Three Empirical Investigations of Information Technology and Innovation

by

Landon Kleis

B.Com, Queen's University, 1996
M.Sc.B., The University of British Columbia, 2004

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

The Faculty of Graduate Studies

(Business Administration)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

September 2011

Abstract

Among the many ways in which information technology has transformed business, the application of IT to the pursuit of innovation offers perhaps the greatest potential. Improved innovation processes allow firms to reduce production costs and offer new products more rapidly and with less risk, leading to competitive advantage. While researchers have studied the link between innovation and productivity, the role of IT in this context has not been established. We seek to understand this relationship by conducting empirical research in three complementary investigations. First, relative to R&D, how much does IT contribute to the creation of innovative knowledge? Second, does IT-enabled innovation contribute to production? Third, do IT and R&D work together to improve the efficiency of capital and labour?

In the first study, we analyze a panel of large U.S. manufacturing firms and find IT is positively associated with innovation output. The relationship between IT, R&D, and innovation is robust across multiple econometric methodologies and is particularly strong in the mid to late 1990s, a period of rapid technological innovation.

The second study incorporates IT-enabled knowledge creation in a model of overall production to compare the effect of IT in the two contexts. Our findings highlight the indirect contribution of IT, through the innovation process, as being more evident than its direct role in production.

The third study extends these findings with a more nuanced model of IT- and innovation-driven production efficiency. We hypothesize that process-oriented R&D further enhances the ability of IT capital to increase productivity by leveraging traditional forms of capital and labour. We estimate these relationships using two panels of US industries for the periods 1987–1998 and 1999–2005. The results indicate qualified support for a synergistic

effect of R&D and IT investment in both samples.

Taken together, these findings establish the innovation-enabling capabilities of IT—an important aspect of IT's contribution to business value. Our results disentangle the direct and indirect effects of IT in the production and innovation contexts. Managerial implications of this research suggest that general IT investments may be leveraged to assist innovation, and IT performance benchmarks could include innovation outcomes.

Preface

Chapter 2 was coauthored with Paul Chwelos, Ronald V. Ramirez, and Iain Cockburn. Paul contributed the genesis of the research question and an expanded version of the IT investment dataset, based on data obtained by CRITO (the Center for Research on Information Technology and Organizations at University of Californa, Irvine). Ron wrote most of the Theory section and portions of the introduction and discussion, and contributed to the overall readability of the paper. Iain contributed advice on econometric adjustments and the 5-year capped citation data. I was responsible for gathering and preparing the data, designing and executing the analysis strategy, presenting the results, and for writing all other parts of the paper not already noted. Chapter 2 is based on work published as:

> Kleis, Landon, Paul D. Chwelos, Ronald V. Ramirez, Iain M. Cockburn. 2011. Information Technology and Intangible Output: The Impact of IT Investment on Innovation Productivity. *Information Systems Research.* Published in Articles in Advance, April 8, 2011.

Chapter 3 was coauthored with Paul Chwelos and Ronald V. Ramirez. Myself, Paul and Ron were responsible for the identification of the research program. I was responsible for designing the research program, prepared the data (as noted above) and performed all of the statistical analyses and the manuscript.

The MFP data used in Chapter 4 was prepared by Zhuo (June) Cheng based on BLS and BEA data sources.

Table of Contents

List of Tables

List of Figures

Acknowledgements

Many kind souls have offered encouragement, guidance and companionship before and during the preparation of this thesis, without which it could not have been accomplished.

I first wish to express my profound gratitude and admiration for the mentors who have shared their talents and goodwill in my development as a researcher and educator. Paul Chwelos, my original supervisor, whose generous nature, wry humour and practical perspective reassured me that I was on the right path, is greatly missed as a mentor and colleague. I hope to live up to his legacy of delight in both teaching and lifelong learning. Al Dexter, my supervisor, has guided me with a steady and reassuring hand through many challenging episodes. His abundant wisdom, experience and warmth have been invaluable in shepherding this work to completion. Barrie Nault, from whom I have been very fortunate to learn a great deal, has graciously shared his considerable knowledge and insight. He has been instrumental in refining and propelling this research, while keeping the bigger picture in view. Hasan Cavusoglu has always demonstrated an eager willingness to help and enthusiastic sense of discovery, and has given me much encouragement. Ron Cenfetelli has been most helpful with advice and camaraderie. Ron Ramirez's sterling character, expository talents and constructive advice are deeply admired.

I also wish to acknowledge the generous financial support of Social Sciences and Humanities Research Council of Canada (SSHRC) Canada Graduate Scholar doctoral fellowships program, as well as the support of my supervisors and the doctoral program at the Robert H. Lee Graduate School at UBC's Sauder School of Business. I am grateful to INFORMS for granting permission to reproduce the article "Information Technology and Intangible Output: The Impact of IT Investment on Innovation Productivity" as a portion of my doctoral dissertation at the University of British Columbia. The IT investment data for both chapters 2 and 3 were obtained from the Center for Research on Information Technology and Organizations (CRITO), which created the database under a grant from the IBM corporation. Neither IBM nor CRITO are responsible for any errors or omissions in this work. To this extent, this research has been supported by grants from the CISE/IIS/CSS Division of the U.S. National Science Foundation and the NSF Industry/University Cooperative Research Center (IUCRC) to CRITO at the University of California, Irvine, as well as SSHRC and the UBC Hampton Fund. Industry sponsors to CRITO include: ATL Products, Boeing, Canon Information Systems, Conexant Systems, IBM, Microsoft, Nortel Networks, Seagate Technology, Sun Microsystems, and Whirlpool.

I am fortunate to have shared time at UBC with many fine colleagues, especially Camille, Sase, and Paul R., who indulged my humours (good and bad) and helped keep academic life in perspective. I reserve a special acknowledgement for Elaine Cho, our doctoral program assistant, whose caring manner and uncommon dedication to her work with graduate students, from application to convocation, has made this journey far more pleasant.

My life is blessed with many dear friends, young and old, near and far, especially Christine, Kevin, and Neil, who have not let schooling get in the way of enduring friendship. Their support and welcome distraction has meant a great deal to me.

Finally, I am profoundly grateful to my wife Marie and both our families for their steadfast support, advice, and faith in my abilities.

For my parents.

Chapter 1

Introduction

Information technology has evolved rapidly from a tool to assist scientific calculation to a ubiquitous aspect of life in industrialized societies. Both governments and industry have embraced IT as a means to improve quality, speed and timeliness in myriad processes, from the sourcing of raw materials to the delivery of the final, sometimes "virtual," product or service. The rise of IT hardware and software as an increasingly dominant form of capital investment in firms has attracted significant attention from researchers, who have attempted to understand and quantify its contribution to all aspects of economic performance. While early investigations focused on productivity and profitability, IT business value research has more recently examined the mechanisms by which firms harness IT to create value. Among these many mechanisms, the application of IT to the pursuit of innovation offers vast potential. Improved innovation processes allow firms to reduce production costs and offer new products more rapidly and with less risk, leading to competitive advantage. While researchers have studied the link between innovation and productivity, the role of IT in this context has not been established. We examine this important relationship by conducting empirical research in three complementary investigations. In doing so, we aim to reveal and quantify a crucial role for IT investments.

We begin at the firm level, where we focus on innovation as an intermediate production process. We propose a IT-augmented knowledge production function, wherein R&D and IT are inputs to knowledge production, and citation-weighted patents are a measure of its output. We analyze a panel of large U.S. manufacturing firms from the period 1987-1997 and find, despite the dominant role of R&D, IT is positively associated with innovation output, with an elasticity of 0.166. The relationship between IT, R&D,

and innovation is robust across multiple econometric methodologies and is particularly strong in the mid to late 1990s, a period of rapid technological innovation. One interesting finding is that IT elasticities are greater for incremental innovations, while highly-cited breakthrough innovations are less responsive to IT investment. Overall, however, our results underscore the role of firm-wide IT investment as an input to the innovation process.

Next, we consider the innovation process in the broader context of the firm-level production function. By incorporating IT-enabled knowledge creation in a model of overall production, we estimate the effect of IT in both its direct and indirect roles. The combined innovation and production model is estimated using structural equation modeling. Our findings show an indirect contribution of IT, through the innovation process, that has a positive impact on production. However, limitations in both the data available and the estimation approach do not allow us to conclude that IT has a direct effect relative to the other inputs.

Finally, we introduce a more nuanced model of IT- and innovation-driven production efficiency at the industry level. We hypothesize that process-oriented R&D further enhances the ability of IT capital to increase productivity by leveraging traditional forms of capital and labour. We use a Cobb-Douglas production function approach with indirect effects to capture the direct and indirect effects of IT, along with R&D, on the other factor inputs, enabling more sophisticated controls for panel error structures. We estimate these relationships using two panels of US industries for the periods 1987–1998 and 1999–2005. The results offer evidence of a synergistic effect of R&D and IT investment. In the first sample, we find that, relative to the traditional Cobb-Douglas model, the output elasticity of IT is lower once the indirect effects of R&D and IT are included. However, the indirect effects of R&D and the combination of R&D and IT are positive. In the second sample, we find an increased estimate of IT elasticity (relative to the Cobb-Douglas model), suggesting that the net effect of IT (augmented by R&D) does improve the efficiency of labour and capital.

Taken together, these findings establish the innovation-enabling capabilities of IT—an important aspect of IT's contribution to business value. Our results disentangle the direct and indirect effects of IT in the production and innovation contexts. While IT continues to demonstrate its role as a key factor of production, our analyses show that IT assets are linked to improved innovation productivity and production efficiency. Having identified this important new role for IT as enabler of innovation, we suggest the managerial implications of this research include the general IT investments may be leveraged to assist innovation, and IT performance benchmarks could be expanded to include innovation outcomes.

Chapter 2

Information Technology and Intangible Output: The Impact of IT Investment on Innovation Productivity

2.1 Introduction

Innovation is a key contributor to a firm's competitive success. Product innovations can enable a company to earn abnormal profits as well as provide an avenue for expansion into new markets and industries (Roberts, 1999; Agarwal and Bayus, 2002). Process innovations create new methods of performing firm activities that can reduce costs or generate new lines of revenue growth (Baily and Chakrabarti, 1988; Dougherty and Hardy, 1996). Together, these benefits motivate firms to invest in the innovation process.

In recent decades, new information technologies and their widespread application have led to three evolutionary changes in the innovation process (Quinn, Baruch, and Zien, 1997). Information technologies such as communication and database applications have helped improve the *management of innovation knowledge*. Researchers, for example, distributed across company research centers can now share knowledge assets across remote geographies and time (Thomke, 2006). *Innovation production* has been improved through IT-based digital methods of design, prototype, and test (Sudarsan, Fenves, and Sriram, 2005; Thomke, 2006). IT-based networks and real-time data flows enable *external innovation collaboration* (Thomke, 2006). Through the outsourcing of innovation production elements (design, prototype development, test, etc.), firms gain access to specialized knowledge and other innovation components that can be incorporated into new products, services, and processes (Chan, Nickerson, and Owan, 2007). The application of information technologies provides the links necessary for effective information sharing and partner monitoring and reduces transaction costs that arise when working with multiple innovation partners in open environments (Dodgson, Gann, and Salter, 2006; Thomke, 2006; Brockhoff, 1992).

In short, through the management of knowledge assets, production support, and interorganizational coordination, information technologies have improved the speed and efficiency of firm innovation. As a result, IT has become essential to product development in firms, especially those in the automobile, consumer products, apparel, and textile industries (Sangiovanni-Vincentelli, 2003; Teresko, 2004; Istook, 2000). Yet, despite the ability of IT to improve the innovation process, innovation remains a costly and risky endeavor. One estimate puts the failure rate of new products at over 90% (Brown and Eisenhardt, 1995). Does this imply that, collectively, information technology has little impact on the firm innovation process? If the innovation creation process remains uncertain and firms run the risk of receiving little or no benefit from these activities, should firms continue to invest in innovation-related IT?

We investigate these research questions by building upon existing preliminary work that considers the relationship between information technology and innovation output (Kleis, Chwelos, Ramirez, and Kraemer, 2003; Han and Ravichandran, 2006). Specifically, we contribute to the literature by applying the knowledge production function (KPF) to (1) evaluate the role of IT in innovation creation across a more extensive time frame; (2) introduce a quality-adjusted measure of innovation output to evaluate the contribution of IT to quality innovation; and (3) conduct an in-depth analysis that provides a richer understanding of the IT innovation relationship. This includes the examination of how the role of IT evolves over our sample period, how the contributions of IT vary across industries that produce versus use IT and, finally, whether IT plays a role in the development of both incremental and breakthrough innovations.

In this study, we utilize a unique data set comprised of annual information on innovation spending research and development (R&D), IT spending, and citation-weighted patents (a quality-adjusted measure of innovation output) for large U.S. firms between 1987 and 1997. We analyze more than 1800 observations using a robust set of econometric methods to test the relationships between IT, R&D, and innovation performance. Our econometric estimates indicate that IT capital has a positive and significant effect on knowledge output. We find a marked increase in the contribution of IT to innovation in the early to late 1990s. However, the evidence does not suggest a significant role for IT in the creation of breakthrough innovations. Rather, as highlighted in existing research, non-IT factors (strategic orientation, organization practices, R&D management, etc.) may hold the key to innovation leadership (Enkel, Gassmann, and Chesbrough, 2009; Majchrzak, Cooper, and Neece, 2004; Malhotra, Majchrzak, Carman, and Lott, 2001; Balachandra and Friar, 1997). Nonetheless, the core result remains: information technology makes a consistent, positive contribution to the innovation creation process.

The remainder of the paper is organized as follows. In the next section, we review theory related to information technology and innovation creation. In § 2.3, we discuss our research design, methodology, and data. We present our empirical analysis and results in § 2.4 and provide a discussion of our findings and a conclusion in §§ 2.5 and 2.6.

2.2 Theory

Firms pursue innovation to build or maintain competitiveness. This is accomplished through the creation of productivity-improving value chain activities (process innovations) or through the extraction of rents generated by the sale of new products or services (product innovations). Preliminary research has identified information technology as a potential contributor to firm innovation efforts. In particular, Kleis et al. (2003) proposed a KPF framework to perform a preliminary test of the IT-innovation relationship. Their analysis found no conclusive results. Han and Ravichandran (2006) test a similar model using panel estimators and find evidence of an indirect IT-innovation relationship (interaction of R&D and IT). Our study builds upon this nascent work and provides a value-added and thorough examination of the role of IT in innovation by using a more extensive time frame, a more complete measure of innovation output, and a more comprehensive evaluation of the direct IT-innovation relationship across multiple contexts. We now explicate the theoretical underpinnings of the role of IT in knowledge production.

2.2.1 Innovation

Technological innovations are created when a new idea or concept is transformed into a product or process for internal or commercial use (Baily and Chakrabarti, 1988). Process innovations are changes to existing processes or the creation of new processes used by an organization to deliver products or services. Product and service innovations are new products or services introduced into the marketplace (Dibrell, Davis, and Craig, 2008). Innovations can arise from raw ideas borne within the firm or result from the adaptation of new knowledge found outside of the firm. This includes basic scientific knowledge generated by corporate and university laboratories as well as inventions spawned by other firms.

Firm innovation proceeds in two stages. First, *research* is conducted to create or determine the efficacy of an invention in addressing some identified problem. Next, the firm undertakes *development* activities related to readying the proposed product or process for its production or application, including design and testing. Because the Federal Accounting Standards Board (FASB) definition of R&D activities mirrors this two-stage process, R&D expense may be considered a good measure of invention and innovation activity in U.S. firms.[1]

[1] FASB Statement of Financial Accounting Standards nos. 2 and 86 provide a definition of research and development. Research is defined as a planned search or critical investigation aimed at the discovery of new knowledge with the hope that such knowledge

To model innovation, we adopt a theoretical model from earlier research in economics and R&D literature (Pakes and Griliches, 1984). The KPF represents the knowledge output generated by a firm as a function of inputs used in its innovation process. We augment the traditional KPF with an information technology input to reflect our notion that information technology contributes to the production of knowledge in a firm. Similar to the "black box" in production theory, the innovation creation process is inherently unobservable.

2.2.2 IT and the Innovation Process

The application of IT contributes to the innovation process through three primary mechanisms. First, information technology contributes to the management of knowledge used in innovation production. Second, information technology enables critical elements of the innovation production process, including opportunity identification, concept development, and innovation design. Third, information technology enables the interorganizational coordination between the focal firm and its external innovation partners.

IT and Innovation Knowledge Management

The management of knowledge is an activity critical to the creation of new innovations. Research and development knowledge, combined with operations knowledge, is used by a firm to develop and produce new products and services (Tanriverdi, 2005). Information technology helps to create an infrastructure for capturing and sharing knowledge across the enterprise on a scale previously unattainable (Tanriverdi, 2005; Majchrzak et al., 2004; Lee and Choi, 2003). The participants involved in the innovation process are interconnected by a knowledge network, sharing, combining, and reusing

will be useful in developing a new product or service, or a new process or technique, or in bringing about a significant improvement to an existing product or process. Development is defined as the translation of research findings or other knowledge into a plan or design for a new product or process or for a significant improvement to an existing product or process whether intended for sale or use. See FASB statements nos. 2 or 86 for full definitions at http://www.fasb.org/st/.

knowledge in the creation of new goods and services (Nerkar and Paruchuri, 2005). Each participant has a specific skill or knowledge that he or she contributes to the innovation process, either directly to the item being created or indirectly through the transfer of knowledge to other actors who then apply it directly in the innovation process. At times, gaps can arise in the knowledge network from links that are missing between participants or from a knowledge element that is missing but is necessary for the innovation process to generate value-added output (Nerkar and Paruchuri, 2005). Information technology helps close these knowledge gaps by enabling the collection of new knowledge assets through improved search capabilities and data mining techniques. Information technology also enables the interconnection of participants who may not be directly involved with the network but who can be temporarily connected to provide a valued knowledge asset that will contribute to the innovation effort. New technologies help to capture internally generated knowledge that can be used throughout the innovation process. The electronic laboratory notebook (ELN), for example, enables scientists to capture and record experiment data electronically versus traditional paper-based lab books, which improves efficiency and eliminates transcription errors (Elliott, 2006). ELNs help to create a central repository of data accessible by other scientists and other drug development systems.

Front-end technologies can also be used to capture innovation-related knowledge. Customer data used in new product development, for example, is captured through company retail websites and through other communication technologies (e.g., e-mail, telecommunication systems) utilized in customer contact interactions (Zahay, Griffin, and Fredericks, 2004). These data enable a firm to develop in-depth knowledge of its customers and create new innovations that better fit customer needs. In the end, this increases the value of a firm's products to its customers and, ultimately, customer satisfaction (Mithas, Krishnan, and Fornell, 2005b,a).

After new knowledge is collected, information technology is critical for the sharing and reuse of knowledge throughout an enterprise (Lee and Choi, 2003). Corporate IP-based networks, telecommunications, and e-mail systems all facilitate the transfer of knowledge between innovation participants.

E-mail, for example, has been shown to be an effective communications and information sharing tool between participants in an R&D network (Rice, 1994). These technologies also provide the infrastructure to facilitate the virtual, interpersonal interactions among R&D teams that ensure the transfer of both explicit and tacit knowledge (Lee and Choi, 2003; Rice, 1994).

External knowledge is also critical for successful innovation. The transfer of external knowledge used in the innovation process can take place through the acquisition of new knowledge, licensing of external innovations, acquisition of firms with unique knowledge, and the hiring of experts with relevant knowledge (Cassiman and Veugelers, 2006). General infrastructure IT (e.g., personal computers (PCs), e-mail, etc.) can assist in this type of knowledge acquisition (Cassiman and Veugelers, 2006). Network and Internet-related communications technologies, for example, have increased the flow of and access to scientific information contained in electronic versions of scholarly journals and public research databases, both of which are important to the R&D process (Kremp and Mairesse, 2004). Computer networks and online access are also critical for the discovery and sharing of competitive and regulatory factors that must be incorporated into new product development innovations (Zahay et al., 2004). IT-based knowledge sharing was the case at Aventis, a major pharmaceutical manufacturer, where the implementation of a chemical biology platform enabled the sharing of knowledge between the virtual community of drug discovery project teams and resulted in a more productive innovation process (Narayanan, Douglas, Schirlin, Wess, and Geising, 2004).

The benefits of IT-enabled innovation have been demonstrated in an academic setting (Hamermesh and Oster, 2002; Agrawal and Goldfarb, 2008). Indeed, the decreasing costs of IT and, subsequently, the costs of communicating and sharing innovation-related knowledge information is motivating an increased use of IT in academic innovation networks (Agrawal and Goldfarb, 2008). E-mail, fax, and telecommunication technologies, for example, have enabled distributed innovation teams to complete new academic research valued by the economics research community (Hamermesh and Oster, 2002). The use of Bitnet, a cooperative U.S. university network predating the Internet, increased the use of multi-institutional partnerships and participation of new actors in university engineering research (Agrawal and Goldfarb, 2008). In summary, whether in a business or academic setting, information technology supports the interactive flow of knowledge between networked participants involved in innovative activity (Swink, 2006; Horn, 2005; Brennan and Dooley, 2005).

IT and Innovation Production

Firms apply and utilize knowledge within an innovation process to produce new products and services. Innovation production has evolved over time from a linear, push-oriented process to a parallel design that incorporates customer and supplier input (Rothwell, 1994). Schilling and Hill (1998) model innovation production as a series of five activities that can occur in parallel and include opportunity identification, concept development, product design, process design, and commercial production. Similarly, Rothwell (1994) models innovation production as six parallel and integrated activities that include marketing, R&D, product development, production engineering, parts manufacturing, and product manufacturing. Although there are slight differences in these models, this research indicates that innovation production involves idea and concept development (new technologies, needs analysis), product development (design), engineering (prototyping), and manufacturing.[2]

Information technology contributes to innovation production in multiple stages. In the idea stage, customer relationship management (CRM) systems act as an information input to innovation production. The information flowing from CRM systems enables a firm to analyze its customers and identify needs that are not being met by current products and services (Nambisan, 2003). This helps the firm generate new product ideas that account for unmet or evolving demand-side factors and contributes to new product success (Rothwell, 1994; Narver, Slater, and MacLachlan, 2004; Mithas et al.,

[2]Consistent with research (Rothwell, 1994; Schilling and Hill, 1998; Tatikonda and Rosenthal, 2000) and the majority of awarded, patented innovations, we focus our process discussion on product innovation.

2005b,a).

Information technology also enables efficient design capabilities. Technology such as computerbased design applications (e.g., CAD/CAM systems) help to digitize a new product's design and make it available throughout the innovation production process. This allows team members to integrate their design efforts, whether located in local or distant R&D centers, from product conception through final assembly (Sudarsan et al., 2005; Gordon, Tarafdar, Cook, Maksimoski, and Rogowitz, 2008; Bartholomew, 2005). Computer-based design software also allows team members to develop virtual prototypes. Engineers can use digital prototypes to run computer simulations to test component compatibility, overall workability, and failure analysis. Digital-based prototypes and simulation, such as those used in automobile, computer, and pharmaceutical manufacturing, not only reduce the cost of traditional wood and clay methods but also allow prototypes to be developed and used much earlier in the innovation production process (Rothwell, 1994). This allows poor designs to be filtered out much earlier in the process and improves overall innovation process efficiency (Thomke, 2006; Rothwell, 1994).

Finally, IT is being used to integrate design and production systems, which enables greater precision and overall product introduction efficiency (Hatch and Mowery, 1998). The use of software-based manufacturability testing, for example, can help designers identify the most efficient ordering of parts that should be used during the manufacturing of final products (Konicki, 2002). CAD systems improve the linkages between design and manufacturing, which helps to integrate the two departments and reduce errors of information transfer and translation (Rothwell, 1994). This serves to minimize manufacturing costs as well as improve the efficiency of production throughput for the new innovation.

IT and External Innovation Collaboration

The production of new innovations involves collaboration between team members working together to create new products, services, or processes.

The distributed team shares a common goal of innovation creation, with each member adding value to the innovation under production (Majchrzak et al., 2004). Traditionally, this process occurred within the firm's boundary and thus required the acquisition and development of innovation-related inputs. However, a shift in innovation practices occurred in the 1980s when firms began to source some of these inputs externally (Rothwell, 1994). As a result, team membership expanded to encompass internal and external participants in both local and geographically distant sites.

Many factors have motivated the inclusion of external partners in innovation. These include the complexity and pace of industrial and technological change, global competition, greater range of available innovation partners, and the mobility and global availability of knowledge workers (Dodgson et al., 2006; Rothwell, 1994; Enkel et al., 2009). The result has been an opening up of the firm's boundary and the creation of an integrated innovation network involving the focal firm, customers, suppliers, and other sources of new knowledge (Dodgson et al., 2006; Rothwell, 1994; Enkel et al., 2009; Christensen, Olesen, and Kjaer, 2005). The movement toward a more open innovation process has had an effect on the propensity of a firm to innovate and has helped improve the fit of new products and the overall efficiency of innovation (Dodgson et al., 2006; Enkel et al., 2009; Schilling and Hill, 1998).

