Following prior work in this vein, I conceptualize an occupation as a collection of tasks (Autor et al., 2003; Felten et al., 2018; Frey and Osborne, 2017; Webb, 2019; Brynjolfsson et al., 2018). Prior work has emphasized the distinction between routine tasks, those which can be fully codified, and non-routine tasks, which include some element of tacit knowledge (Autor et al., 2003; Autor, 2015). As discussed above, computers substituted for labor performing routine tasks in the 1980's and 1990's because computers are designed to perform well defined,

Table 3.5: Diagnostics see greater application of AI than therapeutics post 2010

	AI Patents$_{ij}$
Diagnostic$_j$	0.0606***
	(0.0192)
Post 2010	-0.0105
	(0.0150)
Diagnostic$_j$ $\times$ Post 2010$_i$	0.2672***
	(0.0690)
Secular trend$_i$	0.0067***
	(0.0015)
Constant	-13.359***
	(3.0958)
Control for frequency	Yes
Control for generality	Yes
Occupation FE	Yes
Obs.	21,168
F	9.37
R^2(Adj.)	0.036

Notes: Standard errors in parentheses. Standard errors clustered on medical activity.
* $p<0.10$, ** $p<0.05$, *** $p<0.01$

Figure 3.6: Share of activities granted an AI patent, by activity type



codifiable tasks. By contrast, AI is designed to perform tasks that contain some element of tacit knowledge: an AI is "trained" to be able to recognize a specific individual's face, for example, without the need to explicitly tell the AI every detail and feature of that face – the AI learns those features on its own. This fundamental difference is exactly why advances in AI have become a topic of interest[2]. Furthermore, the preceding analysis provides support for the idea that AI is diffusing to tasks that are non-routine in nature: medical diagnosis. In fact, Autor et al. (2003) explicitly classify medical diagnosis as a non-routine task.

Table 3.6 provides correlational evidence regarding which occupations in health care are being affected the most by the diffusion of AI. I use a linear probability model in all specifications for ease of interpretation, though logistic regression produces qualitatively

[2]A widely used example of a non-routine task in healthcare that an AI ostensibly can replace is that of a radiologist identifying cancerous tumors, which requires years of training to gain the tacit knowledge required for the task.

equivalent results. The first model shows the probability AI will be applied to the tasks of different occupations. I include physicians, nurses, and technicians, with “other” as the baseline group. Model 1 suggests that the tasks of the included occupations do not differ from the baseline group in their propensity for AI application; physician tasks are equally likely as nurse or technician tasks to have AI applied. Because AI technology has diffused to diagnostic tasks at a higher rate than other types of tasks, I next interact the occupations at the high and low end of the education spectrum, Medical Doctor’s and Technicians, with an indicator for whether the task is diagnostic. Doing so provides some suggestive evidence that there are differences in the propensity of AI to be applied to the diagnostic tasks of physicians, relative to technicians. Model 2 indicates that AI is likely to be applied to diagnostic tasks on average, but that this is significantly less so for the diagnostic tasks performed by physicians. By contrast, the diagnostic tasks performed by technicians are more likely to have AI applied to them (Model 3).

These findings suggest that, while AI is more capable than previous technology at handling non-routine tasks, the concerns around which occupations are most likely to be initially impacted might be misplaced. One interpretation of these results is that diagnostic tasks performed by physicians are more complex than the diagnostic tasks performed by technicians, making them more difficult for AI in its current form to handle. For example, a tuberculin (TB) test is a diagnostic task which nurses often perform. Diagnosing TB is not always straightforward (Kiwanuka, 2005), but this is significantly less complex than diagnosing Lyme Disease.

3.5 Discussion and Conclusion

The purpose of this paper is to provide quantitative evidence for the actual diffusion of AI in the past decade, and insight into how the actual diffusion of AI is likely to affect

Table 3.6: Probability of ML application to tasks requiring different levels of education

	(1) I(AI Patents)	(2) I(AI Patents)	(3) I(AI Patents)
Medical doctor$_j$	-0.0153	-0.00249	-0.0142
	(0.0147)	(0.0154)	(0.0147)
Nurse$_j$	-0.00315	-0.00251	-0.00408
	(0.0182)	(0.0185)	(0.0181)
Technician$_j$	0.00284	-0.000254	-0.0222
	(0.0158)	(0.0157)	(0.0151)
Diagnostic$_j$	0.0559***	0.0898***	0.0258**
	(0.0126)	(0.0218)	(0.0108)
Medical doctor$_j$ $\times$ Diagnostic$_j$		-0.0614**	
		(0.0235)	
Technician$_j$ $\times$ Diagnostic$_j$			0.0965***
			(0.0291)
Constant	0.0309**	0.0237	0.0357**
	(0.0141)	(0.0148)	(0.0141)
Control for frequency	Yes	Yes	Yes
Control for generality	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Obs.	21,168	21,168	21,168
F	4.061	3.509	3.770
R^2(Adj.)	0.034	0.039	0.044

Standard errors in parentheses. Standard errors clustered on medical task.

* p<0.10, ** p<0.05, *** p<0.01

the future of occupations in healthcare. In doing so it complements the current body of work speculating how advances in AI will affect the future of work (Frey and Osborne, 2017; Brynjolfsson et al., 2018; Felten et al., 2019; Webb, 2019) by providing empirical support for which tasks are most susceptible to AI. New technologies have sometimes been predicted to have devastating effects on the labor market. While short term adjustments and changes in the composition of occupations are often significant, technology has not yet eliminated the need for workers on a large scale (Autor, 2015). Concerns that AI might do just that are often supported by the observation that AI technology is different than previous technologies, and therefore what we know about the past might not be an appropriate guide for the future. This paper builds on the observation that AI technology is different than prior technology, and provides evidence to argue *how* AI differs, and for whom that is likely to matter.

Specifically, I find that AI has been applied to prediction tasks at a higher rate than non-prediction tasks. This builds on and provides support for the observation that AI is fundamentally a prediction technology (Agrawal et al., 2018). It also supports the notion that AI is different from previous digital innovations, in that it is being applied to tasks that are more complex. I also find that the rate of AI application to diagnostic tasks increases as the performance of AI increases. Finally, the diagnostic tasks performed by physicians have been significantly less affected by these AI applications than the diagnostic tasks of technicians, suggesting that concerns over the automation of high-skill jobs might be misplaced.

One limitation of this paper is that it stops short of analyzing real changes in the ultimate outcome of interest: employment. There are various reasons for this. First, this paper shows that the application of AI (in its most current form) in patents is still relatively new. Patents are the first of many steps in the diffusion process, particularly in healthcare. The devices incorporating these patented technologies must also receive approval from the

FDA, be listed on the approved medical supply list of the hospital, and be purchased and adopted by the physician group/hospital. Second, even after these technologies are in use, adjustments in the labor market will not be instantaneous. Given that the application of AI is still relatively new, we will likely not see direct effects in the labor market for some time.

How the adoption of AI in healthcare will eventually affect demand for various types of labor is an opportunity for future research. The answer will depend on many factors, but will at least partially depend on the extent to which AI technology complements or substitutes for various tasks, the real changes in productivity, and the elasticity of demand for health services. If AI technology acts as a substitute in the tasks performed by medical technicians, then we might see a continuation of what has been called the "hollowing out" of the middle-skill labor market (Dao, Das, Koczan, and Lian, Dao et al.; Autor, 2015). The diagnostic work historically performed by technicians will become increasingly performed by AI. This should improve the marginal productivity of physicians, who often use the output of technicians as inputs into their tasks. These forces combined could lead to lower demand for technicians and higher demand for physicians. Alternatively, AI substituting for medical technicians could lead to higher overall demand for technicians, as was the case for human tellers in the banking industry when banks began adopting ATM's (Bessen, 2015). This resulted because the increase in productivity induced additional branch openings, which required hiring human tellers. The increase in hiring from the improved productivity outweighed the substitution effect. Whether demand for medical technicians rises or falls as AI technology is adopted will also depend on the elasticity of demand for health care services: Highly inelastic demand for health services would imply that even as physician productivity increases, hospitals may not demand more services.

Chapter 4

Walk and Talk: The interplay of work interdependencies, spatial distance, and face-to-face communication

4.1 Introduction

A vast body of scholarly work subscribes to the notion that the *raison d'être* of firms (a.k.a. hierarchies) is to coordinate the integration of knowledge that is dispersed across a number of organizational groups (March and Simon, 1958; Thompson, 1967; Galbraith, 1973; Burton et al., 1984; Bolton and Dewatripont, 1994). This is because firms, relative to markets, offer efficient channels for organizational members to communicate and share knowledge that is vital for the replication of routines (Nelson and Winter, 1982a; Nickerson and Zenger, 2004) and innovation (Grant, 1996; Kogut and Zander, 1992). The mechanisms that influence internal communication are therefore of central concern for scholars who study groups in organizations (March and Simon, 1958; Henderson and Clark, 1990).

One such organizational design feature that has received particular attention is the spatial distance that separates group members who are part of the same firm (Allen, 1977; Van den Bulte and Moenaert, 1998). It has been more than four decades since Allen (1977) published a series of seminal studies showing that the frequency of face-to-face communications between scientists in a corporate lab were highly sensitive to spatial distance. Allen also suggested that communication was affected by whether or not employees shared an organizational bond (e.g. membership in the same group). Allen stressed, however, that "space in most organizations is allocated on a group basis, with people with similar background or people working on the same or similar tasks located near each other" (Allen, 1977).

Very little work has elaborated on Allen's original observation and separated location choices from work interdependencies in order to isolate the impact of spatial distance on communication. We know that low spatial distance reduces communication costs and facilitates the type of casual, face-to face, interactions that help integrate and re-combine knowledge (Allen, 1977; Kabo et al., 2014; Kraut et al., 2014). We also know that high work interdependence increases mutual knowledge (Cramton, 2001) as well as knowledge relevance (Schulz, 2001), thus working interdependently makes it easier for members of different groups to know what knowledge they share in common and how it can be combined (Grant, 1996; Kretschmer and Puranam, 2008). However, we do not know enough about the extent to which work interdependencies interact with spatial distance to shape informal work communication across groups.

