IT-ENABLED BUSINESS PRACTICES: EMPIRICAL INVESTIGATIONS OF PRODUCTIVITY AND INNOVATION

A Dissertation
Presented to
The Academic Faculty

By

Patricia L. Angle

In Partial Fulfillment
Of the Requirements for the Degree
Doctor of Philosophy in IT Management

Georgia Institute of Technology

August, 2019

IT-ENABLED BUSINESS PRACTICES: EMPIRICAL INVESTIGATIONS OF PRODUCTIVITY AND INNOVATION

Approved by:

Dr. Chris Forman
S. C. Johnson College of Business
Cornell University

Dr. Marco Ceccagnoli
Scheller College of Business
Georgia Institute of Technology

Dr. Sabyasachi Mitra
Scheller College of Business
Georgia Institute of Technology

Dr. Eric Overby
Scheller College of Business
Georgia Institute of Technology

Dr. Marius Niculescu
Scheller College of Business
Georgia Institute of Technology

Date Approved: 21 June 2019

Pauca sed matura

—Johann Carl Friedrich Gauss

DEDICATION

I dedicate this work,
as I dedicate my life,
to my Matty
and our beautiful children,
Gabriel and Oliver.

ACKNOWLEDGEMENTS

For Chapter One, I thank conference and seminar participants at the Census Atlanta Research Data Center 2015 conference, WISE 2015, the Center for Economic Studies in 2016, POMS 2016, INFORMS 2016, AOM 2016, AOM OCIS 2017, and CIST 2017 for helpful feedback. I would also like to express gratitude to Kristina McElheran of the University of Toronto and Cheryl Grim of the U.S. Census Bureau for particularly helpful input. Melissa Banzhaf has been a helpful supervisor and wonderful friend through very long days in the RDC. Any opinions and conclusions expressed herein are those of the author and do not necessarily represent the views of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed. Support for this research at the Atlanta RDC from NSF (ITR-0427889) is also gratefully acknowledged.

For Chapter Two, I thank my classmate, Dr. Vijay Venkataraman, for particularly helpful feedback and encouragement.

Within the Institute, there are several people for whom I would like to express my sincere appreciation. My advisors, Dr. Chris Forman and Dr. Marco Ceccagnoli, have individually and collectively been a source of encouragement, challenge, and wisdom. My dissertation committee has always been willing to help make my work better, namely Dr. Sabyasachi Mitra, Dr. Eric Overby, and Dr. Marius Niculescu. Dr. Sri Narasimhan has taught me to know my value even when it means taking chances. Dr. Matt Wimble has served as my mentor throughout this process and, together with his wife April, reminded me that I am capable of becoming a solid researcher and given endless support navigating this new world. I have been blessed with the most wonderful classmates including, but not limited to: Kathryn Holmstrom, Dr. Priyank Arora, Dr. Karthik Nattamai Kannan, Yunyoung Hur, Dr. Amy Briedenthal, Dr. Dana Harrari, and Dr. Laurens Steed. It has been a new and life-changing experience to be among so many smart, kind, supportive, nurturing women. And I would like to mention my debt of gratitude for the encouragement, opportunities, and example provided by Dr. Sandra Slaughter. She was the first woman I have ever known that was powerful, feminine, quick to laugh, and still enormously and universally respected in the field. May eternal light shine upon her.

My parents, Jack and Martha Kelly, have been there when I need them, even when I don't ask. My sister, Laura Rigdon, has been a great conference travel companion, from Phoenix to Orlando and points between. My extended family - Jack and Jen Kelly and kids; Ralph Rigdon; the Wolfgangs – have been nothing but encouraging. Graduate school is hard on friendships; when a friendship survives this process, it's for keeps. Ranjani Chandrasekaran and Tara Goodwin, I can't tell you both how much I love you for sticking with me.

And finally, there are these dudes who live in my house. My husband is my bearded other half and my best friend. He has cooked, cleaned, shopped, and done a million other things to keep us afloat while I sat in front of my computer. Gabriel has not only survived his mom being busy, but thrived. His grades and soccer prospects are better than ever, and I couldn't be prouder of him. Oliver has gone from a baby to a brilliant seven-year-old while I worked on these projects, and I am happy to report that he is as attached to me as ever, despite the amount of time I spend looking at a computer screen.

None of these guys have complained about my work. Not once. They love me that much, and I love them. I hope this work contributes to a better life for them.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

My dissertation centers on the impact of information technology (IT) investments on business processes. I seek to understand the way organizations use software to share information with partners in trade and facilitate innovation. Information-sharing IT and process innovation are complementary under the right circumstances, and understanding why and how the strategic use of software impacts organizations has wide-ranging implications, from supply-chain structure to understanding the contribution of the manufacturing sector to the national economy. The first chapter of my dissertation uses proprietary Census data to investigate the impact of e-selling on total factor productivity (TFP). I find that although large plants see a TFP increase related to e-selling, small plants do not. This highlights the need to understand economies of scale related to IT within organizations. The second chapter of my dissertation is an investigation into complementarities between IT and a firm's research and development (R&D) efforts. While there has been considerable attention paid to IT as a complement to firm capabilities, there is less work examining complementarities between IT and other inputs to innovation. This research represents a novel investigation into the relationship between IT investments and a firm's innovative strategy.

CHAPTER 1: INTRODUCTION

This research centers on the impact of information technology (IT) investments on business processes. I seek to understand the way organizations use software to share information with partners in trade and facilitate innovation. Information-sharing IT and process innovation are complementary under the right circumstances, and understanding why and how the strategic use of software impacts organizations has wide-ranging implications, from supply-chain structure to understanding the contribution of the manufacturing sector to the national economy.

The unbiased measurement of one factor of production among many is not a trivial task, but it is an important one because it provides the answer to a fundamental question: if a firm makes an investment in IT, how will it benefit? While there has been considerable work toward an answer to that question in a variety of contexts, generalizability of these studies has been limited largely to a subset of large firms, at least in the United States. This leaves gaps in our understanding of how IT catalyzes firm growth.

Further, there is a robust body of work that examines the relationship between research and development (R&D) and output, but very little that includes the moderating influence of IT. With the increasing digitization of both goods and services in the U.S. economy, IT is a pervasive input that affects nearly every aspect of R&D, from the identification of demand to the pricing structure of the finished product. Understanding the relationship between these two types of investments, as well as their interaction, and the economic output of a firm is important because the measurement of R&D alone may be biased, particularly when significant IT investments are also made.

With advancements in technology come advancements in data analysis and data-driven decision making. Managers now need to provide evidence that supports requests for budget increases and technology expenditures. The "try it and see" method is a privilege often reserved for very large companies that can absorb the cost. Understanding when and how much various factors contribute to economic growth is an important part of enabling such growth. The strategic magnitude and timing of IT investment may mark the difference between a firm which fails and one which grows.

The following chapters examine the impact of IT on output in two ways. In Chapter Two, I use longitudinal establishment-level data from plants in the U.S. manufacturing sector to measure the elasticity of IT investment in the production function. I find that there exists a statistically significant difference between the total factor productivity of "small" (fewer than about 80 employees) and "large" (greater than or equal to about 80 employees) manufacturing plants when both implement electronic commerce systems. This study marks a new understanding of small plant dynamics and emphasizes the importance of economies of scale when making significant IT investments.

Chapter Three examines the interaction between a firm's stocks of IT and Research and Development (R&D) in a production function with Value Added (VA) as the output variable. I find limited evidence that IT and R&D are substitutes and no evidence that environmental dynamism plays a role in this relationship. This work contributes to the bridging of the gap between the economics of IT and the strategic management lines of literature. It also contributes toward an understanding of the impact of inputs to innovation, both in isolation and in combination.

CHAPTER 2: DOES IT LEVEL THE PLAYING FIELD FOR SMALL ESTABLISHMENTS? EVIDENCE FROM MANUFACTURING

2.1. INTRODUCTION

Research has shown a causal link between information technology and productivity growth in the U.S. manufacturing sector (Dranove et al. 2014; Kundisch, Mittal, and Nault 2014; Aral, Brynjolffson, and Wu 2006; Chang and Gurbaxani 2012; Bartel, Ichniowski, and Shaw 2007; among others). These improvements can manifest in myriad ways. For example, there is evidence that computerization is complementary to other organizational change, and these investments taken together contribute to output growth and productivity increases (Brynjolfsson and Hitt 2003).

These insights and much of the related literature have focused on complementary organizational change such as business process change or workplace practices (Melville, Kraemer, and Gurbaxani 2004). Process change means that the basic assumptions that underlie each process affected by the IT system must be challenged, and the processes must be significantly transformed (Hammer and Champy 2001). Process change is often a costly endeavor, involving coordination both internal and external to the establishment; cultural changes; and continuous improvements to the process design (Kettinger and Grover 1995). Does this imply that IT is only beneficial to large organizations that possess sufficient production scale? There is relatively little evidence in related literature to answer this question, as many studies on the productivity of IT investments rely on data from large, publicly-traded firms. Only recently have large systematically available data sets become available that allow researchers to study IT productivity outside of the largest firms (Tambe and Hitt 2012; Jin and McElheran 2017). It is notable that some of the studies showing a lack of IT payoffs have been those that include smaller organizations (Dedrick, Gurbaxani, and Kraemer 2003).

The moderating effect of economies of scale on the relationship between IT and productivity represents a significant gap in research. Further, most prior results have focused on differential returns for large firms (e.g., Tambe and Hitt 2012; Dedrick et al. 2003; Bloom and van Reenen 2007; Saunders 2011). However, many process change efforts related to the deployment of new IT systems may be localized to particular locations, subprocesses, or products within the firm (Bartel, Ichniowski, and Shaw 2007). Thus, analyses that examine differences in the returns to IT system implementation by firm size may be missing important variation.

Using nonpublic microdata from the U.S. Census Bureau, we explore the relationship between electronic selling (e-selling) and total factor productivity (TFP) in a panel of 11,000 plants in the manufacturing sector. We demonstrate higher returns for IT investments in large plants after controlling for an array of plant-level and firm-level characteristics. We investigate several hypotheses for the differential impact of IT in larger plants, including factors related to differential abilities to deploy complex IT (i.e., whether large plants have a more experienced and capable IT workforce), firm size (i.e., whether large plants tend to exist in large firms), and position within the supply chain (i.e., whether plant size is correlated with whether the plant is in an upstream or downstream industry), among others.

Our findings are consistent with previous research related to heterogeneity in the benefits to firms of IT adoption, in that e-selling is more beneficial to large establishments (for example, Tambe and Hitt 2012). Using a fixed-effects panel regression at the plant-year level, we find that e-selling adoption is associated with a 1.37% increase in TFP, on average, controlling for a range of plant and firm characteristics. However, small plants - those with fewer than 80 employees - exhibit returns to productivity that are statistically indistinguishable from zero, on average. In contrast, large plants in our data set show an average of 2.13% increase in TFP upon adoption of e-selling.

We examine whether our results are robust to an instrumental variables analysis. One instrument exploits the e-commerce adoption behavior of customers, while another relies on local variation of adoption costs. Our instruments demonstrate first stage validity and support a causal interpretation of our findings regarding plant size and TFP.

We explore alternative explanations of our findings. At the plant level, we explore whether plant size is a proxy for plant age, intrafirm shipments, technological sophistication, or position within the supply chain. We also conduct a series of analyses to test whether what we see at the plant level might actually be firm-level effects. Our results remain robust throughout these analyses.

In addition to providing insights into the broader question of where and when do the benefits of IT systems appear, we also address the more specific context of the productivity benefits of interorganizational IT systems. For example, EDI adoption has been shown to benefit both supplier and customer over a range of industries (Mukhopadhyay and Kekre 2002). This benefit may be moderated by resource complementarity between the new IT system and the existing IT capabilities of the company (Zhu and Kraemer 2002). Our study adds to this line of research by showing how the IT value added by relationships between organizations varies by establishment size.

Process change both within and between establishments often necessitates complementary organizational change. A large body of research has explored factors that increase the payoffs to technological investments. For instance, it has been shown that the implementation of IT often occurs hand-in-hand with changes in workforce skills and human resource practices (Bartel, Ichniowski, and Shaw 2007; Bresnahan, Brynjolfsson, and Hitt 2002). We are unable to observe the specific changes to complementary organizational and business process practices that we believe give rise to our findings; rather, our results are consistent with the view that these activities are systematically more valuable in particular settings.

2.2. FRAMEWORK

2.2.1 Business process innovation

Implementing new enterprise IT systems like e-selling requires business process innovation (e.g., Attewell 1992; Bresnahan and Greenstein 1996; Bresnahan, Brynjolfsson, and Hitt 2002; McAfee and Otten 2004) Business process innovation can be both complex and costly, involving changes to decision rights, processes, production timelines, hardware, and software, as well requiring buy-in and adaptation from customers.

Business process innovation frequently requires changes to or investments in organizational complements (Milgrom and Roberts 1990; Brynjolfsson and Milgrom 2013. For one, the deployment of new systems frequently require the presence of complementary software systems that increase their value (Arora, Forman, and Yoon

2010; Aral, Brynjolfsson, and Wu 2006; Aral, Brynjolfsson, and Wu 2012). For example, enterprise systems that involve information-sharing between firms require organizations to first adopt systems that focus on within-firm information sharing like ERP. Perhaps more important, deployment of enterprise systems requires adaptations to people and processes. For example, deployment of such systems may require changes to the worker skills and human resource practices (Bresnahan, Brynjolfsson, and Hitt 2002; Aral, Brynjolfsson, and Wu 2012; Bartel, Ichniowski, and Shaw 2007), changes to business and product strategy (Milgrom and Roberts 1990; Bartel, Ichniowki, and Shaw 2007), as well as the locus of firm boundaries and the firm's relationship with partners (Forman and McElheran 2013; Tambe, Hitt, and Brynjolfsson 2012; Aral, Bakos, and Brynjolfsson 2017).

The need for such complementary changes mean that adoption of enterprise systems involves significant adaptation costs that may accrue during and after initial adoption (e.g., Attewell 1992; McAfee and Otten 2004). In particular, adoption of enterprise software frequently involves co-invention, the complementary innovation to systems, processes, and people to adapt general purpose IT systems to the idiosyncratic needs of a particular organization (Bresnahan and Greenstein 1996). Organizations adopting new enterprise systems who invest significantly in co-invention may obtain the greatest benefits from new systems; however, these benefits will be earned only over time, after a costly period of adaptation. For example, among early potential adopters of client/server systems, large technically sophisticated organizations with frontier IT applications had both the greatest potential benefits and greatest costs of adopting the new systems (Bresnahan and Greenstein 1996).

The presence of significant complementary investments suggests that the benefits and costs from adopting enterprise systems may vary with the scale of output within the plant and firm within which the new system is situated; that is, adoption of enterprise software systems like those that enable e-selling may exhibit significant economies or diseconomies of scale (Clemons, Reddi, and Row 1993; Dedrick et al. 2003; Tambe and Hitt 2012).

Some authors have argued that enterprise software systems would exhibit significant economies of scale (Clemons, Reddi, and Row 1993). As mentioned above, new IT systems often involve significant adaptation and co-invention costs; these costs do not vary with output (Clemons, Reddi, and Row 1993). Further, large plants and firms are frequently early adopters of new technologies, thus such firms may be more likely to have related technological complements (Bresnahan and Greenstein 1996; Brynjolfsson and McElheran 2016). Since data related to transactions can be cheaply and easily transferred within the firm; the benefits of such transactional data transfer relative to other, manual transfer processes will be likely be increasing in output. At the firm level, larger, older, more established firms may have more established business models and more formalized processes that are less likely to change and so may be more amenable to the formal business processes that are often created with enterprise systems (Prahalad and Krishnan 2002). At a strategic level, business process innovation that is embedded within complex business processes in large firms may be more difficult to replicate by competitors (Mata et al. 1995; Melville et al. 2004).

However, the complexity of existing systems in large organizations will also impose additional costs when adopting new enterprise systems. Integrating new systems within the existing installed base is costly and can delay adoption (Bresnahan and Greenstein 1996; Forman 2005). Further, large organizations will face larger adjustment costs because of idiosyncratic work processes and tacit knowledge (Ito 1995). In short, while the presence of pre-existing processes and systems can increase the long run benefits of implementing new systems, they can simultaneously increase the costs to deploying them over the short run. This leads to a prediction of delayed returns to enterprise systems in large organizations, a finding which has been reinforced in the literature (e.g., Brynjolfsson and Hitt 2003; Dranove, Forman, Goldfarb, and Greenstein 2014; McAfee and Otten 2004).

We follow recent literature that has examined the organizational implications of new enterprise technology investments by looking at the productivity implications of adopting a specific technology (Aral, Brynjolfsson, and Wu 2012; Bardhan, Mithas, and Lin 2007; Aral and Weill 2007; Dong, Xu, and Zhu 2009; Rai, Patnayakuni, and Seth 2006), namely e-selling. We discuss the technology behind e-selling and how it motivates our estimation approach in the next section.

2.2.2 E-selling

In this paper we examine the productivity implications of adopting electronic sales software (e-selling). We identify adoption from surveys issued by the US Census Bureau. Specifically, the text of the survey in 2002 asks, “Did any of the amount report in 4 [related to Sales, Shipments, Receipts, or Revenue] include e-commerce sales, shipments, or receipts? (E-commerce sales shipments or receipts are online orders for products from customers where price and/or terms of the sales are accepted or negotiated over an Internet, Extranet, Electronic Data Internet (EDI) network, electronic mail, or other online system. Payment may or may not be made online.)"[1] During our observation period, this survey question was modified to include the coordination of shipments, recognizing the complexity of processes linked to electronic sales activities. We refer to this broad umbrella of software-enabled activities as e-selling.

E-selling is one of a class of supply chain execution technologies such as order management. Like many multi-industry studies that study the adoption of enterprise software (e.g., Aral, Brynjolfsson, and Wu 2006; Aral, Brynjolfsson, and Wu 2012; McAfee and Otten 2004; Bardhan, Mithas, and Lin 2007), we do not observe some details about the specific modules and complementary technologies with which e-selling is used. In this section we describe some of the common benefits and costs of adopting e-selling technology, with the recognition that some of these functionalities may not be adopted by specific plants within our sample.