Information technology is a critical enabler of collaborative innovation by providing the necessary linkages for information exchange with external partners. Infrastructure technologies such as PCs, laptops, data and voice networks, and communications applications (e.g., e-mail) are instrumental to these collaborative efforts (Majchrzak et al., 2004; Enkel et al., 2009). These technologies facilitate the exchange of information to and from external participants in the innovation partnership. An aerospace manufacturer, for example, created specialized Web interfaces to enable company team members to exchange information with external team participants (Malhotra et al., 2001). Within 54 innovation teams across 15 industries including manufacturing, basic infrastructure technologies (e.g., networks, e-mail, NetMeeting, etc.) were utilized as communication infrastructures for the exchange of information across the teams including members external to the firm (Majchrzak et al., 2004). Indeed, from the focal firm perspective, information and communications technologies play a critical connectivity role in the open innovation era (Dodgson et al., 2006; Enkel et al., 2009).

Beyond connectivity, information technology plays an important role in creating an effective partnership between a firm and an external service provider. In today's open innovation process, a focal firm can choose to hire and work with an external organization to outsource innovation activities such as design, component development, and test (Dodgson et al., 2006; Nellore and Balachandra, 2001). In such situations, the use of information technology enhances the relationship between the firm and service provider and contributes to partnership success. Indeed, in IT outsourcing and other collaborative relationships, the deployment of information technology has been shown to have a positive effect on communication, trust, and shared understanding, all of which contributes to the success of the business relationships (Ryssel, Ritter, and Gemunden, 2004; Paulraj, Lado, and Chen, 2008; Malhotra et al., 2001; Ba and Pavlou, 2002). Research in an innovation setting found that interorganizational technologies have a positive effect on external new product development relationships in disk drive manufacturing (Scott, 2000). In the long run, such effective innovation relationships produce value-added inputs to a focal firm's innovation process. Ultimately, these value-added inputs contribute to successful innovation production (Dodgson et al., 2006; Nellore and Balachandra, 2001).

In summary, information technology contributes to the firm innovation process by enabling innovation knowledge management, innovation production, and external innovation collaboration. The end result is a collaborative innovation process, enabled through IT, that creates new value-added innovations in a productive manner. Indeed, Kortum and Lerner (1998) conclude after examining innovation production from the 1990s onward that the rise in innovation output (measured in patenting activity) is mainly because of the application of new information technologies developed during the same time period. Thus, it is reasonable to expect a relationship between information technology and the innovation production process (Lee and Choi, 2003).

2.3 Research Design and Methodology

Our conceptual model of knowledge production, based on the KPF (Pakes and Griliches, 1984), is presented in Figure 2.1. Consistent with existing literature, research and development activity is a primary input to knowledge production in a firm (Hall, Griliches, and Hausman, 1986; Henderson and Cockburn, 1994; Hall and Mairesse, 1995; Madsen, 2008). To incorporate the proposed role of information technology in innovation, we introduce information technology as an additional input to knowledge production.

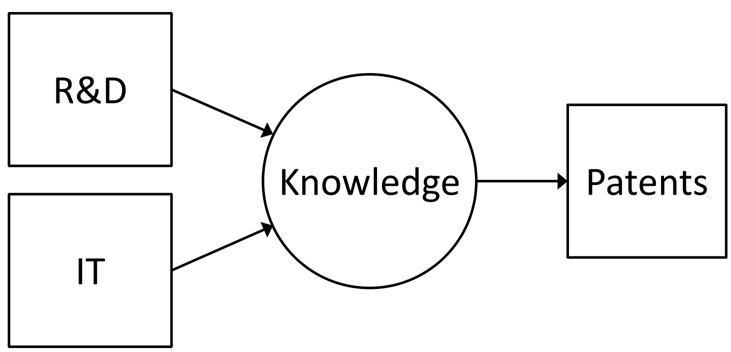


Figure 2.1: Augmented Knowledge Production Function

The output of a firm's knowledge production is new innovation (e.g., new products, services, processes, etc.), with patents being used as the primary, observable output indicator (Griliches, 1990; Hall, Jaffe, and Trajtenberg, 2005). Patents offer many attractive characteristics in this regard. First and foremost, the patent application examination process in the United States imparts a third-party, objective evaluation of an innovation, where the innovation must be novel, nonobvious, and useful, i.e., having a useful purpose and operativeness (USPTO, 2006). In addition, the USPTO standard has been relatively stable over time, making comparisons across years and industries more robust. Second, a patent provides a detailed record of the inventor, the industrial field into which the invention is classified, and citations of prior patents upon which it builds. The enforcement of the last by patent examiners is especially important because it documents a recognition of "prior art" from which the researcher can draw inferences based on the nature and quantity of cited patents.

Our research model can be expressed in Cobb-Douglas form:

$$P = \alpha RD^{\beta_1} IT^{\beta_2}, \tag{2.1}$$

where P is citation-weighted patents awarded, RD is R&D expense in 1993 dollars, and IT is IT capital services in 1993 dollars. The statistical model that we estimate for firm i at time t can be expressed as

$$P_{it} = \alpha RD_{it}^{\beta_1} IT_{it}^{\beta_2} \epsilon_{it}, \tag{2.2}$$

where ϵ_{it} is a multiplicative stochastic error term. For estimation, we transform the model into log-linear form and include controls Z_{it} :

$$\ln P_{it} = \alpha + \beta_1 \ln RD_{it} + \beta_2 \ln IT_{it} + \gamma Z_{it} + \epsilon_{it}. \tag{2.3}$$

We control for firm size, time, and industry. All three have been identified as necessary controls in previous IT and R&D value research (Brynjolfsson and Hitt, 2003; Bresnahan, Brynjolfsson, and Hitt, 2002; Griliches, 1990). Firm size is represented by the log of firm sales in 1993 dollars. Time is represented through the use of year indicator variables. Industry controls are created through the use of Standard Industrial Classification (SIC) codes. Based on earlier research, we utilize a 1.5-digit SIC scheme that aggregates a number of 2-digit SIC industries into a more cohesive and stable 10-industry structure (Bresnahan et al., 2002; Brynjolfsson and Hitt, 2003). Because of the limitations of our sample, only four of these industry groups remain in our analysis: durable, nondurable, process, and high-technology manufacturing.

2.3.1 Data

The estimation of our model utilizes data from three principal sources. First, we use patents as the observable measure of knowledge production output. When considering such a measure, we could utilize raw patent counts as in several existing studies (Pakes and Griliches, 1984; Hall et al., 1986; Crepon, Duguet, and Mairesse, 1998; Han and Ravichandran, 2006). However, the benefits of this type of measure are mitigated by their inability to distinguish the variability in patent quality (Hall, Jaffe, and Trajtenberg, 2002). IT-related research supports the use of quality-based innovation output measures. In particular, Gao and Hitt (2011) consider trademark counts "more appropriate" than raw patent counts because the latter do not incorporate the notion of product variety and differentiation.

Recognizing the limitation of raw patent counts, Hall et al. (2002) developed a patent measure that incorporates a citation-weighting methodology to account for variations in the quality or value of firm innovation. This measure weights each patent by the number of citations received by subsequent patents and then normalizes the citation frequency across certain dimensions. Citation weighting has been used in research concerning the economic value of a firm's innovations (Trajtenberg, 1990; Hall et al., 2005) and innovation productivity and its impact (e.g., Hall and Ziedonis (2001); Thompson and Fox-Kean (2005)) We employ this method to create our dependent variable P.

The raw patent and citation elements of our dependent variable are obtained from the National Bureau of Economic Research (NBER) Patent Citation Database (Hall et al., 2002), which contains nearly three million successful U.S. patent applications made between 1975 and 1999 as well as citations of them by subsequent patents. Although this data set contains many details about each patent, it unfortunately does not distinguish between product and process innovations. The NBER data set contains one row for each patent received by each firm. A separate table contains one row for each citation a patent has received up to the end of 1999. We join these tables to obtain a count of citations for each patent and update the citation count information to 2004 using USPTO data. These patent data are then linked to the Compustat database of publicly traded U.S. firms.[3]

Following the Hall et al. (2002) methodology, we then correct for two underlying factors in our output measure: trends (i.e., inflation) in citation accretion across both time and field[4] and the truncation of accumulated citations for a given patent vintage. Inflation may be addressed by using a fixed-effects benchmarking method to normalize the citations, purging them of year, field, and year-field effects. However, because it is desirable to allow variation among fields, the result is adjusted by dividing the number of citations received by a given patent by the corresponding year-field mean (Hall et al., 2002, p.30). This converts the output data from a count to a continuous-based measure. The truncation effect is addressed by augmenting the NBER data with citation count data from the USPTO to 2004. Hall et al. (2002) suggest that a three-year safety lag is prudent to allow for citation accretion and that a two-year lag for patent application processing should be observed. Taking this into account, the more recent USPTO citation history allows us to use the patent data from 1997 but not from 1998 or 1999 because these years may not contain all of the patent applications that were eventually granted. Finally, we generate a new, alternate citation count based on the citations received within a fixed window of five years from the application date of each patent, and we rescale these citations as described earlier. This removes the vintage effect of older patents for which time has allowed extra citations to accrue. Taken together, these three measures provide a robust assessment of each patent's quality.

When using patents as the basis for an innovation output measure, one must consider whether patents reflect the final innovation output of a firm. The granting of a patent does not guarantee a commercially successful innovation, given the extraordinarily complex task of integrating inventions and existing components into products valued by the marketplace (Fleming and

[3]The patents are matched to the Compustat universe of firms as of 1989. This has the effect of limiting our sample to those firms that belonged to the Fortune 1000 as of 1989.

[4]Hall et al. (2002) define field as the USPTO classification of technological categories for which they provide a simplified taxonomy of 6 fields from more than 10,000 fields in the USPTO classification scheme.

Sorenson, 2003). In addition, some inventions may not meet patentability requirements while, in other cases, firms may choose alternate methods of collecting economic rents from the inventions.[5] This may be reflected in a low patents-to-R&D ratio in some firms. Despite these possibilities, recent evidence from innovation survey data shows broad agreement between the propensity of firms to apply for patents in relation to their innovation activities (Mairesse and Mohnen, 2005; Ramirez and Kleis, 2010). In addition, research indicates that such phenomena, when they do occur, tend to be industry specific (Hall et al., 2002; Baily and Chakrabarti, 1988). To control for some of this environmental patent variation, we include industry and year fixed effects in our econometric analysis.

We obtain data on information technology investment between 1987 and 1998 from the Computer Intelligence Infocorp (CI) database, which details the value of installed IT capital stock at approximately 800 of the Fortune 1000 firms annually. To create the database, CI collects data annually on the quantity of IT hardware in firms (e.g., mainframe, peripherals, minicomputers, personal computer (PC) systems, etc.) using surveys, site visits, physical audits, and telephone interviews. The hardware counts, collected at the site and establishment level, were aggregated to the firm level by CI and a value of a firm's total IT stock was calculated based upon CI's estimate of hardware asset market value.

Because of its detail, the CI data have been used widely in IT business value research (Brynjolfsson, Hitt, and Yang, 2002; Melville, Gurbaxani, and Kraemer, 2007). However, because of a definition change in 1995, we follow the methodology of Chwelos, Ramirez, Kraemer, and Melville (2010), wherein hedonic methods are used to create an IT stock measure that is consistent across the entire 1987–1998 time frame.[6] We then create a measure

[5] As Scotchmer (2004) points out, patents are but one way to protect intellectual property, and they come with the significant trade-off of full disclosure. A firm may decide to use other methods such as "speed to market" or trade secrets depending upon the patenting environment, industry competitiveness, and firm strategy.

[6] Hedonic methods are used to determine the prices of IT components during the 1987–1994 time frame (prior to the definition change in 1995). These prices are applied across the entire 1987-1998 time frame and the resulting technology values are grouped into decentralized and centralized computing categories. These values are then adjusted through

of annual capital services flowing from the IT stock in constant quality-adjusted dollars. These IT measures represent the flow of services from (or, equivalently, payments to) IT assets that are used but not consumed in the production process (Jorgenson and Griliches, 1967). Payments to employees are traditionally measured in a manner similar to R&D expenses. The use of IT capital service variables allows for analysis using both R&D and IT flow variables. Finally, the use of IT capital service measures also accounts for measurement errors that are introduced by stock-based measures (Chwelos et al., 2010).

The yearly value of IT capital services is calculated using the rental price approach used by the Bureau of Labor Statistics (BLS). Rental prices reflect the user cost of capital and are defined as the sum of the rate of return, the rate of depreciation, and the expected rate of asset price change, net of income, and property taxes. Rental prices were created for 12 SIC industries for 2 classes of IT assets, reflecting both decentralized computing (PCs) and centralized computing (central processing equipment) (CPE). Capital services estimates for a firm are then calculated by multiplying the IT capital stock of a firm by the appropriate rental price, given the type of IT capital (PC or CPE) and industry of the firm.

Finally, from the Compustat database, we add a measure of R&D to our data set. R&D activities are recorded annually as R&D expense and are reported in company financial statements. Accounting rules allow firms to expense research, defined as the "planned search or critical investigation aimed at discovery of new knowledge" that is specifically directed at a new or improved output (Oliver, 2003, p.46). Development is defined as transforming "research findings or other knowledge into a plan or design," which can include prototyping and building and operating pilot plants. Declared R&D expense also accounts for in-process research assets and intangibles purchased from other companies. Thus, the R&D expense measures in our

---

the application of hardware-specific price indexes. The adjustment takes into account price changes in similar technology over time are also controls for quality changes in these technologies. Analysis indicates that in both the pre- and post-1994 periods, these adjusted values are highly correlated with the original.

data set include both capital and labor spending put into use by firms to create new product, service, or process innovations. These measures have been widely used as an operationalized representation of the economic inputs to innovation (Pakes and Griliches, 1984; Griliches, 1981; Hall et al., 1986; Hall and Mairesse, 1995; Han and Ravichandran, 2006).

Because a firm's knowledge accumulates (and depreciates) over time, Pakes and Griliches (1984) include lagged R&D spending as part of their model's input. However, subsequent research on the lag structure shows a lack of influence of past R&D spending on future patent output (Hall et al., 1986). Because the impact of R&D on increments to the firm's knowledge stock is almost entirely contemporaneous, current R&D expenditure sufficiently captures the input magnitude.

The merged data set generated from matching the patent, IT, and R&D data sources consists of 260 Fortune 1000 manufacturing firms with an initial total of 1,937 observations between 1987 and 1997. As firms choose to apply for patents relatively early in the R&D cycle (Griliches, 1990), patent application year is matched to the year in which the R&D expense is recorded. To exclude firms with anomalous patent production (perhaps through acquisition of firms with in-process R&D), we further restrict our sample to firms that registered patents in at least 4 years between 1987 and 1997, leaving us with a final count of 201 firms for a total of 1,829 observations. Over the 11-year period, the median firm in the sample has \$2.6 billion in annual sales, \$62.7 million in annual R&D expense, and \$9.6 million in annual IT capital services (Table 2.1).[7] The sample includes 162,381 patents, with the median firm obtaining 24 successful patent applications per year. A total of 1,777,276 citations were recorded through the end of 2004, with the median firm receiving 208 citations on its patents in a given year.[8]

[7] All yearly financial data are indexed to 1993. Sales and R&D data are deflated using industry-specific gross domestic product price indexes from the Bureau of Economic Analysis (BEA). IT data are deflated using the PC price index from Berndt and Rappaport (2001) and the BEA price index for computers and peripheral equipment for all other classes of IT (i.e., mainframes, minicomputers, networking equipment, and computer peripherals), following Chwelos et al. (2010).

[8] Our summary statistics use medians rather than averages because the patent and citation data is heavily skewed, as we discuss in § 2.4.2.

The panel is unbalanced, although over half the firms report patents in all 11 years and almost two-thirds of the firms have at least 10 observations. We restrict our sample to the Fortune 1000 manufacturing sector because the number of firms in the service sector that engage in significant R&D and patenting is small. Over the entire sample, each of the major service sectors (e.g., financial, transportation, etc.) is represented by only four or fewer firms. Finally, as shown in Table 2.2, the manufacturing firms in our sample are categorized into four industries.

2.4 Empirical Analysis

We conduct two sets of analyses. The first involves a thorough examination of our core research questions. We begin with our base Cobb-Douglas specification and then conduct additional analyses to ensure robustness to issues such as unobserved variables, endogeneity, and choice of controls. The second involves further analysis of the impact of information technology on innovation output. This includes an examination of unique time periods including the mid to late 1990s, returns to IT capital in IT-using versus IT-producing industries, and the contribution of IT to highly valued, blockbuster innovations.

2.4.1 Core Analysis

Our initial approach to estimating the KPF is ordinary least-squares (OLS) regression, using the Cobb-Douglas specification given in (3) and including controls for firm size (log of sales), industry (1.5-digit SIC scheme), and year. Because our data set contains repeated observations of the same firm, we can not assume independence of errors within firms. To address this, we perform the OLS estimation with the errors clustered within firms. The clustered errors approach also includes the Huber-White adjustment to control for arbitrary forms of heteroskedasticity. The results of our base regression model are found in Column 1 of Table 2.3. The estimates provide evidence of the relationship between R&D and IT on citation-weighted patent output.

Both inputs have a positive and statistically significant effect ($p < 0.10$), but the effect of R&D is approximately five times as large. The effect of sales (firm size control) is negative, which is consistent with the earlier finding that large firms are not as effective as small firms at producing patents per dollar of R&D (Griliches, 1990). All of the year and industry controls are statistically significant, which supports the expected impact of the competitive environment over time upon the KPF.[9] The model has an R-squared of 61%, indicating that a substantial portion of the variation in patent output is explained by the independent variables.

To determine if the relative contributions of the inputs to the KPF changed over the course of our sample frame, we divide the sample into two periods (1987–1992 and 1993–1997) and estimate the model for each period. These results (Table 2.3, Columns 2 and 3) show that, although the estimated R&D coefficient remains remarkably stable, the IT coefficient is not statistically significant in the first six years of the sample. In contrast to the estimate for the overall sample, the IT coefficient for the latter five years is larger and significant at the 5% level. We will address this finding in more detail in § 2.4.2.

To relax the assumptions of the Cobb-Douglas form, we extend our base analysis and estimate a translog specification of the KPF. The model to be estimated in logs is

$$\begin{aligned} \ln P_{it} = &\alpha + \beta_1 \ln RD_{it} + \beta_2 \ln IT_{it} + \beta_3 \ln RD_{it}^2 \\ &+ \beta_4 \ln IT_{it}^2 + \beta_5 \ln RD_{it} IT_{it} + \gamma Z_{it} + \epsilon_{it}. \end{aligned} \tag{2.4}$$

The estimation results are shown in Column 4 of Table 2.3. The calculated partial elasticity estimates of R&D and IT ($\eta_{R\&D}$ and η_{IT}) are shown in Table 2.3 below the coefficient estimates. The partial elasticities of R&D and IT are similar to the OLS elasticity estimates for the entire sample and the 1987-1992 and 1993-1997 periods (Columns 5 and 6). Because the coefficient estimate for the R&D-IT interaction term is not significant, we

[9] More granular industry controls such as 2- and 3-digit SIC schemes also produce significant control estimates.

can not infer that R&D and IT capital combine in some synergistic way to produce nonlinear increases in patent output.

Because the panel structure of the data violates the OLS assumption of independence of observations, OLS estimates can be inefficient. In addition, because there is added potential for firm-specific omitted variables to persist through time, the problem of omitted variable bias must be addressed. Typically, a fixed-effects or random-effects panel estimator is used to model firm-level time-invariant unobserved variables.[10] We use the Hausman test to compare the estimates from fixed-effects and random-effects variations of our model and find that we can not conclusively reject the null hypothesis, which indicates qualified support for a random-effects approach.

We evaluate several random-effects estimators[11] but find them prone to computational difficulties and unstable estimates of coefficients and standard errors. The basic random-effects estimator requires strong assumptions of exogeneity and a firm-specific error term that is uncorrelated with observables. We find that this estimator fails to produce statistically significant estimates of our principal variables. A more general approach, using the population-averaged random-effects estimator, allows for correlations within a firm to vary over time. A potential disadvantage of this approach is that it can run out of degrees of freedom over longer panels, which is what we encounter. Because of this limitation, we are only able to calculate estimates for subsamples of the first six and the latter five years (Table 2.3, Columns 7 and 8). The results show positive and significant coefficient estimates for R&D in both periods and for IT in the latter period, which echoes the OLS estimates (albeit at a lower magnitude). In light of these difficulties, we propose that the clustered standard errors approach is in the spirit of random-effects estimators in that it allows the error component to contain omitted variables that are orthogonal to the model variables but common

[10]The firm-level fixed-effects (FFE) model is relatively common in IT value research, as in Brynjolfsson and Hitt (1995) who found that FFE explained a sizeable amount of the output elasticity of IT. We find that this approach results in nonsignificant coefficient estimates when applied to our context.

[11]We thank the senior editor and associate editor for their helpful suggestions in this regard.

within the firm. Because the results from the population-averaged random-effects estimator are qualitatively similar to the clustered OLS results, we proceed under the assumption that the latter estimates establish the upper bound of the output elasticities of our model variables. Our secondary analysis will also provide evidence that unobserved variables are not driving the results of our estimations.

These results consistently demonstrate a positive relationship between both IT and R&D with quality-adjusted patent production. However, we need to account for the possibility of endogenous factors that may influence the relationship between the independent variables and the dependent variable. Although it is intuitive to think of R&D as an input to the production of patents, the opposite may also be true because the development and commercialization of a patented invention requires additional resources that will be reflected in the firm's reported R&D expense. In addition, innovations may stimulate further spending on the firm's IT infrastructure or projects related to marketing or implementing the innovation. This reciprocal relationship may cause the independent variables to be correlated with the model error term, which may lead to inconsistent OLS estimates.

To address these issues, we propose a set of instrumental variables to test for endogeneity and to estimate the model using two-stage least-squares (2SLS). The criteria for a good instrument are a high correlation with the endogenous independent variable but no correlation with the error term.[12] For R&D, we use competitors' R&D spending,[13] the BLS wage index for white collar nonsales occupations, and the firm's own one-period lagged R&D. For IT, we use competitors' IT[14] and the firm's own one-period lagged IT capital services. Lagged R&D has been identified in earlier research as a useful instrument (Blundell et al. 1999). The wage index influences R&D spend-

[12]The correlations between R&D and the instruments are as follows: competitors' R&D=0.514; BLS wage index=0.093; and own lagged R&D = 0.990. The correlations between IT capital and the instruments are competitors' IT = 0.148 and own lagged IT capital services=0.986.

[13]Competitors are defined as all firms reporting R&D within the same 4-digit SIC code in the Compustat database.

[14]Competitors are defined as all firms within the same 4-digit SIC code in the CI database.

ing but not patent output. Competitors' spending is expected to influence a firm's R&D and IT investments although industry factors such as technical change (Allred and Swan, 2005) and IT-driven competitive strategies (Sambamurthy, Bharadwaj, and Grover, 2003) are known to influence a firm's investment decisions.

We also validate our instrumental variables using statistical tests. The Hansen J statistic for our main model (Table 2.3, Column 6) tests for overidentification of all the instruments. Under the joint null hypothesis, the instruments are uncorrelated with the error term and the exclusion restrictions are correct. Using the Stata command `ivreg2` (Baum, Schaffer, and Stillman, 2003), the test statistic is $J = 0.469$ ($p = 0.791$), and thus we fail to reject the null hypothesis. Furthermore, we test the orthogonality conditions of the instruments using competitors' R&D and IT. The C statistic (Baum et al., 2003) tests for exogeneity of a selected instrument by comparing the Sargan-Hansen statistics of the model with and without the instrument. The computed C statistics are $C = 0.282$ ($p = 0.596$) and $C = 0.113$ ($p = 0.737$), respectively. Thus, we fail to reject the null hypothesis that both models are valid.

Having selected the instrumental variables, the Durbin-Wu-Hausman test (Davidson and MacKinnon, 1993) can offer evidence of the presence of endogeneity. We conduct the test separately for R&D and IT. In both cases, we reject the null hypothesis that the OLS estimates are consistent. The results of the 2SLS estimation (again using the clustered standard error method), shown in Table 2.4, are comparable to the OLS estimates.[15] This offers some reassurance that endogeneity is not a major concern.[16] We find both IT and R&D have positive and significant coefficients over the full sample and in the latter half of the sample frame. In the 1987–1992 period,

[15]The number of observations in this analysis is slightly lower due to cases that were dropped when firms did not have consecutive years of data for lagged independent and patent output variables.

[16]We acknowledge that using lagged variables as instruments without controlling for firm-level unobservables is not ideal because the instruments may reflect the same endogeneity as the independent variables. Our analysis shows that, in this case, much of the explanatory power of the 2SLS estimation comes from the inclusion of lagged variables. Consequently, we must use caution when interpreting these results.

however, the estimated contribution of IT is not significant.