We address this issue by examining how the frequency of informal work communications between members of different groups is jointly influenced by their relative position within the firm's formal work structure and the spatial distance between them. To do this we focus on the division of labor inside a firm (March and Simon, 1958) and argue that employees who are within more adjacent segments of the production process are increasingly likely to engage in work-related conversations. This is because members of different groups who are proximate in terms of 'work interdependencies' – more of their tasks intersect in the broader production process – are likely to have more relevant information to share with one another. Furthermore, group members are not motionless objects, but rather they move across the office space in patterns that mirror their work interdependencies (Monge et al., 1985).

Along these lines, members of different groups who are spatially distant but close in terms of work interdependencies will be the most motivated to move across the office space to communicate. We expand on this idea and suggest that employees tend to casually 'bump' into each other due to overlaps in boss-determined walking paths and, hence, are frequently exposed to low-cost opportunities to communicate face-to-face. We argue that the formal work structure of the firm constitutes a powerful driver of informal work communications between members of different groups and influences both the flows of relevant information (i.e. increasing the potential benefits for employees to interact) and the motion within an office space (i.e. increasing the opportunities for low-cost interactions).

To shed light on these conceptual arguments, we conduct a field study and examine during a 7-month period the communication patterns of 115 group members located in the headquarters of a financial services company. Our focus is on communications between members of different groups that are unplanned and conducted face-to-face. Archival data are used to capture the structure of work interdependencies and spatial distance between all the participants in our study. Based on our analysis of 13,677 dyadic interactions, we find support for the notion that employees who occupy more adjacent positions within the company's work structure (i.e. who have greater work interdependencies) engage in more frequent unplanned, face-to-face communication with members of other groups compared with those who have fewer work interdependencies. In addition, employees who are further away from one another are less likely to have informal work communication with members of other groups.

We take advantage of a change in the company's formal work structure that occurred 3 months into our study period to examine whether communication between members of different groups who increased in their work interdependencies also increased the frequency with which they communicate. Along these lines, our analysis presents evidence to suggest that work interdependencies can help overcome the barriers that spatial distance imposes on informal work communication by directing employee traffic across the office space. Namely, we find that members of different group who share greater overlap in the walking paths to their respective bosses (which are at least quasi-randomly determined, since office planners did not consider how walking paths of different group members will overlap when assigning seats) engaged more frequently in informal work communication.

4.2 Prior Literature

Allen's (1977) seminal book compiles a number of studies examining the flows of ideas and knowledge within a corporate laboratory. This work focused on the factors that influence communication patterns between employees, such as spatial distance and organizational bonds (e.g. whether a pair of scientists belongs to the same research group). One of the most well-known findings of Allen's work is that communication between employees decreases exponentially with distance. Namely, he finds that the probability that a pair of scientists engages in weekly communication drops from 0.40 to 0.15 when the distance that separates them is increased from 5 to 10 meters and that it nearly breaks down after 30 meters.

A vast body of scholarly work has also examined the relationship between spatial distance and communication patterns (Kiesler and Cummings, 2002; Olson, 2000). The basic premise behind these literature reviews is that, for a number of reasons, spatial distance makes communication more costly. This has partly been attributed to the opportunity cost of the time required to cover the distance that is required to engage in such interactions, but also to the fact that distance reduces the probability of unplanned opportunities to communicate (Chown and Liu, 2015). Spatial distance negatively impacts the frequency and quality of communications, but also makes it difficult for (new) connections to be made.

More recent work in this vein has been concerned with disentangling the endogenous relationship between choices of spatial location and unobserved interdependencies between employees. The logic is that people who are more prone to collaborate or engage in ongoing communication are more likely to collocate next to each other (Allen, 1977). If this is the case, the effects of distance on collaboration are mostly capturing the effects of the location choice, which are naturally confounded with the effects of being within close proximity. Scholars have started to address this shortcoming by randomly assigning the location of employees and/or exploiting natural experiments that help in achieving this same goal. For example, (Hasan and Bagde, 2015) randomized students to dormitories in an Indian college and found that collocation with well-connected students had a lasting effect over the breadth of the social interactions of roommates.

With few exceptions, scholars have studied the influence of spatial distance on employee interactions in settings where work interdependencies between employees are not pronounced. This is far from being a trivial omission, since we have learned from a number of studies that the boundaries of the firm (Monteverde, 1995), business and functional units (Nobel and Birkinshaw, 1998), and teams (Griffin and Hauser, 1992) are often determined with an eye toward streamlining the communication necessary to accomplish the established work goals. A common finding is that employees who need to be in constant communication and who require transferring complex (tacit) knowledge are often brought together under the same organizational unit, be it a firm, a department, or a functional group.

Scholars have also taken a more dynamic approach and argued that communication patterns inside an organization evolve to 'mirror' the technical interdependencies that characterize the work performed by groups in the organization (Henderson and Clark, 1990; Clark and Baldwin, 2000). For example, based on their seminal study of the photolithographic equipment industry, Henderson and Clark (1990) provide a vivid account of how the communication patterns between engineers and managers evolved to follow the technical relationships behind the components that integrate to make a product. That is, if component X interacts with component Y but not with Z, then those group members in charge of developing component X will mostly communicate with those groups developing component Y and not those groups developing component X. The mapping between internal communication and the architecture of the product became so pronounced that, over time, companies failed to recognize important cues that fell outside the communication channels delineated by the technical interdependencies.

4.3 Theory and Hypotheses

A firm can be viewed as a system composed of groups of employees who perform a series of mostly interdependent tasks to obtain a common goal (March and Simon, 1958; Simon, 1978). The interdependencies between the work performed by a pair of group members, say '1' and '2', implies that the value of the tasks carried out by '1' is different when '2' performs its work versus when '2' does not (Galbraith, 1973; Puranam and Raveendran, 2013). Such dyadic interdependencies can be scaled up and reflect the chain of work relations that underlies the division of labor that characterizes a firm.

One of the main problems a firm must solve is to coordinate the integration of the work performed by different groups (Nickerson and Zenger, 2004). There are a number of mechanisms for coordinating the integration of work, such as clustering highly interrelated tasks into modules (Sanchez and Mahoney, 1996; Ethiraj and Levinthal, 2004) and/or facilitating ongoing communication between the groups that share interdependencies. Communication is at the center of the research of several scholars who study how coordination occurs inside organizations and has been referred to as the vehicle by which interdependent actors engage in 'mutual adjustments' (Thompson, 1967), provide each other with 'feedback' (March and Simon, 1958), sustain 'technical dialogue' (Monteverde, 1995), and address ‘exception management’ (Pentland and Rueter, 1994).

Along these lines, scholars have described the internal organization of the firm as a ‘communication network’ by which information is exchanged between interdependent groups to solve complex problems (Bolton and Dewatripont, 1994). Ongoing communication has been viewed as the prime mechanism by which employees who are part of different functional groups adjust their work to changing conditions (Lawrence and Lorsch, 1967) and integrate their specialized knowledge (Grant, 1996; Kogut and Zander, 1992).

Prior work makes a clear distinction between intra-organizational communications that imply face-to-face interactions and those sustained via other means, such as telephone or email (see Olson (2000) for review). These studies suggest that face-to-face communications are the most potent channel by which employees can coordinate (Daft and Lengel, 1986; De Meyer, 1991; Kraut et al., 2014). The reason for this is threefold: First, face-to-face communication facilitates the transfer of complex information. Some information is next to impossible to codify without eroding its meaning, as its transfer often requires ongoing dialogue. Second, it is mainly through personal interactions that employees forge the kind of trusting relationships that are required to exchange valuable knowledge and to develop common language that facilitates coordination in the future. Third, face-to-face communications can sometimes be spontaneous (unplanned), particularly when employees are collocated in the same office space. This opens up the possibility for communication not only to be more opportune, but also to occur more frequently.

We argue that work interdependencies (vertical and horizontal) and spatial distance jointly influence the frequency with which collocated employees engage in face-to-face communication. To do this we assume that the members of the firm are bound by a complex system of work interdependencies designed to streamline the integration of their specialized outputs. That is, the division of labor, the allocation of tasks to employees, and their spatial location within the office space has been designed by a ‘boundedly rational’ designer (March and Simon, 1958). We assume that the designer has some imperfect idea of the optimal settings for these parameters and, hence, employees have to engage in joint search to determine the frequency with which they should engage in face-to-face communication. These choices are going to be influenced by the perceived benefits and costs of engaging in such interactions, which (among other factors) are influenced by the relative position of employees within the system of workflows and the office space.

4.3.1 Spatial Distance and Informal Work Communication

Consistent with prior research, we view the spatial distance that separates members of different groups within an office space as a deterrent of face-to-face communications. Spatial separation introduces a number of costs for those who wish to communicate. The most obvious cost is the time it takes people to walk from their own workspace to that of the person they wish to interact with. This is time that is spent away doing employee work or that is not devoted to interacting with other interdependent colleagues.

Another cost introduced by spatial distance is that of coordinating when to communicate. As denoted by Allen's (1977) early work, proximity has been found to be a powerful element for employees to interact, particularly when these have limited control over how they allocate their time. This is because those who are in close spatial distance can monitor each other's time availability at a lower cost and, hence, find small fragments of downtime to engage in communication. Such coordination devices are as simple as knowing whether a communication partner is at her seat and whether she is taking a break or is between activities and, hence, has time for a conversation. The opportunities to do this are reduced as employees are located further away from each other, since the lack of basic sensorial awareness implies that employees have to set time aside to monitor communication partners to find a moment that is mutually suitable. Thus, as previous studies have found, we expect that:

> Hypothesis 1: Members of different groups who are further apart in spatial distance engage in informal work communication less frequently than members of different groups who are closer in spatial distance.