E-selling can reduce the coordination transaction costs associated with contract negotiation, payment, and monitoring of contract compliance (e.g., Gurbaxani and Whang 1991; Clemons, Reddi, and Row 1993). Without e-selling, supply chain partners need to support a range of order processes, such as mailed and faxed purchase orders and phoned orders. Payment options often include cutting and mailing of checks. Further, delivery verification involves phone calls between supply chain partners. E-selling can reduce these mundane transaction costs by reducing the costs of manual orders, payment, and delivery verification.

The information systems literature has highlighted how supply chain technologies such as those we study can also reduce the costs associated with incentive misalignment

[1] The portion in brackets has been added by the authors.

between supply chain partners and Williamsonian transaction costs (e.g., Malone, Yates, and Benjamin 1987; Gurbaxani and Whang 1991; Clemons, Reddi, and Row 1993). For example, supply chain partners that engage in electronic commerce can share information such as inventory data, sales data, and production and delivery schedules (e.g., Lee and Whang 2000). This can alleviate many costs associated with between-firm supply chain transactions. In particular, supply chain partners who share sales data will receive clearer demand signals relative to those who share only order data; this can reduce costs associated with the "bullwhip effect" (Lee, Padmanabhan, and Whang 1997). More generally, such technologies can be used to monitor and track partner behaviors in order to reduce these agency costs.

By reducing coordination costs and the costs of hold-up and incentive misalignment, e-selling and related technologies can directly reduce the costs of supply chain coordination. This will increase the productivity of the plant. However, these technologies may also have indirect effects. By reducing the costs of coordination with external supply chain partners, adoption of e-selling and related technologies may lead to a shift in production strategy at the plant. Namely, because of the now-lower costs of transacting externally, the plant may redirect its output from internal to external customers. If external prices are higher this will yield higher value-added and greater productivity for the plant.

While there may be significant benefits from adopting e-selling, case studies also emphasize the significant business process innovation that may be involved. For one, adopting e-selling technologies often involve defining standards for data exchange, agreeing on this with partners, and then creating systems that will accept these standards (e.g., McAfee and Otten 2004). During our sample period, it may have involved efforts to develop or modify a website and link data entered into the website into new or existing enterprise applications software (e.g., McAfee 2003). Further, the value of e-selling will be greater when accompanied by complementary software systems such as demand planning systems (McAfee 2003, Raman and Singh 1999). All of these represent substantial fixed costs investments for the plants and firms undertaking them, large or small. As a result of these significant fixed costs of adoption, there may exist economies of scale in e-selling adoption.

While the presence of complements that are frequently available in larger firms may potentially increase the value of e-selling, it also means that the new system must also interface with existing systems and processes that support those complements. An instructive example is tool manufacturer Milacron's e-commerce site Milpro.com (Schultz 1999; Teach 1999; McElheran 2015). Milpro.com matched products to applications and advised on proper machine tool setup; assisted in developing and estimating bids on metalworking jobs; provided account information, online order tracking, and online help; and included metalworking formulas, reference tables, calculators, and material safety data sheets.[2] This functionality required integration with multiple catalog data systems, a range of business systems and EDI servers, transactional database servers, and well as multiple servers supporting hosting of the web site, among others. All told, the firm spent one dollar in consulting and customization for every dollar spent on the e-commerce software license (McElheran 2015; Schultz 1999; Teach 1999). In other words, the presence of complementary systems at Milpro.com increased the

[2] The above is a lightly edited summary of "Milpro.com highlights" from Schultz (1999).

value of the system to users, but the integration challenges also increased the risks and costs of deployment, particularly in the short run.

While the above discussion has focused primarily on how the organizational, process, and system-level benefits and costs to adoption vary with scale, strategic-level considerations may also play a role. As emphasized by McElheran (2015), large e-commerce adopters may also be market share leaders. This may mean that large adopters have the bargaining power to induce customer adoption that some smaller plants lack, implying that large adopters are likely to see greater benefits from digitizing sales. On the other hand, adoption of e-selling will require adoption by supply chain partners, which may carry risks if partners find the costs or risks of adoption too high. In addition, customers may find that that an increase in transparency facilitated by IT reduces their ability to engage in practices such as forward buying and diverting, thus resulting in higher costs of procurement (Clemons and Row 1993).

Because the adoption decisions of partners impose costs and confer benefits on the focal plant, there is the potential for bias in our estimates. We address this below, in our investigation of alternative hypotheses and also in our instrumental variable approach, but we acknowledge that it is a potential limitation of our study.

2.2.3 Prediction

As noted above, increases in scale can have competing effects on the productivity of e-selling investments. To summarize, e-selling investments may benefit from economies of scale, but investments made within the context of larger plants and firms will need to be integrated within legacy environments. This may decrease the benefits of e-selling, particularly in the short run. As a result, we leave it as an empirical question whether the productivity benefits of e-selling will be increasing or decreasing in scale.

Our primary analysis will explore differential productivity benefits of e-selling adoption based on the scale of output of the focal plant prior to our sample period. A challenge for our identification strategy will be that variation in plant scale may be correlated with other things that vary cross-sectionally within the plant and firm. To address this concern, we explore differential benefits of e-selling based on other plant and firm characteristics, including firm size, plant and firm age, the extent of vertical integration within the firm, among others. Our empirical approach will be to examine the robustness of focal measure—differential benefits to e-selling adoption with plant scale—to the inclusion of these alternative measures. If our focal measure is robust to these alternatives, we take it as evidence in favor of varying returns of e-selling adoption with plant scale.

2.3. EMPIRICAL MODEL

2.3.1 Baseline specification

We estimate the marginal returns to TFP from e-selling for manufacturing plants in the U.S. using the fixed-effects regression

$$\ln(TFP_{it}) = \beta_0 + \beta_1 ESELL_{it} + \boldsymbol{\gamma} \boldsymbol{X}_{it} + \alpha_i + \boldsymbol{\tau}_t + u_{it} \quad (1)$$

where $ESELL_{it}$ is an indicator equal to 1 if plant *i* engages in e-selling in year *t*.

$\boldsymbol{X}_{it}$ is a vector that includes a set of time-varying plant- and firm-level characteristics. First, we control for both plant and firm size. Kimberly and Evanisko (1981) found in the context of hospitals, large organizations are strongly correlated (>=.69) with specialization, functional differentiation, and technological innovation. In addition, large firms tend to be able to better absorb the risk associated with adoption of new technologies, compared with small firms (Hannan and McDowell, 1984).

We additionally control for the plant's market share. We operationalize this as the plant's share of its 4-digit NAICS industry total value of shipments in a given year, as in Chang, Fernando, and Tripathy (2009).

The number of products is associated with the marginal cost of production (Milgrom and Roberts 1988). In addition, the number of products produced by the firm has been shown to be associated with revenue-based productivity (Bernard, Redding, and Schott 2010). Thus, we control for the number of products manufactured by a plant in a given year.

A plant that exists within a highly competitive industry works within slimmer operating margins, necessitating good management. Plants operating in industries with more competition might be expected to have better-managed processes and have a higher TFP. Our measure of competition is the Lerner index (defined as $1 - \frac{sales - cost\ of\ materials - wages}{sales}$).

Our data are at the plant-year level. We employ plant fixed effects α_i to control for properties of the plant that we do not expect to change over our sample period, such as the plant's choice of management structure. We also include year indicators, $\boldsymbol{\tau}_t$, in order to control for time-varying factors that affect all plants.

2.3.2 Identification challenge

There are several challenges to identification. In general, there may be unobserved factors that affect both the propensity to e-sell and the TFP of a plant. We probe the robustness of our results to the use of instrumental variables. We instrument for the choice of the plant to e-sell with the average adoption rate in counties in which customers of the plant are located. This relies on the assumption that a plant is more likely to adopt e-selling if its customers adopt e-commerce technologies.

We also instrument for *ESELL* using the adoption decisions of other plants in the same county. This instrument relies on the idea that there may be some time-varying, local factors that influence the cost of adoption but not TFP, directly, once other inputs to production are controlled for. For example, the quality of the local labor market may reduce the uncertainty of installing sales software for all plants in the area. These factors are likely to influence the adoption decisions of plants in the same way as the focal plant but not influence the focal plant's TFP directly.

2.3.3 Plant size

We next seek to show whether there are differential productivity returns to e-selling by plant size. To do this, we interact *ESELL* with indicators of plant size:

$$\ln(TFP_{it}) = \beta_0 + \beta_1 ESELL_{it} + \beta_2 ESELL * LARGE_{it} + \boldsymbol{\gamma X}_{it} + \alpha_i + \boldsymbol{\gamma}_t + u_{it} \qquad (2)$$

where *LARGE* is an indicator for plants above approximately 80 employees (we cannot disclose the actual number due to disclosure restrictions).

Equation 2 relies on the assumption that there are no differential trends in unobservables correlated with e-selling for large and small plants. This is a weaker identification assumption than that used in equation 1. As in our estimation of equation (1), we examine the robustness of our results to the use of instrumental variables. We utilize the same instruments as before, adding an interaction with the variable *LARGE.* We also study differential trends in the timing of benefits to adoption for large and small plants.

2.4. DATA

We use non-public economic data from the U.S. Census Bureau in our study. These data are gathered by the Census Bureau both through surveys and through federal tax records. Our sample is constructed from the 2002 and 2007 Census of Manufactures (CMF), the 2002-2010 Annual Survey of Manufactures (ASM), the 2002 Commodity Flow Survey (CFS), the 1999 Computer Network Use Supplement to the ASM (CNUS), and the 2002-2010 Longitudinal Business Database.

The ASM, CMF, and CNUS surveys ask for information related to technology adoption, total value of shipments, costs of materials and labor, and other expenses and revenue. These data sources have several advantages over other sources. Response to these Census surveys are legally mandatory. In 2014, for example, the Annual Survey of Manufactures' response rate was about 74% (Annual Survey of Manufactures Methodology).

In addition, the surveys cover all 4-digit industries in NAICS codes 31-33, which includes all manufacturing industries. Not only is the data set's breadth extensive, but its depth is, as well. Every plant, no matter its size, has some probability of being sampled each year and is sampled with certainty every five years. In years ending in 2 or 7, this survey is called the Census of Manufactures (CMF). In all other years, this survey is called the Annual Survey of Manufactures (ASM). In this study we use the same variables that are collected in both the ASM and CMF. The difference lies in the plants sampled. The CMF is sent to the entire population of manufacturing plants in the United States. In contrast, the ASM is sent to a rotating sample of establishments (United States Census Bureau 2014). Plants with more than 500 employees or $1 billion in manufacturing shipments are sampled every year (the “certainly sample”). However, smaller plants are chosen with some probability based upon employment and industry on a five-year rolling basis. Although the sizes of the sampled plants are representative of the population size distribution, the identities of the non-certainty plants vary every five years. The plants that are surveyed each year are chosen with size-weighted probability but are otherwise chosen at random within each stratum. Our sample is skewed toward larger plants compared with the population, but it does include a large number of small plants - those with fewer than 80 employees. We perform our analyses across the manufacturing sector.

The U.S. Census maintains an annually-updated database containing identifiers and location information for each manufacturing plant in the U.S., called the Longitudinal Business Database (LBD). Employment information is gathered from the LBD, and the plant identifier is used to link plants across surveys and across years.

We restrict our sample to a balanced panel of those plants which are chosen every year between 2002 and 2010, inclusive. Because large plants are sampled with certainty and plants with fewer than 250 employees are not, the plants in our sample have a larger mean number of employees and are older on average – 292 employees and 20 years old in 2002 - compared with the population of U.S. manufacturing establishments (mean 51 employees and a median between 6 and 10 years old). Because we use a within estimator and need variation in our explanatory variable, *ESELL*, we drop all plants that have adopted e-selling in 2002 or before as in Athey and Stern (2002). Our baseline sample includes approximately 11,000 U.S. manufacturing plants (disclosure requirements prevent us from reporting the exact number of plants), each of which is sampled every year from 2002 to 2010, inclusive. This represents between 3% and 4% of all U.S. manufacturing plants during that period. To provide a frame of reference for the U.S. manufacturing economic output represented in our sample, the plants that are sampled with certainty account for 44% of economic activity in the population in 2002 (Yorgason et al. 2011) and 72% in 2012 (U.S. Census American Factfinder).

Our measure of the contribution of IT to productivity growth is total factor productivity (TFP). In accordance with a long line of research beginning with Solow (1957), we start with the premise that technology accounts for the unexplained residual between inputs and output, after accounting for elasticities of inputs. While IT is an input to the production process, it also contributes to the *way* in which inputs are converted to outputs, such as through organizational practices (Tambe, Hitt, and Brynjolfsson 2012). Using assumptions about technically efficient use of inputs and competitive pricing of outputs, TFP is measured as the residual in the Cobb-Douglas production function, measuring labor, materials, structures, equipment, and energy as inputs and total value shipped, adjusted for inventory, as output (Foster, Grim, and Haltiwanger 2016). Many previous studies have computed TFP, but the computation of TFP within the ASM and CMF data sets takes into account information about inputs and outputs that is much more detailed than that computed in most studies. Further details of our measure of TFP are described in Foster, Grim, and Haltiwanger (2016).

Our key independent variable of interest, *ESELL*, is a binary variable that indicates whether a plant reports accepting orders electronically during a given year, as reported in the ASM. We are agnostic about the particular electronic network used. In the ASM/CMF, the respondent is asked both whether the plant e-sells and the portion of total value shipped this represents. A small number of plants in our sample indicate that they do e-sell, but they also indicate a value of zero for the amount shipped corresponding to e-commerce. This may be due to respondent error, data entry error, or an indication of the capability to e-sell with no actual use. To correct for any such discrepancies, and because this study measures the impact of electronic customer interaction rather than the capability to do so, we designated as e-sellers only those plants that report a value of e-selling shipments greater than zero. That is, if the plant indicates a value of e-selling shipments greater than zero, we set *ESELL* = 1, whether or not the plant indicates "yes" to whether it e-sells.

Descriptive statistics for key variables are shown in Table 2.1. It is important to stress the distinction between e-sellers and adopters here. An e-seller is a plant that e-sells in a particular year. The term "adopter" is used here to connote a plant that e-sells at any point during the sample period. 69% of non-adopters have greater than 80 employees, while 79% of adopters are large plants. In comparison, the average plant size across manufacturing in 2002 was 51 employees (Census employment figures). Because we measure the impact of e-selling on plant performance, we include only plants that had not adopted e-selling by 2002.

Pairwise correlations for the baseline sample in 2002 are shown in Table 2.2. The correlations between ln(TFP) and the two plant size variables, natural log of plant employment and *LARGE*, are statistically insignificant at the 10% level. There is no variation in *ESELL*, since by design no plants e-sell in 2002.

Table 2.1: Descriptive Statistics

	Non-adopters 2002 N = 4500		Adopters 2002 N = 6500		Entire Sample 2002 N = 11,000		Non-adopters 2010 N = 4500		Adopters 2010 N = 6500		Entire Sample 2010 N = 11,000	
VARIABLES	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
Plants per firm	219.8	737.6	270.6	964.9	250.0	880.1	173.5	470.0	147.3	392.5	158.0	425.8
ln(TFP)	1.842	0.595	1.821	0.568	1.830	0.579	1.745	0.660	1.766	0.650	1.757	0.654
ln(plant employment)	4.879	1.190	5.179	1.063	5.057	1.126	4.677	1.252	4.993	1.105	4.864	1.177
ln(firm employment)	7.216	2.200	7.621	2.204	7.457	2.211	7.195	2.286	7.488	2.184	7.369	2.231
ln(relative market share)	-5.910	1.315	-5.607	1.256	-5.730	1.289	-6.036	1.450	-5.695	1.350	-5.833	1.402
Inter-plant transfers (1 = yes)	0.182	0.386	0.225	0.418	0.208	0.406	0.191	0.394	0.260	0.439	0.232	0.422
Fraction of plant's wages for non-production workers	0.366	0.214	0.362	0.202	0.364	0.207	0.379	0.219	0.370	0.207	0.373	0.212
Lerner index	0.507	0.500	0.532	0.499	0.522	0.500	0.973	0.163	0.978	0.148	0.976	0.154

Notes: Adopters are those plants which e-sell at any point during the sample period, 59% of the sample. The large designation is time-invariant. All values are rounded. N is approximate, due to Census non-disclosure requirements.

Table 2.2: Pairwise correlations, baseline sample, 2002 (N=11,000)

	ln(TFP)	large (1 = yes)	plants per firm	ln(plant employment)	ln(firm employment)	ln(relative market share)	inter-plant transfers (1 = yes)	fraction of plant's wages for non-production workers	Lerner index
ln(TFP)	1.0000								
large (1 = yes)	-0.0046	1.0000							
plants per firm	-0.0387***	0.035***	1.0000						
ln(plant employment)	-0.0029	0.725***	0.0863***	1.0000					
ln(firm employment)	-0.0991***	0.260***	0.380***	0.377***	1.0000				
ln(relative market share)	-0.0214**	0.456***	0.109***	0.639***	0.324***	1.0000			
inter-plant transfers (1 = yes)	-0.0865***	0.0780***	0.0413***	0.113***	0.262***	0.121***	1.0000		
fraction of plant's wages for non-production workers	0.0473***	-0.0067	-0.0174*	0.0107	-0.127***	-0.0297***	-0.103***	1.0000	
Lerner index	0.507***	0.0371***	0.0349***	0.0364***	0.0717***	0.0336***	-0.0416***	0.0130	1.0000

The period 2002-2010 saw rapid diffusion of digitization within the manufacturing sector. Plants had already begun to adopt technologies to facilitate processes within the establishment, and now they were beginning to look outside plant boundaries to coordinate the flow of goods through the supply chain. E-selling in our context may be any of a number of technologies used to receive orders and/or coordinate shipments. Because e-selling is used to a) share information between vendor and customer and b) construct production forecasts, it is a measure of IT investment that involves factors both within and outside the establishment affecting TFP. Figure 2.1 shows the diffusion of e-selling within our sample.