Robustness Checks We now verify the robustness of our core results against a number of alternative influences based upon past findings in IT value and KPF research. These are lag effects of IT upon innovation, our method of adjusting patent output for quality, the coarseness of our industry controls, and our adjustments for citation truncation.

First, we test for the influence of lagged IT and R&D on current innovation output. Although it is possible that innovation output responds slowly to variation in KPF inputs, we did not find evidence of a consistent lag structure for either variable. This is likely due to the very stable nature of both R&D expenses and IT capital flows within firms over time. Our results confirm prior research on the slowly-evolving relationship between R&D and patent output (Hall et al., 1986).

Second, we wish to verify that our model is not sensitive to the adjustment methods for patent citations. We re-estimate our model using unadjusted citation counts and unadjusted patent counts. We find that in both cases the results are generally similar to the OLS estimates using citation-weighted patents (Table 2.5). Using a Poisson regression for count data, we find the R&D elasticity to be 0.883 and 0.592 for citations and patents, respectively. The estimates for IT elasticity are 0.370 and 0.452, respectively. All estimates are significant at the 0.10% level. However, these larger estimates do not account for the quality differences among patents or the panel nature of the data, which we have found tends to reduce the magnitude of the IT elasticity estimates.[17]

Third, to address the concern that our 1.5-digit SIC industries insufficiently capture industry-level sources of unobserved influential factors, we re-estimate the model using the 19 2-digit and 73 3-digit SIC industry codes as control variables. We find that the inclusion of this many additional intercepts reduces the significance of our estimate of IT capital below the

[17] We repeated our unweighted patent counts analysis using OLS and 2SLS models. The estimates remain similar in sign, relative magnitude, and significance level to the estimates made using citation-weighted patent counts.

10% level (Table 2.6). Given the lack of efficiency in the clustered standard errors approach, this result is not surprising; as the number of intercepts increases, their share of variance explained rises at the expense of our IT coefficient estimate. However, the industry controls themselves are, for the most part, statistically significant. This effect is robust to 2-digit and 3-digit SIC industry fixed effects.

Finally, all estimations were repeated with the patent citations restricted to a 5-year window. The coefficient estimates and R-squared values are qualitatively similar in all cases. This result reinforces the validity of weighting patent output by citations received, since the results are robust to the length of time each patent is allowed to accrue citations.

2.4.2 Supplementary Analysis

We now explore several aspects of IT value analysis identified in prior literature that add depth to our examination of the IT-innovation relationship.[18] First, we test whether the role of IT evolved in support of innovation over the course of our sample period. Research examining economic growth in the 1990s identified an acceleration in labor productivity and total factor productivity since the mid 1990s. An increase in the rate of decline in technology prices and the deployment of IT in substitution for other more expensive firm inputs are highlighted as key contributors to the resurgence in growth (Jorgenson, 2001). It is possible that research and development is a firm activity that presents opportunities for further deepening in IT investment. Beyond investment motivated by price declines, quality improvements in IT could lead to new IT capabilities that complement other firm inputs (Chwelos et al., 2010) including those used in innovation production.

Adding to this line of research, we examine the contribution of information technology to innovation over the mid to late 1990s. As IT has been highlighted as a source of economic growth throughout the 1990s (Jorgenson, 2001) and given the time frame of our data set, we estimate and compare the contribution of IT to innovation between the time periods of 1987–1992

[18]We thank the associate editor and reviewers for these insightful suggestions.

and 1993–1997.

Accordingly, we perform all of our estimations on both time periods (Table 2.3, Columns 2 and 3, Columns 5 and 8). This analysis highlights the stability of the positive and significant R&D elasticity estimate over the two periods. On the other hand, we find that all estimates show IT elasticity to be larger and stronger in significance during the latter time period when compared with the former. This result is consistent with earlier research showing that investments in information technology provided a higher level of impact during the mid to late 1990s versus its return during the start of the decade (Chwelos et al., 2010; Stiroh, 2002; Gordon, 2000).

Our second supplementary analysis seeks to compare our findings to prior research that examined the contribution of technology to the U.S. economic revival in the 1990s—in particular, whether IT-producing or IT-using industries were the source of U.S. productivity growth. The results of these analyses are mixed, with some research highlighting IT-producing industries as the main contributors to economic growth (Gordon, 2000) while other research identifies both IT-producing and IT-using industries as contributing to the productivity rivalry (Stiroh, 2002; Jorgenson, 2001).

We expand our analysis to examine the relationship between information technology and innovation within these two industry categories. If they did have differential impacts on U.S. productivity growth, is this reflected in the IT-innovation relationship? To test this possibility, we divide our sample according to the 3-digit SIC codes associated with information and communication technology manufacturing. The division resulted in 1,311 IT-using firm observations and 254 IT-producing firm observations. Compared with the entire sample (Table 2.7, Column 1), the results for IT-using firms (Table 2.7, Column 2) are similar, which suggests that firms of this type are benefiting from IT used in the creation of product and process innovations. On the other hand, in IT-producing firms (Table 2.7, Column 3), the estimated IT elasticity is not significant while the estimated R&D elasticity is noticeably larger (1.255% at $p < 0.001$) than that for all firms and for IT-using firms.

There are several possible explanations for the results involving IT-producing firms. First, some IT producers such as Cisco Systems pursue a model of innovation through acquisition. Rather than investing in R&D and related IT similar to traditional industrial-era firms, these IT-producing firms identify and acquire new innovations outside their organizational boundaries. Effectively, traditional innovation investment shifts from the R&D function to the mergers and acquisitions function in some IT producers. As a result, the amount of R&D expensed may be lower while the contribution of R&D to patent output becomes larger relative to their IT capital. An alternative interpretation is that the creation of innovation-intensive IT products is more dependent upon the firm's stock of intangible knowledge rather than the application of tangible information technology assets. Unfortunately, the low number of observations for the IT-producing subsample (some 20 per year) demands that we exercise caution when interpreting these results.

Our third supplementary analysis concerns the contribution of IT to breakthrough or radical innovations. Although rare, such "blockbuster" innovations can create superior value by unleashing market-changing forces. Indeed, some innovation-focused firms such as those in the pharmaceutical industry design their business models around the pursuit of these unique creations (Gilbert, Henske, and Singh, 2003). Research has identified the number of citations a patent receives as an indicator of the value of an innovation, with blockbuster patents accruing far more citations than the average patent (Owen-Smith and Powell, 2003; Hall et al., 2002). The link between IT and blockbuster innovation has not been examined in the IT literature prior to this study. However, university-level research has identified *knowledge flows* and *access to information for external partners* as key factors related to the creation of high-impact patents (Owen-Smith and Powell, 2003). We expect that information technology enables these factors in a firm's pursuit of blockbuster innovations, and that this relationship will be reflected in our data.

The distribution of patents by citation frequency exhibits a very long tail (Figure 2.2), with the median falling at 6.4 citations amid a range from 0 to over 400 among the more than 162,000 patents in our data. By restricting the sample for each firm year to only those patents with citation counts in a certain quartile or above a certain percentile threshold, we may estimate the associated output elasticity of IT and R&D expenditure with innovations ranging from incremental to radical. The results from these restricted samples are reported in Table 2.8.

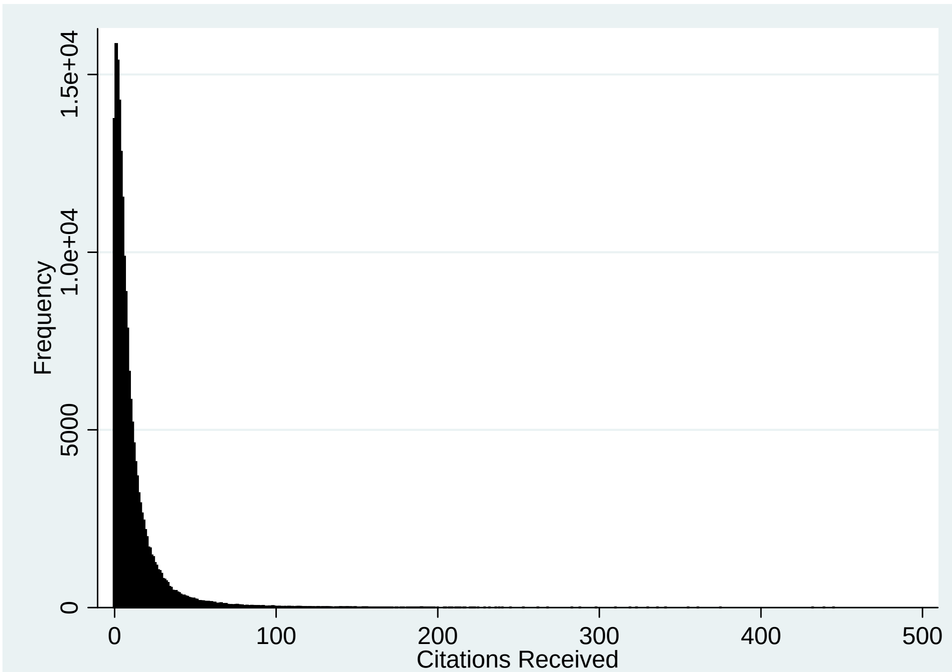


Figure 2.2: Histogram of Citations Received by Individual Patents

By comparing the estimated output elasticities across the columns of Table 2.8, it is evident that the contribution of additional IT capital is greatest for patents in the lowest (1st and 2nd) quartiles of citation frequency but declines when applied to more frequently cited patents. The results suggest that the contribution of IT capital to "blockbuster" patents (in the 90th or greater percentile) is not statistically significant. R&D, on the other hand, has a higher output elasticity estimate for the quartiles that include more frequently cited patents.

These results imply that the effect of general firmwide investment in IT is more likely to facilitate incremental innovation. It appears that IT-enabled improvements in the innovation process, such as data capture, sharing, analysis, and test, etc., are more applicable to building upon current knowledge. In contrast, the contribution of R&D to the production of more innovative patents may highlight the nontechnical nature of the radical creative process. It is possible that data related to such blockbuster ideas are not available in forms than can be captured and analyzed using IT. Rather, such raw and innovative data exist in nonexpressive forms only available to the minds of research and development personnel. Such knowledge does not become available until newly developed ideas have advanced enough to be translated into tangible, binary forms. Such radically new concepts are so novel and unique that they can only be produced with inputs other than information technology.

Overall, our estimates of the innovation elasticity of IT and R&D are stable within the range of 0.14 to 0.28 and 0.65 to 0.90, respectively, and appear robust to different specifications and methods. Furthermore, we show that the contribution of IT to innovation is more evident in specific firms, periods, and innovation quality strata. Taken together, these findings reduce the likelihood that unobserved variables are driving the relationship between IT and innovation output. Although an unknown cause of firm-level heterogeneity could influence our results, such a factor(s) would have to apply only to incremental innovation, IT-using firms, and the period 1993–1997.

The positive and significant output elasticity identified for IT across multiple assumptions and specifications presents strong evidence of the critical role of IT in innovative activity. Although the magnitude of R&D elasticity is larger by comparison, this reflects the central role in innovation creation of R&D-related intangibles (e.g., scientists' knowledge, skill, creativity, etc.) in innovation. It is possible that the point estimate of IT elasticity could be higher and more accurate with the inclusion of other IT spending dimensions (software, labor, etc.). However, if such data were available, we would expect its relative magnitude to remain smaller than that of R&D.

2.5 Discussion

The contribution of IT to firm-level productivity has been demonstrated in prior research, yet much remains unknown about the mechanisms by which this is accomplished. The production of intangible organizational outputs such as knowledge is a key area in which to look for the contribution of IT to a firm's operations and, ultimately, the value of IT. In an effort to uncover the contribution of IT to the creation of new knowledge, we augment the KPF (Pakes and Griliches 1984) to include IT capital input, a variable previously excluded in innovation research. To examine the relationships between IT, innovation activity, and innovation output, we analyze a unique data set containing annual data on R&D expenditures, IT capital flows, and patent citations for 201 Fortune 1000 firms for the years 1987 to 1997.

Our estimates indicate that IT and R&D both play positive and significant roles in innovation production. In our panel of large manufacturing firms, a 1% increase in IT capital services is associated with an increase in (citation-weighted) patent output of 0.166% in the 1987–1997 period. Furthermore, we find evidence that the role of IT in innovation became stronger in the last five years of our sample as firms increased their IT investments dramatically. Beyond capital deepening, the contribution of IT during this time frame is also due, in part, to new IT capabilities being introduced at the time, namely, networking and Internet-based interconnectivity. Indeed, research at the economy level lends credence to this supposition (Jorgenson, 2001; Stiroh, 2002).

Our estimates of the augmented KPF are consistent with earlier research in two key respects: R&D has a positive impact on citation-weighted patent output while firm size is inversely related to patent output. Indeed, the impact on patent output of a 1% increase in R&D is 0.896%, about 5 times larger than that of IT. This larger magnitude is expected because R&D spending is a direct input to innovation process while firmwide IT investment will support the firm's productive and administrative processes in addition to supporting the innovation creation process. The firm size effect is also confirmed to the extent that our OLS and 2SLS estimates of the control variable (sales) are negative (with the exception of the patents with first-quartile citation frequencies).

Our results also demonstrate an important aspect of IT and the innovation creation process: Although firms invest specifically in R&D inputs including dedicated IT spending within R&D programs, the impact of IT on innovation goes beyond R&D-specific IT spending. Specifically, our results suggest that general infrastructure and enterprise technologies of a firm (networks, e-mail, telephony, accounting, and finance ERP modules, etc.) also contribute to the innovation process. Interestingly, we also find that the effect of IT on innovation was strongest for "incremental" patents and nonsignificant for "blockbuster" patents. This suggests that although IT contributes broadly and significantly to innovation whether measured by patents or citations, IT alone does not lead to breakthrough innovations. Rather, breakthrough or radical innovations may be more dependent on other factors such as the tacit knowledge of R&D scientists and engineers.

Finally, we contextualize our results in terms of market value impact. Using a sample of 4,864 publicly traded U.S. firms over the period 1979–1998, Hall et al. (2005) found that if the average number of citations received by a firm's patents increases by one, the market value of the firm increases by 3%. Assuming that this result is generalizable to our sample, we can estimate the marginal impact of IT investment on innovation output and extrapolate it to obtain the expected increase in market value. Using our upper- and lower-bound IT elasticities for the period 1993-1997, we estimate that the median firm in our sample would need to increase its IT capital flows by 39% to 81% ($3.8 to 7.7 million) to obtain one additional citation for each of its patents. Although this amount is rather large at first glance, we note that, between 1993 and 1997, the median firm increased its IT capital services by 26% ($2.5 million). Furthermore, the innovation-driven 3% market value increase predicted by Hall et al. (2005) would amount to $41 million for the median firm in our sample. Thus, our estimates of the return to investment for IT-driven innovation appear to be economically meaningful.

Although our results confirm many of our expectations, we recognize several limitations of our work. First, our data set is not without its shortfalls; the Compustat R&D data can not be broken down into spending by the type of innovation pursued (e.g., product or process). Doing this would allow for a refined analysis to identify the types of innovation efforts that are more likely to pay off in a finished and highly valued innovation. The R&D measure also does not encompass any spending information on informal innovation activity that may be conducted in firms. However, given explicit FASB rules and definitions, annual R&D expenditures must be reported, which thereby constrains off-the-books innovation spending.

Also, although our information technology construct incorporates all types of IT elements including hardware, software, IT skill, and organizational complements, the CI measures are based exclusively on IT hardware. For example, product innovation may require investments in both IT and new marketing skills as a firm moves into new marketplaces or new production processes to manufacture new products (Garcia and Calantone, 2002). This data shortfall limits the identification of exact-point estimates of the impact of IT on innovation output. However, insofar as investments in IT hardware are associated with these other factors, our IT measures provide reasonable insight into the IT construct.

The CI data also can not be broken down into measures of IT that directly support R&D and those that do not. Although the availability of more focused measures of R&D-related IT spending[19] would provide a more comprehensive and precise estimate of the impact of IT on innovation, such data are prohibitively difficult to obtain. However, our data set is not without its merits in relation to innovation activity analysis. The CI data represent a comprehensive, firm-level measure of information technologies primarily devoted to infrastructure. Such technologies (e.g., networking technologies,

[19]R&D-related IT spending measures would provide a more direct match to our IT construct. The ideal measure would include the percentage cost of enterprise infrastructure technologies that support R&D (networks, e-mail, etc.) as well as incremental IT investments made specifically for R&D projects. Project-specific R&D-related IT could include specialized technologies such as CAD/CAM systems used in new product design, data analysis applications such as specialized statistical programs for market research, and customized software for the creation of innovative process-specific activities (e.g., a new transaction process such as Amazon's 1-click).

e-mail applications, databases, etc.) have been explicitly mentioned as enabling collaboration and R&D (Kumar and van Dissel, 1996; Rice, 1994). As a result, the CI data, although not ideal, does offer insight into the contribution of IT to firm innovation.

Patents are a good proxy for measuring the output of firm innovation efforts: however, they represent only one type of outcome associated with innovation and are not guaranteed at that. In addition, as Griliches (1990) summarized, "not all inventions are patentable, not all inventions are patented, and the inventions that are patented differ greatly in ... the magnitude of inventive output associated with them" (p. 1669). Hence, although our data represent patented innovations with adjustments to reflect their quality, we recognize that they do not represent all incremental innovations of a firm.

Beyond facing the aforementioned data challenges, in the spirit of the KPF literature stream (Pakes and Griliches, 1984; Crepon et al., 1998), future research could jointly estimate production and patent functions. In this way, one could examine the direct effects of IT on firm performance while accounting for the indirect effects of IT that take place through new innovations. Market value performance would be a natural first area for examination because a rich, independent literature stream exists in both the IT and R&D areas (see Melville, Kraemer, and Gurbaxani (2004); Hall (2000) for literature surveys).

Future research could also draw on the natural link between innovation and location-specific effects. Although innovation research has long acknowledged the importance of proximity to other innovators and the nature of a firm's external business environment, these factors are a relatively new area of inquiry for IT business value research. In the context of the Internet era as well as the shift of some IT knowledge work overseas, it is possible that today's IT is reshaping the notion of what is "location specific" as applied to innovation.

2.6 Conclusion

Empirical researchers have accumulated significant evidence of the contribution of IT to firm-level productivity. However, the underlying mechanisms by which this effect occurs are not well-understood. By examining intangible outputs such as innovation, we shed new light on the use of IT to create firm value. In particular, we identify IT as a new and effective input to the R&D-driven innovation process. Finally, by using patent citations as a measure of quality-adjusted innovation output, we overcome a major limitation of innovation measurement associated with the KPF. By doing so, we help set the stage for greater understanding of the value of innovation and IT in the overall production context.

Table 2.1: Descriptive Statistics

Variable	Obs.	Median	St. dev.	Min	Max
Sales*	1,829	2,642.93	15,728.26	440.64	166,620.20
R&D expense*	1,829	62.76	784.87	0.46	8,359.77
IT capital services*	1,829	9.68	53.10	0.10	787.00
Patents	1,829	24.00	191.80	1.00	2,405.00
Patent citations received	1,829	208.00	2,423.85	1.00	31,733.00
Patent citations received (adjusted)	1,829	211.09	2,464.27	1.00	36,683.25
5-yr capped citations received	1,820	104.00	1,598.27	1.00	27,082.00
5-yr capped citations received (adjusted)	1,820	105.44	1,496.93	0.79	23,145.31

* Millions, 1993 dollars

Table 2.2: Sample Composition by Industry

	Observations		Firms	
Industry	**N**	**%**	**N**	**%**
Nondurable manufacturing	112	(6.1)	12	(6.0)
Durable manufacturing	571	(31.0)	63	(31.3)
Process manufacturing	517	(28.1)	57	(28.4)
High-tech manufacturing	639	(34.7)	69	(34.3)

The industry classification is motivated by prior IT value research (Bresnahan et al., 2002) and derived from grouping 20 2-digit SIC firms into 10 categories. The firms in our sample represent only 4 of these 10 categories.

Table 2.3: Estimation Results

	1	2	3	4	5	6	7	8
	OLS (Cobb-Douglas) (clustered robust standard errors)			**OLS (translog) (clustered robust standard errors)**			**Random-effects panel (population averaged)**	
	All years	**1987–1992**	**1993–1997**	**All years**	**1987–1992**	**1993–1997**	**1987–1992**	**1993–1997**
$\ln(RD)$	0.896***	0.893***	0.897***	0.885***	0.939***	0.814***	0.766***	0.658***
	(0.0701)	(0.0755)	(0.0860)	(0.0953)	(0.123)	(0.126)	(0.0683)	(0.0709)
$\ln(IT)$	0.166*	0.0970	0.259**	0.267**	0.191	0.377**	0.0880	0.142*
	(0.0946)	(0.0947)	(0.128)	(0.126)	(0.213)	(0.167)	(0.0787)	(0.0755)
$\ln(Sales)$	−0.229**	−0.202*	−0.254*	0.0536	0.0750	0.135	−0.0628	0.0310
	(0.104)	(0.103)	(0.133)	(0.0473)	(0.104)	(0.0905)	(0.0975)	(0.105)
$\ln(RD) \times \ln(IT)$				−0.0205	−0.0232	−0.0472		
				(0.0264)	(0.0460)	(0.0407)		
$\ln(RD)^2$				0.00254	−0.0131	−0.0775		
				(0.0311)	(0.0683)	(0.0736)		
$\ln(IT)^2$				−0.257**	−0.246**	−0.260*		
				(0.112)	(0.119)	(0.140)		
Constant	3.196*	4.598**	2.143	4.891*	5.412**	5.005	2.296	−0.936
	(1.933)	(1.867)	(2.391)	(2.482)	(2.701)	(3.069)	(1.646)	(1.868)
Observations	1,829	1,006	823	1,831	1,008	823	1,006	823
R^2	0.610	0.627	0.596	0.611	0.625	0.598	—	—
η_{RD}	0.896***	0.893***	0.897***	0.896***	0.890***	0.900***	0.766***	0.658***
	(0.0701)	(0.0755)	(0.0860)	0.069	0.078	0.085	(0.0683)	(0.0709)
η_{IT}	0.166*	0.0970	0.259**	0.177*	0.112	0.239*	0.0880	0.142*
	(0.0946)	(0.0947)	(0.128)	0.103	0.116	0.132	(0.0787)	(0.0755)

Dependent variable: log of citation-weighted patents.
Estimates for year and industry indicators omitted from results.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 2.4: 2SLS Estimation Results

	All years	**1987–1992**	**1993–1997**
ln(RD)	0.896***	0.900***	0.889***
	(0.0769)	(0.0850)	(0.0911)
ln(IT)	0.198*	0.114	0.287*
	(0.111)	(0.111)	(0.150)
ln($Sales$)	−0.263**	−0.234**	−0.286**
	(0.114)	(0.117)	(0.142)
Constant	3.472*	4.867**	2.458
	(1.995)	(2.041)	(2.421)
Observations	1,565	788	777
R^2	0.613	0.633	0.599

Dependent variable: log of citation-weighted patents.
Cluster-adjusted robust standard errors.
Estimates for Year and Industry indicator variables omitted from results.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 2.5: Poisson Regressions

	Dependent Variable:	
	Raw citations	**Raw patents**
ln(RD)	0.883***	0.592***
	(0.00115)	(0.00351)
ln(IT)	0.370***	0.452***
	(0.00145)	(0.00474)
ln($Sales$)	−0.393***	−0.152***
	(0.00138)	(0.00440)
Constant	3.572***	−3.581***
	(0.0240)	(0.0775)
Observations	1,839	1,839

Standard errors in parentheses.
Year and Industry indicator variables estimates omitted.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 2.6: Granular Industry Controls

	SIC controls		
	1.5-digit	**2-digit**	**3-digit**
$\ln(RD)$	0.896***	0.901***	1.002***
	(0.0769)	(0.0822)	(0.136)
$\ln(IT)$	0.198*	0.107	0.166
	(0.111)	(0.106)	(0.118)
$\ln(Sales)$	−0.263**	−0.138	−0.291
	(0.114)	(0.123)	(0.181)
Constant	3.472*	2.140	4.034
	(1.995)	(2.282)	(3.531)
Observations	1,565	1,565	1,565
R^2	0.613	0.650	0.717

Standard errors in parentheses.
Year and Industry indicator variables estimates omitted.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 2.7: IT-producing vs. IT-using industries

	All firms	**IT-using 3-digit SIC**	**IT-producing 3-digit SIC**
$\ln(RD)$	0.896***	0.786***	1.255***
	(0.0769)	(0.0938)	(0.304)
$\ln(IT)$	0.198*	0.211*	0.0683
	(0.111)	(0.125)	(0.229)
$\ln(Sales)$	−0.263**	−0.185	−0.359
	(0.114)	(0.136)	(0.326)
Constant	3.472*	1.926	6.039
	(1.995)	(2.381)	(7.134)
Observations	1,565	1,311	254
R^2	0.613	0.6	0.525

Dependent variable: log of citation-weighted patents.
Cluster-adjusted robust standard errors.
Estimates for Year and Industry indicator variables omitted from results.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 2.8: Blockbuster Patents Analysis

	Entire sample	**1st quartile**	**2nd quartile**	**3rd quartile**	**4th quartile**	**90th percentile**	**95th percentile**	**99th percentile**
Lower bound (citation frequency)		0	2.71	6.4	13.6	25.6	37.04	72.58
$\ln(RD)$	0.922***	0.471***	0.496***	0.519***	0.698***	0.636***	0.606***	0.424***
	(0.0311)	(0.0270)	(0.0259)	(0.0257)	(0.0312)	(0.0377)	(0.0461)	(0.0631)
$\ln(IT)$	0.178***	0.249***	0.260***	0.183***	0.119***	0.0351	0.00833	0.0145
	(0.0474)	(0.0403)	(0.0391)	(0.0376)	(0.0449)	(0.0497)	(0.0563)	(0.0676)
$\ln(Sales)$	−0.240***	0.0887**	0.0426	0.0634	−0.142***	−0.117**	−0.155***	−0.159**
	(0.0505)	(0.0422)	(0.0409)	(0.0402)	(0.0473)	(0.0528)	(0.0599)	(0.0731)
Constant	2.649***	−6.562***	−5.765***	−4.756***	−0.246	0.282	1.316	1.823
	(0.786)	(0.662)	(0.633)	(0.630)	(0.727)	(0.819)	(0.964)	(1.143)
R^2	0.626	0.613	0.615	0.610	0.545	0.465	0.376	0.255
Observations	1,809	1,591	1,632	1,629	1,470	1,078	786	348

Dependent variable: log of citation-weighted patents using 5-year capped citation window.
All estimations use cluster-adjusted robust standard errors.
Estimates for Year and Industry indicator variables omitted from results.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Chapter 3

The Knowledge Factory: Innovation and IT Investment in Manufacturing

3.1 Introduction

Research concerning the business value of IT has become increasingly focused on the mechanisms by which firms harness IT to create value. Among intermediate value-creating processes, innovation has long been considered vital to the competitive advantage of firms. Since innovation is tied to the creation and recombination of knowledge, IT has been recognized as playing a central role in the discovery and refinement of new products and manufacturing processes. The goal of this investigation is to evaluate a model that will incorporate IT as a key input to both the direct production process and the intermediate innovation process. The empirical validation and estimation of this model will shed light on the dual role of IT in the production and innovation processes of firms.