4.3.2 Work Interdependencies and Informal Work Communication

For members of different groups, the benefits and costs of engaging in unplanned, face-to-face communication will be associated with their relative position in the system of interdependent workflows that characterizes the firm. We use the term work interdependencies to refer the extent to which a pair of employees share direct work interdependencies (i.e. A depends on the work of B to perform her work and/or vice versa) in the organization.

We expect the benefits of informal work communication to be high when members of different groups have more work interdependencies. This is because they are likely to hold information and knowledge that is both relevant and directly applicable to the work performed by them. For example, when the work of employees is closely related (i.e. as when they are in adjacent segments of the production process), they can exchange information to increase the joint fit between their work outputs. This exchange, on one hand, can help reduce the need to make costly adjustments once employees have completed their work and, on the other hand, has the potential to improve the quality of the integrated output. Either way, employees who are more interdependent in terms of workflow have a better chance to effectively synchronize their actions.

These benefits are likely to be eroded as work interdependencies decrease. This is partly because the information possessed by members of different groups, whereas still related, may not be as directly applicable to the work they do. Employees can give each other a ‘heads-up’ about what is happening in their portion of the workflow; however, this information is likely to be incomplete if it does not incorporate the status of the intermediate pieces. As work interdependencies decrease, this gap is likely to become wider and wider, increasing the disconnect between the information possessed by a pair of people. For this same reason, as the work performed by members of different groups becomes less interdependent, transferring (relevant) information is likely to become costlier. This is because the amount of mutual knowledge (Cramton, 2001) shared by a pair of employees decays as the work they do becomes only marginally related. Transferring information requires providing more context for this to take meaning and, hence, implies more effort on behalf of the employees who are conducting the exchange.

In sum, we expect the net benefits of engaging in informal work communication to increase as work interdependencies increase. Whereas employees who are in relatively adjacent segments of the internal workflow hold information that is mutually relevant and share sufficient context to absorb and act on it, those whose work is less closely related are likely to find the content of information less useful and also the lack of mutual knowledge is prone to make interactions costlier. We therefore predict that:

> Hypothesis 2: Members of different groups who have greater work interdependencies engage in more informal work communication than employees with fewer work interdependencies.

4.3.3 The Interplay of Work Interdependencies, Spatial Distance, and Informal Work Communication

A number of studies suggest that spatial distance is a multidimensional construct that goes beyond the walking distance between the designated places of work of a pair of employees (e.g. seat, cubicle, office) to include the overlap between broader spatial zones in which they do their work. In an early study of the social interactions between WWII GIs enrolled in post-war education, Festinger et al. (1950) provide a detailed account of how communication was highly influenced by whether people shared common spaces such as hallways, staircases, and toilets. People casually collided in these areas, opening up frequent opportunities to engage in ongoing communication. One of the main insights of this study is that, unlike stationary objects (e.g. buildings), employees have motion and, although they may spend a large portion of their time in a particular location, they also circulate throughout a broader space as they carry out their daily routine. Therefore, when examining how spatial distance influences the interactions between a pair of employees, it is just as important to account for the overlap in the broader areas they frequent as it is to account for the spatial distance of their work stations.

The insight provided by Festinger et al. (1950) has been revived in a series of subsequent studies. One of the most recent accounts is the work of Kabo et al. (2014) and colleagues, who study interactions between scientists in a research university and find that employees who have overlapping walking paths within a building are more likely to collaborate in research endeavors. The authors explain this finding by suggesting that the overlap in workspace leads to frequent chance encounters that trigger opportunities for employees to communicate (e.g. meeting in the staircase or the cafeteria). Over time, these chance encounters lead to the formation of common knowledge between employees and, eventually, to more formal collaborations on research projects. That is, the coinciding work paths creates a condition where, for reasons already incorporated into the routine of employees (e.g. walking to the cafeteria), they are exposed to low-cost opportunities to communicate and exchange information with peers with whom the benefits of interacting were previously unknown.

Similar principles are likely to affect the patterns of informal work communication inside hierarchical organizations; however, we expect that, in contrast to settings where employees have more autonomy over their actions (e.g. within colleges or academic research settings), in a firm, workflows constitute the engine of employees' motion within the office space. This idea has been vividly described in a number of studies that suggest that work within a firm is inherently dynamic. (Monge et al., 1985) observed the interactions between employees in a technology firm and found that the spatial distance of employees changed a great deal during the day, as employees moved about through the office space and were only in their work stations (offices) in the early hours of the morning (just after arriving to work) and in the late afternoon (just before heading home). Similarly, Whittaker et al. (1994) describe how engineers and managers at Hewlett-Packard sustained frequent interactions with colleagues as they were on the move. Namely, the authors found that around 31% of an employee's work time was spent having such unplanned work conversations, of which more than a third were held in a public space (e.g. hallway) or on the move (e.g. on the way from an employee's desk to someone else's office).

We build on these findings, and focus on the walking paths of employees that are likely to be present in hierarchical firms. In particular, we argue that the boss is likely to be a source of walking paths in the physical office-space. Our logic is simple: the movement of employees within an office space will be driven by the exchange of information with those who have the greatest work interdependencies. In many cases, this is the boss – whether it is attending scheduled meetings in the office or stopping by when necessary throughout the day. These movements create walking paths that are frequently transited by employees to do their work, thus are shaped by the spatial location of their boss (i.e., very strong work interdependence). In following these routines, employees are likely to 'bump' into members of different groups driven by their own work patterns, generating low-cost opportunities for them to engage in informal work communication.

This overlap in walking paths may be due to the fact that, when members of different groups are simultaneously on the move, they share common space (e.g. hallways) to reach their respective bosses. Alternatively, the overlap may be because an employee's walking path ends in close spatial distance to the work space of the other person's boss. That is, employees may 'bump' into others who are spatially close to the workstations of those whom they traveled to see (who are, presumably, in close proximity to the employee in terms of work interdependence). Since the overlap in walking paths to their respective bosses is likely to generate low-cost opportunities for informal work communication, we expect:

> Hypothesis 3: The larger the overlap between the walking paths of members of different groups to their respective bosses, the more frequently they will engage in informal work communication.

4.4 Methodology

4.4.1 Research Setting

Our study was conducted in a Mexican financial services company, which we will refer to as FinServe, during the first half of 2016. Located in Mexico City, FinServe was founded in 2011 and offers financial solutions to large corporate and government clients. FinServe was an ideal setting for our study because it experienced rapid growth in headcount in late 2015 into the first half of 2016. The rapid growth in headcount caused changes in the organizational structure and work processes of the company. In particular, the company performed a formal reorganization midway through the study, at the end of April 2016. We interviewed employees at all levels of the organization to understand the motivation for the restructuring. These interviews involved two executives directly responsible for the planning and execution of the restructuring, as well as directors and managers who were both directly affected by it and not affected by it. We asked each of them,"What was the purpose of the reorganization?" We found that the initial idea for restructuring the company came from the need for the CEO to reduce the number of his direct reports. This started a conversation among executives about ways to improve efficiency as the company continued to grow. Consistent across the interviews was the idea that the restructuring was a result of the growth of the company. Our informants noted that the rapid growth posed inefficiencies and the organizational change was necessary to accommodate that growth.

In order to facilitate data collection and validation, and also to gain qualitative insights into the nature of the workflow processes, one of the authors was located in the Mexico City office for 2 months (mid-June through mid-August 2016). The author sat with a new product development team each day and frequently interacted with employees at all levels and in virtually every division of the company.

4.4.2 Data Collection

Communication Surveys

A survey instrument was administered in January 2016 (Time 1) and again in July 2016 (Time 2). The survey asked respondent i to report the frequency with which he or she engages in informal work communication with each member (j) of the organization. A pilot survey was sent via email to six employees before the January survey was launched, and the final survey was sent to 157 employees. Of the 163 surveys emailed, 140 yielded responses, for an 86% response rate. Two of the responses were dropped due to technical issues with the responses stemming from the survey software, for a final sample of 138.

The follow-up survey was structured identically to the first survey, with additional questions at the end relating to the reorganization and its impact on the employees' work tasks. The follow-up survey was sent by email to 236 employees and received 189 responses (80%). Of the 138 final responses from the first survey, 19 employees were no longer working in the company at the time of the follow-up survey. Of the remaining 119 who completed the first survey, 115 (97%) completed the follow-up survey. A test comparing the means of observable characteristics (age, tenure, level in the company, etc.) between respondents and non-respondents shows no support for sample selection bias.

Because we are interested in informal work communication across groups, we restrict the sample to 13,677 (out of 15,072 possible) dyads that connect different groups in the organization. Each group is made up of roughly 9 members (M = 8.52, SD = 5.99), and all members of each group report to the same boss. These groups reflect the smallest work units in the organization, and there are strong work interdependencies within each group. By focusing on the informal communication between members of different groups (rather than within the same group), we can gauge the extent to which different groups in the organization interact with one another.

Organizational Characteristics

The Human Resources department of FinServe produces a formal organization chart at the end of each month reflecting the current organizational structure of the company. These were provided to the authors for each month of the study. Additionally, an up-to-date list of personnel is produced twice per month which includes the employee's title, the name of their boss, their division, their sub-division, and the start date of the employee. These two sources were used in combination to create a month-by-month snapshot of the organizational structure of the company. Employee membership in functional groups was validated using interviews when the information on the organizational chart and personnel list did not match.

Spatial Data

FinServe is located on the 12th and 13th floors of an office building in Mexico City. The floorplan was a modern "open" floor space in which the only offices were adjacent to the exterior walls and were reserved for senior management and meeting rooms. The office manager provided architectural drawings that included seating charts for both floors. To collect information about where employees sat, a research assistant performed a census of the employees in March 2016. Employees were shown a map of the office and asked which seats they sat in and for how long, all the way back to their first seat in the office. This method was applied again in July 2016 to get updated information about any employees who had moved and also to gather information on recently hired employees.