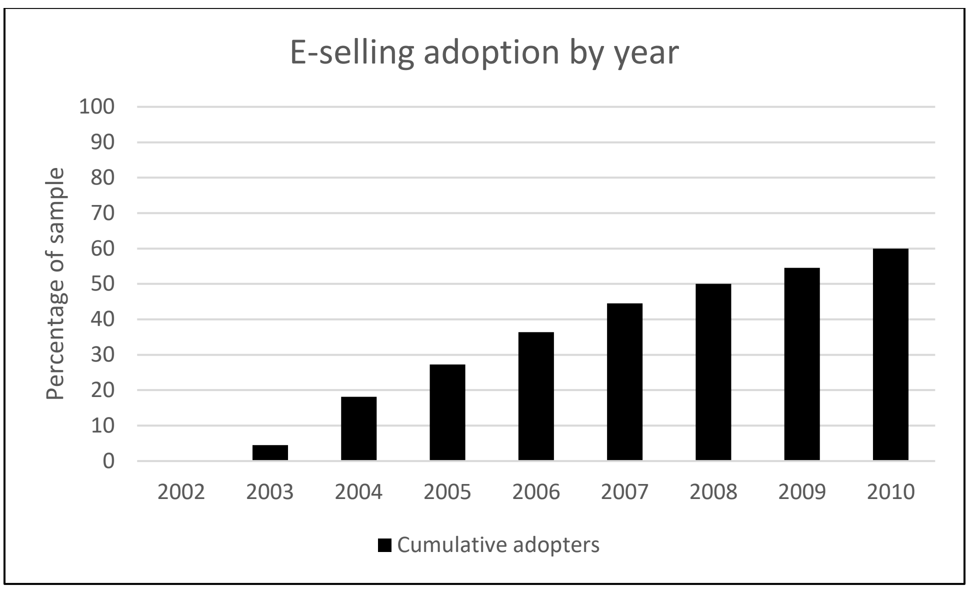

Figure 2.1: Diffusion of e-selling among the plants in our sample

2.5. RESULTS

We begin by demonstrating that there is a relationship between e-selling and TFP and that this relationship differs between small and large plants. We test the robustness of these results to alternative specifications. A series of instruments are employed to substantiate our results, and firm characteristics are discussed.

2.5.1 E-Selling and Total Factor Productivity

In Table 2.3, we show the baseline results across all plants, using fixed effects models with robust standard errors clustered at the plant level. Column 1 reports the correlation between e-selling and TFP without controls. The coefficient on e-selling reveals a statistically significant relationship between a plant's adoption of e-selling and its TFP. In column 2, we add a set of controls to form our baseline specification. The coefficient of e-sell reveals that on average, plants in our sample that adopted e-selling had a TFP that was 1.37% higher than those that didn't. As points of comparison, other studies have found a .05% increase in TFP over three years associated with a 1% increase in IT capital (Han et al. 2011); and the average plant at 90th percentile of TFP has a TFP that is 192% that of a plant at the 10th percentile (Syverson 2011).

We next probe the robustness of our results. To address the concern that TFP and our controls in column 2 are jointly determined, we fix each control at its 2002 value and interact with a linear year trend. Column 3 shows that the correlation between e-selling and TFP remains positive and significant, and of similar magnitude.

Table 2.3: Relationship between e-selling and TFP

	No controls	Baseline	Fixed controls	Industry-relative log TFP	Intensity of e-selling	Lagged effects
	(1)	(2)	(3)	(4)	(5)	(6)
	ln(TFP)	ln(TFP)	ln(TFP)	relative ln(TFP)	ln(TFP)	ln(TFP)
e-sell (1 = yes)	0.0140**	0.0137***	0.0157***	0.0132***		
	(0.0054)	(0.0049)	(0.0054)	(0.0049)		
ln(plant employment)		-0.0510***	0.00375***	-0.0428***	-0.0510***	-0.0513***
		(0.0054)	(0.0008)	(0.0053)	(0.0054)	(0.0054)
ln(firm employment)		0.00554*	-0.000791**	0.00555**		
		(0.0028)	(0.0004)	(0.0028)		
plants per firm		-0.0000183***	0.0000	-0.0000195***	-0.0000183***	-0.0000181***
		(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
ln(relative market share)		0.172***	-0.00136*	0.158***	0.173***	0.172***
		(0.0073)	(0.0007)	(0.0076)	(0.0073)	(0.0073)
inter-plant transfers (1 = yes)		0.0128**	-0.0020	0.0131**	0.0128**	0.0129**
		(0.0061)	(0.0018)	(0.0060)	(0.0061)	(0.0061)
fraction of plant's wages for non-production workers		-0.136***	0.0187***	-0.141***	-0.136***	-0.136***
		(0.0163)	(0.0034)	(0.0162)	(0.0163)	(0.0163)
Lerner index		0.294***	-0.0283***	0.291***	0.294***	0.294***
		(0.0052)	(0.0013)	(0.0051)	(0.0052)	(0.0052)
percentage of shipments e-sold					0.0150**	
					(0.0069)	
e-sell = 1 within two years						0.0117**
						(0.0049)
e-sell = 1 within past 3-4 years						0.0282***
						(0.0073)
e-sell = 1 within past 5-7 years						0.0430***
						(0.0098)
N	99000	99000	99000	99000	99000	99000
R-sq	0.024	0.181	0.041	0.155	0.181	0.182

Notes: Dependent variable in columns 1, 2, 3, 5, and 6 is logged total factor productivity. Dependent variable in column 4 is logged TFP normalized by 3-digit NAICS industry average each year. All specifications include year dummies. Heteroskedasticity-robust standard errors clustered at the plant level in parentheses. N is approximate. *Significant at the 10 percent level. **Significant at the 5 percent level. ***Significant at the 1 percent level.

In column 4, we address the concern that there may be industry-specific unobserved variables that affect the relationship between e-selling and TFP. We demean TFP for each plant by its 4-digit NAICS industry average for each year, as in Foster, Grim, and Haltiwanger (2016). The coefficient of e-selling is still qualitatively similar in magnitude and statistically significant at the 1% level.

Finally, we investigate whether our baseline results are sensitive to the choice of a binary e-selling variable. In column 5, we replace our e-selling dummy with e-commerce revenue as a percentage of total shipments. This regression reveals that an increase in the percentage of output sold through electronic channels from 50% to 75% is associated with a 0.38% increase in TFP.

It has been shown that the returns to IT investment accrue over time, primarily due to business process change that must occur to adapt to the new technology (Brynjolfsson and Hitt 2003). We expect that the returns to e-selling are not immediate. Column 6 presents the estimates for *ESELL* using lagged adoption dummies. We find that averaged across all plants in our sample, some TFP increases occur immediately but continue to accrue over time. For example, immediately after adoption e-selling is associated with a 1.17% increase in TFP; these gains rise to 4.30% five years after adoption. These differences are statistically significantly different from one another at the one percent level.

2.5.2 Plant Size, E-Selling, and TFP

In this section we discuss heterogeneity in our results by plant size. These results constitute the primary findings of our paper. We introduce the variable *LARGE* which is equal to one when plant size exceeds approximately 80 employees in 2002, which represents approximately the 25th percentile of 2002 employment within our sample. This designation is fixed over time, while we employ a separate continuous and time-varying control variable to capture variations in size of the plant by year. In Column 1 of Table 2.4, we repeat our baseline specification (Table 2.3, Column 2) for reference. Column 2 shows the regression of Equation 2, in which we introduce the binary interaction *ESELL*LARGE*, which is equal to 1 if a plant e-sells in a particular year and is large; 0 otherwise. A plant that is large and e-sells has, on average, a TFP that is 3.5 percentage points larger than a small adopter and a TFP that is 2.1% larger than a plant of any size that does not e-sell.

As we did in Table 2.3, in columns 3 and 4 we probe the robustness of our results to the use of time-invariant controls and industry-demeaned TFP. These results indicate that large plants that adopt e-selling have TFP that is 2.14 percentage points to 2.35 percentage points larger than small adopters, results that are similar in magnitude to the results shown in Column 2.

In Column 5, we employ the intensity of use measure in place of the e-selling dummy. The results suggest that for a small plant that increases e-selling from 50% to 75%, TFP will decrease by about 1%. For a large plant that increases e-selling from 50% to 75%, TFP will increase by about 0.66%.

Column 6 shows that large plants continue to see an increase in log TFP both in the short term (less than two years), as well as more than five years after adoption. After two years, the large plants in our sample see an average of 3.20 percentage points higher TFP than small plants. After five years, that difference grows to 4.26 percentage points. These differences are statistically significantly different from one another (t = 74.26).

Table 2.4: Relationship between plant size and TFP returns to e-selling

	Baseline with controls	Plant size	Fixed controls	Industry-relative log TFP	Intensity of e-selling	Lagged effects
	(1)	(2)	(3)	(4)	(5)	(6)
	ln(TFP)	ln(TFP)	ln(TFP)	relative ln(TFP)	ln(TFP)	ln(TFP)
e-sell (yes = 1)	0.0137***	-0.0139	-0.0009	-0.0053		
	(0.0049)	(0.0101)	(0.0112)	(0.0100)		
ln(plant employment)	-0.0510***	-0.0502***		-0.0422***	-0.0500***	-0.0504***
	(0.0054)	(0.0054)		(0.0053)	(0.0054)	(0.0054)
ln(firm employment)	0.00554*	0.00557**		0.00557**	0.00545*	0.00575**
	(0.0028)	(0.0028)		(0.0028)	(0.0028)	(0.0028)
plants per firm	-0.0000183***	-0.0000184***		-0.0000195***	-0.0000183***	-0.0000181***
	(0.0000)	(0.0000)		(0.0000)	(0.0000)	(0.0000)
ln(relative market share)	0.172***	0.172***		0.158***	0.172***	0.172***
	(0.0073)	(0.0073)		(0.0076)	(0.0073)	(0.0073)
inter-plant transfers (yes = 1)	0.0128**	0.0127**		0.0131**	0.0126**	0.0129**
	(0.0061)	(0.0061)		(0.0060)	(0.0061)	(0.0061)
fraction of plant's wages for non-production workers	-0.136***	-0.137***		-0.141***	-0.138***	-0.137***
	(0.0163)	(0.0163)		(0.0162)	(0.0163)	(0.0163)
Lerner index	0.294***	0.294***		0.291***	0.294***	0.294***
	(0.0052)	(0.0052)		(0.0051)	(0.0052)	(0.0052)
large * esell		0.0350***	0.0214*	0.0235**		
		(0.0109)	(0.0122)	(0.0108)		
percentage of shipments e-sold					-0.0355**	
					(0.0159)	
large * percentage of shipments e-sold					0.0619***	
					(0.0170)	
e-sell = 1 within two years						-0.0135
						(0.0104)
e-sell = 1 within past 3-4 years						-0.0025
						(0.0145)
e-sell = 1 within past 5-7 years						0.0091
						(0.0179)
large * e-sell within two years						0.0320***
						(0.0113)
large * e-sell within 3-5 years						0.0388**
						(0.0153)
large * e-sell within 5-7 years						0.0426**
						(0.0188)
N	99000	99000	99000	99000	99000	99000
R-sq	0.181	0.182	0.041	0.155	0.182	0.182

Notes: Dependent variable in columns 1, 2, 3, 5, and 6 is logged total factor productivity. Dependent variable in column 4 is logged TFP normalized by 3-digit NAICS industry average each year. The control variable interactions in column 3 have been removed for concision. All specifications include year dummies. Heteroskedasticity-robust standard errors clustered at the plant level in parentheses. N is approximate. *Significant at the 10 percent level. **Significant at the 5 percent level. ***Significant at the 1 percent level.

2.5.3 Causal Justification

We probe the robustness of our results to the use of instrumental variables. The first utilizes a 2002 Census survey called the Commodity Flow Survey. This survey requires plants to report the zip code to which they ship output for four distinct one-week periods, one per quarter. Plants with greater than 40 shipments in a particular week are asked to provide a logical sample of shipments resulting in between 20 and 40 shipments reported each week. Plants with 40 or fewer shipments in a given week are asked to report all shipments. In this way, we can calculate for each plant the value of goods shipped to each zip code.

We use as our first instrument the adoption behavior of customers. This instrument is based upon the idea that there may exist correlation in technology adoption decisions among supply chain partners that are not correlated with productivity. Because the shipment data in the CFS identifies only the zip code of the customer, we cannot measure whether a plant's customers have adopted e-selling. We can, however, compute the average adoption rate in a particular county, based on zip code. To measure the likelihood that a plant's customers have adopted a means of participating in electronic commerce, we average e-selling adoption over all sampled plants in customer counties, excluding the focal plant if it ships to its own county. The adoption rate in each county is averaged by the number of plants in that county in our sample. This results in a customer adoption percentage that is a simple arithmetic mean of all possible customers of the plant.

The second instrument is a proxy for costs of adoption local to the plant. These costs include things such as telecommunications costs and availability of expertise in installing e-selling systems. If these costs are lower, there will be more plants in the geographical area that adopt e-selling relative to a location in which costs are high. To construct this instrument, we find the arithmetic mean of adoption in the same county as the plant, excluding the focal plant and all other plants owned by the same firm.

In all three of these analyses, the instruments are interacted with the large dummy to instrument for the *ESELL*LARGE* variable. Table 2.5 presents the results of these instrumental variable regressions.

The first stage F-statistics for these regressions range from 72.48 to 4554, which is well above the commonly used threshold of 10 for first stage explanatory power of instruments. In Column 1, we present the just-identified results using our customer adoption instrument. This instrument only exists if our sample includes at least one plant in a zip code to which each focal plant ships output, as reported in the CFS.[3] The coefficient of e-sell is statistically indistinguishable from zero, consistent with the plant size regression results without instruments in Table 2.4, Column 2. The interaction coefficient is significant at the 1% level, but its magnitude is 2.19 times that in the version without instruments. However, the Hausman test retains the null that the coefficients in Column 1 are the same as those in the version without instruments. This may be due to the similarity in magnitude and significance of the control coefficients.

[3] That is, non-missing observations for this analysis will be those where (1) the focal plant exists in the CFS; (2) the focal plant ships to at least one other location; and (3) that location includes at least one plant that was surveyed in the ASM.

Table 2.5: Instrumental variable analysis of Table 2.4, Column 2

	Customer adoption	Adoption by other plants owned by other firms in same county	Both instruments
	(1)	(2)	(3)
FIRST STAGE: Dependent variable is e-sell			
customer adoption	0.519*** (0.0514)		0.426*** (0.0613)
large * customer adoption	0.216*** (0.0202)		0.277*** (0.0407)
county adoption		-0.0579*** (0.0106)	0.0862*** (0.0318)
large * county adoption		0.134*** (0.0093)	-0.0677** (0.0339)
partial R^2	0.0080	0.0029	0.0076
f-statistic	162.3	120.5	72.48
*FIRST STAGE: Dependent variable is e-sell * large*			
customer adoption	-0.362*** (0.0183)		-0.361*** (0.0576)
large * customer adoption	1.103*** (0.0189)		1.066*** (0.0382)
county adoption		-0.574*** (0.0094)	-0.0138 (0.0299)
large * county adoption		0.791*** (0.0083)	0.0327 (0.0319)
partial R^2	0.0811	0.0981	0.0797
f-statistic	1773	4554	819.8
SECOND STAGE: Dependent variable is TFP			
e-sell	-0.0482 (0.0816)	-0.3780* (0.2090)	-0.0112 (0.0837)
e-sell * large	0.0768*** (0.0263)	0.155*** (0.0381)	0.0731*** (0.0265)
overidentification test (p-value)	n/a	n/a	0.731
Hausman test (p-value)	0.191	0.000	0.103
R^2	0.202	0.121	0.200
N	45000	94000	43000

Notes: Dependent variable is logged total factor productivity. All specifications include controls and year dummies. Heteroskedasticity-robust standard errors clustered at the plant level in parentheses. Controls are the same as in Table 2.3. N is approximate. *Significant at the 10 percent level. **Significant at the 5 percent level. ***Significant at the 1 percent level.

In Column 2, using the local county instrument, the results are qualitatively similar but the estimated second stage coefficient for *ESELL*LARGE* is greater. The regression results imply that the productivity benefits of adopting e-selling are 15.5 percentage points greater for large plants than for small plants. The results in this column suggest that adoption of e-selling by small plants is associated with lower productivity. In sum, these results show a 22.5% lower TFP for large adopting plants compared with non-adopters of all sizes; however, this value is not statistically different from zero at the 10% level. The sample is smaller than that in Table 2.4 because the instrument only exists for plants that are not the sole in-sample plant in its county. In this regression, the Hausman test rejects the null that the coefficients in Column 2 are the same as those in Table 2.4, Column 2.

Column 3 shows the results using both instruments. The second stage results in this column are similar to those in Column 1. The overidentification test has a p-value of 0.731. The Hausman test indicates that the coefficients in Column 3 are statistically the same as in Table 2.4, Column 2. However, once again the controls change very little.

In all, our instrumental variables analyses are consistent with our main finding that there exist differential TFP returns to e-selling based on plant size. Large plants see a statistically and economically significant advantage to e-selling relative to small plants and relative to non-adopters.

2.5.4 Alternative hypotheses to economies of scale

In this section, we examine the findings presented in section 5.2 to investigate potential reasons for differential benefits to e-selling by plant size. These results are presented in Table 2.6. Column 2 of Table 2.4 is repeated in the first column for ease of comparison.

It is possible that the difference in change in TFP we see between small and large plants that is related to e-selling is due to experience with systems that a plant gains as it grows and deepens IT capital investments over time. This relative inexperience may result in the misalignment of small plant capability with its IT investment efforts (Melville and Fichman 2014). It has been shown that industries that produce or heavily use IT have higher productivity gains than those that do not (Stiroh 2002). If plant size is a proxy for other IT capabilities within the firm, then interacting prior IT investments with e-selling should absorb some of the differential benefits of e-selling based on plant size.

To capture prior IT investments at the plant level, we employ a 1999 supplemental survey to the ASM called the Computer Network Use Survey (CNUS). This survey had a response rate of 82%, which represented 10.8% of all manufacturing plants. The purpose of the CNUS was to understand how data networks were used in various processes within manufacturing plants. In order to match the CNUS data to our sample, the plants in our sample must have been chosen in the 1999 ASM sample and be in the 82% of plants that responded to the survey. After matching our baseline sample with the CNUS, about 7,100 plants remain. This compares to approximately 11,000 plants in our baseline sample in Table 2.1.[4]

[4] For further details on the CNUS, see McElheran (2015).