3.2 Background

Productivity is an economic concept that expresses the ratio between the quantity of output produced by some input or combination of inputs. Economists model and estimate the nature of production of firms and economies using production economics, which represents the process of generating economic value as a mathematical formulation relating inputs to outputs. Such production functions are used to model this relationship at the firm, industry, or economy level. Productivity is considered an important indicator of economic growth and general prosperity.

The classical factor inputs in production functions are labour and capital. Formally, the aggregate production function in a Cobb-Douglas form is:

$$Y = AK^{\alpha}L^{\beta}, \tag{3.1}$$

where Y is output in GDP, K is capital and L is labour. The exponent terms, α and β, represent the share elasticity of the factors. The constant A represents technological change–also known as multifactor productivity or the Solow residual, since it explains the growth left over after accounting for labour and capital as having arisen from structural shifts, mainly due to technology (Lee, Gholami, and Tong, 2005). Although other specifications of production functions have been advanced, researchers continue to find the Cobb-Douglas form useful in empirical analyses due to the ease of interpreting the regression estimates. By taking the log of both sides of the equation, one may interpret the regression coefficients of labour and capital as output elasticities in percentage form.

3.2.1 IT Productivity Research

In the late 1980s, a number of economists and IS researchers began to address the question of IT productivity at the economy level and found disturbingly small, and even negative, results. Nobel laureate Robert Solow attracted widespread attention with his 1989 remark that "one sees computers everywhere except the productivity statistics." By 1993 this so-called "productivity paradox" was considered an economic issue of great importance because, despite massive technological improvements in computing power and widespread large-scale investment in IT, much of the literature found this capital build-up had a negligible-to-negative relationship with productivity at the overall and at the factor levels. In his widely-cited review of IT productivity literature, Brynjolfsson (1993) suggested measurement error as a primary reason as to why the results to date had been so disappointing. His survey included four studies at the economy level and fourteen at the firm level, conducted between 1983 and 1993 and published in business and economics journals and related conference proceedings. Typically, such studies used the production function approach, with IT capital modeled as a separate variable from "conventional" capital (denoted Z in equation (3.2), below),

$$Y = AK^{\alpha}L^{\beta}Z^{\gamma}. \tag{3.2}$$

Brynjolfsson identified several difficulties with this approach: the direction of causality is unknown, and the results are sensitive to the assumptions of functional form and measurement (such as price deflators and industry categorizations). However, the research also showed that traditional inputs had the expected (positive) influence on output, affirming the production function approach and presenting some reason to suspect that measurement error may not be entirely to blame. Other research around this time showed that the nature of IT investment varied widely across firms and that these differences could explain why earlier studies found mixed results. Brynjolfsson and Hitt (1995) analyzed a panel of firms that had been used in previous studies and found that firm fixed effects helped explain as much as half the productivity gains ascribed to IT. Weill (1992) pointed out the sensitivity of the IT-productivity relationship to the form of IT: gains from strategic IT investments were competed away, while transactional IT investments appeared to have a stronger connection to improved firm performance.

Recent summaries of subsequent research findings (Dedrick, Gurbaxani, and Kraemer, 2003; Kohli and Devaraj, 2003; Melville et al., 2004) have largely put to rest the macroeconomic productivity paradox, but at the micro level the consensus is more cautious and qualified, contingent on complementary investments and environmental moderators. Questions remain as to factors explaining the variation in the contribution of IT to productivity at the firm and industry levels, as well as in the form that productivity takes. At the industry level, recent research suggests that gains in labour productivity (in response to IT investment) have been mostly positive between the 1989–95 and 1995–99 periods (Council of Economic Advisors, 2001), but changes in this contribution are different across industries. This is likely related to the second question, which is whether researchers are correctly modeling the form of productivity improvements that can be attributed to IT. In contrast to standard measures of output, IT may be contributing more to intermediate processes such as the deployment of new business models, gathering of market intelligence, and innovation. Indeed, this was the conclusion of Barua, Kriebel, and Mukhopadhyay (1995) who found evidence of a relationship between IT capital investments and intermediate-level indicators such as capacity utilization and inventory turnover.

3.2.2 Knowledge Production Function

Among economic models of innovation, one of the most compelling is that proposed by Pakes and Griliches (1984). Similar to the neoclassical production function (NPF) given in 3.1, the knowledge production function (KPF) represents the knowledge output generated by a firm as a function of the inputs used within its innovation process. The KPF, represented by the lower half of Figure 3.1, links the neoclassical production function in the upper half (with inputs denoted $X_{1..j}$, with error terms $v_{1..j}$, and several potential measures of output, such as Value Added or productivity, denoted by $Z_{1..j}$) to the influence of knowledge on these production processes. In the KPF, the knowledge created by a firm, denoted K, is a function of R&D expenditure R (subject to a stochastic disturbance u reflecting the uncertainty of knowledge creation). This new knowledge is inherently unobservable. However, patented inventions, P, can serve as an observable indicator of the production of new knowledge. Pakes and Griliches estimated their specification of the KPF and found a positive and statistically significant contribution of R&D to patent output. Due to limitations of data and computation, however, it was not feasible to estimate the larger model at the time of their publication.

The use of patents as indicators of inventive activity has proven fruitful to researchers due to several factors (see Griliches (1990) and Hall et al. (2005) for extensive reviews). First and foremost, the patent application

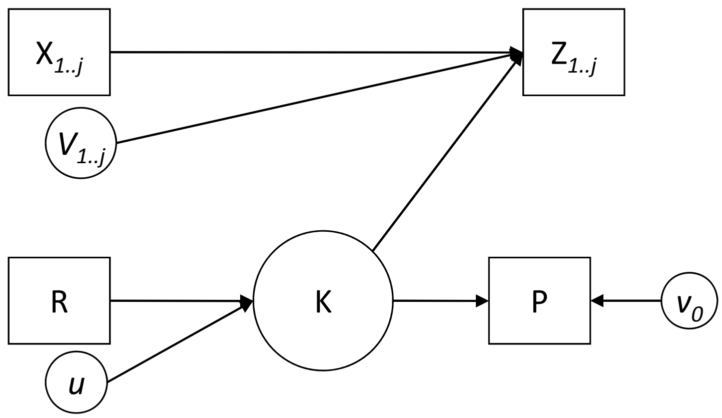


Figure 3.1: Microeconomic Model of Production including Knowledge Production Function (Pakes and Griliches, 1984)

examination process imparts objectivity. In the US, patent law requires the process or product to be novel and non-obvious, as well as "useful," i.e. having a useful purpose and operativeness (USPTO, 2006). This standard has been relatively stable over time, making comparisons across years and industries more robust. Second, a patent also provides a detailed record of the inventor, the industrial field into which the invention is classified, and citations of prior patents upon which it builds. The enforcement of the latter by patent examiners is especially important since it documents a recognition of "prior art," allowing the researcher to draw inferences based on the nature and quantity of cited patents.

Although patents are a rich and robust indicator of innovation output, there are two principal shortcomings that have been discussed in the literature. First, patents do not measure all of a firm's inventive output (see § 2.3.1 for a discussion of this aspect of patent data and related literature). The second shortcoming is the influence exerted by regulatory and competitive climate on the decision of firms to pursue patents. Researchers investigating the determinants of patenting have noted the roles of industry concentration and rates of technological progress in patenting, and the differences of these on the propensity to register process versus product patents (Lunn, 1987). Arora and Ceccagnoli (2006) show that more effective protection of patents provides, not surprisingly, an increased incentive for firms to seek patentable innovations. Differences in incentives may arise at the firm level if the firm possesses specialized complementary assets that make it more difficult to license an innovation.

3.3 Model

The objective of this research is to integrate IT capital as a distinct input from labour and capital in the original overall model, and then estimate the path coefficients to understand the simultaneous contribution of IT capital on both indicators of patents and output. The model proposed by Pakes and Griliches (1984) resembles a structural equation model (SEM). Knowledge is a latent (unobservable) construct, to which R&D may be considered a formative indicator and patent output a reflective indicator. In the SEM tradition, this is a multiple-indicators multiple-causes (MIMIC) model because there are no endogenous latent constructs. The updated structural equation model for this research, an operationalized Knowledge Production Function, is presented in Figure 3.2. The upper portion of the model (representing the neoclassical production function) does not contain a latent variable, but it is straightforward to consider the X variables as formative indicators of a new latent construct we shall call Productive Resources (PR). Thus, PR represents an index of the firm's available resources with which it produces output. A path from Knowledge to PR maintains the original model's hypothesis that innovation (positively) influences the firm's productivity.

The top half of the structural equation model represents the IT-augmented neoclassical production function (equation (3.2)). The traditional inputs labour, L, conventional (non-IT) capital, K, and IT capital, Z, are modeled as formative indicators of the PR latent variable. Output, Y, is measured by Value Added. As the single output indicator of PR, Y is considered to have no measurement error. The lower half of the model represents the KPF, with

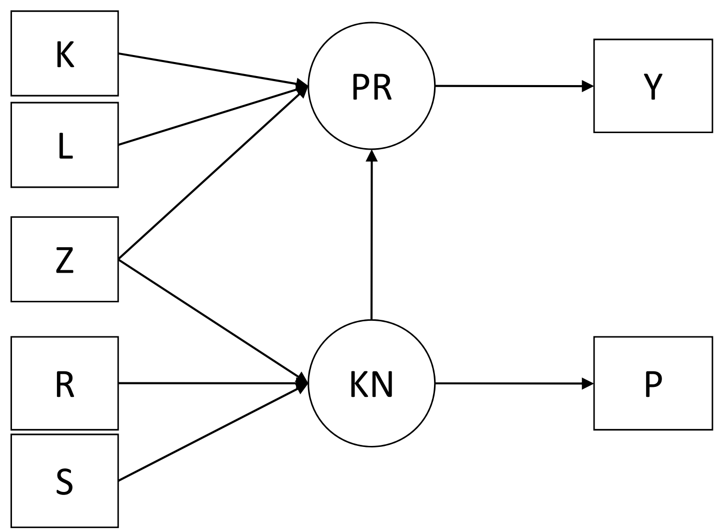


Figure 3.2: Structural Equation Model

the latent construct Knowledge (now denoted KN) having formative indicators Z and R, and output indicator P, citation-weighted patent output. As a control for firm size, the number of employees (denoted S) is included because of the theoretical and empirically-verified diseconomies of scale on innovation. Griliches (1990), for example, found smaller firms (having less than 1,000 employees or spending less than two million dollars per year on R&D) in his sample of publicly-traded companies obtained more patents in proportion to their size than larger ones.

There are several unusual aspects to the measurement model. First, guidelines for SEM research recommend having at least three indicators per construct. In our basic model, we are not able to provide three indicators due to data limitations. Second, our hypothesis is that IT capital influences both direct production and intermediate knowledge production. Therefore the indicator Z has paths leading to both the PR and KN constructs. Computationally, this has an attenuating effect on the loading of the indicator on both paths. Finally, the R&D factor, R, is a subset (in some unobservable combination) of all three production function inputs L, K, and Z

(e.g. scientists, research labs, and computer equipment). This results in some double-counting of the firm's overall labour, non-IT capital and IT capital inputs since some portion of each of these is also included in R&D. The double-counting of these inputs in the knowledge-creation process may result in inflated measurement errors for the estimation of the path loadings.

3.4 Data

The estimation of the model utilizes data gathered from three principal sources, as detailed in the previous chapter of this dissertation. A summary of these sources and the data adjustment process is given here. First, we obtain the dependent variable for the KPF from the NBER Patent Citation Database (Hall et al., 2002), which details the successful patent applications between 1975 and 1999 for a large number of U.S. firms. This database has updated by including citations up to 2004 to allow patents near the end of the sample period enough time to receive adequate citation counts. Second, we obtain data on IT capital from the Computer Intelligence Infocorp database, which gives details on the value of installed IT at nearly 800 Fortune 1000 firms from the period 1987–1999. Finally, the R&D expense, labour expense, (non-IT) capital book value and annual Value Added output for the firms in the union of the first two data sets is obtained from Standard and Poors Compustat, a database of financial and stock market information.

The NBER Patent Citation database (Hall et al., 2002) records, for each successful patent application, the date of application and the date the patent was granted. Using the application date is more appropriate for our purpose since R&D expense tends to be recorded early in the patenting process.

As mentioned earlier, patents are not a perfect indicator of innovative activity. Insofar as they do measure an actual innovation, however, an additional concern is that patents do not account for variations in the quality or value of the innovation. We address this through the use of citation-weighting methodology proposed by Hall et al. (2002). Since the citation of prior patents is enforced by the USPTO, the number of citations a patent receives by subsequent patents can be taken as an indicator of its economic value (Hall et al., 2005; Trajtenberg, 1990). To prevent older patents from accumulating more citations at sample time than new patents, we generated a citation count based on the citations received within a fixed window of five years from the application date of each patent.

All variables are expressed in “flow,” rather than “stock,” terms. The yearly value of IT and non-IT capital services is calculated using the rental price approach used by the U.S. Bureau of Labor Statistics. Although non-IT capital rental prices have been used extensively in production economics research, the history of difficulty with price deflators for IT capital requires a specialized approach. Based on the IT capital stock data provided by Computer Intelligence, Chwelos et al. (2010) create a measure of annual capital services flowing from this stock of IT in constant quality-adjusted dollars. Rental prices reflect the cost of capital and include the rate of return, depreciation and expected rate of asset price change, net of income and property taxes. Value Added output is calculated by subtracting non-labour expenses from sales. Finally, R&D expense is an item reported in firms’ financial statements. To improve tractability, the natural log of all variables has been taken.

The balanced panel of 123 firms includes observations from 1993 to 1997. Summary statistics are reported in Table 3.1. The 1993–1997 period was chosen because it captures the rise of the internetworking era of computing, in which communications technologies enabled vast improvements, using other forms of IT, in co-ordination of teams and operations on a global scale. The sample includes 72,565 patents, with the median firm obtaining 36 successful patent applications per year. A total of 680,649 citations were recorded by these patents to the end of 2004, with the median patent receiving 230 citations. The data are restricted to the manufacturing sector because the number of firms in the service sector that engaged in significant R&D and patenting was very small.

3.5 Methods and Results

The partial least-squares (PLS) approach is considered a primarily exploratory approach, which is appropriate here since although theory has established the model form, the constructs have not been tested in the operationalized model. Previous research investigating the KPF and neoclassical production functions has predominantly employed least-squares regression, especially with panel (i.e. cross-sectional time series) data. Our research model incorporates both a NPF and the KPF, and since common unobservables are likely to affect both processes with the firm, analyzing the production model as a system offers the advantage of explaining variation in both processes in a single estimation. We use a component-based PLS approach implemented in the SmartPLS program (version 2.0(beta)) (Ringle, Wende, and Will, 2005).

We estimate the model in two variations. First, we model the production of goods and knowledge in a single year and repeat the analysis for period 1993–1997. Second, we reformulate the model to allow for multiple observations of the same firm over the 5-year period in a single estimation. This will allow for a closer comparison to existing KPF and IT capital research using similar data.

The objective of the first round of analysis is to estimate the single-year model for each year in the sample. This produces five sets of results, from 1993 to 1997. In order to estimate the single-year MIMIC model in the SmartPLS software package, two latent variables are created to represent each construct (Productive Resources: PR and PR'; and Knowledge: KN and KN') in the model (Figure 3.3). This is necessary because SmartPLS will not allow a construct to have both formative and reflective indicators. However, the result of the estimations is equivalent since the two latent variables represent the same construct and there are no other paths that influence either of the constructs. The path between them tests the validity of the formative indicators (Chin, 1998).

To analyze the model in this form, an additional estimation stage is required. The first stage (above) validates the formative model for the NPF

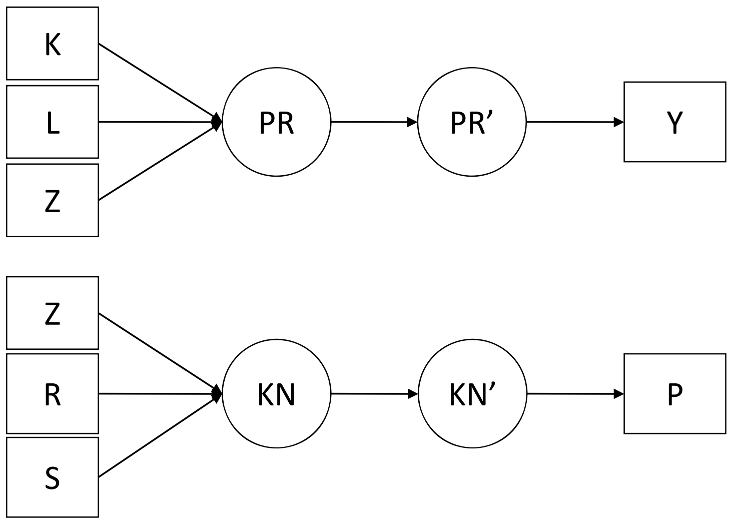


Figure 3.3: First-stage model: MIMIC Validation

and (separately) the formative model for the KPF. Having established the indicators are valid formative measures of the constructs, a second estimation is performed using the latent variable scores (*PRscore* and *KNscore*) from the first estimation to predict the influence of KN on PR, as seen in Figure 3.4. If the entire model were to be estimated in one stage, the MIMIC validation becomes unreliable since the path from KN to PR acts as another influence on PR. In other words, it becomes difficult to interpret the path weights as either an indication of formative indicator validity or prediction from one construct to the other. By performing the prediction separately, we isolate the validation and prediction estimates.

The estimation results from this two-stage single-year model are reported in Table 3.2. For the first-stage estimation of the validation coefficients, Chin (1998) suggests this coefficient should be at least 0.8 and preferably above 0.9. The PR construct (path coefficient from PR to PR') is validated for all years 1993 to 1997, with validation path coefficients in the range of 0.975 to 0.980. The validity coefficient for the KN construct (KN to KN') closely approaches Chin's cut-off, in the range of 0.750 to 0.795, with the exception

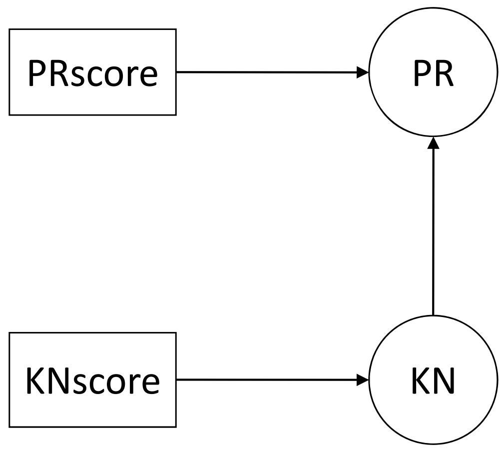


Figure 3.4: Second-stage model: Predictor Estimation

of 1996 (0.602).

The estimated indicator weights for K, L and R are significant at the 5% level (t-values greater than 1.96) for all years in the sample. While the indicator weights for Z on PR are not significant in any of the sample years, the weights of Z on KN are significant for the years 1993–1995. Because the output indicators Y and P are modeled as singular indicators, they are considered error-free measures and their loadings are reported as 1.00, accordingly.

In addition to examining each formative indicator's weight, it is useful to consider its *loading*: the bivariate correlation of an indicator to its construct (Cenfetelli and Bassellier, 2009). These are interpreted as the absolute (rather than relative) importance of an indicator. Notably, the loadings for all indicators are consistently high (Table 3.3), including the loadings for Z. On the Productive Resources construct, Z loadings are consistently in the 0.81 to 0.85 range, and in the 0.66 to 0.81 range for the Knowledge construct. This indicates that while the *unconditional* effect of IT capital on Productive Resources is high, the *partialled* effect is not significant once the model controls for the effects of non-IT capital and labour. This is likely a result of multicollinearity, given the high correlation between Z and L (0.79 to 0.86) and Z and K (0.71 to 0.75). Thus, while we cannot predict the outcome of a marginal change in the amount of IT capital, we also cannot dismiss altogether its importance to the firm's Productive Resources.

We examine several measures of reliability that pertain to this model specification. The latent constructs, PR and KN, capture a large portion of the variation in the model. For all sample years, the R-squared for PR is consistently above 95%, while the R-squared for the KN construct is between 56% and 63%, with the exception of 1996 (36%). Due to the MIMIC structure of the latent variables, additional measures of construct reliability (Average Variance Extracted (AVE) and Cronbach's alpha) are not computed by SmartPLS. Finally, the path coefficient between KN and PR is strong and significant at the 1% level for all sample periods, ranging from 0.623 to 0.721.

3.5.1 Multiple-Observations Model

The second set of estimates, below, are from the multi-year variation on the original model. In this multiple-observations model, presented in Figure 3.5, each formative indicator (K, L, etc.) from the original model is implemented as a latent variable with a cluster of formative indicators composed of three years of observations for that variable for 1993–1995 (e.g. k_{93}, k_{94}, k_{95}). The original latent constructs (representing Knowledge and Productive Resources) become endogenous constructs, with reflective (output) indicators for value-added output and patent citations for 1995–1997, respectively. The objective of modeling the KPF in this way is to incorporate additional information on the composition of the firm's Knowledge and Productive Resources. By accumulating a series of input flows over time, we create constructs that represent stocks of the resources used in both knowledge production and direct production in subsequent periods. The labour construct will thereby approximate the stock of human capital—skills and tacit knowledge—built up over several years through wage expenses. Similarly, the non-IT and IT capital constructs will capture the firm's stock of equipment that, over time, is allocated to provide the most efficient deployment. The firm's overall stock of intellectual property is also built up over time, as reflected in the accumulations of R&D expenditures and patent citation output.