Employee Characteristics

FinServe also provided access to employee application materials. These application packets provided details such as the employee's birthday, their place of birth, their educational background, and their prior work experience.

4.4.3 Variable Construction and Measurement

Dependent Variable

Our variable of interest, informal work communication, comes from the survey on communication. The survey asked each employee employee (i) to report the frequency with which they communicate with each employee (j) in the company by chance to discuss work-related issues. This instrument has been used by numerous researchers studying unplanned face-to-face communication (Allen, 1977; Conrath, 1973; Festinger et al., 1950; Kraut et al., 2014; Monge et al., 1985). Each question was rated on a Likert scale from 1 to 5, 1 indicating that they never communicate and 5 indicating daily communication. The frequency of communication in Time 2 is used as our main dependent variable.

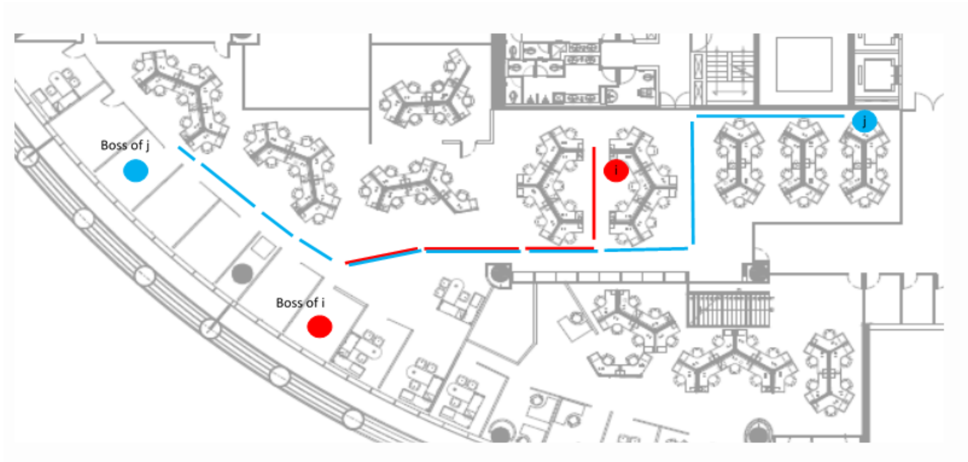
Independent Variables

Spatial distance. The distance between each pair of workstations was measured manually by drawing walking paths on the floor plans and calculating the shortest path between them. The measurements were converted from feet to steps assuming 2.5 feet per step. These paths and measurements were verified in person by one of the authors and updated according to the observed paths used in the office during the qualitative phase of the study.

Work interdependencies. This is measured as how close two employees are in the organizational structure. The logic is that in hierarchical firms, the division of labor captures which sets of work activities are most closely related. So groups within the same unit have more work interdependencies compared with groups in another unit, and groups within the same division have more work interdependencies compared with groups in another division. Using the organization chart, we computed work interdependencies on a 10-pt scale, with 1 indicating unrelated activities (e.g., two employees are as far apart as possible in the chart) and 10 indicating highly related activities (e.g., two employee are in the same supervisory group, within the same unit, and within the same division).

Walking Path Overlap. To measure the interplay of work interdependencies and spatial distance, we observe the walking path a group member must take to speak face-to-face with his or her boss. To calculate this, we divided the major walking paths in the office into segments and recorded which segments each employee passes through on their way to their direct boss. Next, we found which walking segments were mutual within each dyad. Finally, using the length of the overlap segments, we calculated total overlap in path distance divided by the total path distance from employee i to the boss of i. Figure 4.1 provides an illustration.

Figure 4.1: Boss Walking Path Overlap



Controls

A number of controls were included to account for employee characteristics as well as characteristics of the dyad. Specifically, we wanted to account for shared characteristics of i and j, indicating some social or other similarity. For employee characteristics, we include controls for age, gender, tenure in the company, level in the company, and dummies for division, for both i and j. For similarities between i and j, we control for their absolute difference in age, their absolute difference in tenure, whether they are in the same functional group (e.g. HR, Accounting, etc.), whether they attended the same university, and whether they studied the same topic in their undergraduate education. Tables 4.1 and 4.2 provide the summary statistics and correlations for all variables in Time 2.

Table 4.1: Descriptive Statistics

	N	Mean	StDev
Dependent Variable			
Informal work communication$_{t_2}$	13,677	1.43	1.02
Independent Variables			
Spatial distance$_{t_2]}$	13,677	84.94	48.93
Work interdependencies$_{t_2}$	13,677	3.91	1.66
Boss walking path overlap$_{t_2}$	13,677	0.06	0.20
Controls			
Same function	13,677	0.06	0.23
Same level	13,677	0.16	0.37
Level of *i*	13,677	4.74	1.61
Level if *j*	13,677	4.88	1.86
Same gender	13,677	0.54	0.50
Gender of *i* (male=1)	13,677	0.63	0.48
Gender of *j* (male=1)	13,677	0.66	0.47
Same university	13,677	0.05	0.21
Same field of study	13,677	0.10	0.30
Age difference	13,677	9.96	7.67
Tenure difference	13,677	1.31	1.23
Age of *i*	13,677	35.94	8.93
Age of *j*	13,677	36.00	8.78
Tenure of *i*	13,677	2.31	1.30
Tenure of *j*	13,677	2.30	1.23

Table 4.2: Correlations between variables

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**
1. Informal work communication, time 2																		
2. Boss walking path overlap, time 2	0.24																	
3. Work interdependencies, time 2	0.41	0.3																
4. Spatial distance, time 2	-0.21	-0.22	-0.28															
5. Same function	0.41	0.43	0.51	-0.28														
6. Same level	0.03	0.04	0.00	-0.04	0.05													
7. Level of i	0.14	-0.06	0.34	-0.07	0.00	-0.03												
8. Level of j	0.16	0.00	0.39	-0.12	-0.02	-0.06	-0.01											
9. Same gender	0.04	0.03	0.09	0.01	0.06	0.02	0.06	0.05										
10. Gender of i	0.05	0.00	0.09	0.04	0.04	0.01	0.18	0.00	0.32									
11. Gender of j	0.02	0.07	0.13	0.01	0.03	0.00	0.00	0.21	0.25	-0.01								
12. Same university	0.05	0.03	0.05	0.03	0.04	0.02	0.01	0.01	0.02	0.01	0.00							
13. Same field of study	0.12	0.1	0.17	-0.1	0.19	0.03	-0.01	0.03	0.07	0.06	0.06	0.09						
14. Age difference	-0.02	0.06	0.09	0.15	-0.02	-0.04	0.03	0.04	0.05	0.12	0.11	0.02	0.00					
15. Tenure difference	-0.03	-0.03	0.06	0.11	-0.01	0.00	0.02	0.02	0.04	0.09	0.07	-0.01	0.01	0.06				
16. Age of i	0.04	0.02	0.15	0.1	0.00	-0.02	0.44	0.00	0.05	0.16	0.00	0.02	-0.01	0.25	0.08			
17. Age if j	0.08	0.1	0.2	0.05	0.00	-0.03	0	0.5	0.04	0.00	0.15	0.02	0.01	0.23	0.07	-0.01		
18. Tenure of i	0.02	0.00	0.05	0.12	-0.01	-0.02	0.04	0.00	0.02	0.07	0	0.02	-0.01	0.07	0.49	0.15	0.00	
19. Tenure of j	0.04	0.00	0.08	0.08	-0.01	-0.02	0.00	0.11	0.03	0.00	0.12	0.02	0.01	0.07	0.45	0.00	0.18	-0.01

4.5 Results

As noted above, our analysis focuses on frequency of informal work communication in Time 2. However, we restrict the sample to only those employees who responded to both surveys, and as noted above, to pairs of employees in different groups. Estimates are reported using ordinary least squares regression. We acknowledge that non-independence among observations is a well-known problem when studying dyadic communication. This is because a dyad containing employee i will be correlated with all other dyads with employee i, and likewise for employee j. Additionally, our data contain both observation Y_{ij} and Y_{ji} , which are also correlated. In order to correct for the potentially underestimated standard errors, we implement the `clus_nway.ado` package developed by Kleinbaum et al. (2013) in STATA and cluster on each employee in the dyad and the dyad itself (see also Cameron et al. (2011)). This results in three-way cluster-robust standard errors for each of our analyses.

4.5.1 Main Results

We report our results for Hypotheses 1 and 2 in Table 4.3. We include a baseline model first to show that, as expected and in line with prior studies, similarities between employees i and j greatly influences communication. For members in different groups, being in the same functional unit (Human Resources, Accounting, etc.) has by far the largest effect. All else equal, being in the same functional unit can be the difference between never having informal work communication and having it weekly. Having the same undergraduate training and being in the same level of the company also had significant effects. Hypothesis 1 is supported in Models 1 and 2, with increased spatial distance being strongly associated with a decrease in informal work communication. In Model 2 we test Hypothesis 2 and find that the relationship between work interdependencies and informal work communication is positive and significant, as expected.

As mentioned above, one of the benefits of our setting is the variation in work interdependencies over the period of study. This is convenient because it allows us to examine changes in the structure of the workflow in the same firm over the same time period. Additionally, any variation in spatial distance was not directly associated with the changes in the formal organization of the company due to the growth of the company. Below, we examine the differential impact of changes in work interdependencies on the changes in informal work communication.

As a result of the reorganization, 28% of our sample changed divisions due to the reorganization. This reorganization resulted in changes in the work interdependencies between group members across the company. For example, prior to the reorganization the project management group reported to a manager, who reported to a director, who in turn reported to the vice president of the division. However, after the reorganization the project management group reported directly to the vice president in their new division, which increased the work interdependencies between themselves and the vice president.

We estimate the impact of changes in work interdependencies on the changes in informal work communication and report the results in Table 4.4. Specifically, we find that having closer work interdependencies significantly increases the frequency of two employees' informal work communication. But decreasing the work interdependencies between two employees has no impact on the change in informal work communication.