Table 2.6: Relationship between plant size and returns to e-selling relative to plant characteristics

	Plant size	IT-intensive control	Inter-plant transfer control	Plant age	Capital intensity	Relative capital intensity	Position within supply chain
	(1)	(2)	(3)	(4)	(5)	(6)	(6)
	ln(TFP)	ln(TFP)	ln(TFP)	ln(TFP)	ln(TFP)	ln(TFP)	ln(TFP)
e-sell	-0.0139 (0.0101)	-0.0170 (0.0133)	-0.0105 (0.0102)	-0.0118 (0.0109)	-0.0171* (0.0104)	-0.0414*** (0.0145)	-0.0188* (0.0104)
large*e-sell	0.0350*** (0.0109)	0.0359** (0.0109)	0.0368*** (0.0109)	0.0357*** (0.0110)	0.0332*** (0.0110)	0.0311*** (0.0113)	0.0327*** (0.0110)
e-sell * high investments		-0.0046 (0.0028)					
e-sell * IPT			-0.0195** (0.0089)				
e-sell * old plant				-0.0041 (0.0088)			
e-sell * high capital intensity					0.00803 (0.0071)		
esell * high relative capital intensity						0.0339*** (0.0120)	
esell * downstream							0.0225** (0.0092)
N	99000	64000	99000	99000	99000	99000	99000
R-sq	0.182	0.185	0.182	0.182	0.182	0.182	0.186

Notes: Dependent variable is logged total factor productivity. All specifications include year dummies. Heteroskedasticity-robust standard errors clustered at the plant level in parentheses. N is approximate. *Significant at the 10 percent level. **Significant at the 5 percent level. ***Significant at the 1 percent level.

In one question of the CNUS survey, respondents were asked whether they use the following computer-networked applications to manage production: computer-aided design/computer-aided engineering (CAD/CAE), production design, production scheduling, production monitoring, test and acceptance of product, and R&D outsourcing. All but the last of these indicate the use of IT for internal purposes, rather than between establishments. Using the first five, we construct a measure of internal IT experience using principal component analysis. The vector of largest variance, which we label "high investments," is used as an indication of the IT sophistication of the plant relative to the rest of the sample. Other methods of operationalizing high investments, such as simple investment counts and inclusion of other technologies, do not significantly change our results.

In Column 2, we interact this measure of IT sophistication with *ESELL* to measure whether plants with more experience with IT within the establishment see higher TFP returns to e-selling. Plants with prior experience related to these technologies do not expertise differential benefits from adopting e-selling. Further, the coefficient on the interaction *ESELL* and *LARGE* differs little between columns 1 and 2.

We next examine differential returns to e-selling based on supply chain relationships. It is widely believed that communication technologies like e-selling will be particularly effective at reducing the costs of coordinating with external trading partners (e.g., Malone, Yates, and Benjamin 1987; Gurbaxani and Wang 1991; Ray, Wu, and Konana 2009; Forman and McElheran 2013). As a result, the productivity benefits of e-selling are likely to be smaller for plants that ship a high proportion of output to other plants within the same firm - that is, plants with a high proportion of interplant transfers (IPT). If IPT is correlated with plant size—as may be the case, for example, if large plants tend to produce final products rather than intermediate goods—then our results may reflect in part correlation between plant size and position within the supply chain. We create a dummy variable that indicates whether the plant reports *any* IPT shipments in the ASM/CMF in the focal year. (Using Census data, Atalay, Hortacsu, and Syverson (2014) report in a sample of 65,700 establishment-years that almost half of establishments do not have any IPT.) Column 3 shows that small plants with IPT > 0 do, in fact, see fewer benefits from e-selling. However, it continues to be the case that large plants experience productivity benefits from e-selling: the coefficient estimate for *LARGE* * *ESELL* changes little between columns 1 and 3.

Older plants may be less likely to seek out innovations such as those we study, in part due to costs associated with reconfiguring old processes (Banker et al. 2006; McElheran 2015). In Column 4, we investigate whether our measure of plant size captures variance in age by interacting *ESELL* with a variable that indicates whether the age of a plant is above the 25th percentile as of 2002. We find that the relationship between plant size, e-selling, and TFP persists.

Another phenomenon that may be confounded with plant size is capital intensity, defined as stock of computing hardware. Our dummy is equal to 1 if a plant possesses a greater stock of hardware than its industry (4-digit NAICS) average in a particular year; 0 otherwise. To the extent that large plants possess larger stocks of computing machinery and related equipment, they may possess greater facility with the implementation of new systems and may broadly may possess complements that increase the value of adopting IT systems. We test whether capital intensity explains some of the variation in our *LARGE*ESELL* coefficient in Column 5. We find that e-selling does not appear to be more valuable in high-capital environments after controlling for plant size.

To address the concern that small plants in high-capital environments may be influencing these results, we normalize computer stock by total value of shipments each year in the regression shown in Column 6. The normalized IT stock dummy is equal to 1 if a plant possesses a greater share of hardware, relative to its own output, than its 4-digit industry mean. We find here that in high relative capital environments, e-selling is correlated with a 3.39% increase in ln(TFP), after controlling for plant size. The incremental effect of e-selling for large plants is still consistent to that in earlier regressions. What does change, however, is the effect of e-selling for all plants. When we control for relative capital intensity, e-selling is associated with a *decrease* in TFP across the sample. This indicates that the stock of hardware within a plant, *relative to sales*, has bearing on its ability to successfully implement e-selling for all plants. It does not, however, explain the difference in outcomes between small and large plants.

As has been noted elsewhere, the business process innovation required for adopting e-selling technology is likely to be greater in a business-to-business (B2B) than in a business-to-consumer environment (B2C) (McElheran 2015). B2B transactions typically involved both more complicated products and more complicated transactions, which increase the complexity of digitizing these processes. If position along the value chain is correlated with plant size, then our estimated differential effects by plant size may be capturing position within the value chain. In Column 7, we present the results of the regression in which we control for position within the supply chain. A plant is classified as either an upstream or a downstream plant in two steps. First, we identify the commodities of the plant as “final use” or “input” commodities, as defined by the 2002 Bureau of Economic Analysis input/output table. Second, we designate a plant as downstream if at least 50% of its total value of shipments are designated as final use commodities; upstream otherwise. While downstream e-sellers see a larger increase in productivity relative to other types of adopters, this does not have a significant impact on the large plant coefficient.

We see evidence, though these plant-level regressions, that a host of factors related to the productivity of an establishment do not affect the differential effects enjoyed by large plants. An interesting take-away from this set of regressions is the finding that relative capital intensity reduces the value of e-selling for both small and large plants, in approximately equal measure. The incremental value of e-selling for large plants relative to small plants, though, remains the same.

2.5.5 The role of firm characteristics

In Table 2.7 we investigate whether the effects of plant size may be capturing heterogeneous returns to e-selling by parent firm characteristics. As in our other tables, Column 1 repeats the baseline results for ease of comparison.

First we allow heterogeneity in the effects of e-selling based on firm size. Large plants may be disproportionately found within large firms, and so our plant size regression results may differential effects of e-selling for large firms. Tambe and Hitt (2012) show that the larger firms in their sample see higher returns from IT investments that persist over time. In Column 2, we explore heterogeneity in the effects of *ESELL* by firm size. We create an indicator for large firm size that is based on the 25^{th} percentile of firm size within our sample in 2002. Our results are robust to variations of this threshold. This designation is constant throughout the sample period, therefore we interact the large firm dummy with *ESELL*. Column 2 shows that large firms with plants that e-sell see a 0.91 percentage point smaller increase in TFP from adopting e-selling than small firms, but the results are not statistically significant. The coefficient on *LARGE*ESELL* (i.e., differential effects from e-selling based on plant size) remains qualitatively similar to that in other specifications.

Table 2.7: Differences in the productivity implications of e-selling relative to firm characteristics

	Plant size	Firm size	Multiunit firm	Firm age	IT-intensive firm
	(1)	(2)	(3)	(4)	(5)
	ln(TFP)	ln(TFP)	ln(TFP)	ln(TFP)	ln(TFP)
e-sell	-0.0139 (0.0101)	-0.0091 (0.0111)	-0.00416 (0.0213)	0.000693 (0.0118)	-0.0145 (0.0136)
e-sell * large plant	-0.0350*** (0.0109)	0.0370*** (0.0109)	0.0362*** (0.0110)	0.0377*** (0.0109)	0.0325** (0.0142)
esell * large firm		-0.0091 (0.0082)			
esell * multiunit firm			-0.0126 (0.0107)		
esell * old firm				-0.0223** (0.0095)	
esell * high firm investments					-0.0038 (0.0106)
N	99000	99000	99000	99000	99000
R-sq	0.182	0.182	0.182	0.182	0.185

Notes: Dependent variable is logged total factor productivity. All specifications include year dummies. Heteroskedasticity-robust standard errors clustered at the plant level in parentheses. N is approximate. *Significant at the 10 percent level. **Significant at the 5 percent level. ***Significant at the 1 percent level.

We similarly explore differences in our results based on whether the plant is a member of a multi-establishment firm. For example, establishments in multi-establishment firms might have access to resources beyond the focal establishment. Moreover, multi-establishment status could capture features of the firm related to firm size or organizational structure. The results in Column 3 show there are no differential benefits to e-selling from being a plant within a multi-establishment firm, although the differential benefits of e-selling for large plants are robust to the inclusion of this new variable.

Another measure of a plant's access to expertise is the age of its parent firm. Older firms may have lower returns to adopting new IT systems if, for example, they have an installed base of older IT systems (Bresnahan and Greenstein 1996). We test the extent to which our main result related to plant size might be explained by the age of the parent firm (Column 4). In contrast with plant age, firm age does affect the payoffs to productivity from e-selling. We define old firms as those above the 25th percentile of age, within sample in 2002. Across the sample on average, e-sellers owned by old firms see a decrease in TFP of about 2.2% over those owned by young firms. However, the comparative advantage of large plants over small ones remains relatively unchanged. These results are robust to other age thresholds.

In an attempt to measure the IT expertise of the firm, we construct a variable constructed by principle component analysis using the same five technologies as in Table 2.6, Column 2. These technology counts are aggregated in two steps. First, the number of the technologies in use, out of five, is counted for each plant. Next, a sum of these counts is taken among all child plants of a particular firm. This firm-level measure of IT intensity is used to compute the principal vector of variance between firms. We note that the source of these data, the 1999 ASM, was a survey, and so this measure is likely not to include data from all of the other plants in the firm. (We note that this is different from other variables, such as age, number of establishments, and size, which are based on Census years in which all U.S. plants are sent surveys and so we have nearly-complete data from which to calculate firm-level measures.) The data show no statistically significant effect of firm technology adoption on the relationship between e-selling and productivity (Column 5). This may be due to measurement error because of incomplete data with which to compute these measures. As elsewhere, the differential effects of e-selling for large plants is robust to the inclusion of this alternative measure.

In short, these results show that there remain significant differences in the productivity benefits of e-selling based on plant size, even controlling for differential returns based upon firm characteristics. This set of regressions provides evidence that plant characteristics are not proxies for the characteristics of the firms that own them. Further, they underscore the importance of plant-level data when studying the influence of IT investments on productivity.

2.6. DISCUSSION

2.6.1 Summary

A well-developed line of literature has investigated the role of IT in the productivity of manufacturing plants. We extend this literature by using a unique data set to study productivity differences by plant size in a way that has been difficult to accomplish using prior data sets. Small plants constitute the majority of plants in the U.S. manufacturing sector. It is important, therefore, to understand when inferences drawn from large plants and firms are generalizable to other contexts.

In this study, we find that the benefits of e-selling are increasing with plant size. Our estimates of e-selling suggest a total factor productivity boost of 2.1% for large plants over non-adopters, which is 3.5 percentage points higher TFP than that of plants with fewer than 80 employees. We find that these results are robust to variations in variable and sample construction. Further, instrumental variable analysis supports a causal interpretation of the relationship between e-selling and TFP. We test whether factors such as human capital, plant age, capital intensity, and position within the supply chain might explain the differences in productivity benefits between small and large plants, and they do not. Our analyses suggest that there are economies of scale related to the adoption of e-selling.

2.6.2 Implications

Much of the multi-industry productivity studies in the IT value literature is drawn from surveys of large firms. Advancing recent work that has highlighted differences in the benefits to IT investment by firm size (Tambe and Hitt 2012), we show considerable differences in the productivity benefits of e-selling investments by plant size. While we provide evidence of the factors that are not explaining this variance and show a role for economies of scale, data limitations limit our ability to isolate the specific mechanisms that drive our results. Future research might focus on the source of these differences.

By providing evidence of differential benefits by plant size, our results also have implications for practitioners. The costs of implementing technologies such as e-selling might be higher than expected and persist over a period of several years. This does not mean that small plants should refrain from IT adoption; it implies that expectations and financial planning ought to reflect the particular context in which the plant operates.

2.6.3 Limitations and directions for future research

We have provided a range of analyses to bring a causal interpretation to our findings. However, as in any study of the productivity benefits of IT, we must be aware of the effect that omitted variables have on our estimates. Because our focus is on differential benefits of e-selling by plant size—rather than the benefits of e-selling per se—our identification assumptions are somewhat weaker than if our focus had been on the average effects of e-selling adoption.

This provides several opportunities for further exploration. For example, one area of further exploration may be related to managerial quality. Future work should investigate the quality of managers and more generally the role of human capital in the payoffs from IT implementation.

We have provided evidence of the benefits of studying the productivity benefits of IT at the plant level. However, as is common in plant-level studies of productivity, our approach does not allow for plant level decisions related to inputs to be directly correlated with one another across business units within the same firm (Banker et al. 2006; Bharadwaj 2000). This is an open area for future research.

Further, while we believe that our data are some of the best available for studying productivity at the plant level, there remains the potential for measurement error to influence our estimates. This was particularly the case for some of our instruments. Because we have only customer zip codes, and not customer plant identifiers, we can only measure a likelihood that the customer has e-commerce capabilities.

Finally, the particular technology used to electronically accept orders is not observed in our data. Because some technologies, such as EDI, involve significant coordination with the customer and others, such as web-based order forms, involve relatively little, it would be useful to control for the type of technology used. As we explain in Section 1, we take care to distinguish the coordination technologies that we study here from other types of IT. Further disaggregation of technologies, such as between search and order completion technologies, would be desirable in the future (Mishra, Konana, and Barua 2007).

Our analysis of firm effects reveals a surprising lack of influence of firm characteristics on the TFP of plants we study. This suggests that the researcher should consider data at the plant data when possible. Care should be taken to interpret our findings in the context of manufacturing, but an investigation of establishment size in other sectors should be undertaken in order to assess the generalizability of this study.

2.7. REFERENCES

Aral, S., Bakos, Y., and Brynjolfsson, E. 2017. “Information Technology, Repeated Contracts, and the Number of Suppliers,” *Management Science* (64:2), pp. 592-612 (doi:10.1287/mnsc.2016.2631).

Aral, S., Brynjolfsson, E., and Wu, D.J. 2006. “Which Came First, IT or Productivity? Virtuous Cycle of Investment and Use in Enterprise Systems.” Available at SSRN (http://dx.doi.org/10.2139/ssrn.942291).

Aral, S., Brynjolfsson, E., and Wu, L. 2012. "Three-Way Complementarities: Performance Pay, Human Resource Analytics, and Information Technology,” *Management Science* (58:5), pp. 913-931 (doi:10.1287/mnsc.1110.1460).

Aral, S., Weill, P. 2007. "IT Assets, Organizational Capabilities, and Firm Performance: How Resource Allocations and Organizational Differences Explain Performance Variation," *Organization Science* (18:5), pp. 763-780.

Arora, A., Forman, C., and Yoon, J.W. 2010. "Complementarity and Information Technology Adoption: Local Area Networks and the Internet," *Information Economics and Policy* (22:3), pp. 228-242.

Atalay, E., Hortacsu, A., and Syverson, C. 2014. "Vertical Integration and Input Flows," *American Economic Review* (104:4), pp. 1120-1148.

Athey, S. and Stern, S. 2002. "The Impact of Information Technology on Emergency Health Care Reforms." *RAND Journal of Economics* (33), pp. 399-432.

Attewell, P. 1992. "Technology Diffusion and Organizational Learning: The Case of Business Computing," *Organizational Science* (3:1), pp. 1-19.

Banker, R. D., Bardhan, I. R., Chang, H. H., and Lin, S. 2006. "Plant information systems, manufacturing capabilities, and plant performance,” *MIS Quarterly* (30:2), pp. 315-337.

Bardhan, I., Mithas, S., and Lin., S. 2007. "Performance Impacts of Strategy, Information Technology Applications, and Business Process Outsourcing in U.S. Manufacturing Plants." *Production and Operations Management* (16:6), pp. 747-762 (doi: 10.3401/poms).

Bartel, A., Ichniowski, C., and Shaw, K. 2007. "How does information technology affect productivity? Plant level comparisons of product innovation, process improvement, and worker skills,” *Quarterly Journal of Economics* (122:4), pp. 1721-1758 (doi: 10.1162/qjec.2007.122.4.1721).

Bernard, A. B., Redding, S., and Schott, P. K. 2010. "Multiple-Product Firms and Product Switching," *American Economic Review* (100:1), pp. 70-97 (doi: 10.1257/aer.100.1.70).

Bharadwaj, A. S. 2000. "A Resource-Based Perspective on Information Technology Capability and Firm Performance: An Empirical Investigation." *MIS Quarterly* (24:1) pp. 169-196.

Bloom, N., and van Reenen, J. 2007. "Measuring and Explaining Management Practices across Firms and Countries," *The Quarterly Journal of Economics* (122:4), pp. 1351-1408.

Bresnahan, T., and Greenstein, S. 1996. "Technical progress and co-invention in computing and in the uses of computers," *Brookings Papers on Economic Activity*, pp. 1-83 (doi: 10.2307/2534746).

Bresnahan, T. F., Brynjolfsson, E., and Hitt, L. M. 2002. "Information technology, workplace organization, and the demand for skilled labor: Firm-level evidence," *Quarterly Journal of Economics* (117:1), pp. 339-376 (doi: 10.1162/003355302753399526).

Brynjolfsson, E., and Hitt, L. M. 2003. "Computing productivity: Firm-level evidence," *Review of Economics and Statistics* (85:4), pp. 793-808 (doi: 10.1162/003465303772815736).

Brynjolfsson, E. and McElheran, K. 2016. "The Rapid Adoption of Data-Driven Decision-Making," *American Economic Review* (106:5), pp. 133-139.

Brynjolfsson, E. and Milgrom, P. 2013. "Complementarity in Organizations," in *The Handbook of Organizational Economics*, R. Gibbons and J. Roberts (eds.), Princeton University Press, pp. 11-55.

Bureau of Economic Analysis Input-Output Accounts Data. Available at https://www.bea.gov/industry/io-annual.htm.

Chang, H, Fernando, G. D., and Tripathy, A. 2009. "Productivity Spillovers along the Supply Chain," Working Paper.