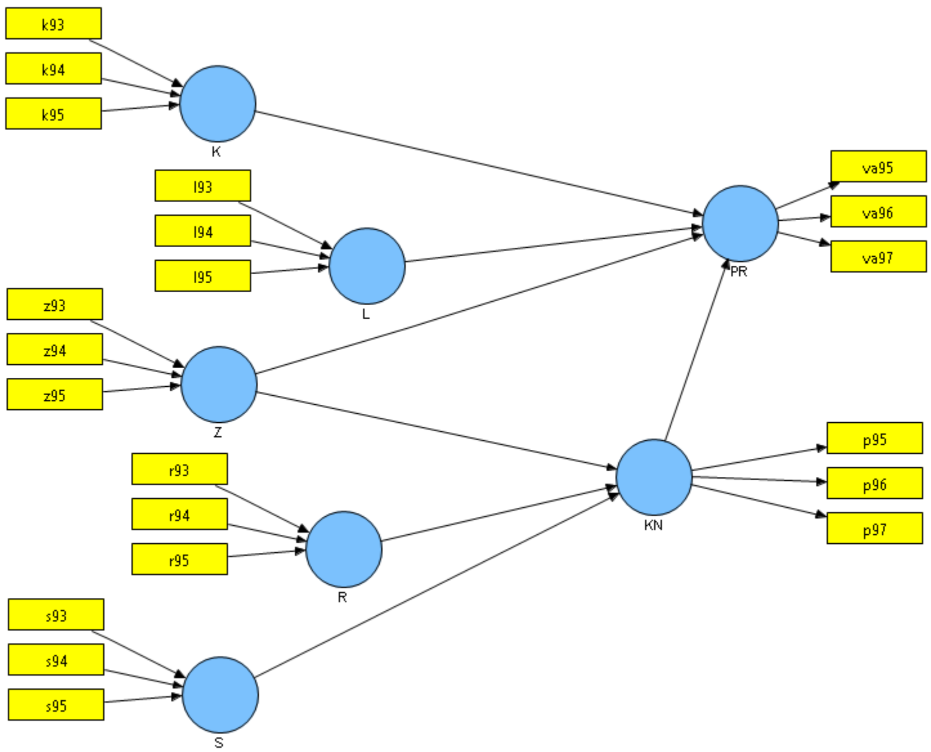


Figure 3.5: Multi-year Model

The indicator weights on the first-order latent variables are reported in Table 3.4. Among the inputs to the model, R&D is the only indicator that has significant weights for all years. Labour, non-IT capital, IT capital and the control variable Size exhibit a pattern where the weights on 1995 have higher t-values than 1993 and 1994. The input indicators with weights significant at the 5% or better level are: k_{94}, k_{95}, l_{95}, r_{93}, r_{94}, and r_{95}. The output indicators, $va_{95} \cdots va_{97}$ and $p_{95} \cdots p_{97}$, are significant at the 1% level in all years.

Although the number of indicators with insignificant (or negative) weights is curious at first glance, it is likely that multicollinearity is responsible. The indicators for both IT and non-IT capital are driven by the same slowly-evolving capital stocks in the firm; similarly, labour and size (sales) will also be driven by related underlying phenomena. A non-significant path weight or negative path weight in this context can result from a high correlation of, for example, r_{94} with the other indicators on the construct R, and should be seen as the effect of the indicator after having controlled for the other indicators' influence on the construct (Cenfetelli and Bassellier, 2009). The negative sign on the k_{94} and r_{94} indicators should not be interpreted as a deleterious effect of non-IT capital or R&D in 1994.

Indicator loadings are reported in Table 3.5. All indicators have high loadings of at least 0.88. In contrast to the single-year model, the indicator loadings in this model are less informative since they represent the correlations of three observations of an indicator of the same type to a formative construct. As such, we would expect high loading values since firms tend not to make fundamental changes to their factor mix (or, indeed, R&D budget) over a short period of time.

Support for construct validity is high, with AVE (Table 3.7) for the latent variables being higher than their respective correlation to all other variables (Table 3.8). The reliability of the PR and KN constructs is also very good, indicated by Cronbach's alpha (0.996 and 0.951, respectively) and R-squared for the constructs (95.5% and 65.2%, respectively).

The path coefficients are reported in Table 3.6. The paths between the input and output constructs are significant at the 1% level with the exception of the $Z \rightarrow KN$ path, which is significant at the 10% level. The path weights are generally consistent with prior regression-based estimates, showing that labour and non-IT capital have the largest influence on production output and that R&D is a strong predictor of Knowledge. However, there is a small, negative path weight on $Z \rightarrow PR$, significant at the 5% level, which was not expected.

Finally, the path coefficient from Knowledge to Productive Resources is 0.106 and significant at the 1% level. In relative terms it is not as influential as the conventional NPF inputs (0.379 for non-IT capital and 0.655 for labour), but it nonetheless confirms the hypothesized relationship from the KPF to the NPF.

3.6 Discussion

The estimation results from both variations of the model shed light on the relative contribution of IT capital vis-à-vis the traditional factors of production in an integrated model of knowledge and conventional production. Controlling for the simultaneous processes of knowledge-creation and conventional production, IT capital's role in direct production appears limited. However, we find some confirmation that IT capital works to increase the firm's stock of Knowledge, which in turn has an influence on the firm's overall Productive Resources. An advantage of this modeling approach is that the formative indicators composing the constructs are factors of production; thus, the model makes actionable predictions that can inform managerial decision-making.

In the single-year model, results for the five year period 1993 to 1997 roughly confirm earlier production function research. The indicator weights give an indication of the relative importance of a factor given the effect of the other factors, similar to regression coefficients. The results show strong contributions from K and L in the NPF (i.e. relating to the PR construct) and R in the KPF (the KN construct). The lack of statistically significant weights for the Z indicators in the neoclassical production function are not surprising given past research has found this effect to be small in comparison to K and L, and sensitive to industry and firm differences. Further, as mentioned earlier, we expect the indicator weights for Z to be attenuated due to the shared paths from Z to KN and PR.

In the knowledge production function, the indicator weights for Z and S on KN are positive and significant for 1993–1995 but not 1996–1997. This could indicate a shift in the relative importance of Z in the KPF in the post-1995 period. However, a broader perspective must acknowledge that insignificant weights are a relative measure, indicating that Z is unable to make a measurable contribution once the influence of the other factors has been accounted for. The strong indicator loadings suggest that all factors, including IT capital, are making strong absolute contributions to the latent constructs, and that these contributions are very consistent throughout the sample time frame.

A very important finding is the strong, positive and significant path coefficient from KN to PR across all five sample years. This indicates that knowledge is making a strong contribution to the production process, taking into account the use of all the formative indicators. The path weight is interpreted as the predicted increase in Productive Resources: for each unit increase in Knowledge, the predicted increase in PR is estimated between .649 to .707 standard deviations.

While the results across the years are roughly consistent, there are some anomalies. The PR construct estimates and R-squared remain stable, but the KN construct has low R-squared and a non-significant validity coefficient ($KN \rightarrow KN'$) for the 1996 sample year. Since these return to more typical values for 1997, it may be that anomalous events distorted the usual relationship between the indicators and the latent constructs related to Knowledge. Second, it is notable that the path weight on $Z \rightarrow PR$ is increasing over time, while for $Z \rightarrow KN$ it increases to 1994, decreases in 1995, and is not significant afterwards. This may indicate that, over time, firms emphasized the application of IT capital resources for production efficiencies while de-emphasizing the role of IT in knowledge-generation.

The multiple-year model estimation produces results that confirm the conventional factors of production for the NPF and R&D for the KPF, but do not support a role for IT capital. The path from KN to PR is positive and significant, and the formative constructs for labour (L) and non-IT capital (K) retain their strong explanatory power with significant paths on the output construct (PR). The path coefficient from Z to PR is significant at the 5% level, and has a small negative magnitude. This unexpected result may be due to multicollinearity with other inputs. In the latent variable correlation scores (Table 3.8), the Z construct has a high correlation with K and L (along with R and S in the KPF). However, the path coefficients from the IT capital construct (Z) to PR are not statistically significant. Further, the indicator weights for $Z_{93} \cdots Z_{95}$ are not significant, calling into question the validity of the Z construct. Overall, the results from this estimation suggest that, controlling for the other inputs, firms' accumulated IT resources from the 1993–1995 period did not explain the Productive Resources or Knowledge outputs in the 1995–1997 period.

3.6.1 Limitations and Future Research

The SEM techniques used here have not been widely applied to econometric data, and panel data in particular. One of the shortcomings of the SEM approach is that it does not allow one to control for firm fixed effects or industry effects. Both effects have featured prominently in the IT productivity literature. In past research using patent statistics, there is evidence of clear differences across industries in their propensity to patent (Griliches, 1990). This variation arises from the different stages in technological evolution, general industry maturity, level of competition and the degree to which patents are seen as effective methods for securing intellectual property rights. For similar reasons, propensity to patent varies from year to year, as well as the employment of various factors of production according to unobserved environmental changes, such as tax policy.

Future extensions to this line of inquiry could harness additional indicators of knowledge-creation in order to model KN. As mentioned earlier, one of the deficiencies of patents as indicators of knowledge is that they do not capture many less-formal forms of innovation—many of which may be supported extensively by IT. One potential indicator that has been recently explored is the number of trademarks obtained by a firm (Gao and Hitt, 2011). Firms often register new trademarks when adding new product lines as a result of innovations. If this and several other reliable measures of knowledge-accretion could be found, a more robust structural equation model could be estimated. Such analysis could provide further insights about the measurement and structure of the knowledge-creation process.

3.7 Conclusion

Information Technology has provided industry with ever-improving tools to capture, analyze and communicate information—-activities which are common to both production and innovation processes. Accordingly, as firms increase their IT investment, it is expected that their general productivity, as well as their returns to innovation, should improve. By applying structural equation modeling techniques we test a model that incorporates IT as a distinct form of capital and estimate its simultaneous contribution to both areas. Our results indicate this effect, for the 1993–1995 period, to be mediated through the knowledge production process. This provides qualified support for the proposition that IT creates value at intermediate stages of production. Future applications of this technique, using expanded data, will shed light on both the viability of the model and patterns of influence across different manufacturing industries.

Table 3.1: Summary statistics: entire sample

	Variable	Median	Std. Dev.	Min	Max
Y	Value-Added Output	1,614.28	7,143.52	256.95	62,882.66
K	Non-IT capital	102.71	682.32	8.71	6,620.23
L	Labour	954.14	3,646.02	106.40	31,916.19
Z	IT capital	23.71	67.12	0.66	463.81
R	R&D	94.70	973.73	4.49	8,359.77
S	Employees	20,100.00	77,706.95	2,335.00	750,000.00
P	Patent Citations	134.66	2,068.26	0.00	23,145.30

Balanced panel, 615 observations, 1993-1997.
All measures in millions of 1993 dollars, except Employees (in persons) and Patent Citations (in units).

Table 3.2: Single-Year Model: Path Weights

Indicator	Construct	1993	1994	1995	1996	1997
K	PR	0.305***	0.299***	0.349***	0.407***	0.363***
		(5.924)	(6.867)	(9.228)	(12.192)	(8.530)
L		0.754***	0.741***	0.696***	0.620***	0.676***
		(14.366)	(12.905)	(14.483)	(14.038)	(12.111)
Z		−0.029	−0.005	0.000	0.028	0.015
		(0.741)	(0.104)	(0.005)	(0.805)	(0.381)
Z	KN	0.443***	0.460***	0.267*	0.146	0.096
		(3.103)	(3.128)	(1.715)	(1.074)	(0.612)
R		0.987***	0.891***	0.940***	1.142***	1.045***
		(8.626)	(6.993)	(10.322)	(7.994)	(9.282)
S		−0.463***	−0.372***	−0.213	−0.380	−0.168
		(3.211)	(2.674)	(1.277)	(1.519)	(0.957)
	PR→PR’	0.978***	0.975***	0.978***	0.980***	0.976***
		(160.472)	(163.167)	(172.238)	(203.943)	(168.407)
	KN→KN’	0.785***	0.785***	0.795***	0.602***	0.750***
		(16.802)	(17.686)	(21.483)	(5.916)	(16.011)
	KN→PR	0.665***	0.700***	0.721***	0.623***	0.704***
		(10.704)	(12.476)	(12.573)	(8.769)	(12.814)
PR R^2		0.956	0.951	0.957	0.960	0.953
KN R^2		0.616	0.632	0.632	0.362	0.563

t-values in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table 3.3: Single-Year Model: Indicator Loadings

Indicator	Construct	1993	1994	1995	1996	1997
K	PR	0.920	0.913	0.914	0.924	0.907
L		0.987	0.986	0.979	0.969	0.974
Z		0.854	0.841	0.831	0.824	0.811
Z	KN	0.779	0.656	0.807	0.714	0.750
R		0.969	0.968	0.989	0.975	0.995
S		0.632	0.656	0.685	0.573	0.666

Table 3.4: Multi-year Model: Outer Weights

	K	KN	L	PR	R	S	Z
k93	0.926						
	(1.290)						
k94	-2.362^{**}						
	(2.025)						
k95	2.425^{***}						
	(4.076)						
l93			-0.420				
			(0.943)				
l94			0.604				
			(1.196)				
l95			0.812^{***}				
			(3.930)				
p95		0.404^{***}					
		(134.890)					
p96		0.296^{***}					
		(13.829)					

Continued on next page

Table 3.4 – continued from previous page

	K	KN	L	PR	R	S	Z
p97		0.368^{***}					
		(73.565)					
r93					2.657^{***}		
					(3.511)		
r94					-4.371^{***}		
					(4.097)		
r95					2.680^{***}		
					(3.623)		
s93						0.862	
						(0.524)	
s94						-0.876	
						(0.477)	
s95						1.017	
						(1.241)	
va95				0.338^{***}			
				(734.692)			
va96				0.336^{***}			
				(1502.295)			

Continued on next page

Table 3.4 – continued from previous page

	K	KN	L	PR	R	S	Z
va97				0.329***			
				(557.592)			
z93							0.142
							(0.343)
z94							0.325
							(0.408)
z95							0.539
							(0.998)

t-values in parentheses.

*** p<0.01, ** p<0.05, * p<0.1

Table 3.5: Multi-year Model: Outer Loadings

	K	KN	L	PR	R	S	Z
k93	0.957						
k94	0.969						
k95	0.990						
l93			0.970				
l94			0.985				
l95			0.998				
p95		0.959					
p96		0.878					
p97		0.953					
r93					0.957		
r94					0.938		
r95					0.954		
s93						0.960	
s94						0.960	
s95						0.996	
va95				0.994			
va96				0.997			
va97				0.993			
z93							0.977
z94							0.995
z95							0.995

Table 3.6: Multi-year Model: Path Coefficients

	KN	PR
K		0.379***
		(9.151)
KN		0.106***
		(3.693)
L		0.655***
		(11.297)
R	0.777***	
	(10.114)	
S	−0.208	
	(1.458)	
Z	0.222*	−0.086**
	(1.661)	(2.166)

t-values in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table 3.7: Multi-year Model: Reliability Measures

	AVE	R Squared	Cronbach's Alpha
KN	0.867	0.652	0.923
PR	0.990	0.955	0.995

Table 3.8: Multi-year Model: Latent Variable Correlations

	K	KN	L	PR	R	S	Z
K	1.000						
KN	0.491	1.000					
L	0.817	0.507	1.000				
PR	0.901	0.572	0.948	1.000			
R	0.687	0.798	0.691	0.770	1.000		
S	0.811	0.506	0.986	0.934	0.680	1.000	
Z	0.757	0.615	0.820	0.804	0.729	0.834	1.000

Chapter 4

Productivity, IT and Innovation

4.1 Introduction

The modern pursuit of innovation depends heavily upon information technology. Most of the new things and new ways of doing things we now encounter, if not entirely consisting of IT, have been shaped by it in some way. Advances in computing power and applications are not only evident in everything from smartphones to cars, but are inextricably involved in the development of new drugs, logistics services and mass customization processes. Yet we do not fully understand the relationship between IT and innovation. In its ubiquity, IT plays multiple roles: input to innovation, output of innovation (as a class of innovative products), and enabler of innovation-driven changes. Making the distinction between these three roles is critical to understanding the economic impacts of IT.

Some research has posited IT as an antecedent to innovation, arguing firms use IT to innovate more quickly by automating, communicating and managing data, and enabling new methods of exploration (Thomke, 1998). Conversely, as the output of the innovation process, IT is seen as a new form of capital, displacing labour and complementing traditional capital (Chwelos et al., 2010). Another challenge facing researchers is determining which industries are originating such technical advances, and which are realizing the benefits (Cheng and Nault, 2007, 2011). Finally, the advent of e-commerce, e-procurement and e-services has enabled firms to take advantage of IT-enabled process innovations, such as mass customization, global supply chains, and vendor-managed inventories.

While IT-driven innovation has gained prominence, firms continue to invest in established research and development (R&D) approaches to innovation with the goal of improving their processes and products. As they implement such innovations in their manufacturing operations, we may expect new opportunities will arise for IT to contribute to the firm's productivity. Consider a firm that develops an innovative production process: once implemented, it must reallocate some of its factor inputs (labour, non-IT capital, IT capital) to accommodate the new process. At this stage, we hypothesize that IT plays an enabling role in the implementation of the process innovation. IT has been shown to improve flexibility of production within an existing manufacturing setup (Bartel, Ichniowski, and Shaw, 2007), and to enable new possibilities for production scaling and new product introduction (Gao and Hitt, 2011) that could facilitate the successful implementation of a process innovation. If this is the case, we should see that the impact of R&D on total factor productivity (TFP) has some relationship with the productivity of IT, and that IT and R&D *together* have an indirect relationship with the other inputs.

4.2 Innovation and IT Productivity Research

4.2.1 IT Productivity Literature

Existing IT productivity research has established the role of IT as a distinct form of capital investment. Evidence from empirical studies suggests IT contributes to productivity by substituting for less-efficient labour and, in some cases, capital (Dewan and Min, 1997; Mohammad, Zhang, Cheng, and Nault, 2009; Chwelos et al., 2010). Using industry-level data over a 30-year period, Hu and Quan (2005) addressed the question of causality: whether IT investments lead to improved industry performance, or vice-versa. Employing the Granger causality model they found evidence that IT investment does lead to productivity growth (measured in GDP per employee) in most industries. Further, the authors found evidence of feedback from these productivity gains to subsequent increases in IT spending.

While there is a clear role for IT in automation, it is the capacity for transforming business that has offered the greatest ongoing potential for productivity gains. David (1990) proposed that IT is a general purpose technology (GPT) which, like electricity or the steam engine, creates the conditions for widespread technological change due to its flexibility and stimulation of innovation. Empirical research in this arena has investigated the ability of IT to enable new organizational forms and improve the quality and variety of a firm's products (Brynjolfsson and Hitt, 2000; Brynjolfsson et al., 2002; Gao and Hitt, 2011). One way of measuring this transformative capacity from a production theory perspective is to consider the augmented productive capacity of the traditional labour and capital inputs owing to IT capital. Using this approach, Mittal and Nault (2009) found evidence of such scaling in a study of US manufacturing industries from 1953–2000, with a greater effect over time and in the most IT-intensive industries.

In sum, researchers have made the case that IT capital has both a direct effect, altering the mix of factor inputs, and an indirect effect, by augmenting the other inputs to production by changing the way the firm does business (Mittal and Nault, 2009). Measuring the extent of this success and the mechanisms by which it is manifested has, however, proven difficult. Notably, the rapid price declines and quality improvements in IT lead to potential productivity mismeasurement across industries (Cheng and Nault, 2007; Han, Kauffman, and Nault, 2010). In this research, we propose to further disentangle the direct and indirect effects of IT and how these relate to innovation and productivity. Despite a great deal of interest in both IT business value and innovation, few researchers have focused on this interaction. In the context of patent output, Kleis, Chwelos, Ramirez, and Cockburn (2011) found no evidence of an interaction between R&D and IT at the firm level. However, Bardhan, Krishnan, and Lin (2010) found an interaction effect for R&D and IT that relates positively to Tobin's q. To the best of our knowledge, there has been no prior research on the R&D-IT interaction relationship in the context of productivity at the industry level.

4.2.2 Innovation and Productivity

Considerable research has attempted to model and measure the relationship between R&D and productivity. Griliches (1988, ch.15) summarized the general approach taken by researchers: extending the production function approach to include research and development activities.

$$
\begin{aligned}
Q &= Tf(C, L) \\
T &= g(K, O) \\
K &= \sum w_i R_{t-i}
\end{aligned}
$$

> "Where Q is output, C and L are capital and labour input respectively, T is the current level of (average) technological accomplishment (total factor productivity), K is a measure of accumulated and still productive (social or private) research capital ("knowledge"), O represents other forces affecting productivity, R_t measures the real gross investment in research in period t, and w_i's connect the levels of past research to the current state of knowledge." (Griliches, 1988, p.246)

Let us consider each of Griliches' equations as a way of organizing the literature. The first equation concerns production and technological change: how does productivity advance beyond increased quantities of the inputs? The answer is through technology. In the growth accounting framework, the factor-neutral technology constant, in essence, sets an intercept for the firm's ability to efficiently transform inputs into outputs. It is a residual because it is not modeled explicitly; it is the leftover portion of productivity that cannot be explained by changes in the quantities of the factor inputs.

Extensive research in economics has sought to better understand the measurement and sources of changes in production technology. Researchers have long argued that knowledge drives the technological state of firms (Griliches, 1994; Nelson, 1982). However, there are many difficulties with this framework, not the least of which is the measurement of innovation inputs and outputs (spillovers, R&D, patents); the appropriability of innovations (Griliches, 1979); and the measurement of quality of output (Hulten, 1992; Siegel, 1994).

In the foregoing framework, R&D enters the production function via a system of equations, and is not considered a factor input in the classical sense of capital and labour (the first equation). Although some researchers have taken the approach that R&D should be treated as a stock of knowledge that is a factor input, beginning with Griliches (1979) and continuing to the present (for reviews, see Mairesse and Sassenou (1991); Hall, Mairesse, and Mohnen (2009)), other researchers have included knowledge stock as part of the technology residual (Lichtenberg and Siegel, 1991; Siegel, 1997).

The accumulation of R&D expenditures into a stock of knowledge is, however, common to these approaches. There are three main reasons why it is appropriate to construct such a measure. First, stock is the more appropriate concept in production because the innovation process exploits an accumulated stock of knowledge. This knowledge stock changes over time in response to many internal and external factors, of which R&D investments are the most directly observable. A given year's R&D expenditure can be compared to annual additions to capital stock, which may renew or expand some portion of the capacity of capital to generate output, but do not explain the overall capital capacity of the firm. There are several notable differences between capital and R&D stocks, such as resale value, observability, rate of obsolescence, spillovers and appropriability. One difference relevant to this discussion is that, while the relationship between capital and production is largely deterministic, innovation is a stochastic process. There is no certainty that a stock of R&D knowledge will produce an expected output of viable innovations. The stock produces some proportion of innovations that are viable, but since the "effective" portions of the R&D stock cannot be isolated, we must consider the entire R&D stock as a kind of productive stock. Second, the knowledge stock concept facilitates the lag structure linking R&D spending with innovation output. Intuitively, there is some delay between the R&D investment and its effect on performance. However, there is good reason to believe this lag varies across industry and time. A more general way to incorporate heterogeneity in the lag is to represent the total knowledge available to an industry using an R&D stock variable. Finally, the knowledge stock concept is germane to industry-level analysis due to intra-industry knowledge spillovers. For example, a firm may create an innovation, implement it internally, and later decide to license it to another firm in the same industry. There may thus be a lag between R&D expenditure, implementation and spillover. A stock will more appropriately make this R&D investment available to the eventual output of the industry than would a flow measure of R&D. In addition, research suggests the receiving firm's R&D expenditure is complementary to spillovers by creating absorptive capacity (Cohen and Levinthal, 1990; Cockburn and Henderson, 1998).

In Griliches' second equation, we suppose that T (technology) originates from knowledge and "other factors." A persistent feature of this literature has concerned research spillovers and how they add to the available stock of knowledge. Knowledge and R&D spillovers, in particular, are of interest (Scherer, 1982; Griliches and Lichtenberg, 1984; Griliches, 1992). Although measurement of spillover effects has proven difficult, researchers have suggested several factors that are believed to influence this process. The importance of industry concentration and R&D intensity is a primary factor (Acs and Audretsch, 1991, ch.1). Another important concept is that the flow of spillovers is proportional to the flow of trade between industries (Los and Verspagen, 2002). Using either the Input-Output matrix or by analyzing technology flow through patent-tracing, "borrowed R&D" comes from purchases, and accounts for more of the return-to-R&D than the industry's own R&D (Nadiri, 1993). While there is evidence of some substitution effect of spillover R&D for own R&D, the findings generally indicate firms need to spend on R&D in order to appropriate spillovers, as mentioned in the previous paragraph.

Finally, in the third equation, Griliches proposes that the genesis of knowledge is the R&D process. Many researchers have addressed the construction of measures for the stock of knowledge and its rate of accretion/depreciation.[20] The R&D “knowledge production function” (Griliches, 1979) has featured prominently in this literature. However, even “private” R&D can be subject to external characteristics and influences, through contract R&D and government-funded research. Howells (1999) provides some examples of the “push” and “pull” factors that influence a firm’s choice to outsource R&D: complexity of research (i.e. costs and risks); government policies (including government–funded R&D and defense–related R&D) which likely vary by industry. Pisano (1990) studies the decision to do R&D internally or contract out in the context of Transaction Cost theory. With a sample of 92 biotechnology R&D projects, the probability of contracting R&D is modeled as a function of variables such as size, rivalry and asset specificity.