Table 4.5 reports the results of our test of Hypothesis 3. Here we see that the degree to which the walking paths of both i and j to their respective managers overlap has a significant impact on the frequency with which they have informal work communication. One of the beneficial characteristics of the measure of walking-path overlap is that it is at least somewhat exogenous in our model. This is because it is highly unlikely that the office manager took into account the relative positions of i and j to their bosses when she was

Table 4.3: OLS Results of Spatial Distance and Work Interdependencies on Informal Work Communication

	Baseline	(1)	(2)
Spatial distance		-0.00***	-0.00***
		(0.00)	(0.00)
Work interdependencies			0.12***
			(0.02)
Same function	1.84***	1.73***	1.30***
	(0.11)	(0.11)	(0.13)
Same level	0.07**	0.06**	0.06**
	(0.03)	(0.03)	(0.03)
Same gender	0.00	0.00	0.00
	(0.02)	(0.02)	(0.02)
Same university	0.10*	.012**	0.11*
	(0.06)	(0.06)	(0.06)
Same field of study	0.15***	0.13***	0.09*
	(0.05)	(0.05)	(0.05)
Age difference	-0.00*	-0.00	-0.00**
	(0.00)	(0.00)	(0.00)
Tenure difference	-0.09***	-0.09***	-0.09***
	(0.02)	(0.02)	(0.02)
Controls (for both i and j)			
Division	Yes	Yes	Yes
Age	Yes	Yes	Yes
Gender	Yes	Yes	Yes
Tenure	Yes	Yes	Yes
Level	Yes	Yes	Yes
Education	Yes	Yes	Yes
Obs.	13,677	13,677	13,677
R^2(Adj.)	0.25	0.26	0.35

Standard errors in parentheses.

Standard errors clustered at the level of individual i, individual j, and the dyad.

* p<0.10, ** p<0.05, *** p<0.01

Table 4.4: OLS Results of Change in Work Interdependencies on Change in Informal Work Communication Between Time 1 and Time 2

	(3)	(4)
Farther	-.013	-0.13
	(0.11)	(0.11)
Closer	0.69***	0.69***
	(0.20)	(0.20)
Δ Spatial Distance		0.00
		(0.00)
Obs.	13,677	13,677
R^2(Adj.)	0.00	0.00

Standard errors in parentheses.
Standard errors clustered at the level of individual i, individual j, and the dyad.
* p<0.10, ** p<0.05, *** p<0.01

making seating assignments.

As a whole, this analysis provides support for the idea that the common walking paths of an employee, which are largely determined by the hierarchical nature of the organization, may provide low-cost opportunities for communicating with members of other groups.

4.6 Discussion and Conclusion

Over the last 40 years we have learned a great deal about the effects of spatial distance on communication patterns. One of the goals of this study was to further enrich our understanding of spatial distance and the mechanisms underlying its effect on communication. Another goal is to highlight the importance of taking into account firm structure when studying communication patterns (Allen, 1977). This is particularly salient for group communication researchers with an interest in firms, which are by nature hierarchical. Though a limitation of this study is that it provides evidence from only one firm over a relatively short period, we believe these findings should be generalizable to a wide range of com-

Table 4.5: OLS Results of Overlapping Walking Path to Boss on Informal Work Communication

	(5)	(6)
Boss walking path overlap		0.25***
		(0.08)
Spatial distance	-0.00***	-0.11***
	(0.00)	(0.02)
Work interdependencies	0.12***	0.00***
	(0.02)	(0.00)
Same function	1.30***	1.23***
	(0.13)	(0.13)
Same level	0.06**	0.06**
	(0.03)	(0.03)
Same gender	0.00	0.01
	(0.02)	(0.02)
Same university	.011*	0.10*
	(0.06)	(0.06)
Same field of study	0.09*	0.09*
	(0.05)	(0.05)
Age difference	-0.00**	-0.00**
	(0.00)	(0.00)
Tenure difference	-0.09***	-0.09***
	(0.02)	(0.02)
Controls (for both *i* and *j*)		
Division	Yes	Yes
Age	Yes	Yes
Gender	Yes	Yes
Tenure	Yes	Yes
Level	Yes	Yes
Education	Yes	Yes
Obs.	13,677	13,677
R^2(Adj.)	0.25	0.25

Standard errors in parentheses.

Standard errors clustered at the level of individual *i*, individual *j*, and the dyad.

* $p<0.10$, ** $p<0.05$, *** $p<0.01$

panies and industries. Further work is needed to verify whether the results are robust to different settings. Additionally, we do believe that the reorganization provides semi-exogenous variation in the organizational structure, but we are cautious about making any causal claims.

This study contributes directly to two additional streams of research. First, we contribute to the literature on organizational design by suggesting that work interdependencies act as a motor that drives motion in physical space and triggers chance encounters that lead to frequent informal work communication even between employees who don't share immediate work bonds. Second, our work reminds those who examine the effects of spatial distance on a number of outcomes (e.g. innovation, productivity) of the importance of taking into account the system of work relationships that bind the interactions of the participants they study. Whereas being cognizant of such connections may be more pressing for those interested in intra-organizational distance, attention is also warranted when the object of study is inter-organizational collaborations. For example, firms that don't share a work bond may not engage in ongoing communication even when they are collocated, others that are further apart but share work ties (e.g. are part of the same vertical value chain) may do so more frequently.

The findings reported here are also relevant for practitioners. Over the years, more and more organizations view the layout of the workplace as an important lever to improve collaboration between employees and, consequently, as a means to improve efficiency and innovation (West Jr and Wind, 2007; Chan et al., 2007). Our study suggests that careful design of the location of employees within an office is as important as the architectural design of such space. On the one hand, putting group members (who have strong work interdependencies) together can increase efficiency within the group, but may limit opportunities for group members to have casual collisions with members of other groups in the organization. On the other hand, spreading these group members across the floor plan of the organization may increase opportunities for informal work communication, but at the same time may undermine coordination among members. Finally, we introduce the idea that where the boss sits can also impact informal work communication among members of different groups, by shaping how employees move about the office. When members in different groups have overlap in their walking paths to their respective bosses (even if members are sitting with their fellow group members), we show that there is more likely to be informal work communication among different groups. These benefits need to be weighed against the potential costs of a boss not being able to closely monitor the group, thus there is a balance to be struck regarding how far a boss should sit away from the group.

Chapter 5

Conclusion

This dissertation makes three contributions to the broad literature on technological change. The first essay proposes and finds support for the notion that new ventures are more innovative than established firms because they work on different problems to begin with. Analysis suggests that in the U.S. medical device industry, new ventures are more likely than established firms to undertake invention in less explored areas because of differences in complementary capabilities. The second essay contributes to the current debate on the affect of artificial intelligence on labor by providing some of the first empirical evidence of the actual diffusion of AI. The analysis suggests that while AI differs from prior technology in its ability to perform non-routine tasks, the rate of application of AI to high skill occupations is significantly less than to middle skill occupations. One interpretation of this is that AI is following a predictable course by continuing the trend of hollowing out middle skill occupations. The last chapter investigates the internal structure of an innovating financial services firm. Analysis from a seven month field study suggests that boundary spanning communication - an essential input into innovative - is influenced by a nuanced interaction between the physical and organizational structures of the firm. Together, these essays provide new insights into factors influencing the rate and direction of inventive activity.

Appendices

Appendix A: Example patents for various atherectomy devices

Patent US 8,551,128

Abstract The invention provides a rotational atherectomy system, device and method comprising a flexible, elongated, rotatable drive shaft with an abrasive section within a pre-curved section of the drive shaft. The device may further comprise a concentric or eccentric enlarged diameter section that is at least partially covered with abrasive material to comprise the abrasive section. The abrasive section may further comprise an abrasive crown or burr mounted to the drive shaft. The pre-curved drive shaft allows smaller diameter and/or massive abrasive regions to be used while sweeping larger diameters during high-speed rotation. The pre-curved region is substantially straightened for insertion into vasculature and placement adjacent stenosis by insertion of the guide wire. Removal of guide wire proximally from the pre-curved region allows the drive shaft to return to its pre-curved form for ablation. Reinsertion of the guide wire beyond the pre-curved region straightens the drive shaft for ease of removal.

Patent US 9,308,016

Abstract Devices, systems, and methods are employed to perform an atherectomy in an identified region to restore patency to arterial lesions. A bioactive material is introduced into the identified region before, after or during performing the atherectomy. The bioactive material can be introduced, e.g., on a balloon coated with the bioactive material, which is expanded in contact with the identified region to deliver the bioactive material. The bioactive material can be, e.g., at least one of a restenosis-inhibiting agent, a thrombus-inhibiting agent, and an anti-inflammatory agent.

Appendix B: Patent CPC codes used to define medical patents

CPC Main Groups

A61B: Diagnosis; Surgery; Identification

A61F: Filters Implantable Into Blood Vessels; Prostheses; Devices Providing Patency To, Or Preventing Collapsing Of, Tubular Structures Of The Body, Etc

A61H: Physical Therapy Apparatus, E.G. Devices For Locating Or Stimulating Reflex Points In The Body; Artificial Respiration; Massage; Etc

A61N: Electrotherapy; Magnetotherapy; Radiation Therapy; Ultrasound Therapy

A61C: Dentistry; Apparatus Or Methods For Oral Or Dental Hygiene

A61J: Containers Specially Adapted For Medical Or Pharmaceutical Purposes; Devices Or Methods Specially Adapted For Bringing Pharmaceutical Products Into Particular Physical Or Administering Forms; Devices For Administering Food Or Medicines Orally; Baby Comforters; Devices For Receiving Spittle

A61K: Preparations For Medical, Dental, Or Toilet Purposes

A61L: Methods Or Apparatus For Sterilising Materials Or Objects In General; Disinfection, Sterilisation, Or Deodorisation Of Air; Chemical Aspects Of Bandages, Dressings, Absorbent Pads, Or Surgical Articles; Materials For Bandages, Dressings, Absorbent Pads, Or Surgical Articles

A61M: Devices For Introducing Media Into, Or Onto, The Body

G16H: Healthcare Informatics, I.E. Information And Communication Technology [Ict] Specially Adapted For The Handling Or Processing Of Medical Or Healthcare Data

Appendix C: Medical Procedure Survey

Each respondent was provided with a list of approximately 200 medical procedures along with their descriptions from the Unified Medical Language System. Respondents were asked to score the following questions:

1. This activity is too vague to be meaningful (T/F)
2. How familiar are you with this activity? (1=not familiar, 5=very familiar)
3. Is this typically performed outside of the hospital? (E.g. a clinic or dentist office)
4. In approximately how many different units in the hospital will this activity typically be performed?
5. What type of physician or employee is most likely to carry out this procedure? (Medical Doctor, Nurse, Technician, Other)
6. On average, how many times per week is this activity performed across the entire organization?