Chang, Y.B. and Gurbaxani, V. 2012. "Information Technology Outsourcing, Knowledge Transfer, and Firm Productivity: an Empirical Analysis," *MIS Quarterly* (36:4), pp. 1043-1063.

Clemons, E., Reddi, S., and Row, M. 1993. "The Impact of Information Technology on the Organization of Economic Activity: The 'Move to the Middle' Hypothesis," *Journal of Management Information Systems* (10:2), pp. 9-35.

Clemons, E. and Row, M. 1993. "Limits to Interfirm Coordination through Information Technology: Results of a Field Study in Consumer Packaged Goods Distribution," *Journal of Management Information Systems* (10:1), pp. 73-95.

Dedrick, J., Gurbaxani, V., and Kraemer, K. 2003. "Information Technology and Economic Performance: A Critical Review of the Empirical Evidence," *ACM Computing Surveys* (35:1), pp. 1-28.

Dong, S., Xu, S., Zhu, K. 2009. "Information Technology in Supply Chains: The Value of IT-Enabled Resources Under Competition," *Information Systems Research* (20:1), pp. 18-32 (doi: 10.1287/isre.1080.0195).

Dranove, D., Forman, C., Goldfarb, A., and Greenstein, S. 2014. "The Trillion Dollar Conundrum: Complementarities and Health Information Technology," *American Economic Journal-Economic Policy* (6:4), pp. 239-270 (doi:10.1257/pol.6.4.239).

Fichman, R. 2000. "The Diffusion and Assimilation of Information Technology Innovations," in *Framing the Domains of IT Management: Projecting the Future… Through the Past*, Robert Zmud (ed.), Pinnaflex Educational Resources, pp. 105-128.

Fichman, R. and Melville, N. 2014. "How Posture-Profile Misalignment in IT Innovation Diminishes Returns: Conceptual Development and Empirical Demonstration," *Journal of Management Information Systems* (31:1), pp. 203-239 (doi: 10.2753/MIS0742-1222310109).

Forman, C. 2005. "The Corporate Digital Divide: Determinants of Internet Adoption," *Management Science* (51:4), pp. 641-654.

Forman, C. and McElheran, K. 2013. "The Digital Erosion of Firm Boundaries: Complementarities between IT Use and Production Chain Organization in U.S. Manufacturing," Working paper.

Foster, L., Grim, C., and Haltiwanger, J. 2016. "Reallocation in the Great Recession: Cleansing or Not?" *Journal of Labor Economics*, (34:1), pp. S293-S331 (doi:10.1086/682397).

Granger, C. W. 1969. "Investigating Causal Relations by Econometric Models and Cross-Spectral Methods," *Econometrica: Journal of the Econometric Society*, (37:3), pp. 424-438.

Gurbaxani, V., and Whang, S. J. 1991. "The Impact of Information-Systems on Organizations and Markets," *Communications of the ACM* (34:1), pp. 59-73 (doi: 10.1145/99977.99990).

Hammer, M. and Champy, J. *Reengineering the Corporation: A Manifesto for Business Revolution*. HarperCollins Publishers, Inc., 2001.

Han, K., Chang, Y. B., and Hahn, J. 2011. "Information Technology Spillover and Productivity: The Role of Information Technology Intensity and Competition," *Journal of Management Information Systems* (28:1), pp. 115-145 (doi: 10.2753/Mis0742-1222280105).

Hannan, T. H., and Mcdowell, J. M. 1984. "The Determinants of Technology Adoption - the Case of the Banking Firm." *Rand Journal of Economics* (15:3), pp. 328-335 (doi: 10.2307/2555441).

Ito, Harumi. 1995. "The Structure of Adjustment Costs in Mainframe Computer Investment," Working Paper, Stanford University.

Jin, W., and McElheran, K. 2017. "Economies Before Scale: Survival and Performance of Young Plants in the Age of Cloud Computing," Working Paper.

Kettinger, W., and Grover, V. 1995. "Special Section: Toward a Theory of Business Process Change Management," *Journal of Management Information Systems* (12:1), pp. 9-30.

Kimberly, J. R., and Evanisko, M. J. 1981. "Organizational Innovation - the Influence of Individual, Organizational, and Contextual Factors on Hospital Adoption of Technological and Administrative Innovations." *Academy of Management Journal* (24:4), pp. 689-713 (doi: 10.2307/256170).

Kundisch, D., Mittal, N., and Nault B. 2014 "Using Income Accounting as the Theoretical Basis for Measuring IT Productivity," *Information Systems Research* (25:3), pp. 449-467 (doi: 10.1287/isre.2014.0534).

Lee, H. L., Padmanabhan, V., and Whang, S. 1997. "Information distortion in a supply chain: The bullwhip effect." *Management Science* (43:4), pp. 546-558 (doi: 10.1287/mnsc.43.4.546).

Lee, H.L. and Whang, S. 2000. "Information Sharing in a Supply Chain," *International Journal of Manufacturing Technology and Management* (1:1), pp. 79-93.

Malone, T. W., Yates, J., and Benjamin, R. I. 1987. "Electronic Markets and Electronic Hierarchies," *Communications of the ACM* (30:6), pp. 484-497.

Mata, F., Fuerst, W., and Barney, J. 1995. "Information Technology and Sustained Competitive Advantage: A Resource-based Analysis," *MIS Quarterly* (19:4), pp. 487-505.

McAfee, A. 2003. "Moore Medical Corporation," Harvard Business School Case 9-601-142.

McAfee, A. and Otten, M. 2004. "IBM: Ordering Midrange Computers in Europe." Harvard Business School Publishing.

McElheran, K. 2015. "Do Market Leaders Lead in Business Process Innovation? The Cases of E-business Adoption." *Management Science* (61:6), pp. 1197-1216 (doi:10.1287/mnsc.2014.2020).

Melville, N., Kraemer, K., and Gurbaxani, V. 2004. "Information technology and organizational performance: An integrative model of IT business value." *MIS Quarterly* (28:2) pp. 283-322.

Milgrom, P. and Roberts, J. 1988. "Communication and inventory as substitutes in organizing production." *The Scandinavian Journal of Economics* (90:3), pp. 275-289.

Milgrom, P. and Roberts, J. 1990. "The Economics of Modern Manufactruing: Technology, Strategy, and Organization," *The American Economic Review* (80:3), pp. 511-528.

Mishra, A.N., Konana, P., and Barua, A. 2007. "Antecedents and Consequences of Internet Use in Procurement: An Empirical Investigation of U.S. Manufacturing Firms," *Information Systems Research* (18:1), pp. 103-120.

Mukhopadhyay, T. and Kekre, S. 2002. "Strategic and Operational Benefits of Electronic Integration in B2B Procurement Processes." *Management Science* (48:10), pp. 1301-1313.

Prahalad, C. and Krishnan, M. 2002. "The Dynamic Synchronization of Strategy and Information Technology," *MIT Sloan Management Review* (43:4), pp. 24-33.

Rai, A., Patnayakuni, R., and Seth, N. 2006. "Firm Performance Impacts of Digitally Enabled Supply Chain Integration Capabilities," *MIS Quarterly* (30:2), pp. 225-246 (doi: 10.2307/25148729).

Raman, A. and Singh, J. 1999. i2 Technologies, Inc. Harvard Business School Case Study 9-699-042.

Ray, G., Wu, D., and Konana, P. 2009. "Competitive Environment and the Relationship Between IT and Vertical Integration," *Information Systems Research* (20:4), pp. 585-603.

Saunders, A. 2011. "Has Information Technology Leveled the Competitive Playing Field?" Essays on Information Technology and Intangible Capital, Massachusetts Institute of Technology.

Schultz, B. 1999. "Tooling Along Online: Metalworking Products Maker Milacron Wins Our Debut E-Com Innovator Award," *Network World*, https://web.archive.org/web/20060317003900/http://www.networkworld.com/ec/0222innovator.html.

Solow, R. M. 1957. "Technical change and the aggregate production function." *The Review of Economics and Statistics* (39:3), pp. 312-320.

Stiroh, K. J. 2002. "Information Technology and the US Productivity Revival: A Review of the Evidence," *Business Economics* (37:1), p. 30.

Syverson, C. 2011. "What Determines Productivity?" *Journal of Economic Literature* (49:2), pp. 326-365.

Tambe, P., and Hitt, L. M. 2012. "The Productivity of Information Technology Investments: New Evidence from IT Labor Data." *Information Systems Research* (23:3), pp. 599-617 (doi:10.1287/isre.1110.0398).

Tambe, P., Hitt, L. M., and Brynjolfsson, E. 2012. "The Extroverted Firm: How External Information Practices Affect Innovation and Productivity," *Management Science* (58:5), pp. 843-859.

Teach, E. 1999. "Setting up Cybershop," *CFO,* http://ww2.cfo.com/technology/1999/10/setting-up-cybershop.

United States Census Bureau. 2014. Annual Survey of Manufactures Methodology.

United States Census Bureau. American Factfinder, https://factfinder.census.gov.

Von Hippel, E. 1986. "Lead Users: A Source of Novel Product Concepts," *Management Science* (32:7), pp. 773-907 (doi: 10.1287/mnsc.32.7.791).

Yorgason, D., Bridgman, B., Cheng, Y., Dorfman, A. H., Lent, J., Liu, Y. K., Miranda, J., and Rumburg, S. 2011. "Cutoff Sampling in Federal Surveys: An Inter-Agency Review," *Proceedings of the American Statistical Association, Section on Government Statistics*.

Zhu, K. and Kraemer, K. 2002. "e-Commerce Metrics for Net-Enhanced Organizations: Assessing the Value of e-Commerce to Firm Performance in the Manufacturing Sector," *Information Systems Research* (13:3), pp. 275-295.

CHAPTER 3: THE INTERPLAY OF INFORMATION TECHNOLOGY AND R&D AND ITS IMPLICATIONS FOR INNOVATION

3.1 INTRODUCTION

The management of knowledge is a driver of the organization of innovative activity within a firm. The source of such activity is often customer demand, but that demand is often characterized by variation in frequency and extent of change. With the advent of IT-enabled innovation, ideas are more quickly and easily passed between employees and divisions. This allows quick identification of demand, broadening the pool of possible solutions, and raising the potential payoff. However, codified and digitized knowledge can facilitate undesirable knowledge spillovers. It can therefore be difficult to determine the appropriate level of information availability and cross-functional integration within the firm. Previous research has studied the links between IT and knowledge spillovers (e.g. Forman et al, 2008; Tambe and Hitt, 2013), dynamicism (Teece, Pisano, and Shuen, 1997), demand-led innovation (Arora, Cohen, and Walsh, 2015), and R&D knowledge management (Ceccagnoli, Venturini, and van Zeebroeck, 2018). However, to our knowledge, none has examined the role of knowledge management systems in the payoffs from innovation in a dynamic environment.

We find that in our sample, across U.S. industries and over a five-year panel covering 2004-2008, IT and R&D are substitutes. This may indicate that many firms choose to decentralize knowledge in the pursuit of innovation rather than employ knowledge integration and risk spillovers to competitors. Testing the moderating effect of variation in industry demand over time, we find that IT and R&D are substitutes in stable environments. In dynamic environments, however, we find evidence of neither substitution nor complementarity.

The importance of this work lies in the fact that with increasing digitization, the ability to identify demand, assess capability fit, and quickly execute a solution are critical components of competitive advantage. We assert that the payoffs to innovation depend heavily upon these capabilities, which often necessitate the use of information systems and R&D investments in tandem.

3.2.LITERATURE

3.2.1 Key constructs

We begin our literature review by establishing the definitions of our key constructs- dynamic environments and dynamic capability - and our perspective of IT. Dynamic environments are those in which technological change is rapid enough such that time-to-market and strategic decision timing are vital and future demand and competition is difficult to assess (Teece, Pisano, and Shuen 1997). Dynamic capability is the firm's capacity to adjust competencies and innovate as a response to such an environment. In this study, we take the view of IT as a knowledge management system which allows the firm to coordinate efforts across departments or with collaborators external to the firm, store codified information, and analyze demand through simple analytics or advanced forecasting techniques. A knowledge management system may manifest as an Intranet chat forum or a procedural database, for examples.

3.2.2 IT-enabled Knowledge Spillovers

This study primarily builds upon prior work related to IT-related knowledge spillovers as well as dynamic capabilities. The IT-related knowledge spillover literature has focused primarily on the flow of knowledge through access to specialized knowledge (Dedrick et al 2003; Tambe and Hitt, 2013). Adoption of knowledge management systems incurs costly restructuring and process reconfiguration, more so if the firm has significant investments in existing integrated processes and structures (Utterback and Abernathy, 1975). The adoption and integration of new technologies, or skill-biased technical change, shifts the desired labor market toward employees with tacit or difficult-to-learn knowledge (Bresnahan, Brynjolfsson, and Hitt 2002). This knowledge may be of people, processes, or products and represents a desired inflow.

Besides the acquisition of human capital through hiring, inflows may be induced in other ways. External knowledge may be accessed by participation in common interest groups like conferences or online common-interest sites. Huang, et al (2019) show that participation in an online discussion forum can facilitate knowledge spillovers. This effect is moderated negatively if the information is difficult to learn and positively with prior IT investments. In addition, spillovers can arise from research collaborations or from the study of existing patents (Duguet, 2006).

Although information technologies may aid in the functions of research and development, the link between IT and output is, in the IT-enabled innovation literature, unclear. While there is evidence that IT may strengthen a firm's absorptive capacity, that evidence does not support the notion that this realized absorptive capacity results in an increase in commercialized innovations, except through the development of patents

(Joshi et al 2010). How, then, is it possible that IT and R&D are complements? There is a body of evidence that successful IT-enabled innovation depends strongly on complementary organizational capabilities (e.g. Ashurst, Freer, Ekdahl, and Gibbons 2012; Bresnahan, Brynjolfsson, and Hitt 2002; Melville et al 2004; Pavlou and el Sawy 2006). These complementary capabilities, not measured in the Joshi study, may explain why some firms are able to leverage IT in the successful introduction of products to market, particularly because of the coordination involved between business functions and along the supply chain.

3.2.3 Environmental Dynamism

In dynamic environments, the knowledge necessary for innovation is quickly changing or draws upon a wide range of expertise domains. Knowledge management systems can be an important link between such disparate knowledge domains (Yoo, Henfridsson, and Lyytinen 2010). For example, physicians use decision support systems (DSS) to draw upon medical knowledge outside their own expertise in order to quickly diagnose and innovate to create a treatment plan tailored to an individual patient.

Experience with KMS can positively influence innovation if the firm possesses NPD competencies. This advantage is amplified in the presence of environmental dynamism (Pavlou and el Sawy, 2006). Dewan et al (2000) found that when dynamism is demand-led, the costs of gathering and analyzing demand can be reduced through the use of cookies, a relatively basic Internet 2.0 technology. This suggests that the ability to leverage technologies to identify demand may be beneficial to the innovative process. Further, information systems allow a firm to access a greater number and variety of sources of knowledge, as well as provide the infrastructure for more accurate, faster, and more relevant knowledge inflows (Overby et al, 2006).

According to a survey by Arora, Cohen, and Walsh (2015), demand accounted for over a quarter of manufacturing innovation. The uncertainty introduced by demand fluctuations, however, presents a choice to the firm. These fluctuations cause what Bloom (2007) calls a "caution effect", in which R&D responsiveness to changes is reduced. Some firms find that the best response to environmental dynamism is to take no action at all, particularly if they are older firms with an established reputation (Stieglitz, Knudsen, and Becker 2016). On the other hand, while riskier, the payoff to innovation can be higher, moderated by the market share of the firm (Prajogo, 2015). For firms willing to respond to dynamism, flexibility is a strong predictor of performance (Anand and Ward, 2004).

3.2.4 The Consequences of Knowledge Flows and Restrictions in Innovation

The tension between, on one hand, the protection of intellectual property arising from R&D, and on the other, the risk of inducing competition via knowledge leakage, informs the firm's response in a dynamic environment. In one sense, the purpose of knowledge integration and the aim of IP protection via patenting are at odds with one another. Venturini, Ceccagnoli, and van Zeebroeck (2018) found that in industries with a higher likelihood of undesirable knowledge spillovers (outflows), knowledge integration is a less likely strategy.

Even the protection of innovative output can itself induce knowledge spillovers. For example, patents represent knowledge outflows by nature of legally-mandated public design disclosure. As a result, some firms choose to only patent innovations which can be more easily protected from imitation (Moser, 2005), or when the product could be reverse-engineered by one who is able to examine the product (Hall and Harhoff, 2012). On the other hand, Ouellette (2012) found that patents are often not examined closely by competitors due to the difficulty to a non-attorney to understand them. Some variation in the value of patents, the "patent premium", has been shown (Arora, Ceccagnoli, and Cohen, 2008), with the manufacturing sector realizing the largest gains from patenting.

3.2.5 The Gap

Table 3.1 represents a survey of related empirical studies. By no means is it exhaustive; rather, it is intended to highlight the range of subjects and methodologies in the IT-enabled innovation and strategic management of innovation lines of literature. Only four of these studies consider the interactions between IT and R&D. Of those three, only two consider complementarity between the two - Wang, Wang, and Li (2017) and Hall, Lotti, and Mairesse (2013).

In the Wang et al study, there are three major differences from ours. First, their dependent variable, Tobin's Q, measures the extent to which a firm might be over- or undervalued, based upon its assets. This depends upon the market value of the firm, which can be influenced by a number of subjective measures, including, for example, predictions by shareholders or news media of future events. In short, Tobin's Q is a forward-looking measure. Our dependent variable, added value, is a measure of the difference between sales and operations costs. Since these are events that have already happened, this is a backward-looking measure. Both are valid, but the interpretation of the interaction of IT and R&D investments will differ. Second, the Wang et al study relies upon the interaction term of a Cobb-Douglas specification to establish complementarity between the explanatory variables. Although we establish a baseline using the Cobb-Douglas form, we employ a translog specification to establish whether complementarity is present. We feel that this is a more appropriate approach, as we do not expect our data to exhibit constant returns to scale. Third, the Wang et al sample is from the population of Chinese firms. This is a critical point because between China and the United States, there may be systematic differences in, for examples, organizational structure or external incentives. Their measure, therefore, is difficult to interpret as it applies to U.S. firms, so we include it here as robustness to our primary measure of dynamism.