4.2.3 Summary

In this research we will extend the IT productivity research, drawing on the innovation literature, by investigating the indirect effects of R&D and IT on output (via non-IT capital and labour). In doing so, we will be able to offer evidence of the presence and size of these effects and compare them to the estimates from traditional approaches to productivity analysis. We will estimate if these indirect effects exist, to what degree they vary over time and across industries, and to what extent they are salient. We will distinguish between the indirect effects of R&D and IT, and estimate the interaction of the two, which represents the synergistic effect of R&D in concert with IT upon the traditional factor inputs. In addition to our analysis of this aspect of IT value, we offer new insights into the influence of innovation and how these translate to productivity at the industry level.

4.3 Modeling the R&D–IT Productivity Link

We hypothesize that a knowledge stock of process-oriented R&D improves the efficiency of the factor inputs by means of an indirect effect. This can

[20] Hall (2005) explores some of the difficulties in estimating the true rate of R&D depreciation.

be expressed in the following generalized form of a production function:

$$Y = f(h_1K, h_2L, h_3Z),$$

where output, Y, is a function of non-IT capital K, labour L, and IT capital Z, and the terms h_1, h_2 and h_3 represent the indirect (augmenting) effects on the inputs. In other words, h_3, for instance, can be characterized as a scaling factor applied to Z, so that h_3Z represents the effective quantity available for production.

If we assume a standard Cobb-Douglas functional form, we have:

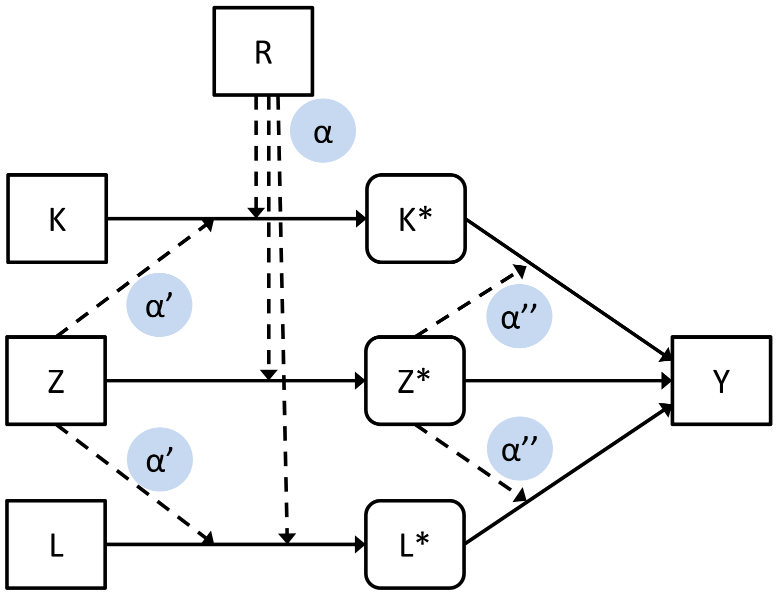
$$Y = A[h_1K]^{\beta_1}[h_2L]^{\beta_2}[h_3Z]^{\beta_3}, \tag{4.1}$$

where A is the technological change parameter, also known as total factor productivity. Subscripts for industry and time are omitted for ease of exposition.

We hypothesize that h_3, the indirect effect on Z, is a function of the knowledge stock of process-oriented R&D, denoted R below. However, past research offers evidence that Z creates indirect effects on K and L. Building on this, we model the indirect effects on K and L (h_1 and h_2, respectively) as being driven by both R&D and the *R&D-augmented*-IT capital. This structure allows R&D stock to enter the production function in a novel way. Since it is not directly involved in production, modeling R&D as a factor input can be problematic. Rather, in this model we conceive of it as a parallel business endeavour in which the industry invests and receives a dividend in the form of a scaling factor on the industry's effective quantity of the factor inputs. While some innovation may be applied directly to non-IT capital and labour, we propose that the capabilities of IT to improve input efficiencies also provide a mechanism by which process innovations may be implemented in production. In other words, the efficiency improvements from R&D on IT capital "flow through" to improve the efficiency of non-IT capital and labour, along with the conventional indirect effect of R&D on non-IT capital and labour.

Figure 4.1 presents this diagrammatically, with the dashed lines representing indirect effects (labeled α, α', α'', to be introduced below) and solid lines representing direct effects. The measured amounts of K, L, Z, R and Y are depicted in square boxes, and the effective amounts of K^*, L^* and Z^* in rounded boxes.

Figure 4.1: Indirect Effects Model of R&D



The structure of the scaling factors may be expressed as:

$$
\begin{aligned}
h_1 &= f_1(R, Z(R)) \\
h_2 &= f_2(R, Z(R)) \\
h_3 &= f_3(R).
\end{aligned}
$$

Now, let us define these terms with specific exponential forms:

$$\begin{aligned} h_1 &= e^{(\alpha_1 R + (\alpha_1^{'} + \alpha_1^{''} R) Z)} \\ h_2 &= e^{(\alpha_2 R + (\alpha_2^{'} + \alpha_2^{''} R) Z)} \\ h_3 &= e^{\alpha_3 R}. \end{aligned}$$

With these equations, we model the indirect effect of R&D through IT as a multiplicative effect. Although IT and R&D each drive productivity improvements in labour and non-IT capital, they produce an additional effect in combination.

We now adopt the assumption that the indirect effect of Z acts upon K and L in the same way, i.e. for a given investment in IT capital, the effective amounts of non-IT capital and labour are changed by the same factor. This generalization allows us to estimate a general effect of Z (and, separately, ZR) on both K and L. Although the magnitudes of the indirect effects may be different, this is not essential to our research question. Let us impose the following constraints, based on the assumption that the indirect effects of $Z(R)$ and Z are independent of K and L : $\alpha_1^{'} = \alpha_2^{'} = \alpha^{'}$ and $\alpha_1^{''} = \alpha_2^{''} = \alpha^{''}$. Now,

$$h_1 = e^{(\alpha_1 R + (\alpha^{'} + \alpha^{''} R) Z)} \tag{4.2}$$

$$h_2 = e^{(\alpha_2 R + (\alpha^{'} + \alpha^{''} R) Z)} \tag{4.3}$$

$$h_3 = e^{\alpha_3 R}. \tag{4.4}$$

Substituting the terms into the Cobb-Douglas production function (equation (4.1)), we have:

$$Y = A[e^{(\alpha_1 R + \alpha^{'} Z + \alpha^{''} ZR)} K]^{\beta_1} [e^{(\alpha_2 R + \alpha^{'} Z + \alpha^{''} ZR)} L]^{\beta_2} [e^{\alpha_3 R} Z]^{\beta 3}.$$

Taking the log of both sides and collecting the terms, we have:

$$\begin{aligned} y &= a + \beta_1 k + \beta_2 l + \beta_3 z + (\beta_1\alpha_1 + \beta_2\alpha_2 + \beta_3\alpha_3)R \\ &\quad + \alpha^{'}(\beta_1 + \beta_2)Z + \alpha^{''}(\beta_1 + \beta_2)ZR \\ &= a + \beta_1 k + \beta_2 l + \beta_3 z + \theta_1 R + \theta_2 Z + \theta_3 ZR, \end{aligned} \tag{4.5}$$

where $\theta_1 = \beta_1\alpha_1 + \beta_2\alpha_2 + \beta_3\alpha_3$, $\theta_2 = \alpha^{'}(\beta_1 + \beta_2)$, and $\theta_3 = \alpha^{''}(\beta_1 + \beta_2)$. Lowercase letters are used for variables where the natural log has been taken. Note the estimates of $\alpha^{'}$ and $\alpha^{''}$ are recoverable, but α_1, α_2 and α_3 are not separately recoverable. Thus we will be able to estimate the indirect effect of R on K, L and Z collectively, and the indirect effect of Z and ZR on K and L, collectively.[21]

4.3.1 Interpreting the Coefficients

A useful property of the Cobb-Douglas production function is that the main-effect coefficients, β_1 and β_2, are interpreted as the output elasticities for each input. Thus, β_1 is the expected percentage change in Y for a percentage change in K, and β_2 is the equivalent for L. However, our main interest in this research is whether the estimates of the indirect effects of IT capital and R&D, and their interaction, are significant and positive. By specifying the production function with both direct and indirect effects, we are able to separately estimate the direct contribution of IT capital to production, along with its indirect impact on labour and non-IT capital. We expect this to be positive in confirmation of the existing literature (Mittal and Nault, 2009). We also expect to find a positive indirect effect of R&D. Although θ_1 and θ_2 represent a part of these effects, the overall indirect effect of R and Z also depends on ZR. Therefore we compute the estimate of the partial elasticity for R, and the output elasticity of Z (which will include all direct

[21]We may impose $\alpha_1 = \alpha_2 = \alpha_3 = \alpha$, i.e. the indirect effect of R&D is the same on non-IT capital, IT capital and labour. Under this restriction, we could identify all the parameters in the model, as $\theta_1 = \alpha(\beta_1 + \beta_2 + \beta_3)$. However, it is reasonable to expect that Z and R have differing indirect effects on K and L (i.e. $h_1 \neq h_2$), and comparing their magnitudes will be a useful exercise toward understanding their roles in actualizing process innovations.

and indirect effects of IT capital).

The partial elasticity of R is derived from the production function (equation (4.5)), and is given by:

$$\begin{aligned} \frac{\partial \ln Y}{\partial R} &= (\beta_1\alpha_1 + \beta_2\alpha_2 + \beta_3\alpha_3) + \alpha^{''}(\beta_1 + \beta_2)Z \\ &= \theta_1 + \theta_3 Z. \end{aligned} \tag{4.6}$$

Since R is not in log form, the partial elasticity expresses the percentage change in Y given a *unit* change in R (in our estimates, the unit is billions of dollars). Since the elasticity is dependent on the level of Z, it would be evaluated at the mean or median value of Z.

We estimate the output elasticity of IT, given by:

$$\frac{\partial \ln Y}{\partial \ln Z} = \beta_3 + \theta_2 Z + \theta_3 ZR,$$

which expresses the expected percentage change in Y given a percentage change in Z, including both direct and indirect effects. This may be compared to the output elasticity of Z in the Cobb-Douglas specification to see the change induced by adding indirect effects to the estimation.

Finally, we test three additional hypotheses regarding the statistical significance of the indirect coefficients. First, we want to know if the indirect effect of Z is the same as ZR. If this is true, then the overall indirect effect of $Z(R)$ upon both K and L is the same. To determine this, we test for equality of the indirect effects of Z on K and L: under the null, $\alpha^{'} = \alpha^{''} = 0$ (or $\theta_2 = \theta_3$). If either coefficient is not equal to 0, then we also reject the hypothesis that Z has no indirect effects on K and L. The third hypothesis tests if all the indirect effects (arising from both R and $Z(R)$) are the same upon all inputs K, L and Z. To determine this, we specify the test as follows: under the null, $\theta_1 = \theta_2 = \theta_3$.

4.4 Data

Using data from U.S. government agencies, we construct two datasets: 1987–1998 and 1998–2005, based on the Standard Industrial Classification (SIC) and North American Industry Classification System (NAICS) standards, respectively. Due to the change from SIC to NAICS in 1998, the data are not directly comparable across eras. Both datasets are constructed in a similar manner. We obtain multifactor productivity (MFP) data from the Bureau of Labor Statistics (BLS).[22] Research and development expenditures are obtained from the National Science Foundation (NSF) survey of industrial research and development (National Science Foundation, 2009), which reports domestic private-sector R&D expenditures (from both company and federal government sources) at the industry level. All data is publicly available on the respective agencies' websites. Summary statistics are reported in Table 4.1.

4.4.1 Dataset I: 1987–1998

Dataset I is a balanced panel of 276 observations (23 manufacturing industries) for the years 1987–1998. The following paragraphs explain the construction of this dataset.

MFP Data The MFP data is the same as that used in Cheng and Nault (2007, 2011), which was originally collected by the BLS. The BLS tracks gross output, capital, labour hours, and intermediate inputs (energy, materials and services) for all domestic industries. For the period 1987–1999, there are 140 industries in the data at the 3-digit SIC level. The variables consist of output and intermediate purchases in millions of nominal dollars. Capital equipment is in millions of nominal dollars, while labour is measured in millions of hours.

In order to prepare the data for analysis we perform several manipulations. First, we drop any observation where any data element is missing.

[22] In Dataset II we obtain data on value added from the Bureau of Economic Analysis (BEA), but the remaining components of the MFP data are obtained from the BLS.

Second, our research is primarily concerned with manufacturing industries, so we drop service and government industries. Third, to create separate amounts of IT capital stock and non-IT capital stock, we identify five asset types that comprise IT and related equipment(computers & related equipment, office equipment, communication, instruments, photocopy & related equipment) in the capital stock data, which are expressed in constant 1987 dollars. The remaining 25 categories become non-IT capital stock.[23] Fourth, value added is calculated as gross output less intermediate inputs, each of which is deflated to constant 1987 dollars (using the output and intermediate input deflators, respectively).

R&D Data The NSF Survey of Industrial Research and Development has been an ongoing effort since the late 1950s to provide data about corporate R&D for policymakers, researchers and industry. The survey is based on a sample that seeks to include all for-profit companies that perform R&D, and automatically includes all companies with 1000 or more employees (the 1987 sample size was 154,000 companies). The NSF updates the sample annually to ensure that firms previously excluded on the basis of not having R&D activity are included if they begin to undertake R&D. The survey is administered by the Bureau of the Census and firms are required by law to respond. Individual surveys are sent to establishments (i.e. prominent research labs or manufacturing installations) but are aggregated by company, and the company is categorized into the SIC designation based on the type of business which had the highest dollar share of payroll. Additional details of the NSF's sampling technique are available on its website.[24]

For the purposes of measuring company and federally-funded R&D, the data covers the period 1967–1998. It is organized by SIC; however, individual SIC codes at the 2-digit and 3-digit level are sometimes combined into aggregated categories. The NSF table H-3 reports company-funded R&D

[23]The BLS classifies productive capital into five classes: structures, equipment, land, rental residential capital and inventories. In our definition of capital and non-IT capital, the last three classes are excluded.

[24]http://www.nsf.gov/statistics/srvyindustry/

spending at U.S. locations.[25] The amount of R&D is reported in millions of nominal dollars. We use the GDP deflator to convert these amounts to constant 1987 dollars.[26]

Since the R&D data are aggregated at different SIC levels, some adjustments are required in order to merge with the MFP and other data from BLS. Of the 31 NSF industry groups, 4 are at the 2-digit level, 5 are combinations of 2-digit industries, 6 are at the 3-digit level, and 10 are combinations of 3-digit industries. To preserve the largest number of observations, we use this scheme as the basis for aggregating MFP data. Wherever possible, we match 3-digit industries from both sources. Where one source aggregates several 3-digit industries, or reports only the 2-digit industry, we replicate this aggregation in the other source. Finally, we drop any industries missing an observation in either the first or last year. When combined with MFP data, this yields an integrated data set of 276 observations with 23 industries over the years 1987–1998.[27] Some industries do not have R&D data for certain years due to NSF confidentiality policies. One industry is missing an observation for 1989, and 5 industries are missing observations in 1991. We use linear interpolation to generate these observations. Table 4.2 lists the industry groups by which the dataset is organized.

The NSF data provides the annual industry expenditure on R&D. However, as previously discussed, the concept of a stock of R&D-related knowledge is more useful when modeling the innovation process and its relationship with productivity. Most researchers have used the perpetual inventory method with an assumed depreciation rate (δ) of 15% to create a measure

[25] Although NSF collects and reports separate amounts for total (including federal funds) and contract R&D performed outside the firm, many of these observations are obscured in the public reports to avoid identifying individual firm responses. This would reduce our potential sample size to only 132 complete observations.

[26] In an attempt to improve the measurement of R&D investment over time, the BEA has produced an R&D input price index for the years 1987–1999. However, the correlation to the GDP index is 0.9995. Given this close correlation and the unavailability of the R&D index for our second dataset, we chose to use the GDP deflator.

[27] It is also important that R&D data is available from 1982 for all 23 industries, which we later use construct our measure of R&D stock.

of R&D stock. The initial amount of R&D stock is calculated as:

$$Stock_{t_0} = \frac{Expenditure_{t_0}}{(1-\delta)},$$

where $Expenditure_{t_0}$ is the expenditure in a base year. For each subsequent year, the stock is computed:

$$Stock_t = (1-\delta)Stock_{t-1} + Expenditure_t.$$

Ideally, one should begin calculating the stock several years prior to the first year for which the measure is used in analysis in order to smooth out any unusual movements in the annual R&D expense. For Dataset I, we begin with the R&D data for 1982 and apply this formula to compute a stock for 1983–1998.

4.4.2 Dataset II: 1998–2005

Dataset II consists of a balanced panel of 200 observations (25 industries). It is a continuation of the first dataset, but differs in two material ways. The industries, although comparable in number, are aggregated into larger groups. It also includes non-manufacturing industries, such as Information and Finance, Insurance and Real Estate.

MFP Data Other than being organized by NAICS, the MFP data and their construction are comparable to those in the first dataset. Labour hours, IT and non-IT capital stock are obtained from the BLS. To calculate value added, we subtract intermediate inputs from gross output by 3-digit NAICS industry. These amounts, expressed in constant 2005 dollars, are obtained from the BEA.

R&D Data As with Dataset I, we obtain R&D data from the NSF's annual R&D in Industry reports that present the results of their Survey of Industrial Research and Development. Beginning with 1999, the industries are grouped by NAICS. In order to facilitate the transition to NAICS, the

NSF also provides bridge tables for 1997 and 1998. Thus, NAICS-based data are available from 1997 to 2008. We use 1997 as the base year for our R&D stock calculation. We omit 2006–2008 due to changes made by the NSF to the industry aggregations in 2006 and again in 2008, which result in the data not being comparable to earlier years.

As with the earlier SIC-based reports, the NSF does not report R&D on a consistent level of NAICS industry aggregation. Some industries are aggregated to the 2-digit level, while others are broken down to the 4-digit level. To facilitate matching to the MFP data, we choose the 3-digit level where available, using the 2-digit level where necessary. Table 4.3 lists the industry groups by which the dataset is organized.

We obtain the amount of company- and non-federally-funded R&D from NSF table A-7 for 1998, NSF table 12 for 1999–2003, and NSF table 1 for 2004–2005. These tables report R&D performed within the company's U.S. facilities, funded (predominantly) by the company or outside organizations excluding the federal government. Because company-funded R&D is unavailable for two industries (NAICS 312 and 324) in 1999, the missing observations are generated by linear interpolation. Two major industries, wholesale trade and retail trade, are unavailable prior to 2002 so we exclude these industries. In all cases, the amount of R&D is reported in millions of nominal dollars. We use the BEA GDP deflator (as published August 20, 2009) to convert these amounts to constant 2005 dollars.

Lastly, we create the R&D stock variable as before, with the exception of the fixed rate of depreciation. The standard 15% rate is used for most industries but for three industries in particular, the BEA recommends different rates on the basis of a review of the literature (Mead, 2007). These depreciation rates are given in Table 4.3.

4.5 Analysis

We present our analysis in three parts. We first describe the estimation strategies and econometric adjustments appropriate to industry-level panel data. Next, we estimate several variations of our research model using three panel regression approaches. Finally, we check for the robustness of our main results by exploring subsamples and decompositions of our data, and by testing an alternative model.

Except where noted, the dependent variable in all estimations is the natural log of value added. The variables are first rescaled to billions of dollars to improve the readability of the estimates. Where variables are shown in lowercase, the natural log has been taken. In all estimations, we assume a multiplicative error structure on the production function (in its log form).

4.5.1 Econometric Adjustments

Ordinary least-squares (OLS) regression techniques are sensitive to a number of violations of underlying assumptions with respect to error structure, which can bias the estimation results. In the present study we address the violation of the assumption of independent, identically distributed (IID) errors. Since our data includes repeated observations of the same industries through time, we expect time-invariant unobservables exist and require corresponding adjustment in our estimations. Two common approaches for panel regressions are fixed effects and random effects. Fixed effects control for a portion of the error that does not vary within each industry over time by separating the industry-specific portion of the error term from the total error. The random effects panel estimator uses a matrix-weighted average of the fixed-effects and between-effects panel model estimates.

While the IT productivity literature has featured fixed effects in a number of studies, the approach is not without its drawbacks. By removing time-invariant industry unobservables, a potentially large share of the overall variation within panels is eliminated, leaving relatively little for the model variables to explain. This can also reduce the impact of variables that vary only slightly through time, which is often the case with R&D spending. Recent studies using substantially similar data (Cheng and Nault, 2011; Han et al., 2010) have also found problems with fixed effects. This is the case in our attempts to estimate the model using fixed effects, which produce implausible coefficient estimates and low R^2. A less restrictive approach is to use coarser fixed effects at the sector level. We use a 1.5-digit SIC scheme of four manufacturing sector-level controls, but find it fails to produce reasonable estimates with our Dataset I. Similarly, using MFP data similar to our Dataset II, Han et al. (2010) found that using sector-level fixed effects (classifying their 60 industries into 13 sectors) inflated the standard errors of the estimate for IT capital compared to other estimation strategies. Cheng and Nault (2011) had more success with sectoral dummies, albeit with a much larger sample of industries. Using two different sectoral dummies schemes they found similar results to their baseline regression (using generalized least squares (GLS)).

We explore a number of alternatives to the fixed effects approach to controlling for industry-level unobservables: random effects, panel-corrected OLS and GLS estimators. The random effects estimator assumes the industry-specific component of the error term arises from a random process that is not correlated with the regressors. An advantage of this approach over the fixed effects approach is that it can accommodate time-invariant unobservables, as well as cross-sectional unobservables. One drawback is that it is not possible to control for heteroskedastic error structures using this approach (a concern which we discuss in further detail below).

While not directly intended to address unobservables, the GLS estimator can be used with several techniques to account for problems that may arise from cross-sectional time series data, including unobservables. The adjustment for heteroskedastic panels addresses the panels having different error variance. Autocorrelation can also be a problem, which is addressed by estimating a portion of the error term arising from serial correlation within panels. Allowing estimates of the autocorrelation parameter to vary across panels incorporates some of the industry-level unobservable factors insofar as they contribute to autocorrelation.

Dataset I tests Heteroskedasticity and autocorrelation are common in panel data, and prior literature in this area gives us reason to suspect there may be autocorrelation patterns in the error term. We use the likelihood ratio (LR) test to understand if heteroskedasticity is present in our data. The LR test compares the log-likelihood of the model under the assumptions of heteroskedastic and homoskedastic errors. The null hypothesis is that a nonhetereoskedastic model is nested in a heteroskedastic model. The test statistic (χ^2=444.76) is highly significant, so we reject the null hypothesis. An alternative test is the modified Wald test for groupwise heteroskedasticity in time-series feasible GLS model (implemented in the Stata 11 software package as `xttest3`). Under the null hypothesis, $\sigma_i^2 = \sigma^2$ for all i. The test statistic (χ^2 =41873.53), is highly significant at any reasonable level. Thus we conclude there is strong evidence for heteroskedasticity in the sample and we must correct for it.

We also expect autocorrelation to be present, since some of our measures rely on smoothing procedures (price deflators and R&D knowledge stock) and because industry reactions to economic shocks may correlate through time. To test for first-order autocorrelation of errors (AR(1)), we use the Wooldridge test. Since the test is a post-estimation procedure we run it following our estimation of both the Cobb-Douglas and full model specifications using GLS. Under the null hypothesis, there is no first-order autocorrelation. The F-statistics are F(1,22)=40.951 and 57.612, respectively, with p-values well below 0.001. Thus we may reject the null hypothesis of no autocorrelation.

We also test for whether the amount of autocorrelation is widely distributed across our industries. If it is, the panel-specific AR(1) (denoted PSAR1) adjustment can allow for this by estimating a different AR(1) ρ coefficient each group (industry). To determine if this is appropriate we perform a Likelihood Ratio test of PSAR1 vs. AR1. This tests whether, under H_0, $\rho_1 = \ldots = \rho_{25}$; under H_1, the correlations are not equal. The difference in likelihoods (PSAR1-AR1) for the indirect effects model in equation (4.5) is 144.863. The critical value of χ^2 with (i-1=22) d.f. = 40.289 at 1% significance. As an additional confirmation, we test the Cobb-Douglas specification and find the difference in likelihoods is 138.052. Therefore we reject the null hypothesis that the ρ coefficients are equal to one another, and conclude that there is evidence to support using a panel-specific autocorrelation parameter.

Dataset II tests The foregoing tests are repeated on Dataset II. The LR test statistic (χ^2=364.54) is highly significant, so we reject the null hypothesis. The modified Wald test statistic (χ^2 =39658.04) is highly significant at any reasonable level. Thus we conclude that, as with Dataset I, there is strong evidence for heteroskedasticity in Dataset II and we must correct for it. The Wooldridge test for first-order autocorrelation of errors returns F-statistics of $F(1,24)=28.131$ and 27.428, respectively, with p-values well below 0.001. Thus we may reject the null hypothesis of no autocorrelation. The difference in likelihoods for the LR test of PSAR1 vs. AR1 with the indirect effects model is 59.026. For verification, the same calculation for the Cobb-Douglas specification is 43.870. The critical value of χ^2 with (i-1=24) d.f. = 42.98 at 1% significance. Thus, as with Dataset I, we reject the null and conclude PSAR1 is recommended.