Appendix D: Example Medical Activities

Auscultation

Autopsy

Colostomy

Cystoscopy

Dental Scaling

Enterostomy

Gastric Bypass

Glossectomy

Heart Transplantation

Iontophoresis

Laminectomy

Laparoscopy

Lipectomy

Liver Transplantation

Pericardiectomy

Renal Dialysis

Sclerotherapy

Skin Transplantation

Spinal Fusion

Thoracotomy

Tracheotomy

Vasectomy

Appendix E: Patenting on Medical Procedures

My analysis focuses on medical devices used in the process of performing a medical procedure. It should be noted that medical procedures themselves may also be patented according to United States patent laws. However, as of 1997, physicians infringing on a patented medical procedure are not held liable. United States patent law "deprives the patentees of remedies for infringement by a medical practitioner's performance of a medical activity" (Meier, 2015). In other words, while a procedure may be patented, there are no consequences for infringement. This law was enacted after Pallin v. Singer (1996), a case in which Singer, an ophthalmologist, infringed on a patented cataract surgical procedure. Pallin lost the case and the patent claims were invalidated, setting the stage for a controversy in congress over whether medical procedures should be patentable in the United States. This was ultimately decided with the passage of the 1997 Omnibus Consolidated Appropriations Act. For a summary of these events see Meier (2015). Thus, during my sample period there are no legal benefits to patenting a medical procedure, greatly reducing the incentive of firms to patent the procedure itself.

Appendix F: Dictionary of AI Terms

AI Search Terms

Machine Learning, Unsupervised Learning, Learning Algorithm, Reinforcement Learning, Supervised Learning, Semi-supervised Learning, Statistical Learning, Convolutional Neural Network, Feedforward Neural Network, Hopfield Neural Network, Recurrent Neural Network, Deep Boltzmann Machine, Restricted Boltzmann Machine, Multilayer Perceptron, Multi-layer Perceptron, Artificial Neural Network, Deep Neural Network, Fuzzy Neural Network, Multilayer Neural Network, Multi-layer Neural Network, Deep Belief Network, Stacked Auto-encoder, Stacked Autoencoder, Agent Architecture, Brute-force Search, Capsule Neural Network, Cluster Analysis, Means Clustering, Cognitive Architecture, Cognitive Computing, Computational Intelligence, Decision Tree, Random Forest, Greedy Algorithm, Evolutionary Algorithm, Expert System, Feature Learning, Feature Selection, Feature Engineering, Intelligent Agent, Kernel Method, Principal Component, Knowledge-based System, Naive Bayes.

AI Patent Classes (Cooperative Patent Classification code)

Section G06N, Section G06K Group 7, Section G06K Group 9, Section G06K Group 11, Section G05B Group 13.

Appendix G: Example AI Patents

Patent US 9,538,925

Abstract A method and system for determining fractional flow reserve (FFR) for a coronary artery stenosis of a patient is disclosed. In one embodiment, medical image data of the patient including the stenosis is received, a set of features for the stenosis is extracted from the medical image data of the patient, and an FFR value for the stenosis is determined based on the extracted set of features using a trained machine-learning based mapping. In another embodiment, a medical image of the patient including the stenosis of interest is received, image patches corresponding to the stenosis of interest and a coronary tree of the patient are detected, an FFR value for the stenosis of interest is determined using a trained **deep neural network** regressor applied directly to the detected image patches.

Note: Emphasis added to highlight the application of AI technology.

Appendix H: Survey Instrument

Based on the work you have done in the last 3 months, please answer the following 6 questions. You can choose answers from a menu with options on a scale from 1 to 5, where 1 is the lowest (Never) and 5 is the highest (Daily). You only have to answer the questions that correspond to the people with whom you have interaction (the default answer is 1, which means that you do not maintain interaction). The questions are designed to capture three types of interactions you have with other FinServe employees ...

1. How often do you talk by chance with the following people, on topics related to work

(Dependent variable)

2. To what degree do you depend on the following people for information and / or materials you need to do your job?

3. To what degree do the following people depend on you for information and / or materials they need to do their job? (Alternative measure of work distance)

Appendix I: Seating Census

Reason(s) for changing seats (check all that apply)

1. Office remodeling
2. Company reorganization in April
3. Only seat available
4. To be with team
5. Growth of the company
6. Entire department moving
7. Team moving
8. Moved to be closer to someone you work with
9. Other
10. Do not know

Note: Items (1), (3), and (5) were marked as exogenous in our analysis.

Bibliography

Abbott, B. Google ai beats doctors at breast cancer detection—sometimes. *The Wall Street Journal*.

Acs, Z. J. and D. B. Audretsch (1988). Innovation in large and small firms: an empirical analysis. *The American economic review*, 678–690.

Agarwal, R., R. Echambadi, A. M. Franco, and M. B. Sarkar (2004). Knowledge transfer through inheritance: Spin-out generation, development, and survival. *Academy of Management journal 47*(4), 501–522.

Agrawal, A. K., J. S. Gans, and A. Goldfarb (2018). Prediction, judgment and complexity: A theory of decision making and artificial intelligence. Technical report, National Bureau of Economic Research.

Akst, D. (2013). What can we learn from past anxiety over automation? *The Wilson Quarterly*.

Alarie, B., A. Niblett, and A. H. Yoon (2016). Using machine learning to predict outcomes in tax law. *Can. Bus. LJ 58*, 231.

Allen, T. J. (1977). Managing the flow of technology: Technology transfer and the dissemination of technological information within the r & d organization(book). *Research supported by the National Science Foundation. Cambridge, Mass., MIT Press, 1977. 329 p.*

Argyres, N. S. and B. S. Silverman (2004). R&d, organization structure, and the development of corporate technological knowledge. *Strategic Management Journal 25*(8-9), 929–958.

Arntz, M., T. Gregory, and U. Zierahn (2016). The risk of automation for jobs in oecd countries.

Arora, A., S. Belenzon, and A. Patacconi (2018). The decline of science in corporate r&d. *Strategic Management Journal 39*(1), 3–32.

Arora, A., W. M. Cohen, and J. P. Walsh (2016). The acquisition and commercialization of invention in american manufacturing: Incidence and impact. *Research Policy 45*(6), 1113–1128.

Arora, A. and A. Gambardella (1995). The division of innovative labor in biotechnology. *Sources of medical technology: Universities and industry*, 188–206.

Associates, G. R. (1982). The relationship between industrial concentration, firm size, and technological innovation. *Small Business Administration*.

Autor, D. H. (2015). Why are there still so many jobs? the history and future of workplace automation. *Journal of economic perspectives 29*(3), 3–30.

Autor, D. H., L. F. Katz, and A. B. Krueger (1998). Computing inequality: have computers changed the labor market? *The Quarterly journal of economics 113*(4), 1169–1213.

Autor, D. H., F. Levy, and R. J. Murnane (2003). The skill content of recent technological change: An empirical exploration. *The Quarterly journal of economics 118*(4), 1279–1333.

Baker, T. and R. E. Nelson (2005). Creating something from nothing: Resource construction through entrepreneurial bricolage. *Administrative science quarterly 50*(3), 329–366.

Bayus, B. L. and R. Agarwal (2007). The role of pre-entry experience, entry timing, and product technology strategies in explaining firm survival. *Management Science 53*(12), 1887–1902.

Berman, E., J. Bound, and Z. Griliches (1994). Changes in the demand for skilled labor within us manufacturing: evidence from the annual survey of manufactures. *The Quarterly Journal of Economics 109*(2), 367–397.

Berman, E., J. Bound, and S. Machin (1998). Implications of skill-biased technological change: international evidence. *The quarterly journal of economics 113*(4), 1245–1279.

Bessen, J. (2015). *Learning by doing: the real connection between innovation, wages, and wealth*. Yale University Press.

Bodenreider, O. (2004). The unified medical language system (umls): integrating biomedical terminology. *Nucleic acids research 32*(suppl_1), D267–D270.

Bolton, P. and M. Dewatripont (1994). The firm as a communication network. *The Quarterly Journal of Economics 109*(4), 809–839.

Bresnahan, T. F. (1999). Computerisation and wage dispersion: an analytical reinterpretation. *The Economic Journal 109*(456), 390–415.

Bresnahan, T. F., E. Brynjolfsson, and L. M. Hitt (2002). Information technology, workplace organization, and the demand for skilled labor: Firm-level evidence. *The quarterly journal of economics 117*(1), 339–376.

Brown, N. and T. Sandholm (2018). Superhuman ai for heads-up no-limit poker: Libratus beats top professionals. *Science 359*(6374), 418–424.

Brynjolfsson, E. and A. McAfee (2011). *Race against the machine: How the digital revolution is accelerating innovation, driving productivity, and irreversibly transforming employment and the economy*. Brynjolfsson and McAfee.

Brynjolfsson, E., T. Mitchell, and D. Rock (2018). What can machines learn, and what does it mean for occupations and the economy? In *AEA Papers and Proceedings*, Volume 108, pp. 43–47.