The Hall, Lotti, and Mairesse differs from ours in setting – Italian manufacturing firms only – and methodology – they use CDM. However, the most significant difference between their study and lies in the results of their analyses. Hall et al find evidence of neither substitutability nor complementarity in their sample. We find evidence of substitutability, which is moderated by demand-side dynamism and robust to several specifications.

Most of the IT-enabled spillover literature is concerned with the flow of tacit knowledge, primarily through management of human capital and labor mobility. In addition, IT is usually considered an input to production, rather than an asset whose interaction with other inputs depends heavily on the environment in which it is employed. There is evidence that uncertainty caused by dynamic environments is a moderating factor in the relationship between IT, R&D, and output (Wang, Wang, and Li 2017). This suggests that neglecting to consider turbulence obscures some of the heterogeneity which might explain returns to IT investment. Ours is one of the first studies to consider the relationship between IT-enabled knowledge flows, R&D investments, and productive output in such environments.

In the strategic management literature, the focus on R&D tends to be on knowledge appropriation and knowledge spillovers, largely ignoring the impact of knowledge management through the use of IT. On one hand, such management systems enable identification and analysis of demand, collaborative research and development, and increased responsiveness. On the other hand, KMS requires knowledge that is codified and thus easily transferrable, which facilitates undesired leakage. Therefore, the impact of IT on R&D strategies is unclear. This represents a significant gap in both the IT and strategic management literatures.

Table 3.1: A sample of related empirical approaches

Study	Model	Dependent variable	RHS variables	Key finding
Abdih Joutz 2006	knowledge production fn	TFP	patent stock	supports existence of within-firm knowledge spillovers over time
Arora et al 2008	innovation production fn	R&D, patent activity	patent effectiveness	heterogeneous value added to innovation by patent across industries
Bloom et al 2013	C-D type production fn	quantity	industry/technology spillovers	technology spillovers stronger than misappropriation
Brynjolffson and Hitt 1995	C-D and translog	va	IT stock	quantify relationship between IT and productivity using several estimation methods
Ceccagnoli et al 2014	translog	product pipeline	R&D, in-licensing investments	identifies heterogeneities in R&D and in-licensing interaction by firm
Forman and van Zeebroeck 2012	DID	patents	Internet adoption	Internet adoption associated with increased patents from geographically-dispersed teams
Gao and Hitt 2012	GLS	trademarks	IT capital stock	IT associated with higher trademark holdings
Hall et al 2013	CDM model	labor productivity	ICT and R&D	both R&D and ICT associated with both innovation and productivity
Hall and Mairesse 1995	C-D type production fn	lva/employee	R&D capital	measurement insights related to returns to R&D investment
Huang et al 2019	translog	value added	knowledge stock	prior IT investments increase IT knowledge inflows
Joshi et al 2010	SEM	patents, innovations	IT	IT-enabled knowledge capabilities are positively related to firm innovation
Kleis et al 2012	C-D type production fn	patents	IT, R&D	IT positively associated with innovative output
Martin and Nguyen-Thi 2015	CDM model	va/employee	R&D, ICT	reveal heterogeneity in R&D/IT interaction related to innovative output
Mohnen et al 2006	generalized Tobit	share of innovative sales	four measures of R&D	demonstrates a new measure of innovation
Neuhausler et al 2011	OLS	Tobin's Q, ROI	patent activity	forward patent citations and patent family size positively influence firm market value
Pavlou and el Sawy 2006	SEM	competitive advantage	IT leveraging competence	IT-leveraging competence builds competitive advantage
Syverson 2011	C-D type production fn	TFP	capital, labor, materials	describes heterogeneity in productivity benefits of IT
Tambe and Hitt 2013	C-D	va	IT labor flows	movement of human capital influences IT knowledge spillovers
Wang et al 2017	OLS	Tobin’s Q	IT investment, R&D, dynamism	R&D positively moderates relationship between IT and Tobin’s Q
Wu et al 2017	C-D type production fn	sales	analytics skills, innovation	analytic abilities and not innovation practices are complementary with process improvement

3.3 THEORY

3.3.1 The innovative process

The innovative process is defined by Price and Bass as a "complex feedback-type information processing system" (1969). A number of studies have described the nature of innovation as an explore-versus-exploit decision (Wu, Hitt, Lou working), the leverage of functional competencies (Pavlou and el Sawy 2006), a response to demand (Dewan, Jing, and Seidmann 2000), the generation of options (Sambamurthy, Bharadwaj, and Grover 2003), and the leverage of intellectual capital (Meso and Smith 2000). While taking individual perspectives, these studies possess the characteristic belief that the management of knowledge is at the core of innovation. Thus it is important to understand the role that knowledge management systems and other information technologies play in the innovative process.

The complexity of the innovative process described by Price and Bass lies in two basic notions: one, the uncertainty of sources and content of new information, and two, the tendency of innovation to occur across departmental and firm boundaries. Including the processing of feedback, we shall consider the ways in which IT enables the innovative process in these three areas.

3.3.2 Uncertainty

Innovation can be a response to uncertainty in demand. Demand may vary over time in terms of volume or product preference. Uncertainty can be ameliorated by the use of information systems that allow a firm to detect changes in demand and analyze them (Dewan, Jing, and Seidmann 2000). If the firm has sufficient substantive capabilities to meet the demand, then fit is established. An example of the identification of demand is the use of cookies in Internet searches (Dewan, Jing, and Seidmann 2000). These cookies can be used to describe target markets and identify clusters of similar products. Miller and Friesen (1983) found that the firms that respond best to uncertain environments tend to be those with the best analytical capability. Spreadsheets and accounting systems, as well as more complex statistical programs and customizable enterprise resource planning software, are all tools that can enable innovation through analytical reduction of uncertainty. The identification of demand and fit assessment are the beginning to advanced research. The presence of analytical IT in firms performing advanced research may facilitate the early stages of research. In this way, the two functions may be complements.

Supply-side uncertainty may be an antecedent of innovation, as well. The use of new technologies or new processes on the supply side may lower the cost of inputs to production, inducing the firm to find ways to incorporate the new products or to use more of the lower-cost products. Innovation of new products to make use of the new input may make use of computer-aided design, computer-aided machinery, and digital prototyping (Joshi and Lauer 1998, Loch, Terwiesch, and Thomke 2001). Economies of scale dictate that scaling production in response to lower input costs improve the productivity of an organization (Angle and Forman, 2019). This may benefit larger firms more than smaller ones, such that complementarities between IT and R&D exist more strongly for larger firms. Further, implementing technological change can be costly, particularly if the technology consists of machinery or other hardware. The Angle and Forman study suggests that smaller organizations face the challenge of absorbing the costs of such changes. These improvements to productivity may necessitate a reduction in R&D investment, thus empirically resulting in substitutability between IT and R&D.

3.3.3 Collaboration

Collaborative technologies such as the Internet reduce the cost of communicating with geographically dispersed teams (Forman and van Zeebroeck 2012). Online messages boards are another significant source of inflow of ideas to the firm (Huang, Ceccagnoli, Forman, and Wu 2019). The Internet also provides a platform for crowdsourcing of innovative ideas. For example, Lego maintains a site (https://ideas.lego.com) that solicits ideas for new building sets and produces some of the designs for purchase in stores.

Collaboration becomes more integral to the innovative process with the complexity of the product. Complex products, such as jet engines, often necessitate technical expertise from disparate knowledge domains. Hobday (1998) describes a complex product as "a temporary coalition of organisations which usually cuts across the boundaries of single supplier firms." The digitization of knowledge lowers the cost of sharing information across firm boundaries, enabling the innovation of complex products (Yoo, Henfridsson, and Lyytinen 2010). The more complex the product, the more likely that IT is to complement existing R&D resources. In contrast, with simple products, the risk of knowledge misappropriation may outweigh the benefits of collaboration. In this case, we expect that IT and R&D may be substitutes.

3.3.4 Feedback

As the firm begins to choose between possible design paths, a degree of path dependence is established, which then further constrains the ability of the firm to explore alternative designs later in the innovative process (Teece, Pisano, and Shuen 1997). Therefore, feedback and analysis early in the process are important. Once a direction has been decided, there is often a period of learning by doing. Particularly when the innovation is driven by technological change, this period of self-generated feedback increases the absorptive capacity of the firm and thus its ability to respond to future technological change (Cohen and Levinthal 1990).

Customers can aid in co-creation, testing of prototypes, and offering new design path ideas (Nambisan 2002). The usefulness of this customer feedback, particularly if the innovation involves significant customer co-invention, increases to the extent that it can be easily shared and analyzed. In March 2019, the United States Government Accountability Office, recognizing the importance of feedback in the innovative process, noted that the U.S. Department of Homeland Security "is not well positioned to integrate the results and share lessons learned because limited R&D customer feedback information is collected and analyzed." The USGAO thus recommended an accounting system be adopted to "align processes and information sources for collecting R&D project data."

The earlier in the design process that feedback is collected, the earlier a design path can be established, allowing the deployment of resources toward other objectives, such as the exploration of alternative design paths in later stages of development. The impact of R&D on a firm's output is likely to be increased in the presence of information technologies, such as feedback forms on websites, shared documents between collaborators, and stored "lessons learned" from earlier projects.

3.3.5 Tension

Sambamurthy, Bharadwaj, and Grover (2003) found that the relationship between IT and innovation is mediated by agility, digital options, and entrepreneurial alertness. The functions of R&D within a firm may depend in part upon the ability to identify entrepreneurial opportunities and the firm's ability to pivot toward those opportunities. If a firm possesses the dynamic capabilities of alertness and agility, then the marginal dollar of IT should increase the value of a given R&D investment through feedback, analytic ability, and lowered collaboration costs.

On the other hand, when IT enables the integration of knowledge within the firm, there may be lower payoffs to R&D efforts due to, for example, leakage to competitors. There is empirical evidence that knowledge integration and R&D efforts can be substitutes in the manufacturing sector (Ceccagnoli, van Zeebroeck, and Venturini 2018). When tacit knowledge is codified and digitized, it becomes explicit knowledge. In this form, it is no longer a competitive-advantage-sustaining resource (Randeree 2006; Barney 1991). Substitutability between IT and R&D may also occur when the firm does not possess sufficient capability to translate IT-enabled insights into a design path. The marginal dollar spent on R&D may be, in this case, spent in pursuit of suboptimal design or fruitless avenues of research.

3.3.6 Conditions leading to substitution

The digitization of information about products, processes, design paths, or customer demand may facilitate appropriation by competitors. In addition, if the product complexity is low, that information may be more easily used to copy an innovation. If the chance of knowledge leakage is high, and that knowledge is of a nature that can be easily understood by competitors, then IT which facilitates knowledge flows may decrease the returns to R&D.

3.3.7 Conditions leading to complementarity

The payoffs to R&D may be enhanced by systems which facilitate the identification of demand, analysis of design paths, and transfer of information between collaborators. In addition, the IT-enabled ability to analyze design decisions and digitally prototype inventions can increase the value added attributable to R&D through the facilitation of ex ante design expectations and reducing the likelihood of costly design revisions.

There are likely to be some degree of both benefits and costs to the employment of IT in the functions of R&D. Complementarity does not imply that only benefits are present, only that they outweigh the negative consequences. For example, while there exists the possibility of unwanted knowledge spillovers, leaked information does not necessarily become usable knowledge by competitors. Tacit knowledge primarily resides within the human capital of the firm. The information stored in spreadsheets, shared in discussion forums, and used to create computational models, among other things, is explicit. When information leakage occurs, the receiving firm must possess the substantive capabilities to assimilate it, combine it with its own tacit knowledge, and use the resulting ideas for competitive advantage (Cohen and Levinthal 1990). Therefore, information leakage alone is not sufficient to lose the advantages provided by information technologies in the research and development functions. In this way, a firm’s relevant tacit knowledge and dynamic capability should be expected to positively moderate the relationship between IT, R&D, and added value.

3.3.8 Dynamism

Dynamic environments can be defined by “rate of change and innovation in the industry” and “uncertainty or unpredictability of the actions of competitors and customers” (Miller Friesen 1983). Many research studies have been performed in this setting, which has been described as [a] “high velocity environment”, “volatility”, “environmental dynamism”, and “rapid technological change” (Eisenhardt and Bourgeois 1988; Kurz and Senses 2016; Jiao et al 2013; Prajogo 2015; Teece, Pisano, and Shuen 1997).

Of these terms, only “rapid technological change” indicates the source of the dynamism. However, if the ability of the firm to respond to changes in its environment depends upon organizational complements, then one must consider the source of those changes to understand what organizational complements are necessary.

Corporate strategy change in dynamic environments, however, is not always beneficial to the firm (Stieglitz et al, 2016; Prajogo, 2015). For example, some markets reward risk-taking while others reward reliable service or product offerings. However, In the Miller and Friesen study, the firms that responded well to environmental dynamism were those with the ability to better analyze the environment (that is, involve more factors in decision making) and innovate. This suggests that there exists heterogeneity in firms’ ability to capitalize on rapidly shifting demand.

The growth in the sophistication of analytical software packages allows a firm to consider more variables and make predictions for a greater number of scenarios in a shorter amount of time, and with less programming experience, than earlier methods. They are also more expensive. The firm that spends more on computation power, i.e.

hardware and IT labor, will potentially be better able to respond to environmental dynamism through analytic capability. Indeed, Pavlou and el Sawy (2006) found that dynamism amplifies the effect of IT-related competence. This would suggest that the complementarity of IT and R&D increases in the presence of environmental dynamism.

Another, different, response to uncertainty in a dynamic environment is what Bloom (2007) calls the “caution effect” and Stieglitz et al call “inertia.” In this case, firms often choose not to respond to dynamism, choosing instead to exploit existing capabilities (Stieglitz, Knudsen, and Becker 2015). From that perspective, we might expect that firms choose an exploitive strategy, spending less on IT investments, R&D, or both. Barnett and Pontikes (2006) showed that over time, a firm that survives an intensely competitive environment can suffer from attempting to enter a new market. Prajogo (2015) also found that competitive environments negatively moderate the effect of product innovation on firm performance. One key insight of both studies lies in the importance of fit between the innovation and existing firm capabilities. Substitutability between IT and R&D, that is, may occur in the presence of market-specific competencies that do not translate well to new endeavours.

3.4 DATA

3.4.1 Sources

We draw from two primary data sources: the Harte Hanks Computer Intelligence Technology database (HH) and Standard & Poor’s Compustat Annual Fundamentals database (Compustat). HH collects information related to technology investments and human capital from establishments throughout North America and Europe. Relevant to this study are the number of PCs, servers, employees, and IT employees. Compustat is a collection of data gathered from a number of sources, including Securities and Exchange Commission filings, press releases, shareholder reports, and others.[5] Our data related to sales, operating expenses, and non-IT assets are gathered from Compustat.

Our data set comprises an unbalanced panel from 2004 to 2008. It includes the 698 firms that are included in both HH and Compustat during that time period (N = 3301). Summary statistics for the pooled sample, first year, and last year of our study are shown in Table 3.2, and correlations between these variables are shown in Table 3.3.

Table 3.2: Descriptive Statistics

	Pooled (N = 3301)			2004 (n = 627)			2008 (n = 590)		
Variable	Mean (sd)	Min	Max	Mean (sd)	Mi n	Max	Mean (sd)	Min	Max
ln(value added)	7.10 (1.50)	0	11.35	6.68 (1.84)	0	10.77	7.19 (1.69)	0	11.35
ln(IT stock)	3.83 (1.57)	0	7.49	3.52 (1.76)	0	7.28	4.28 (1.13)	0.61	7.48
ln(R&D stock)	2.22 (3.01)	0	10.37	1.83 (2.58)	0	9.08	2.64 (3.32)	0	10.37
ln(non-IT capital)	7.44 (1.86)	0	12.69	7.22 (1.98)	0	12.46	7.71 (1.77)	0	12.63
ln(non-IT labor)	5.55 (1.26)	0.03	9.25	5.32 (1.36)	1.7 6	9.20	5.62 (1.21)	0.09	9.25

[5]For more information, see Dai (2012).

Table 3.3: Correlations between variables

	ln(value added)	ln(IT stock)	ln(R&D stock)	ln(non-IT capital)	ln(non-IT labor)
ln(value added)	1.00	0.43	0.24	0.54	0.57
ln(IT stock)	0.43	1.00	0.12	0.13	0.71
ln(R&D stock)	0.24	0.12	1.00	0.22	0.28
ln(non-IT capital)	0.54	0.13	0.22	1.00	0.33
ln(non-IT labor)	0.57	0.71	0.28	0.33	1.00

Notes: Pearson method of correlation is used. N = 3301. All correlations are statistically significant at the 1% level.

3.4.2 IT artifact

We are interested in the interplay between technologies that facilitate information sharing and the effort of a firm to conduct proprietary research. Our proxy for the ability of information to flow between employees or divisions of a firm is the firm's stock of personal computers, servers, and IT labor. Because software provides the infrastructure by which some firm knowledge is stored and transferred, we use IT labor investments as a measure of both the firm's expertise in facilitating knowledge flows, i.e. networking, and its investments in software. This method assumes that a large portion of labor expenses are for the purposes of software development (Huang, Ceccagnoli, Forman, and Wu 2019).

Hardware, such as PCs and servers, facilitate the ability of a firm to share knowledge, through networking and the storage of codified knowledge. The more heavily a firm invests in computers and other hardware, the more networked the employees are likely to be with one another. The stock of hardware for a firm in a given year is collected from the Harte Hanks CI Technology Database. We aggregate the number of personal computers and servers over all sites to the parent firm level. PC and server prices for 2004 and 2005 were obtained from the Gartner Dataquest Global PC Annual Forecast. The prices in subsequent years were assumed to follow a linear price decrease. The market value of PCs and servers were converted to 2005 values and multiplied by the number of PCs plus servers for each firm.

IT labor falls into two categories in the 2000 BLS Standard Occupational Classification system: Computer and Mathematical Occupations and Architecture and Engineering Occupations. Although network occupations are closely related, they are split between computer scientists and computer engineers. Since both are likely to impact a firm's store and transfer of technical knowledge, we collect data on the wages and total compensation for both major categories of employment. These values, obtained from the BLS Occupational Employment Statistics series, are averaged over the two employment categories (15 and 17). The IT labor cost is computed as the total compensation per employee multiplied by the number of employees in the firm as reported in the CI database site data for each year. Because the CI database reports the range of IT employees into which each firm fits, each range is converted to its mean value for the purposes of calculating total labor cost. The firm's stock of IT is calculated as the sum of PC market value, server market value, and three times IT labor expenses in a given year, as in Brynjolfsson and Hitt (1995).