In summary, our tests conclude that both Datasets I and II warrant econometric adjustments for groupwise heteroskedasticity and panel-specific autocorrelation.

4.5.2 Estimation Results

Several random effects panel estimators can correct for both heteroskedasticity and autocorrelation. We estimate the model using three estimators: the generalized least-squares (GLS) regression with heteroskedastic panels and panel-specific autocorrelation (hereafter labeled He+PSAR1), the panel-corrected standard errors (PCSE) regression with AR1 (labeled He+PCSE), and a random-effects estimation with first-order autoregressive errors (labeled AR1). In the He+PSAR1 and He+PCSE estimates we employ the adjustment for heteroskedastic panels. We control for cross-sectional temporal unobservables by adding a dummy variable for each year to all regressions (for brevity we omit these coefficient estimates from the results tables).

He+PSAR1 In light of the foregoing tests of panel error structure, we begin with the GLS estimator with adjustments for heteroskedastic error and panel-specific autocorrelation. The results are presented in Tables 4.4 and 4.5. In these estimations we obtain a reasonable result from the Cobb-Douglas specification in column (1): all three inputs are positive and significant, with magnitudes relatively close to the findings of other researchers on similar samples. In column (2) we add the indirect effect of IT by estimating the θ_2 parameter from equation (4.5), representing the indirect effect of Z. The negative and significant estimate is unexpected, given the positive finding of Mittal and Nault (2009) with this specification. However, we note that our estimate of Z with this specification is not significant in any of our other estimations (with both Dataset I and II). We also find a nonsignificant estimate of Z when using the more general He+AR1 (GLS estimation with adjustments for heteroskedasticity and first-order autoregressive error) adjustment with this model, suggesting the estimate may be an artifact of the PSAR1 controls. In column (3) we add the indirect effect of R&D by estimating the θ_1 parameter. This is a simplified variation of our full model, without the interaction of R&D and IT indirect effects. The estimates are similar to the previous results for the k, l, z and Z parameters, but the coefficient estimate for R is not significant. Finally, we estimate our full research model in column (4). The estimated coefficients are positive and significant for k, l, and z, while the indirect effect of Z is negative and significant. The indirect effect of R is not significant, but the interaction term θ_3 is positive and significant. This suggests that the while the indirect effect of IT on the other inputs may not be positive, there is a positive indirect effect when combined with R&D. Because of the interaction (and, in the case of Z, the overall direct effect), the net effects need to be evaluated according to the elasticities in equations (4.6) and (4.7), while the indirect effects may be evaluated in the context of the scaling factors in equations (4.2) and (4.3).

With Dataset II (table 4.5), the Cobb-Douglas parameter estimates are qualitatively similar to those in Dataset I. The θ_2 parameter, however, is not significant in columns (2) and (3), while the θ_1 parameter is negative and significant in column (3). In the full model in column (4), all coefficient estimates are significant. The estimates for Z and R are negative, while the interaction term ZR is positive.

He+PCSE An alternative estimation approach that adjusts for autocorrelation and heteroskedasticity is panel-corrected standard errors estimation. The PCSE estimator uses OLS regression but adjusts the estimated standard errors for correlation within panels. It is theoretically more efficient than fixed-effect OLS estimators under heteroskedasticity and autocorrelation while being less susceptible to underestimating the variability of estimates in finite samples than GLS methods (Beck and Katz, 1995). We run the estimation using the adjustments for heteroskedastic panels and AR(1) autocorrelation structure.[28]

The results are presented in Table 4.6 and Table 4.7 for Datasets I and II, respectively. In Table 4.6, the first column shows the Cobb-Douglas results, where we note that the estimate of k is not significant while the significant estimate for z seems unusually high at 0.38. The addition of the Z and R variables in columns (2) and (3) have little impact on the factor input coefficient estimates, while the indirect coefficient for Z remains non-significant. However, in column (3) the estimate for R is marginally significant at the 10% level. For the full model in column (4), all variables, with the exception of l, are significant. In Dataset II, the direct coefficient estimates k, l, and z are positive and significant in all model variations. However, the only significant indirect coefficient estimate is that for R in the full model (column (4)). Interpretation of the overall effect of Z depends upon the computed value of the elasticity, to be discussed below.

AR1 Finally, we report our estimations using the random effects AR1 approach on Dataset I. In Table 4.8, column (1) reports the standard Cobb-Douglas production function as a base model to check for reasonableness. For all model variations in columns (1) through (4), the sign, significance and magnitude of the direct and indirect effect estimates are very similar to those

[28] Although a PSAR1 option exists, the panel-specific error correction will include contemporaneous correlation at the panel level.

found using the He+PCSE estimator, with the exception of the estimate for k which is not significant. In Dataset II, Table 4.9 shows all model variations have k and l have positive and significant estimates (again similar to the He+PCSE estimates) but z is not significant. None of the indirect effects have significant coefficient estimates, except for the interaction term in the full model in column (4). In sum, the AR1 estimator produces similar estimates with somewhat larger standard errors, perhaps due to the coarser treatment of autocorrelation error structures compared with He+PSAR1 or the lack of adjustment for heteroskedasticity.

Elasticities

The output elasticities of labour, non-IT capital and IT capital are relatively straightforward to calculate and interpret in our model. However, since we introduce R&D through an indirect effect in our model, we calculate a partial elasticity that has a slightly different interpretation. We evaluate three aspects of the direct and indirect effects in our model. First, we evaluate the output elasticity of Z and compare this with the earlier estimates. Second, we compute the Z and $Z(R)$ components of the scaling factors ($\alpha^{'}$ and $\alpha^{''}$), to the extent made possible by our model. Third, we compute the partial elasticity of R to estimate the influence of the indirect effect of R&D on the factor inputs.

In Dataset I, for the full model (Table 4.4, column (4)), the estimated output elasticity η_Z is 0.169 when evaluated at the mean (.168 at the median), with a standard error of 0.035. This does not suggest an improvement over the estimated direct effect of 0.322. Rather, we might infer that once the indirect effects of R&D and IT capital are included in the model, the true elasticity of IT capital is lower than the traditional Cobb-Douglas elasticity. This follows from the estimates of the scaling factor components $\alpha^{'}$ (-.0573) and $\alpha^{''}$ (0.00138); the indirect effect of Z is negative, while for ZR it is positive. Thus it would appear that IT capital has a small, negative indirect effect on K and L, which is not entirely mitigated by the interaction of IT with R&D. However, R&D appears to have the expected positive marginal effect on output through its collective indirect effects on the factor inputs (α). The partial elasticity estimate for R is 0.101 with SE of 0.044 (t=2.29), although the model specification does not allow us to decompose this effect into different amounts for K, L, and Z. In our estimations using PCSE and AR1, the elasticities follow the same pattern of sign and significance.

In Dataset II, for the full model (Table 4.5, column (4)), the estimated output elasticity η_Z is .229 when evaluated at the mean (.261 at the median), with a standard error of 0.036. This exceeds the estimated Cobb-Douglas output elasticity of 0.172, indicating that the true impact of IT capital is larger than first estimated, once the indirect effects are included in the model. The estimate of $\alpha^{'}$ is -0.00365, and the estimate of $\alpha^{''}$ is 0.0000279. Thus, the indirect effect of Z is negative, but the interaction ZR (along with the larger direct effect of Z) results in a larger net output elasticity. This result is again confirmed by the PCSE and AR1 estimates. In contrast, the partial elasticity estimate for R is -0.071, with standard error of 0.031 (t=-2.25). In the PCSE and AR1 estimations, the partial elasticity of R is not significant. Thus, it seems that R&D on its own has a negative effect (if any) on the factor inputs collectively, although we cannot isolate the scaling effect on each input.

Hypothesis Tests We test the hypotheses as discussed in section 3.3. We use the He+PSAR1 results for both datasets. For the first test, we compare the indirect effect of Z and $Z(R)$. Under the null, $\alpha^{'} = \alpha^{''} = 0$ (or $\theta_2 = \theta_3$). For Dataset I, we reject the null ($\chi^2 = 9.52, p < 0.01$). For Dataset II, we reject the null ($\chi^2 = 16.21, p < 0.001$). Thus we reject the null hypothesis that the indirect effect of Z and $Z(R)$ are indistinguishable. We also reject the hypothesis that Z has no indirect effect (alone or in conjunction with R&D) on K and L.

Next, we test the hypothesis that all the indirect effects (arising from both R and $Z(R)$) are the same upon all inputs K, L and Z ($\theta_1 = \theta_2 = \theta_3$). For Dataset I, we reject the null ($\chi^2 = 12.0, p < 0.01$). For Dataset II, we reject the null ($\chi^2 = 24.01, p < 0.001$). Thus we conclude that the indirect effects of R, Z, and ZR are different from one another.

4.5.3 Robustness Checks on Sample Partitions

We now check the robustness of our results by performing supplementary analyses. First, we analyze a subset of Dataset II (restricted to the manufacturing sector) and compare it with Dataset I, in which only manufacturing industries are included. Summary statistics for this subsample are listed in Table 4.1, and Table 4.3 shows the sector membership of each industry. Second, we separate our sample into IT-intensive and non-IT-intensive industries to see if the indirect effects differ between the two groups.

Table 4.10 reports the estimation results when the Dataset II sample is restricted to the 16 manufacturing industries. Using the He+PSAR1 estimator, the Cobb-Douglas model estimate of z in column (1) is not significant, as with the full model estimation in column (4). In the full model we note that k is no longer significant, and none of the indirect effect estimates are significant. As a result the elasticities are also not significant: the output elasticity of Z evaluated at the mean is 0.144 (at median, 0.187), but is not significant at the 10% level (standard error of 0.192). The partial elasticity of R at the mean is -0.0730 and is not significant at any meaningful level. We thus cannot conclude the earlier results are robust to the manufacturing sector in Dataset II. However, the estimates are highly sensitive to changes in the model, suggesting the estimation approach is hindered by the smaller number of industries representing the same (manufacturing) sector of the economy.

We next partition the sample by IT intensity. We compute a measure of mean IT intensity (Z/K) for each industry over the sample years in each dataset and use a scree plot to identify the break point between the intensive and non-IT-intensive industries. In both datasets, nine industries are considered IT-intensive (indicated in Tables 4.2 and 4.3 with the † symbol). We present the analyses of our full model on each dataset in tables 4.11 and 4.12. In columns (1) and (2) we estimate the model using the IT- and non-IT-intensive subsamples, respectively, using He+AR1. We repeat this in columns (3) and (4) using He+PSAR1. We include both AR1 and PSAR1 estimation approaches because the additional parameters estimated in the latter can be problematic with smaller sample sizes.

For the IT-intensive group of industries, some of the estimates are out of line with expectations and may be artifacts of the smaller sample size (108 and 72 observations in Dataset I and II, respectively). In Dataset I, (Table 4.2, columns (1)), the estimate for k is not significant under He+AR1, while under He+PSAR1 (column (3)), the k estimate is negative and significant. Similarly, in Dataset II, the estimate for k is not significant or negative and significant for the IT-intensive group (Table 4.2, columns (1) and (3), respectively).

The estimates for the non-IT-intensive industries display some evidence of the estimation problems described above, and we proceed cautiously given the only slightly larger sample size. In Dataset I, both AR1 and PSAR1 approaches (Table 4.2, columns (3) and (4)) result in a non-significant estimate of one of the inputs (k and z, respectively). The ZR estimate is positive in both cases, but significantly smaller than in the full sample. In Dataset II (Table 4.3, columns (3) and (4)), the ZR estimate is positive but significant only under He+AR1 estimation, and only at the 10% level. Given the unstable and sometimes nonsensical results, in addition to the potential for estimation problems with the smaller sample sizes in these analyses, we find it difficult to make inferences based on the IT-intensity sample splits.

Finally, in columns (5) and (6) we estimate the full sample using a dummy variable ("intensive1") to indicate the industries' membership in either group. In Dataset I (Table 4.2), the estimations under both He+AR1 and He+PSAR1 report k as not significant, while the coefficient for z is much larger than Cobb-Douglas, and only under AR1 are the indirect effects significant for Z and ZR (negative and positive, respectively). In Dataset II (Table 4.3), both He+AR1 and He+PSAR1 estimations return significant estimates for the direct effects and Z, but not R and ZR. In both datasets, the dummy variable estimate is negative and significant, though the interpretation of this result is not meaningful beyond its qualitative implication (that the intercept for the IT-intensive industries is lower than that for other industries).

Variable Coefficients Model

An alternative approach to the indirect effects model presented in equation (4.1) is to define the exponents of the Cobb-Douglas form such that R&D forms a component of the exponent rather than a scaling factor on the quantity of the input. This model has the advantage of being more straightforward to interpret, but does not feature the indirect effect of Z on the other inputs nor the interaction $Z(R)$ of the indirect effects model. We define the model as follows:

$$\begin{aligned} Y &= AK^{\alpha}L^{\beta}Z^{\gamma}, \\ \text{where} & \\ \alpha &= a_k + b_k R \\ \beta &= a_l + b_l R \\ \gamma &= a_z + b_z R. \end{aligned}$$

All variables share the same definitions as the foregoing indirect effects model. Taking the log of both sides yields:

$$y = a + a_k k + b_k kR + a_l l + b_l lR + a_z z + b_z zR. \tag{4.7}$$

We present our estimates of this model in Tables 4.13 and 4.14. For Dataset I, using He+AR1 estimation in column (3), we find positive and significant coefficients for the direct effects, but none of the interaction terms are significant. Using the panel-specific estimator (He+PSAR1), k is not significant, l is very large, and only the lR interaction is significant (negative). The output elasticity of z is 0.205 at the mean (0.185 at the median) using He+AR1, with a standard error of 0.057. Using PSAR1, the elasticity of z is 0.256 at the mean (0.255 at the median), with a standard error of 0.043. These compare roughly to the Cobb-Douglas elasticity of 0.212.

In Dataset II, He+PSAR1 estimation (column (4)) returns positive and significant coefficient estimates for the direct effects. Non-IT capital and labour are similar in magnitude to their Cobb-Douglas counterparts, while

IT capital is somewhat larger. The interaction of kR is negative and significant, while the zR term is positive and significant. The output elasticity of z is 0.229 at the mean (0.187 at the median) using He+AR1, with a standard error of 0.0532. Using PSAR1, the elasticity of z is 0.343 at the mean (0.263 at the median), with a standard error of 0.0391. Compared to the Cobb-Douglas elasticity of 0.234, the variable coefficient model produces a similar or noticeably larger elasticity, depending on the choice of AR1 or PSAR1.

Overall, it is difficult to draw conclusions from the variable coefficients model results. For Dataset II, the results confirm and extend the Cobb-Douglas form while lending some support to the positive impact of R&D on IT capital, at the apparent expense of the effective quantity of non-IT capital. On the other hand, the results from Dataset I seem unreasonable using He+PSAR1 estimation, while the He+AR1 estimator gives results qualitatively similar to the Cobb-Douglas specification, finding no indirect effect of R&D.

4.6 Discussion and Conclusion

In this research, we propose a model of the indirect effects of IT and R&D on production. The results of our analyses shed some light on the relationship between innovation and IT, and how they combine to influence production efficiencies at the industry level.

In the first dataset, covering the years 1987–1998, the elasticity of IT capital in the indirect effects model is lower than that of the Cobb-Douglas model. Decomposing this net effect, we find that IT capital has a slightly deleterious indirect effect on non-IT capital and labour that outweighs the positive indirect effect of the interaction of R&D and IT. Nevertheless, the estimated net marginal contribution of IT capital to productivity remains positive. We find the net effect of R&D knowledge stock is positive, indicating R&D investments enable the factor inputs to perform more efficiently. Furthermore, part of this net effect arises from the interaction of R&D with IT.

In the second dataset (1998–2005), the net marginal contribution of IT capital is larger in the full model than in the Cobb-Douglas model, suggesting that the production efficiencies are enhanced by the indirect effects of IT. More specifically, since the indirect effect of IT alone is negative, it is the interaction indirect effect of IT with R&D that creates the overall increase in net effect. However, for R&D, the estimated net effect is negative or nonsignificant. Thus it appears that, for our second sample, IT drives productivity enhancements in the presence of R&D, but R&D on its own does not. We are unable to determine if this is due to the different nature of the sample (inclusion of non-manufacturing industries, grouping by NAICS) or other factors, such as a shift in the way that industries use IT and R&D to improve production efficiency. There is reason to believe that widespread improvements in productivity in the late 1990s and early 2000s arose from IT and complementary organizational investments and practices that helped firms and industries capitalize on earlier IT capital investments (Brynjolfsson and Hitt, 2003; Jorgenson, Ho, and Stiroh, 2008). Thus, part of this effect may owe to process innovation efficiencies captured by the indirect effects in our model.

These results contribute to our understanding of how and when industries have increased the efficiency of inputs through investment in IT and R&D. Although neither IT capital nor R&D are individually found to augment the effective quantity of non-IT capital and labour, the scaling factor of IT capital, *in combination* with R&D, is positive. This supports our hypothesis that process innovations are enabled by IT investments. Our analysis of the first sample reinforces earlier research findings regarding the returns to R&D. The results from our second sample, however, raise the question of whether IT now dominates R&D in the role of augmenting the efficiency of factor inputs.

4.6.1 Limitations

Several limitations of our data should be noted. First, the level of aggregation of industries is not at a consistent level, due to the limitations of the NSF R&D data. Ideally we would prefer to match the MFP data at the 3-digit SIC or 4-digit NAICS code with R&D data at the same level, but this is not possible. We presume that having observations that represent a more homogeneous grouping of firms would produce results with more robust findings. Second, the lack of a long history of NAICS-based R&D data limits the generation of the R&D stock variable in Dataset II to a shorter smoothing period. Third, we are unable to separate the portion of the factor inputs that are dedicated to R&D, so these amounts are double-counted. Fourth, our measure of R&D does not distinguish between process- and product-related innovation. Although our hypotheses focus on production efficiencies arising from process innovation, product innovation commands a significant share of R&D spending in many industries (Cohen and Klepper, 1996). Thus our use of R&D amounts is a rough proxy for the resources dedicated to innovations which aim to affect efficiency, while an unknown proportion is devoted to product improvement. We also note that a number of industries were eliminated from our analyses in both datasets due to the NSF's intentional withholding of R&D survey data. While it is possible that this may bias our results, it occurs only where a very small number of firms perform R&D within an industry. As such, these industries may be unusual in structure and therefore unsuitable for inclusion in our sample.

In addition to the data limitations, our conclusions regarding the different nature of indirect effects between the samples must be taken cautiously. While it is tempting to conclude that a change has taken place in the way industries implement process innovations, the comparison of the two sets of results is hindered by differences in the makeup of the samples. The observations in the second sample represent very different industries. As can be noted from the summary statistics in Table 4.1, the industries are much larger due to the higher level of NAICS aggregation. Further, the inclusion of non-manufacturing industries, such as finance and information, introduces different characteristics with respect to measurement of outputs and their methods of production compared to manufacturing industries. In our supplemental analysis we restrict the second dataset to manufacturing industries, but differences in the aggregation of industries and the small number of observations prevent a straightforward comparison.

4.6.2 Future Research

Extensions to this research could explore several dimensions of the R&D knowledge stock, which can be split into internal, contract research, and government-funded research. In addition, the NSF R&D survey has expanded to include additional data on the breakdown of spending on different R&D activities. This may make it possible to remove the labour, IT capital and non-IT capital components of R&D from the factor inputs, resulting in a more accurate estimate of our model.

Another area for further exploration is the effect of R&D spillovers on innovation and the indirect effect of IT. The scaling effect of R&D and IT may not be isolated within firms but extend across an industry and between industries. Economists have found empirical support for R&D spillovers (Griliches and Lichtenberg, 1984; Griliches, 1992), and more recently, spillover phenomena have also captured the attention of Information Systems researchers (Stiroh, 2002; Hitt and Tambe, 2006). Such industry-level effects are worth investigating given the nature of IT, which lends itself to rapid communication, replication and discovery.

Table 4.1: Summary Statistics

	N	Mean	Std. Dev.	Min	Max
Dataset I: 1987–1998					
(millions of 1987 dollars)					
Value added	276	44,944.83	48,406.65	1,674.66	505,614.00
Non-IT capital	276	74,960.50	57,521.65	2,897.80	223,860.90
Labour (millions of hours)	276	1,509.49	1,048.20	119.80	3,972.50
IT capital	276	6,073.78	5,932.29	258.50	39,195.20
R&D stock	276	14,648.24	14,443.52	915.41	60,509.25
R&D expense	276	2,526.91	2,656.55	51.90	11,330.51
Dataset II: 1998–2005					
(millions of 2005 dollars)					
Value added	200	342,022.70	524,093.70	30,097.00	2,606,450.00
Non-IT capital	200	473,052.90	605,868.20	20,340.00	2,584,289.00
Labour (hours)	200	7,262.45	11,839.93	259.30	60,110.00
IT capital	200	56,754.75	82,130.94	832.00	381,533.00
R&D stock	200	39,927.73	62,945.78	238.57	241,671.90
Dataset II: 1998–2005 - Manufacturing Sector only					
(millions of 2005 dollars)					
Value added	128	88,077.37	50,818.17	30,097.00	193,132.00
Non-IT capital	128	153,246.50	92,431.69	20,340.00	330,068.00
Labour (hours)	128	2,171.33	1,147.28	259.30	4,769.40
IT capital	128	15,193.91	16,291.97	832.00	62,336.00
R&D stock	128	46,689.48	71,480.34	238.57	241,671.90

Table 4.2: Dataset I Industries

Industry	SIC Codes	Major Industry	Subordinate Industry
2†	13, 29	Petroleum refining and extraction	
4	20, 21	Food, kindred, and tobacco products	
5	22, 23	Textiles and apparel	
6	24, 25	Lumber, wood products, and furniture	
7	26	Paper and allied products	
8†	27, 31, 39	Other manufacturing industries	
10†	281–82, 286	Chemicals and allied products	Industrial chemicals
11†	283		Drugs and medicines
12†	284–85, 287–89		Other chemicals
13	30	Rubber products	
14	32	Stone, clay, and glass products	
99*	331–32, 3398–99	Primary metals	Ferrous metals and products
99*	333–36		Nonferrous metals and products
18	34	Fabricated metal products	
20	351–56, 358–59	Machinery	Other machinery, except electrical

† = IT-intensive industry

* The two Primary Metals industries are combined into a single industry to facilitate matching to the MFP industry aggregation

Continued on next page

Table 4.2 – continued from previous page

Industry	SIC Codes	Major Industry	Subordinate Industry
21	357		Office, computing, and accounting machines
23	361–64, 369	Electrical equipment	Other electrical equipment
24†	365		Radio and TV receiving equipment
25	366		Communication equipment
26	367		Electronic components
28	371	Transportation equipment	Motor vehicles and motor vehicles equipment
30†	373–75, 379		Other transportation equipment
32†	381–82	Professional and scientific instruments	Scientific and mechanical measuring instruments
33†	384–87		Optical, surgical, photographic, and other instruments

† = IT-intensive industry

Table 4.3: Dataset II Industries

Industry	Sector	NAICS Codes	Description	δ
1	NM	21	Mining, extraction, and support activities	0.15
2†	NM	22	Utilities	0.15
3	NM	23	Construction	0.15
5	M	311–312	Food, beverage and tobacco products	0.15
7	M	313–316	Textiles, apparel, and leather	0.15
8	M	321	Wood products	0.15
9	M	322–323	Paper, printing, and support activities	0.15
10	M	324	Petroleum and coal products	0.15
11†	M	325	Chemicals	0.11
12	M	326	Plastics and rubber products	0.15
13	M	327	Nonmetallic mineral products	0.15
14	M	331	Primary metals	0.15
15	M	332	Fabricated metal products	0.15
16†	M	333	Machinery	0.15
17†	M	334	Computer and electronic products	0.165

† = IT-intensive industry

Sector M = Manufacturing, NM = Non-manufacturing

δ = R&D annual depreciation rate

Continued on next page

Table 4.3 – continued from previous page

Industry	Sector	NAICS Codes	Description	δ
18	M	335	Electrical equipment, appliances, and components	0.15
19†	M	336	Transportation equipment	0.18
20	M	337	Furniture and related products	0.15
21	M	339	Miscellaneous manufacturing	0.15
24†	NM	48–49	Transportation and warehousing	0.15
25†	NM	51	Information	0.15
26	NM	52–53	Finance, insurance, and real estate	0.15
27†	NM	54	Professional, scientific, and technical services	0.15
28†	NM	621–623	Health care services	0.15
29	NM	55, 56, 61, 624, 71, 72, 81	Other non-manufacturing	0.15

† = IT-intensive industry

Sector M = Manufacturing, NM = Non-manufacturing

δ = R&D annual depreciation rate

Table 4.4: GLS He+PSAR1 Regression Results

Dataset I: 1987–1998

VARIABLES	(1) Cobb-Douglas	(2) Indirect Z	(3) Uninteracted	(4) Full Model
k	0.135***	0.185***	0.118***	0.178***
	(0.0516)	(0.0548)	(0.0416)	(0.0431)
l	0.710***	0.634***	0.703***	0.650***
	(0.0542)	(0.0570)	(0.0448)	(0.0474)
z	0.246***	0.299***	0.301***	0.322***
	(0.0330)	(0.0384)	(0.0380)	(0.0549)
Z		-0.0162***	-0.0125**	-0.0474***
		(0.00483)	(0.00554)	(0.0153)
R			0.00217	-0.00231
			(0.00252)	(0.00357)
ZR				0.00114**
				(0.000475)
Constant	2.468***	2.312***	2.495***	2.383***
	(0.178)	(0.183)	(0.142)	(0.155)
Observations	276	276	276	276
η_Z (mean)	0.246	0.201	0.225	0.169
η_Z (median)	0.246	0.223	0.243	0.168
$\eta_K + \eta_L + \eta_Z$ (mean)	1.091	1.02	1.046	0.997

Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1
Dependent variable: Value Added.
Year dummies omitted from results.