Burton, R. M. et al. (1984). *Designing efficient organizations: Modelling and experimentation*, Volume 7. North Holland.

Cameron, A. C., J. B. Gelbach, and D. L. Miller (2011). Robust inference with multiway clustering. *Journal of Business & Economic Statistics 29*(2), 238–249.

Carrell, D. S., R. E. Schoen, D. A. Leffler, M. Morris, S. Rose, A. Baer, S. D. Crockett, R. A. Gourevitch, K. M. Dean, and A. Mehrotra (2017). Challenges in adapting existing clinical natural language processing systems to multiple, diverse health care settings. *Journal of the American Medical Informatics Association 24*(5), 986–991.

Catalini, C. (2017). Microgeography and the direction of inventive activity. *Management Science 64*(9), 4348–4364.

Chan, D. C. and M. J. Dickstein (2019). Industry input in policy making: Evidence from medicare. *The Quarterly Journal of Economics 134*(3), 1299–1342.

Chan, J. K., S. L. Beckman, and P. G. Lawrence (2007). Workplace design: A new managerial imperative. *California Management Review 49*(2), 6–22.

Chatterji, A. K. and K. R. Fabrizio (2016). Does the market for ideas influence the rate and direction of innovative activity? evidence from the medical device industry. *Strategic management journal 37*(3), 447–465.

Chown, J. D. and C. C. Liu (2015). Geography and power in an organizational forum: Evidence from the us senate chamber. *Strategic Management Journal 36*(2), 177–196.

Choy, G., O. Khalilzadeh, M. Michalski, S. Do, A. E. Samir, O. S. Pianykh, J. R. Geis, P. V. Pandharipande, J. A. Brink, and K. J. Dreyer (2018). Current applications and future impact of machine learning in radiology. *Radiology 288*(2), 318–328.

Christensen, C. (1997). The innovators dilemma. *Harvard Business School Press, Cambridge, Mass.*

Clark, K. and C. Baldwin (2000). Design rules, the power of modularity, vol. 1.

Clough, D. R., T. P. Fang, B. Vissa, and A. Wu (2019). Turning lead into gold: How do entrepreneurs mobilize resources to exploit opportunities? *Academy of Management Annals 13*(1), 240–271.

Cockburn, I. M., R. Henderson, and S. Stern (2018). The impact of artificial intelligence on innovation. Technical report, National Bureau of Economic Research.

Cohen, W. M. (2010). Fifty years of empirical studies of innovative activity and performance. In *Handbook of the Economics of Innovation*, Volume 1, pp. 129–213. Elsevier.

Cohen, W. M. and S. Klepper (1992). The tradeoff between firm size and diversity in the pursuit of technological progress. *Small Business Economics 4*(1), 1–14.

Cohen, W. M. and S. Klepper (1996). A reprise of size and r & d. *The Economic Journal 106*(437), 925–951.

Cohen, W. M., R. R. Nelson, and J. P. Walsh (2000). Protecting their intellectual assets: Appropriability conditions and why us manufacturing firms patent (or not). Technical report, National Bureau of Economic Research.

Conrath, D. W. (1973). Communications environment and its relationship to organizational structure. *Management Science 20*(4-part-ii), 586–603.

Corredoira, R. A., B. D. Goldfarb, and Y. Shi (2018). Federal funding and the rate and direction of inventive activity. *Research Policy 47*(9), 1777–1800.

Cramton, C. D. (2001). The mutual knowledge problem and its consequences for dispersed collaboration. *Organization science 12*(3), 346–371.

Daft, R. L. and R. H. Lengel (1986). Organizational information requirements, media richness and structural design. *Management science 32*(5), 554–571.

Dao, M., M. Das, Z. Koczan, and W. Lian. The hollowing out of middle-skilled labor share of income.

Davenport, T. H. and K. Dreyer (2018). Ai will change radiology, but it won't replace radiologists. *Harvard Business Review*, 1–5.

De Meyer, A. (1991). Tech talk: How managers are stimulating global r&d communication. *MIT Sloan Management Review 32*(3), 49.

Dierickx, I. and K. Cool (1989). Asset stock accumulation and sustainability of competitive advantage. *Management science 35*(12), 1504–1511.

Dippel, E. J., P. Makam, R. Kovach, J. C. George, R. Patlola, D. C. Metzger, C. Mena-Hurtado, R. Beasley, P. Soukas, P. J. Colon-Hernandez, et al. (2015). Randomized controlled study of excimer laser atherectomy for treatment of femoropopliteal in-stent restenosis: initial results from the excite isr trial (excimer laser randomized controlled study for treatment of femoropopliteal in-stent restenosis). *JACC: Cardiovascular Interventions 8*(1 Part A), 92–101.

Doms, M., T. Dunne, and K. R. Troske (1997). Workers, wages, and technology. *The Quarterly Journal of Economics 112*(1), 253–290.

Eggers, J. and A. Kaul (2018). The problem you know: An empirical examination of firms' recombination strategies. Working Paper.

Elfenbein, D. W., B. H. Hamilton, and T. R. Zenger (2010). The small firm effect and the entrepreneurial spawning of scientists and engineers. *Management Science 56*(4), 659–681.

Ethiraj, S. K. and D. Levinthal (2004). Modularity and innovation in complex systems. *Management science 50*(2), 159–173.

Felin, T. and T. R. Zenger (2015). Crossroadsstrategy, problems, and a theory for the firm. *Organization Science 27*(1), 222–231.

Felten, E., M. Raj, and R. C. Seamans (2019). The effect of artificial intelligence on human labor: An ability-based approach. In *Academy of Management Proceedings*, Volume 2019, pp. 15784. Academy of Management Briarcliff Manor, NY 10510.

Felten, E. W., M. Raj, and R. Seamans (2018). A method to link advances in artificial intelligence to occupational abilities. In *AEA Papers and Proceedings*, Volume 108, pp. 54–57.

Festinger, L., K. Back, S. Schachter, H. H. Kelley, and J. Thibaut (1950). Theory and experiment in social communication.

Fisher, L. (2016). Who decides the future of work. *Communications of the ACM*.

Frey, C. B. and M. A. Osborne (2017). The future of employment: How susceptible are jobs to computerisation? *Technological forecasting and social change 114*, 254–280.

Fuster, A., P. Goldsmith-Pinkham, T. Ramadorai, and A. Walther (2018). Predictably unequal? the effects of machine learning on credit markets. *The Effects of Machine Learning on Credit Markets (November 6, 2018)*.

Galbraith, J. (1973). Designing complex organizations.

Gera, S., W. Gu, and Z. Lin (2001). Technology and the demand for skills in canada: an industry-level analysis. *Canadian Journal of Economics/Revue canadienne d'économique 34*(1), 132–148.

Grant, R. M. (1996). Toward a knowledge-based theory of the firm. *Strategic management journal 17*(S2), 109–122.

Griffin, A. and J. R. Hauser (1992). Patterns of communication among marketing, engineering and manufacturing—a comparison between two new product teams. *Management science 38*(3), 360–373.

Hall, B. H. and R. H. Ziedonis (2001). The patent paradox revisited: an empirical study of patenting in the us semiconductor industry, 1979-1995. *rand Journal of Economics*, 101–128.

Hasan, S. and S. Bagde (2015). Peers and network growth: Evidence from a natural experiment. *Management Science 61*(10), 2536–2547.

Haugeland, J. (1985). Artificial intelligence: the very idea.

He, J., S. L. Baxter, J. Xu, J. Xu, X. Zhou, and K. Zhang (2019). The practical implementation of artificial intelligence technologies in medicine. *Nature medicine 25*(1), 30.

He, K., X. Zhang, S. Ren, and J. Sun (2015). Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In *Proceedings of the IEEE international conference on computer vision*, pp. 1026–1034.

Helfat, C. E. (1997). Know-how and asset complementarity and dynamic capability accumulation: the case of r&d. *Strategic management journal 18*(5), 339–360.

Henderson, R. (1993). Underinvestment and incompetence as responses to radical innovation: Evidence from the photolithographic alignment equipment industry. *The RAND Journal of Economics*, 248–270.

Henderson, R. and I. Cockburn (1993). Scale, scope and spillovers: the determinants of research productivity in the pharmaceutical industry. Technical report, National Bureau of Economic Research.

Henderson, R. M. and K. B. Clark (1990). Architectural innovation: The reconfiguration of existing. *Administrative science quarterly 35*(1), 9–30.

Hestness, J., S. Narang, N. Ardalani, G. Diamos, H. Jun, H. Kianinejad, M. Patwary, M. Ali, Y. Yang, and Y. Zhou (2017). Deep learning scaling is predictable, empirically. *arXiv preprint arXiv:1712.00409*.

Jaffe, A. B. (1986). Technological opportunity and spillovers of r&d: evidence from firms' patents, profits and market value.

Kabo, F. W., N. Cotton-Nessler, Y. Hwang, M. C. Levenstein, and J. Owen-Smith (2014). Proximity effects on the dynamics and outcomes of scientific collaborations. *Research Policy 43*(9), 1469–1485.

Kiesler, S. and J. N. Cummings (2002). What do we know about proximity and distance in work groups? a legacy of research. *Distributed work 1*, 57–80.

Kiwanuka, J. P. (2005). Interpretation of tuberculin skin-test results in the diagnosis of tuberculosis in children. *African health sciences 5*(2), 152–156.

Kleinbaum, A. M., T. E. Stuart, and M. L. Tushman (2013). Discretion within constraint: Homophily and structure in a formal organization. *Organization Science 24*(5), 1316–1336.

Klepper, S. and K. L. Simons (2000). Dominance by birthright: entry of prior radio producers and competitive ramifications in the us television receiver industry. *Strategic Management Journal 21*(10-11), 997–1016.

Klepper, S. and S. Sleeper (2005). Entry by spinoffs. *Management science 51*(8), 1291–1306.

Kogut, B. and U. Zander (1992). Knowledge of the firm, combinative capabilities, and the replication of technology. *Organization science 3*(3), 383–397.