3.4.3 Research and development

Research and development expenses are collected in the Compustat Annual Updates – Fundamentals series, available at https://wrds-web.wharton.upenn.edu, variable XRD. Since the Compustat data come from Security and Exchange Commission filings which are required for all publicly-traded firms, a firm not reporting any R&D expenditures is likely to have none. A stock measure is created using a 15% yearly depreciation rate, as in Hall, Jaffe, and Trajtenberg (2005).

To measure complementarity, we interact IT and R&D stock. We are interested not in whether a firm innovates per se, but rather the economic impact of its choice in investments in the inputs to innovation that we study. If we chose to use a patent count, for example, we would only be able to tell the extent to which increasing expenditures in a particular category raises the likelihood of innovation. What we want to know, though, is how IT affects the productivity of a firm - are IT expenditures complementary with R&D in the output of a firm? If so, how does this complementarity affect the value added by production?

3.4.4 Value added

Our measure of output is value added (VA). VA is a measure of production output that has less variance across industries than sales (Brynjolfsson and Hitt 2003; Dewan and Min 1997; Huang et al 2019). VA is calculated by subtracting the cost of materials from firm sales. Sales and cost of materials are obtained from the Compustat database. Price deflators for sales are given by the BEA Gross Output and Related series, and those for materials are given by the BLS Producer Price Index.

3.4.5 Controls

We also control for non-IT capital and non-IT labor, since both are likely to influence the output of the firm. For example, product testing does not fall under our classification of IT activities, but prototyping is a common part of the innovative process. We measure non-IT capital as deflated total capital minus deflated computer capital, as in Huang et al

(2019). To construct non-IT capital, we calculate average age of capital stock (Compustat item PPEGT) as the quotient of accumulated depreciation, depreciation, and amortization (Compustat item DPACT) and depreciation and amortization (Compustat item DP). Firm age is used to apply industry-year-level deflators, which are given in the BLS Detailed Capital Measures series, to the value of capital stock. Subtracting IT stock from capital stock gives non-IT capital.

Total labor and related expenses are either gathered from Compustat (item XLR), when available. Otherwise, labor expenses are calculated as the number of employees in the CI database multiplied by total employee compensation in private, goods-producing industries (Bureau of Labor Statistics Employer Costs for Employee Compensation Historical Listing). Next, IT labor is subtracted from total labor to calculate non-IT labor.

3.4.6 Dynamism

We define dynamism as the unexplained portion of an industry's deviation from average sales growth over a four-year window, following the form of Kurz and Senses (2016). This approach captures the unexplained, time-variant, deviation in demand for a particular industry. Consider, for example, one industry that exhibits dynamism in our study: oil and gas extraction. Demand growth over time might depend in part upon household economic growth, which is relatively predictable and common across industries. There may also be a component of demand that does not vary much over time but affects oil and gas extraction differently, and predictably, than other industries. An example of this is the growth of alternative energy sources. Finally, there is a component of demand change that cannot be predicted and varies over time. This may be weather affecting offshore drilling activity, an oil rig explosion, or a Federal subsidy encouraging exploration of natural gas sources. This component varies not only in frequency, but in magnitude as well.

Mathematically, we begin by aggregating total sales to the industry level by addition over NAICS 2-digit industries. Next, we define growth rate of industry *j* at time *t* as follows:

$$\gamma_{jt} = s_{jt} - \mathrm{s}_{\mathrm{j(t-1)}} = \phi_j + YEAR_t + \nu_{jt} \tag{1}$$

Industry fixed effects (ϕ_j) are employed to isolate time-invariant idiosyncrasies of each industry that might affect sales growth. Year dummies are also added to control for time-varying effects that affect all industries in a similar manner. This regression yields a residual, ν_{jt}, which indicates the unexplained, time-varying, difference between the growth of a particular industry and the average growth across industries.

To mathematically define dynamism, we employ the Kurz and Senses "residual approach", which is to calculate the standard deviation of the residual from equation (1) above:

$$\sigma_j^w = \sqrt{\frac{1}{w}\Sigma\nu_{jt}^2} \tag{2}$$

where *w* represents the number of years over which growth rate is calculated; we use four years. This is a better measure than simply measuring the residual, since we have taken the difference of squares in the specification in Equation 1. This will better capture the average deviation between any two years than an arithmetical average. This variable is a measure of variance in growth rates of demand. Finally, we define a dynamic industry as one in which σ_j^w is greater than the aggregate median.

3.5 EMPIRICAL APPROACH

3.5.1 Baseline model

We employ a panel approach derived from the Cobb-Douglas production function, in which IT and R&D are inputs and innovations are the outputs. This firm-year-level regression takes the following form:

$$va_{it} = b_1 it_{it} + b_2 rd_{it} + b_3 it_{it} S * rd_{it} + \boldsymbol{\gamma} \boldsymbol{X}_{it} + year_t + u_i + e_{it}, \qquad (3)$$

where va_{it} is the value added from sales of firm *i* in year *t*, *it* is the stock of IT hardware and labor in firm *i* in year *t*; *rd* is the stock of R&D expenditures; ***X*** is our matrix of controls; *year* is our year control; *u* is our time-invariant error term for firm *i;* and *e* is the error term of our dependent variable. Because random effects are statistically inconsistent in our sample (χ^2=87.53, df=5, p=0.0000), fixed effects are employed (random effects for the baseline are also reported). Further analyses show evidence of heteroskedacticity, thus an HC3 heteroskedasticity consistent covariance matrix appropriate for small samples is used (Long and Ervin, 2000).

3.5.2 Translog specification

The translog specification is a less restrictive form of the Cobb-Douglas production function that relaxes the assumptions of constant returns to scale and homotheticity. It is conceivable that IT and R&D exhibit increasing (or decreasing) returns to scale, and that the complementarity between IT and R&D varies between low and high levels of spending. We adopt the log-log form of the translog specification, which is defined

$$VA_{it} = IT_{it}^{\alpha_1} RD_{it}^{\alpha_2} e^{S+\beta_1 (lnIT_{it})^2 + \beta_2 (\ln RD_{it})^2 + \gamma lnIT_{it} lnRD_{it} + u_{it}} \qquad (4)$$

where S is the exogenous component of the production function and u_{it} is the error term. Taking the natural log of both sides of the equation[2] gives the regression equation

$$va_{it} = \alpha_1 it_{it} + \alpha_2 rd_{it} + \beta_1 (it_{it})^2 + \ \beta_2 (rd_{it})^2 + \gamma it_{it} rd_{it} + S + u_{it} \tag{5}$$

We denote logged variables in lower case and levels in upper case. In our study context, S takes the form $\boldsymbol{\rho X}$, where **X** is an array of covariates. Taking the second derivative with respect to *IT* and *RD* gives the following (subscripts are omitted for clarity):

$$\frac{\partial^2 va}{\partial IT \partial RD} = \frac{va}{IT \cdot RD}(Z_1 Z_2 + \gamma) \tag{6}$$

where

$$Z_1 = \alpha_1 + 2\beta_1 it + \gamma rd$$

and

$$Z_2 = \alpha_2 + 2\beta_2 it + \gamma rd$$

Since $\frac{va}{IT \cdot RD}$ is always positive in our data, the sign of the cross-partial derivative is given by $(Z_1 Z_2 + \gamma)$. We calculate Equation 6 in two steps: first, we estimate Equation 5 to obtain the coefficients α_1, α_2, β_1, β_2, and γ. Second, we use mean values of va, IT, and R&D, along with the coefficients, to compute Equation 6.

3.5.3 Identification

The level of R&D and IT investments are choices that may be endogenously related to a firm's output. Although in our model we treat these investments as exogenous, in fact the decision to introduce a new product, to meet new demand for example, may necessitate new investments in hardware, software, and research. Another way to look at this is that the decision to conduct research often comes with the expectation that it will be monetized.

The endogeneity of IT depends in part upon the purpose for which it is implemented. If a system is adopted primarily for the purpose of product development, then we face the problem of simultaneity. On the other hand, if an ERP system is adopted for the purpose of tracking inventory and improving the production process, and a new marketable use for it is discovered through experience, then the innovation is likely a by-product of the IT investment and not the other way around. Likewise, if IT is adopted for the purpose of basic research and results in an invention that is brought to market, then we can have a reasonable assumption of exogeneity.

Similarly, the increase in R&D investments in order to develop a product to meet customer demand faces simultaneity issues. Whereas, in the case of a firm that has a productive R&D division due to highly skilled scientists and makes a discovery in the course of general research, and then determines how best to market that product, the causality is primarily in one direction.

There exists the possibility that there are unobserved factors that influence both a firm's decisions to invest in R&D and IT **and** its decision to introduce a new product to market. For example, a firm may have an intrinsic propensity to develop and market new products. Or there may be competitive reasons within an industry to do both. Similarly, consumer demand may influence both our dependent and independent variables. Therefore, omitted variable bias is a concern in our study, and we cannot interpret our results as causal.

We only observe large firms. There may be a difference in effects of our explanatory variables on innovation between firms which were included in both Standard & Poor's and Compustat's databases, and those which were not included in both. If so, this selection issue would present bias in our estimates if we were attempting to explain drivers of output for the population of U.S. firms. Because we are primarily interested in understanding the existence of complementarity between IT and R&D in firms that are major contributors to U.S. economic output, our coefficient estimates may be treated as such and are therefore not likely to be biased.

3.6 RESULTS

3.6.1 Baseline

We establish a baseline panel regression in Table 3.4. Individually, both IT stock and R&D stock are associated with a statistically significant increase in value added. The point estimate in Column 1 indicates that on average in our sample, an increase of 1% in IT stock is associated with a 0.149% increase in value added. At the mean values of IT stock and value added in this sample, that is equivalent to an increase in IT stock of about $1.1 million, and the increase in value added is approximately $4.5 million.

The regression in Column 2 shows an increase in value added of 0.196% is associated with an increase of 1% in R&D. At mean values, that is about $5.6 million increase in R&D and $6.0 million in value added.

Table 3.4: IT and R&D show no evidence of complementarity across sample, controlling for industry and scale of operations

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	IT only	R&D only	Direct effects	With interaction	Baseline	Random effects	Binary IT and R&D
ln(IT stock)	0.149***		0.137***	0.196***	0.096**	0.138***	0.160**
	(0.029)		(0.030)	(0.041)	(0.039)	(0.029)	(0.065)
ln(R&D stock)		0.196***	0.164***	0.294***	0.126***	0.146***	0.078
		(0.033)	(0.032)	(0.040)	(0.044)	(0.025)	(0.082)
ln(IT) * ln(R&D)				-0.026***	-0.017***	-0.016***	-0.163**
				(0.007)	(0.006)	(0.005)	(0.074)
ln(non-IT cap)					0.116**	0.358***	0.107**
					(0.050)	(0.029)	(0.051)
ln(non-IT labor)					0.198***	0.337***	0.229***
					(0.053)	(0.039)	(0.052)
N (n)	3301 (698)	3301 (698)	3301 (698)	3301 (698)	3301 (698)	3301 (698)	3301 (698)
R-sq within	0.017	0.009	0.023	0.028	0.398	0.080	0.396
R-sq between	0.285	0.075	0.188	0.172	0.263	0.766	0.227
R-sq overall	0.186	0.057	0.140	0.130	0.302	0.557	0.276

Notes: Standard errors are in parentheses. Heteroskedasticity-robust standard errors are used in all regressions. Columns 5-7 include controls for 2-digit NAICS industry and year. Statistical significance indicated by * (0.10), ** (0.05), *** (0.01). N indicates total number of observations in the unbalanced panel; n is the number of unique firms.

Including both IT and R&D stock in the specification shown in Column 3, we see that the magnitudes of both decrease, about 0.1 percentage point for IT stock and about 0.3 percentage points for R&D stock.

In Column 4, we introduce the interaction between (the logs of) IT stock and R&D stock. If the two are complements, we should expect to see a positive coefficient on the interaction term. Here we see that complementarity is not supported in our data. In contrast, there is evidence here of substitutability.

In Column 5, we run the full Cobb-Douglas specification, which includes stock of non-IT capital and flow of non-IT labor (flow is used here because it is a relatively smooth input to the production function over time). Thus, we include both in the regression shown in Column 5. In addition, we control for both the industry over time, by interacting 2-digit NAICS industry with year, and the year itself. The coefficients on R&D stock and IT stock remain positive and statistically significant, and the evidence of substitution remains. On average in our data, as the stock of R&D increases, the relationship between IT stock and value added decreases. Similarly, as the stock of IT increases, the relationship between R&D stock and value added weakens.

Random effects are shown in Column 6. If the unobserved firm effects are, in fact, uncorrelated with the right-hand-side variables, then a random effects specification is more efficient and thus desirable. This regression estimates a 0.138% increase in value added with the increase of 1% in IT stock (about $419,934 at the mean value of VA). A 1% increase in R&D stock is associated with a modest 0.146% increase in value added (about $444,278).

To aid interpretation, we also test binary indicators of the level of IT and R&D investment (Column 7). We set the "high IT investment" threshold at the mean of own NAICS 2-digit industry, which in our pooled sample ranges from $28 million to $336 million. Likewise, "high R&D investment" indicates an amount greater than own-industry mean. Since over 60% of our sample has no R&D investment, for some industries "high R&D investment" is any amount greater than zero[6]. The range of mean-by-industry R&D stock in our sample is $0 to $1.45 billion. The coefficient of the interaction is easier to interpret here: in firms that invest heavily in R&D, value added decreases by about 16% when they also invest heavily in IT.

3.6.2 Translog

In our baseline regressions, we use a functional form based on the Cobb-Douglas production function. In Table 3.5, we relax the Cobb-Douglas restrictions using the translog specification of the production function.

Table 3.4, Column 4 is repeated here in Column 1 for ease of comparison. Column 2 shows the analogous translog specification. While the interaction term does not change, the translog specification is used here primarily as a means to calculate the cross-partial derivative, $\frac{\partial^2(VA)}{\partial(IT)\partial(RD)}$. A negative cross-partial derivative, as seen in Column 2, indicates concavity of the production function. That is, as one input increases for a given value of VA, the other decreases, indicating substitutability between the inputs. The translog specification, in both fixed and random effects regressions, shows evidence of substitutability between IT and R&D investments.

[6]Those industries are construction, transportation and warehousing, finance and insurance, real estate and rental and leasing, educational services, health care and social assistance, and other services (except public administration).

Table 3.5: Translog specifications

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	FE, Cobb-Douglas	FE, no controls	FE, Cobb-Douglas	FE, translog	RE, no controls	RE, Cobb-Douglas	RE, translog
ln(IT stock)	0.196***	0.274***	0.096**	0.039	0.058	0.138***	-0.086
	(0.041)	(0.093)	(0.039)	(0.088)	(0.071)	(0.029)	(0.058)
ln(R&D stock)	0.294***	0.186***	0.126***	0.122**	-0.094***	0.146***	-0.054**
	(0.040)	(0.071)	(0.044)	(0.057)	(0.035)	(0.025)	(0.027)
ln(IT) * ln(R&D)	-0.026***	-0.026***	-0.017***	-0.018***	-0.027***	-0.016***	-0.031***
	(0.007)	(0.007)	(0.006)	(0.006)	(0.005)	(0.005)	(0.004)
ln(non-IT capital)			0.116**	0.122**		0.358***	0.315***
			(0.050)	(0.050)		(0.029)	(0.023)
ln(non-IT labor)			0.198***	0.202***		0.337***	0.324***
			(0.053)	(0.054)		(0.039)	(0.037)
N (n)	3301 (698)	3301 (698)	3301 (698)	3301 (698)	3301 (698)	3301 (698)	3301 (698)
R-sq within	0.028	0.029	0.398	0.398	0.021	0.080	0.078
R-sq between	0.172	0.193	0.263	0.284	0.423	0.766	0.714
R-sq overall	0.130	0.146	0.302	0.318	0.290	0.557	0.522
Cross-partial derivative		-0.000396		-0.000261	-0.000529		-0.000727

Notes: Standard errors are in parentheses. Heteroskedasticity-robust standard errors are used in all regressions. Columns 2 and 4 include controls for 2-digit NAICS industry and year. Statistical significance indicated by * (0.10), ** (0.05), *** (0.01). N indicates total number of observations in the unbalanced panel; n is the number of unique firms.

3.6.3 Dynamism

We test whether the substitutability between IT and R&D across our sample is moderated by dynamism within the firm's industry. Results for this series of analyses are shown in Table 3.6. Our baseline specification from Table 3.4, Column 5 is repeated here for ease of comparison.

In Column 2, we see that in the fixed effects regression, adding our measure of dynamism has virtually no effect on any of the other coefficients, nor is it statistically significant. The random effects specification (Column 3, shows similar results).

Table 3.6: Substitution is evident in stable environments.

	(1) Baseline	(2) Fixed effects	(3) Random effects	(4) Split, FE, DYN=0	(5) Split, FE, DYN=1	(6) Split, RE, DYN=0	(7) Split, RE, DYN=1
ln(IT stock)	0.096**	0.097**	0.140***	0.018	0.150***	0.088*	0.157***
	(0.039)	(0.038)	(0.029)	(0.062)	(0.050)	(0.052)	(0.037)
ln(R&D stock)	0.126***	0.130***	0.152***	0.096**	-0.093	0.134***	0.061
	(0.044)	(0.042)	(0.025)	(0.051)	(0.398)	(0.034)	(0.102)
ln(IT)*ln(R&D)	-0.017***	-0.017***	-0.016***	-0.009	-0.021	-0.015**	-0.020
	(0.006)	(0.006)	(0.005)	(0.009)	(0.024)	(0.007)	(0.025)
it*rd*DYN		-0.022	-0.022				
		(0.044)	(0.019)				
ln(non-IT cap)	0.116**	0.116**	0.355***	0.107**	0.113	0.372***	0.330***
	(0.050)	(0.050)	(0.029)	(0.050)	(0.088)	(0.042)	(0.041)
ln(non-IT labor)	0.198***	0.198***	0.340***	0.295***	0.073	0.353***	0.393***
	(0.053)	(0.053)	(0.039)	(0.051)	(0.103)	(0.051)	(0.052)
N (n)	3301	3301	3301	2031	1270	2031	1270
	(698)	(698)	(698)	(426)	(272)	(426)	(272)
R-sq within	0.398	0.398	0.080	0.431	0.388	0.213	0.051
R-sq between	0.263	0.266	0.767	0.512	0.076	0.792	0.737
R-sq overall	0.302	0.305	0.557	0.497	0.167	0.682	0.457

Notes: Standard errors are in parentheses. Heteroskedasticity-robust standard errors are used in all regressions. All regressions include controls for 2-digit NAICS industry and year. Statistical significance indicated by * (0.10), ** (0.05), *** (0.01). N indicates total number of observations in the unbalanced panel; n is the number of unique firms.