Kraut, R. E., C. Egido, and J. Galegher (2014). Patterns of contact and communication in scientific research collaborations. In *Intellectual teamwork*, pp. 163–186. Psychology Press.

Kretschmer, T. and P. Puranam (2008). Integration through incentives within differentiated organizations. *Organization Science 19*(6), 860–875.

Langlotz, C. P. (2019). Will artificial intelligence replace radiologists?

Lawrence, P. R. and J. W. Lorsch (1967). Differentiation and integration in complex organizations. *Administrative science quarterly*, 1–47.

Lee, J. and N. Berente (2012). Digital innovation and the division of innovative labor: Digital controls in the automotive industry. *Organization Science 23*(5), 1428–1447.

Leonard-Barton, D. (1992). Core capabilities and core rigidities: A paradox in managing new product development. *Strategic management journal 13*(S1), 111–125.

Lerner, J. (2006). The new financial thing: The origins of financial innovations. *Journal of Financial Economics 79*, 223–255.

Levy, F. and R. J. Murnane (1996). With what skills are computers a complement? *The American Economic Review 86*(2), 258–262.

Lieberman, M. B. (1989). The learning curve, technology barriers to entry, and competitive survival in the chemical processing industries. *Strategic Management Journal 10*(5), 431–447.

Liew, C. (2018). The future of radiology augmented with artificial intelligence: a strategy for success. *European journal of radiology 102*, 152–156.

Machin, S. and J. Van Reenen (1998). Technology and changes in skill structure: evidence from seven oecd countries. *The Quarterly Journal of Economics 113*(4), 1215–1244.

Makadok, R., R. Burton, and J. Barney (2018). A practical guide for making theory contributions in strategic management. *Strategic Management Journal 39*(6), 1530–1545.

Mann, K. and L. Puttmann (2018). Benign effects of automation: New evidence from patent texts. *Available at SSRN 2959584*.

Mansfield, E. (1981). Composition of r and d expenditures: relationship to size of firm, concentration, and innovative output. *The Review of Economics and Statistics*, 610–615.

Manyika, J., M. Chui, J. Bughin, R. Dobbs, P. Bisson, and A. Marrs (2013). *Disruptive technologies: Advances that will transform life, business, and the global economy*, Volume 180. McKinsey Global Institute San Francisco, CA.

March, J. G. and H. A. Simon (1958). Organizations.

Marx, M. and A. Fuegi (2019). Reliance on science in patenting. *Available at SSRN*.

McElheran, K. (2019). Economic measurement of ai. *NBER Working Paper*.

Meier, B. J. (2015). New patent infringement liability exception for medical procedures, the; note. *Journal of Legislation 23*(2), 265.

Mitchell, W. (1989). Whether and when? probability and timing of incumbents' entry into emerging industrial subfields. *Administrative science quarterly*, 208–230.

Monge, P. R., L. W. Rothman, E. M. Eisenberg, K. I. Miller, and K. K. Kirste (1985). The dynamics of organizational proximity. *Management Science 31*(9), 1129–1141.

Monteverde, K. (1995). Technical dialog as an incentive for vertical integration in the semiconductor industry. *Management Science 41*(10), 1624–1638.

Nelson, C. and S. Winter (1982a). Organizational capabilities and behavior: An evolutionary theory of economic change.

Nelson, R. and S. Winter (1982b). An evolutionary theory of economic change (belknap, cambridge, ma). *NelsonAn Evolutionary Theory of Economic Change1982*.

Nelson, R. R. (1982). The role of knowledge in r&d efficiency. *The quarterly journal of economics 97*(3), 453–470.

Nelson, R. R. et al. (1962). *The rate and direction of inventive activity: Economic and social factors*. Princeton University Press.

Nickerson, J. A. and T. R. Zenger (2004). A knowledge-based theory of the firm—the problem-solving perspective. *Organization science 15*(6), 617–632.

Nobel, R. and J. Birkinshaw (1998). Innovation in multinational corporations: Control and communication patterns in international r&d operations. *Strategic management journal 19*(5), 479–496.

Norris, J., P. Schuber, and C. Tolman (2014). Trends in healthcare investments and exits 2014. *Santa Clara, CA: Silicon Valley Bank Financial Group*.

Olson, J. (2000). Distance matters. *Human-Computer Interaction 15*.

Penrose, E. (1959). The theory ofthe growth ofthe firm. *JohnW iley& Sons, New York*.

Pentland, B. T. and H. H. Rueter (1994). Organizational routines as grammars of action. *Administrative science quarterly*, 484–510.

Peteraf, M. A. (1993). The cornerstones of competitive advantage: a resource-based view. *Strategic management journal 14*(3), 179–191.

Puranam, P. and M. Raveendran (2013). Interdependence and organization design. In *Handbook of economic organization*. Edward Elgar Publishing.

Quillen Jr, C. D. and O. H. Webster (2005). Continuing patent applications and the us patent and trademark office-updated. *Fed. Cir. BJ 15*, 635.

Roessner, J. D. (1981). Innovation, competition, and government policy in the semiconductor industry.

Rosenberg, N. and R. R. Nelson (1994). American universities and technical advance in industry. *Research policy 23*(3), 323–348.

Sanchez, R. and J. T. Mahoney (1996). Modularity, flexibility, and knowledge management in product and organization design. *Strategic management journal 17*(S2), 63–76.

Scherer, F. M. (1965). Firm size, market structure, opportunity, and the output of patented inventions. *The American economic review 55*(5), 1097–1125.

Schmookler, J. (1962). Economic sources of inventive activity. *The Journal of Economic History 22*(1), 1–20.

Schulz, M. (2001). The uncertain relevance of newness: Organizational learning and knowledge flows. *Academy of management journal 44*(4), 661–681.

Schumpeter, J. (1942). Creative destruction. *Capitalism, socialism and democracy 825*, 82–85.

Scott Morton, F. M. (1997). Entry decisions in the generic pharmaceutical industry. *NBER Working Paper* (w6190).

Shih, A. J., Y. Liu, and Y. Zheng (2016). Grinding wheel motion, force, temperature, and material removal in rotational atherectomy of calcified plaque. *CIRP Annals 65*(1), 345–348.

Silver, D., J. Schrittwieser, K. Simonyan, I. Antonoglou, A. Huang, A. Guez, T. Hubert, L. Baker, M. Lai, A. Bolton, et al. (2017). Mastering the game of go without human knowledge. *Nature 550*(7676), 354–359.

Simon, H. A. (1978). Information-processing theory of human problem solving. *Handbook of learning and cognitive processes 5*, 271–295.

Sørensen, J. B. and T. E. Stuart (2000). Aging, obsolescence, and organizational innovation. *Administrative science quarterly 45*(1), 81–112.

Stern, S. (2004). Do scientists pay to be scientists? *Management science 50*(6), 835–853.

Teece, D. J. (1986). Profiting from technological innovation: Implications for integration, collaboration, licensing and public policy. *Research policy 15*(6), 285–305.

Teodoridis, F. (2017). Understanding team knowledge production: The interrelated roles of technology and expertise. *Management Science 64*(8), 3625–3648.

Thompson, J. D. (1967). Organizations in action; social science bases of administrative theory. 1967. *New York*.

Topol, E. J., F. Leya, C. A. Pinkerton, P. L. Whitlow, B. Hofling, C. A. Simonton, R. R. Masden, P. W. Serruys, M. B. Leon, D. O. Williams, et al. (1993). A comparison of directional atherectomy with coronary angioplasty in patients with coronary artery disease. *New England Journal of Medicine 329*(4), 221–227.

Tripsas, M. and G. Gavetti (2000). Capabilities, cognition, and inertia: Evidence from digital imaging. *Strategic management journal 21*(10-11), 1147–1161.

Van den Bulte, C. and R. K. Moenaert (1998). The effects of r&d team co-location on communication patterns among r&d, marketing, and manufacturing. *Management Science 44*(11-part-2), S1–S18.

Van Norman, G. A. (2016). Drugs, devices, and the fda: part 2: an overview of approval processes: Fda approval of medical devices. *JACC: Basic to Translational Science 1*(4), 277–287.

Webb, M. (2019). The impact of artificial intelligence on the labor market. *Available at SSRN 3482150*.

West Jr, A. P. and Y. Wind (2007). Putting the organization on wheels: Workplace design at sei. *California Management Review 49*(2), 138–153.

Whittaker, S., D. Frohlich, and O. Daly-Jones (1994). Informal workplace communication: What is it like and how might we support it? In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 131–137. ACM.

Wu, S. T., H. Liu, D. Li, C. Tao, M. A. Musen, C. G. Chute, and N. H. Shah (2012). Unified medical language system term occurrences in clinical notes: a large-scale corpus analysis. *Journal of the American Medical Informatics Association 19*(e1), e149–e156.

Wuest, T., D. Weimer, C. Irgens, and K.-D. Thoben (2016). Machine learning in manufacturing: advantages, challenges, and applications. *Production & Manufacturing Research 4*(1), 23–45.

Yepes, A. J. J., L. Plaza, J. Carrillo-de Albornoz, J. G. Mork, and A. R. Aronson (2015). Feature engineering for medline citation categorization with mesh. *BMC bioinformatics 16*(1), 113.

Ziedonis, R. H. (2004). Don't fence me in: Fragmented markets for technology and the patent acquisition strategies of firms. *Management science 50*(6), 804–820.

Biography

David P. Hall completed his undergraduate education at Brigham Young University–Idaho, where he graduated *Cum Laude* with a Bachelor of Science (Economics) degree in 2010. He received a Master of Business Administration from Brigham Young University in 2014 and a Doctor of Philosophy in Business Administration (Strategy) in 2020 from the Fuqua School of Business, Duke University. During his studies at Fuqua, David was awarded the Fuqua School of Business Doctoral Fellowship (2014-2019), and was selected as a Strategic Research Fund Dissertation Scholar by the Strategic Management Society (2018).