However, the split sample yields a bit more insight. We divide the sample into low- and high-dynamism subsamples (DYN = 0 and 1, respectively) in the specifications shown in Columns 4-7. The data reveal that in relatively industries with relatively stable demand, the substitution effect between IT and R&D are consistent with the baseline regression results. However, in dynamic environments, there is evidence for neither complementarity nor substitution. This is true in both the fixed effects specifications (Columns 4 and 5) and random effects specifications (Columns 6 and 7).

To test robustness of these results, we construct an alternative measure of dynamism as described in Li Ye (1999) and Wang et al (2017). The measure exploits the variation in sales over time for each industry, normalized by the industry's mean sales. This "dynamism index" is constructed in several steps. First, for each industry, yearly sales is regressed on year to obtain the standard error of the coefficient on year. Next, the mean sales over our five-year period is calculated. Finally, the standard error of the year coefficient is divided by the mean sales to construct the dynamism index. This method has the advantage of simplicity, but the index itself does not have an intuitive scale and is thus difficult to interpret. To aid in interpretation, we create a binary variable that is equal to 1 if the industry has a higher index than the median; 0 otherwise. A comparison of dynamic industries using the Kurz-Senses and Li-Ye constructs is shown in Table 3.7.

Results of the series of analyses using the dynamism index are shown in Table 3.8. The same set of regressions as in Table 3.6 yield similar results. The main difference is the significance of the interaction coefficient in Column 4. This provides further evidence that the substitution effect IT and R&D investments occurs primarily in relatively stable industries.

Table 3.7: Comparison of dynamism measures

NAICS 2-digit code	Industry name	Dynamic (Kurz-Senses)	Dynamic (Li-Ye)
11	Agriculture, Forestry, Fishing and Hunting		✓
21	Mining, Quarrying, and Oil and Gas Extraction	✓	✓
22	Utilities	✓	
23	Construction	✓	
31	Manufacturing		
32	Manufacturing		
33	Manufacturing		
42	Wholesale Trade		
44	Retail Trade	✓	
45	Retail Trade	✓	✓
48	Transportation and Warehousing	✓	
51	Information		
52	Finance and Insurance	✓	
53	Real Estate and Rental and Leasing	✓	✓
54	Professional, Scientific, and Technical Services		✓
56	Administrative and Support and Waste Management and Remediation Services		✓
61	Educational Services	✓	✓
62	Health Care and Social Assistance		✓
71	Arts, Entertainment, and Recreation		✓
72	Accommodation and Food Services		
81	Other Services (except Public Administration)	✓	✓

Table 3.8:Alternative dynamism construct

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Baseline	Li, FE	Li, RE	Split, FE, DYN=0	Split, FE, DYN=1	Split, RE, DYN=0	Split, RE, DYN=1
ln(IT stock)	0.096**	0.096**	0.126***	0.089**	0.143	0.108***	0.206***
	(0.039)	(0.038)	(0.029)	(0.041)	(0.115)	(0.032)	(0.073)
ln(R&D stock)	0.126***	0.124***	0.155***	0.121***	0.139	0.153***	0.059
	(0.044)	(0.043)	(0.026)	(0.044)	(0.292)	(0.028)	(0.091)
ln(IT) * ln(R&D)	-0.017***	-0.017***	-0.017***	-0.016**	-0.013	-0.016***	-0.008
	(0.006)	(0.006)	(0.005)	(0.007)	(0.020)	(0.005)	(0.017)
it*rd*DYN		0.013	-0.008				
		(0.022)	(0.010)				
ln(non-IT cap)	0.116**	0.116**	0.350***	0.162**	-0.015	0.349***	0.351***
	(0.050)	(0.050)	(0.029)	(0.066)	(0.049)	(0.033)	(0.060)
ln(non-IT labor)	0.198***	0.199***	0.370***	0.211***	0.108	0.381***	0.318***
	(0.053)	(0.053)	(0.029)	(0.057)	(0.158)	(0.042)	(0.093)
N (n)	3301	3301	3301	2797	504	2797	504
	(698)	(698)	(698)	(590)	(108)	(590)	(108)
R-sq within	0.398	0.398	0.058	0.410	0.355	0.063	0.048
R-sq between	0.263	0.259	0.765	0.333	0.108	0.771	0.726
R-sq overall	0.302	0.299	0.550	0.358	0.172	0.567	0.463

Notes: Standard errors are in parentheses. Heteroskedasticity-robust standard errors are used in all regressions. All regressions include controls for 2-digit NAICS industry and year. Statistical significance indicated by * (0.10), ** (0.05), *** (0.01). N indicates total number of observations in the unbalanced panel; n is the number of unique firms.

3.7 DISCUSSION

The body of research on IT-enabled innovation skews toward finding complementarity with R&D, but we find that among Fortune 1000 firms, IT and R&D investments exhibit a substitution effect with value added. Our finding is important for several reasons. First, the marginal dollar of IT investment may not result in an increase in value added (or profit, from a manager's perspective) if R&D plays a central role in the firm's strategy. This may reflect the assertion of Hall and Mairesse (2006) that innovating firms are more concerned with preventing unwanted knowledge leakage and enabling knowledge flows to collaborators. An innovative firm may be more careful in the technologies they employ, choosing to restrict investments to those that are better understood. This reduces the risk to the firm of unintended knowledge flows.

Second, this work highlights the role of innovative strategy in IT investment decisions. The focus in much of the prior literature has been on the productivity benefits of IT, but the tension between desired and undesired knowledge spillovers is less understood. This friction is particularly important for firms that tend to innovate.

Our work presents a new finding within both the IT-enabled strategy literature and the knowledge spillover literature. It also represents a contribution toward bridging the gap between information systems and strategic management. It is becoming increasingly difficult to consider the strategy of a firm without considering the implications of investing in information technology, since the digitization of both products and processes necessitate additional choices for investment for the firm.

We show evidence that IT are R&D are substitutes across our sample, and mixed evidence that they are substitutes in low dynamism environments. In our construction, this low dynamism means that the change in demand from one year to the next is a smooth function, relative to other industries. The decision to invest in IT is often made to combat uncertainty, but our data reveal that such an approach may be naïve in certain circumstances. Future research should focus on the identification of sources of complementarity between IT and R&D, as well as understanding sources of heterogeneity in dynamic environments.

For both the researcher and the manager, our work brings focus to factors that might moderate the benefit of information systems. While IT may be attractive as an aid to the collaborative R&D process, we find evidence that knowledge management should be a priority of firms that innovate to ensure the appropriate availability and confidentiality of digitized knowledge.

3.8 REFERENCES

Abdih, Y. and Joutz, F. (2006). Relating the knowledge production function to total factor productivity: an endogenous growth puzzle. IMF Staff Papers, 53(2): pp. 242-271.

Anand, G. and Ward, P. (2004). Fit, flexibility and performance in manufacturing: coping with dynamic environments. Production and Operations Management, 13(4): pp. 369-384.

Angle, P. and Forman, C. (2019). Does IT level the playing field for small establishments? Evidence from manufacturing. Working paper. Available at https://ssrn.com/abstract=3246142.

Arora, A.; Ceccagnoli, M.; and Cohen, W. (2008). R&D and the patent premium. International Journal of Industrial Organization, 26: pp. 1153-1179.

Arora, A; Cohen, W.; and Walsh, J. (2014). The acquisition and commercialization of invention in American manufacturing: incidence and impact. NBER working paper series. Available at http://www.nber.org/papers/w20264.

Ashurst, C.; Freer, A.; Ekdahl, J.; and Gibbons, C. (2012). Exploring IT-enabled innovation: A new paradigm? *International Journal of Information Management*, 32: pp. 326-336.

Barnett, W. and Pontikes, E. The Red Queen, success bias, and organizational inertia. Management Science, 54(7): pp. 1237-1251.

Barney, J. (1991). Firm resources and sustained competitive advantage. Journal of Management, 17(1): pp. 99-120.

Bloom, N. (2007). Uncertainty and the dynamics of R&D. NBER working paper series. Available at http://www.nber.org/papers/w12841.

Bloom, N.; Schankerman, M.; and van Reenen, J. (2013). Identifying technology spillovers and product market rivalry. Econometrica, 81(4): pp. 1347-1393.

Bresnahan, T; Brynjolfsson, E.; and Hitt, L. (2002). Information technology, workplace organization, and the demand for skilled labor: firm-level evidence. Quarterly Journal of Economics, 117(1): pp. 339-376.

Brynjolfsson, E. and Hitt, L. (1995). Information technology as a factor of production: the role of differences among firms. Economics of Innovation and New Technology, 3: pp. 183-200.

Brynjolfsson, E. and Hitt, L. (2003). Computing productivity: firm-level evidence. Review of Economics and Statistics, 85(4): pp. 793-808.

Ceccagnoli, M.; Higgins, M.; and Palermo, V. (2014). Behind the scenes: sources of complementarity in R&D. Journal of Economics & Management Strategy, 23(1): pp. 125-148.

Ceccagnoli, M.; van Zeebroeck, N.; and Venturini, R. (2018). Cross-functional knowledge integration, patenting and firm's performance. Working paper. Available at https://ssrn.com/abstract=2286693.

Cohen, W. and Levinthal, D. (1990). Absorptive capacity: a new perspective on learning and innovation. Administrative Science Quarterly, 35(1): pp. 128-152.

Dai, R. (2012). International accounting databases on WRDS: comparative analysis. Available at SSRN: https://ssrn.com/abstract=2938675.

Dedrick, J.; Gurbaxani, V.; and Kraemer, K. (2003). Information technology and economic performance: a critical review of the empirical evidence. ACM Computing Surveys (CSUR), 35(1): pp. 1-28.

Dewan, R.; Jing, B.; and Seidmann, A. (2000). Adoption of Internet-based product customization and pricing strategies. Journal of Management Information Systems, 17(2): pp. 9-28.

Dewan, R. and Min, C. (1997). The substitution of information technology for other factors of production: a firm level analysis. Management Science, 43(12): pp. 1660-1675.

Duguet, E. (2006). Innovation height, spillovers and TFP growth at the firm level: evidence from French manufacturing, 15(4-5): pp. 415-442.

Eisenhardt, K. and Bourgeois, L. (1988). Politics of strategic decision making in high-velocity environments: toward a midrange theory. Academy of Management Journal, 31(4) pp. 737-770.

Forman, C.; Goldfarb, A.; and Greenstein, S. (2008). Understanding the inputs into innovation: do cities substitute for internal firm resources? Journal of Economics and Management Strategy, 17(2): pp. 295-316.

Forman, C. and van Zeebroeck, N. (2012). From wires to partners: how the Internet has fostered R&D collaborations within firms. Management Science, 58(8): pp. 1549-1568.

Gao, G. and Hitt, L. (2012). Information technology and trademarks: implications for product variety. Management Science, 58(6): pp. 1211-1226.

Hall, B. and Harhoff, D. (2012). Recent research on the economics of patents. Annual Review of Economics, 4(1): pp. 541-565.

Hall, B.; Jaffe, A.; and Trajtenberg, M. (2005). Market value and patent citations. RAND Journal of Economics, pp. 16-38.

Hall, B.; Lotti, F.; and Mairesse, J. (2013). Evidence on the impact of R&D and ICT investments on innovation and productivity in Italian firms. Economics of Innovation and New Technology, 22(3): pp. 300-328.

Hall, B. and Mairesse, J. (1995). Exploring the relationship between R&D and productivity in French manufacturing firms. Journal of Econometrics, 65: pp. 263-293.

Hall, B. and Mairesse, J. (2006). Empirical studies of innovation in the knowledge-driven economy. Economics of Innovation and New Technology, 15(4-5): pp. 289-299.

Hobday, M. (1998). Product complexity, innovation industrial organization. Research Policy, 26: pp. 689-710.

Huang, P.: Ceccagnoli, M.; Forman, C.; Wu, D. (2019). IT knowledge spillovers, absorptive capacity, and productivity: evidence from enterprise software. Working paper. Available at https://ssrn.com/abstract=2243886.

Jiao, H.; Alon, I.; Koo, C.; and Cui, Y. (2013). When should organizational change be implemented? The moderating effect of environmental dynamism between dynamic capabilities and new venture performance. Journal of Engineering and Technology Management, 30: pp. 188-205.

Joshi, K.; Chi, L.; Datta, A.; and Han, S. (2010). Changing the competitive landscape: continuous innovation through IT-enabled knowledge capabilities. *Information Systems Research*, 21(3): pp. 472-495.

Joshi, K. and Lauer, T. (1998). Impact of information technology on users' work environment: a case of computer aided design (CAD) system implementation. Information & Management, 34(6): pp. 349-360.

Kurz, C. and Senses, M. (2016). Importing, exporting and firm-level employment volatility. Journal of International Economics, 98: pp. 160-175.

Kleis, L., Chwelos, P., Ramirez, R., and Cockburn, Iain. (2012). "Information Technology and Intangible Output: the Impact of IT Investment of Innovation Productivity," *Information Systems Research*, 23(1), pp. 42-59.

Loch, C.; Terwiesch, C.; Thomke, S. Parallel and sequential testing of design alternatives. Management Science, 47(5): pp. 663-678.

Martin, L. and Nguyen-Thi, T. (2015). The relationship between innovation and productivity based on R&D and ICT use: an empirical analysis of firms in Luxembourg. Revue Économique, 66(6): pp. 1105-1130.

Melville, N.; Kraemer, K.; and Gurbaxani, V. (2004). Information technology and organizational performance: an integrative model of IT business value, *MIS Quarterly*, 28(2), pp. 283-322.

Meso, P. and Smith, R. (2000). A resource-based view of organizational knowledge management systems. Journal of Knowledge Management, 4(3): pp. 224-234.

Miller, D. and Friesen, P. (1983). Strategy-making and environment: the third link. Strategic Management Journal, 4(3): pp. 221-235.

Mohnen, P.; Mairesse, J.; and Dagenais, M. (2006). Innovativity: a comparison across seven European countries. Economics of Innovation and New Technology, 15(4-5): pp. 391-413.

Moser, P. (2005). How do patent laws influence innovation? Evidence from nineteenth-century world fairs. American Economic Review, 95: pp. 1214-1236.

Nambisan, S. (2002). Designing virtual customer environments for new product development: toward a theory. Academy of Management Review, 27(3): pp. 392-413.

Neuhausler, P.; Frietsch, R.; Schubert, T.; Blind, K. (2011). Patents and the financial performance of firms – an analysis based on stock market data. Fraunhofer ISI Discussion Papers Innovation Systems and Policy Analysis, 28.

Ouellette, L. (2012). Do patents disclose useful information? Harvard Journal of Law & Technology, 25(2): pp. 545-607.

Overby, E.; Bharadwaj, A.; and Sambamurthy, V. (2006). Enterprise agility and the enabling role of information technology. European Journal of Information Systems, 15: pp. 120-131.

Pavlou, P. and el Sawy, O. (2006). From IT leveraging competence to competitive advantage in turbulent environments: the case of new product development. *Information Systems Research*, 17(3), pp. 198-227.

Prajogo, D. (2015). The strategic fit between innovation strategies and business environment in delivering business performance. International Journal of Production Economics, 171: pp. 241-249.

Price, W. and Bass, L. (1969). Scientific research and the innovative process. Science, 164(3881): pp. 802-806.

Randeree, E. (2006). Knowledge management: securing the future. Journal of Knowledge Management, 10(4): pp. 145-156.

Sambamurthy, V.; Bharadwaj, A.; and Grover, V. (2003). Shaping agility through digital options: reconceptualizing the role of information technology in contemporary firms. MIS Quarterly, 27(2): pp. 237-263.

Stieglitz, N.; Knudsen, T.; and Becker, M. (2016). Adaptation and inertia in dynamic environments. Strategic Management Journal, 37(9): pp. 1854-1864.

Syverson, C. (2011). What determines productivity? Journal of Economic Literature, 49(2): pp. 326-365.

Tambe, P. and Hitt, L. (2013). Job hopping, information technology spillovers, and productivity growth. Management Science, 60(2): pp. 338-355.

Teece, D.; Pisano, G.; and Shuen, A. (1997). Dynamic capabilities and strategic management. Strategic Management Journal, 18(7): pp. 509-533.

Utterback, J. and Abernathy, W. (1975). A dynamic model of process and product innovation. OMEGA, the International Journal of Management Science, 3(6): pp. 639-656.

Venturini, R.; Ceccagnoli, M.; and van Zeebroeck, N. (2019). Knowledge integration in the shadow of tacit spillovers: empirical evidence from US R&D labs. Research Policy, 48(1): pp. 180-205.

Wang, Y.; Wang, T.; and Li, X. (2017). Does R&D create additional business value through IT? Chinese Management Studies, 11(2): pp. 194-208.

Wu, L.; Hitt, L.; and Lou, B. (2017). Data analytics skills, innovation and firm productivity. The Wharton School Research Paper, 86.

Yoo, Y.; Henfridsson, O.; and Lyytinen, K. (2010). The new organizing logic of digital innovation: an agenda for information systems research. Information Systems Research, 21(4): pp. 724-735.

VITA

Patricia Angle was born in Royal Oak, Michigan (née Kelly). She currently resides in Powder Springs, Georgia, with her husband and two sons. This work represents the six years she spent in graduate school studying Information Technology Management in the Scheller College of Business, Georgia Institute of Technology. Prior to Georgia Tech, she earned Bachelor's and Master's degrees in Aerospace Engineering at the University of Michigan, Ann Arbor.

During the course of study at Georgia Tech, Dr. Angle presented at many conferences, including the Workshop on Information Systems and Economics in 2013 (best student paper finalist) and the Conference on Information Systems and Technology in 2017. She has taught courses on network systems and analytics and was nominated for a teaching award in 2018.

Outside her research and teaching interests, Dr. Angle enjoys running for her buddy Annemarie as part of the IRun4 organization, watching her son Gabe play soccer, cuddling with her son Ollie, and doing anything with her husband Matt.