Table 4.5: GLS He+PSAR1 Regression Results

Dataset II: 1998–2005

	(1)	(2)	(3)	(4)
VARIABLES	Cobb-Douglas	Indirect Z	Uninteracted	Full Model
k	0.338***	0.370***	0.283***	0.266***
	(0.0380)	(0.0412)	(0.0476)	(0.0448)
l	0.520***	0.546***	0.477***	0.506***
	(0.0279)	(0.0288)	(0.0381)	(0.0336)
z	0.172***	0.101**	0.285***	0.315***
	(0.0330)	(0.0474)	(0.0480)	(0.0437)
Z		0.000641	-0.000458	-0.00282***
		(0.000494)	(0.000452)	(0.000701)
R			-0.00228***	-0.00364***
			(0.000764)	(0.000919)
ZR				2.15e-05***
				(5.86e-06)
Constant	2.036***	1.984***	2.155***	2.190***
	(0.139)	(0.136)	(0.173)	(0.164)
Observations	200	200	200	200
η_Z (mean)	0.172	0.138	0.259	0.229
η_Z (median)	0.172	0.115	0.275	0.261
$\eta_K + \eta_L + \eta_Z$ (mean)	1.03	1.054	1.019	1.001

Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1
Dependent variable: Value Added.
Year dummies omitted from results.

Table 4.6: PCSE He+AR1 Regression Results

Dataset I: 1987–1998

VARIABLES	(1) Cobb-Douglas	(2) Indirect Z	(3) Uninteracted	(4) Full Model
k	0.0819	0.0726	0.110	0.0906
	(0.162)	(0.157)	(0.133)	(0.135)
l	0.646***	0.665***	0.669***	0.636***
	(0.191)	(0.193)	(0.169)	(0.173)
z	0.380***	0.350***	0.260***	0.539***
	(0.0727)	(0.0927)	(0.0774)	(0.108)
Z		0.00512	0.00177	-0.0911***
		(0.0131)	(0.0119)	(0.0278)
R			0.00981*	-0.0142*
			(0.00525)	(0.00817)
ZR				0.00298***
				(0.000879)
Constant	2.511***	2.554***	2.388***	2.670***
	(0.554)	(0.542)	(0.458)	(0.475)
Observations	276	276	276	276
R-squared	0.820	0.828	0.847	0.856
η_Z (mean)	0.380	0.381	0.271	0.338
η_Z (median)	0.380	0.374	0.268	0.290

Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1
Dependent variable: Value Added.
Year dummies omitted from results.

Table 4.7: PCSE He+AR1 Regression Results

Dataset II: 1998–2005

	(1)	(2)	(3)	(4)
VARIABLES	Cobb-Douglas	Indirect Z	Uninteracted	Full Model
k	0.317***	0.324***	0.247***	0.248***
	(0.0590)	(0.0640)	(0.0685)	(0.0675)
l	0.439***	0.438***	0.391***	0.405***
	(0.0647)	(0.0655)	(0.0678)	(0.0678)
z	0.227***	0.209***	0.328***	0.353***
	(0.0609)	(0.0797)	(0.0850)	(0.0886)
Z		0.000319	1.13e-05	-0.00126
		(0.000730)	(0.000690)	(0.00124)
R			-0.00212	-0.00302*
			(0.00130)	(0.00157)
ZR				1.25e-05
				(9.72e-06)
Constant	2.110***	2.112***	2.333***	2.299***
	(0.217)	(0.217)	(0.240)	(0.235)
Observations	200	200	200	200
R-squared	0.968	0.968	0.969	0.969
η_Z (mean)	0.227	0.227	0.329	0.325
η_Z (median)	0.227	0.215	0.328	0.329

Standard errors in parentheses.

*** p<0.01, ** p<0.05, * p<0.1

Dependent variable: Value Added.

Year dummies omitted from results.

Table 4.8: Random Effects AR(1) Regression Results

Dataset I: 1987–1998

	(1)	(2)	(3)	(4)
VARIABLES	Cobb-Douglas	Indirect Z	Uninteracted	Full Model
k	0.113	0.104	0.115	0.0944
	(0.113)	(0.114)	(0.0979)	(0.0937)
l	0.604***	0.626***	0.652***	0.614***
	(0.120)	(0.123)	(0.105)	(0.100)
z	0.386***	0.335***	0.272**	0.561***
	(0.0884)	(0.112)	(0.107)	(0.118)
Z		0.0102	0.00435	-0.0940***
		(0.0137)	(0.0127)	(0.0240)
R			0.00790*	-0.0180***
			(0.00426)	(0.00673)
ZR				0.00318***
				(0.000671)
Constant	2.392***	2.433***	2.375***	2.683***
	(0.395)	(0.398)	(0.347)	(0.338)
Observations	276	276	276	276
η_Z (mean)	0.386	0.397	0.298	0.366
η_Z (median)	0.386	0.382	0.292	0.310

Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1
Dependent variable: Value Added.
Year dummies omitted from results.

Table 4.9: Random Effects AR(1) Regression Results

Dataset II: 1998–2005

	(1)	(2)	(3)	(4)
VARIABLES	Cobb-Douglas	Indirect Z	Uninteracted	Full Model
k	0.449***	0.450***	0.432***	0.418***
	(0.108)	(0.109)	(0.104)	(0.105)
l	0.417***	0.417***	0.429***	0.442***
	(0.0901)	(0.0900)	(0.0859)	(0.0866)
z	0.105	0.0990	0.111	0.155
	(0.0958)	(0.107)	(0.110)	(0.113)
Z		0.000105	8.44e-05	-0.00138
		(0.000739)	(0.000739)	(0.00113)
R			0.000345	-0.00110
			(0.00113)	(0.00143)
ZR				1.48e-05*
				(8.58e-06)
Constant	1.782***	1.786***	1.824***	1.836***
	(0.437)	(0.437)	(0.409)	(0.411)
Observations	200	200	200	200
η_Z (mean)	0.105	0.105	0.115	0.128
η_Z (median)	0.105	0.101	0.112	0.129

Standard errors in parentheses.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$
Dependent variable: Value Added.
Year dummies omitted from results.

Table 4.10: Manufacturing Industries GLS He+PSAR1 Regression Results

Dataset II: 1998–2005

	(1)	(2)	(3)	(4)
VARIABLES	Cobb-Douglas	Indirect Z	Uninteracted	Full Model
k	0.387***	0.325**	0.307*	0.337
	(0.111)	(0.161)	(0.162)	(0.210)
l	0.476***	0.364***	0.306***	0.491***
	(0.0997)	(0.111)	(0.101)	(0.112)
z	0.0215	0.214	0.273**	0.0843
	(0.0792)	(0.131)	(0.137)	(0.205)
Z		-0.00645	-0.00330	0.00280
		(0.00597)	(0.00704)	(0.0107)
R			-0.00177	-0.000240
			(0.00123)	(0.00259)
ZR				-3.44e-05
				(6.07e-05)
Constant	2.059***	2.032***	2.089***	2.148***
	(0.370)	(0.563)	(0.567)	(0.680)
Observations	128	128	128	128
η_Z (mean)	0.0215	0.116	0.222	0.0651
η_Z (median)	0.0215	0.169	0.249	0.0999

Standard errors in parentheses.

*** p<0.01, ** p<0.05, * p<0.1

Dependent variable: Value Added.

Year dummies omitted from results.

Table 4.11: IT-intensity Sample Partition Regression Results

Dataset I: 1987–1998

	(1)	(2)	(3)	(4)	(5)	(6)
	IT-intensive	non-IT-intensive	IT-intensive	non-IT-intensive	IT-intensive dummy	
VARIABLES	He+AR1	He+AR1	He+PSAR1	He+PSAR1	He+AR1	He+PSAR1
k	−0.0408	0.189	−0.184*	0.249***	0.0600	0.0285
	(0.0952)	(0.129)	(0.104)	(0.0913)	(0.0906)	(0.0662)
l	0.410***	0.399***	0.840***	0.380***	0.495***	0.670***
	(0.0757)	(0.130)	(0.145)	(0.0708)	(0.0804)	(0.0646)
z	0.649***	0.235**	0.432**	0.0512	0.440***	0.403***
	(0.119)	(0.108)	(0.185)	(0.0687)	(0.0823)	(0.0787)
Z	$-1.07e{-}05$	$-3.47e{-}05$	$4.25e{-}05$	$6.94e{-}06$	$-3.22e{-}05$**	$-2.28e{-}05$
	$(2.45e{-}05)$	$(2.27e{-}05)$	$(3.04e{-}05)$	$(2.02e{-}05)$	$(1.41e{-}05)$	$(1.79e{-}05)$
R	$1.21e{-}06$	$-1.56e{-}05$**	$8.75e{-}06$	$2.35e{-}06$	$-9.00e{-}06$	$-5.55e{-}06$
	$(7.60e{-}06)$	$(7.61e{-}06)$	$(7.89e{-}06)$	$(5.28e{-}06)$	$(5.50e{-}06)$	$(5.15e{-}06)$
ZR	$-4.04e{-}10$	$2.64e{-}09$***	$-1.97e{-}09$**	$2.27e{-}09$***	$1.16e{-}09$**	$4.28e{-}10$
	$(7.75e{-}10)$	$(6.40e{-}10)$	$(9.67e{-}10)$	$(5.97e{-}10)$	$(4.67e{-}10)$	$(5.62e{-}10)$
intensive1					−0.432***	−0.299*
					(0.149)	(0.162)
Constant	2.326***	3.865***	2.554***	4.551***	3.009***	2.277***
	(0.619)	(1.191)	(0.605)	(0.954)	(0.725)	(0.475)
Observations	108	168	108	168	276	276
Industries	9	14	9	14	23	23

Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Dependent variable: Value Added. Year dummies omitted from results.

Table 4.12: IT-intensity Sample Partition Regression Results

Dataset II: 1998–2005

VARIABLES	(1) IT-intensive He+AR1	(2) non-IT-intensive He+AR1	(3) IT-intensive He+PSAR1	(4) non-IT-intensive He+PSAR1	(5) IT-intensive dummy He+AR1	(6) IT-intensive dummy He+PSAR1
k	−0.201**	0.226**	−0.334***	0.364***	0.201***	0.282***
	(0.0901)	(0.100)	(0.0747)	(0.0740)	(0.0534)	(0.0606)
l	0.261***	0.332***	0.233***	0.555***	0.479***	0.549***
	(0.0979)	(0.0488)	(0.0872)	(0.0595)	(0.0491)	(0.0473)
z	0.488***	0.511***	0.478***	0.295***	0.382***	0.315***
	(0.151)	(0.119)	(0.126)	(0.109)	(0.0782)	(0.0734)
Z	−0.00271**	−0.00230	−0.00329***	$-8.32e{-05}$	−0.00146*	−0.00118*
	(0.00136)	(0.00195)	(0.00122)	(0.00168)	(0.000768)	(0.000707)
R	−0.00348***	−0.00755*	−0.00655***	−0.0113***	0.000193	−0.00133
	(0.00112)	(0.00458)	(0.00121)	(0.00414)	(0.000811)	(0.00125)
ZR	$2.15e{-05}$***	0.000183*	$2.99e{-05}$***	$6.19e{-05}$	$8.86e{-06}$	$9.40e{-06}$
	($7.42e{-06}$)	(0.000110)	($6.91e{-06}$)	($9.51e{-05}$)	($6.55e{-06}$)	($6.26e{-06}$)
intensive1					−0.284**	−0.452***
					(0.116)	(0.143)
Constant	4.702***	2.215***	5.832***	1.800***	2.354***	2.075***
	(0.731)	(0.329)	(0.665)	(0.240)	(0.178)	(0.198)
Observations	72	128	72	128	200	200
Industries	9	16	9	16	25	25

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Dependent variable: Value Added. Year dummies omitted from results.

Table 4.13: Variable Coefficients Model Regression Results

Dataset I: 1987–1998

VARIABLES	(1) Cobb-Douglas He+AR1	(2) Cobb-Douglas He+PSAR1	(3) VC Model He+AR1	(4) VC Model He+PSAR1
k	0.177	0.135***	0.237***	0.0137
	(0.111)	(0.0516)	(0.0677)	(0.0396)
kR			-0.000687	-0.000695
			(0.00233)	(0.00182)
l	0.618***	0.710***	0.599***	0.909***
	(0.119)	(0.0542)	(0.0866)	(0.0571)
lR			-0.00334	-0.0168***
			(0.00542)	(0.00419)
z	0.212***	0.246***	0.141***	0.253***
	(0.0700)	(0.0330)	(0.0509)	(0.0405)
zR			0.00442	0.000214
			(0.00366)	(0.00323)
Constant	2.390***	2.468***	2.198***	2.884***
	(0.413)	(0.178)	(0.244)	(0.145)
Observations	276	276	276	276
Industries	23	23	23	23

Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1
Dependent variable: Value Added.
Year dummies omitted from results.

Table 4.14: Variable Coefficients Model Regression Results

Dataset II: 1998–2005

	(1)	(2)	(3)	(4)
	Cobb-Douglas	Cobb-Douglas	VC Model	VC Model
VARIABLES	He+AR1	He+PSAR1	He+AR1	He+PSAR1
k	0.261***	0.338***	0.272***	0.281***
	(0.0409)	(0.0380)	(0.0476)	(0.0351)
kR			-0.000642	-0.00231***
			(0.000568)	(0.000649)
l	0.509***	0.520***	0.574***	0.511***
	(0.0311)	(0.0279)	(0.0430)	(0.0406)
lR			-0.00146	-0.000976
			(0.000969)	(0.00145)
z	0.234***	0.172***	0.170***	0.229***
	(0.0334)	(0.0330)	(0.0448)	(0.0329)
zR			0.00150*	0.00284***
			(0.000907)	(0.00110)
Constant	2.278***	2.036***	2.319***	2.301***
	(0.153)	(0.139)	(0.177)	(0.135)
Observations	200	200	200	200
Industries	25	25	25	25

Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1
Dependent variable: Value Added.
Year dummies omitted from results.

Chapter 5

Conclusion

To further the understanding of how IT creates value, this research examines its role in the innovation context and presents three major findings. First, we find that IT contributes positively to innovation-related knowledge production at the firm level, as evidenced by a positive elasticity to patent output. Second, expanding on this, we find the IT-augmented knowledge production process contributes to the overall production process. Third, we find the direct and indirect effects of IT are present at the industry level. Importantly, the indirect effect is synergistic with innovation-related investments, indicating an innovation-enabling role for IT. These findings bring to light a new role for IT in value-creation. In the context of overall IT business value research, our work complements other researchers' efforts to quantify the intangible value arising from IT investments, such as increased product variety and strategic flexibility. We also contribute a new IT-centred dimension to research concerning innovation and productivity.

The research in this work draws on approaches well-established in productivity and innovation literature. Our patent and production data are obtained from US government agencies, and are thus prepared using transparent procedures, and have been scrutinized by auditors. The patent data is quality-adjusted by citation-weighting to distinguish between incremental and breakthrough innovations. However, some limitations must be noted. First, our data is limited in detail, size and scope by the sources from which we have obtained it. Regrettably, both input and output measures of innovation are limited to formal innovation activities. Further, there is insufficient detail on the composition of R&D by type of input (labour, capital, IT capital). Our analysis approaches also introduce limitations. In particular, the SEM approach has no ability to control for fixed effects or panel error structures, both of which feature in the literature relating to IT productivity. On the other hand, regression approaches (such as those used in chapters 2 and 4) cannot prove causality and have limitations with respect to small sample sizes and the influence of unobservable effects. We have, where possible, taken measures to mitigate these shortcomings, but prudence demands we interpret the results with some caution. Finally, innovation research has noted the influence of location-specific factors, such as government R&D incentives and the availability of research-related skilled labour and technology. Thus, there are limitations to the generalizability of our results to other countries, given our sample data draws exclusively from US firms and industries.

We believe this research creates a foundation upon which future research into the IT-innovation relationship may build. Improved measures, such as greater detail and scope of innovation outputs (including informal innovation), will become available with expanded industrial R&D survey programs of the National Science Foundation and statistical agencies in other countries. Although it is challenging to keep pace with the changes IT has introduced to innovation (e.g. open-source software and crowd-sourced innovation competitions), the increased availability of some kinds of transactional and network-oriented data offer some promise to future research endeavours. Another avenue for extending this research involves the examination of intra-industry spillovers of innovation, which could explain some of the industry heterogeneity with respect to innovation dynamics.

Bibliography

Acs, Zoltan J., David B. Audretsch. 1991. Innovation and technical change: An overview. Zoltan J. Acs, David B. Audretsch, eds., *Innovation and Technical Change: An International Comparison*. University of Michigan Press, 1–23.

Agarwal, R., B.L. Bayus. 2002. The market evolution and sales takeoff of product innovations. *Management Science* **48**(8) 1024–1041.

Agrawal, Ajay, Avi Goldfarb. 2008. Restructuring research: Communication costs and the democratization of university innovation. *American Economic Review* **98**(4) 1578–1590.

Allred, B. B., K. S. Swan. 2005. The mediating role of innovation on the influence of industry structure and national context on firm performance. *Journal of International Management* **11**(2) 229–252.

Arora, A., M. Ceccagnoli. 2006. Patent protection, complementary assets, and firms' incentives for technology licensing. *Management Science* **52**(2) 293–308.

Ba, S., P. Pavlou. 2002. Evidence of the trust building technology in electronic markets: Price premiums and buyer behavior. *MIS Quarterly* **26**(3) 243–268.

Baily, M.N., A.K. Chakrabarti. 1988. *Innovation and the Productivity Crisis*. The Brookings Institution, Washington, D.C., Washington, D.C.

Balachandra, R., J.H Friar. 1997. Factors for success in R&D projects and new product innovation: a contextual framework. *IEEE Transactions on Engineering Management* **44**(3) 276–287.

Bardhan, Indranil, Vish V. Krishnan, Shu Lin. 2010. Complementarity Effects of R&D and Information Technology on Firm Market Value. Workshop on Information Systems Economics, December 11-12, 2010, St. Louis, MO.

Bartel, Ann P., Casey Ichniowski, Kathryn L. Shaw. 2007. How Does Information Technology Really Affect Productivity? Plant-Level Comparisons of Product Innovation, Process Improvement and Worker Skills. *Quarterly Journal of Economics* **122**(4) 1721–1758.

Bartholomew, D. 2005. Manufacturers nibbling on PLM. *Industry Week* **254**(4) 63.

Barua, Anitesh, Charles H. Kriebel, Tridas Mukhopadhyay. 1995. Information technologies and business value: An analytic and empirical investigation. *Information Systems Research* **6**(1) 3–23.

Baum, C. F., M. E. Schaffer, S. Stillman. 2003. Instrumental variables and GMM: Estimation and testing. *Stata Journal* **3**(1) 1–31.

Beck, Nathaniel, Jonathan N. Katz. 1995. What to do (and not to do) with time-series cross-sectional data. *American Political Science Review* **89**(3) 634–647.

Berndt, E. R., N.J. Rappaport. 2001. Price and quality history of desktop and mobile computers: A quarter-century historical overview. *American Economic Review* **91**(2) 268–273.

Brennan, A., L Dooley. 2005. Networked creativity: a structured management framework for stimulating innovation. *Technovation* **25** 13881399.

Bresnahan, Timothy F., Erik Brynjolfsson, Lorin M. Hitt. 2002. Information technology, workplace organization, and the demand for skilled labor: Firm-level evidence. *Quarterly Journal of Economics* **117**(1) 339–376.

Brockhoff, K. 1992. R&D cooperation between firms—A perceived transaction cost perspective. *Management Science* **38**(4s) 514–524.

Brown, S.L., K.M. Eisenhardt. 1995. Product development: Past research, present findings, and future directions. *Academy of Management Review* **20**(2) 343–378.

Brynjolfsson, Erik. 1993. The productivity paradox of information technology. *Communications of the ACM* **36**(12) 66–77.

Brynjolfsson, Erik, Lorin M. Hitt. 1995. Information technology as a factor of production: the role of differences among firms. *Economics of Innovation and New Technology* **3** 183–199.

Brynjolfsson, Erik, Lorin M. Hitt. 2000. Beyond Computation: Information Technology, Organizational Transformation and Business Performance. *Journal of Economic Perspectives* **14** 23–48.

Brynjolfsson, Erik, Lorin M. Hitt. 2003. Computing productivity: Firm-level evidence. *Review of Economics and Statistics* **85**(4) 793–808.

Brynjolfsson, Erik, Lorin M. Hitt, Shinkyu Yang. 2002. Intangible Assets: Computers and Organizational Capital. *Brookings Papers on Economic Activity* **2002**(1).

Cassiman, Bruno, Reinhilde Veugelers. 2006. In search of complementarity in innovation strategy: Internal R&D and external knowledge acquisition. *Management Science* **52**(1) 68–82.

Cenfetelli, Ronald T., Geneviève Bassellier. 2009. Interpretation of Formative Measurement in Information Systems Research. *MIS Quarterly* **33**(4) 689–708.

Chan, Tat, Jack A. Nickerson, Hideo Owan. 2007. Strategic management of R&D pipelines with co-specialized investments and technology markets. *Management Science* **53**(4) 667–682.

Cheng, Zhuo June, Barrie R. Nault. 2007. Industry Level Supplier-Driven IT Spillovers. *Management Science* **53**(8) 1199–1216.

Cheng, Zhuo June, Barrie R. Nault. 2011. Relative Industry Concentration and Customer-Driven IT Spillovers. *Information Systems Research* Published in Articles in Advance, April 8, 2011.

Chin, W.W. 1998. The partial least squares approach for structural equation modelling. G.A. Marcoulides, ed., *Modern Methods for Business Research*. Lawrence Erlbaum Associates, 295–336.

Christensen, J., M. Olesen, J. Kjaer. 2005. The industrial dynamics of open innovation—evidence from the transformation of consumer electronics. *Research Policy* **34**(10) 1533–1549.

Chwelos, Paul, Ronald V. Ramirez, Kenneth L. Kraemer, Nigel Melville. 2010. Does Technological Progress Alter the Nature of Information Technology as a Production Input? New Evidence and New Results. *Information Systems Research* **21**(2) 392–408.

Cockburn, Iain M., Rebecca M. Henderson. 1998. Absorptive capacity, coauthoring behavior, and the organization of research in drug discovery. *Journal of Industrial Economics* **46**(2) 157182.

Cohen, Wesley M., Steven Klepper. 1996. Firm Size and the Nature of Innovation within Industries: The Case of Process and Product R&D. *The Review of Economics and Statistics* **78**(2) 232–243.

Cohen, Wesley M., Daniel A. Levinthal. 1990. Absorptive capacity: A new perspective on learning and innovation. *Administrative Science Quarterly* **35**(1) 128–152.

Council of Economic Advisors. 2001. *The Economic Report of the President*, chap. The annual report of the Council of Economic Advisors. U.S. Government Printing Office, Washington, DC.

Crepon, B., E. Duguet, J. Mairesse. 1998. Research, innovation and productivity: An econometric analysis at the firm level. *Economics of Innovation and New Technology* **7** 115–158.

David, Paul A. 1990. The Dynamo and the Computer: An Historical Perspective on the Modern Productivity Paradox. *American Economic Review* **80**(2) 355–361.

Davidson, R., J. G. MacKinnon. 1993. *Estimation and Inference in Econometrics*. Oxford University Press, New York.

Dedrick, J., V. Gurbaxani, K.L. Kraemer. 2003. Information technology and economic performance: A critical review of the empirical evidence. *ACM Computing Surveys* **35**(1) 1–28.

Dewan, Sanjeev, Chung K. Min. 1997. The Substitution of Information Technology for Other Factors of Production: A Firm Level Analysis. *Management Science* **43**(12) 1660–1675.

Dibrell, Clay, Peter S. Davis, Justin Craig. 2008. Fueling innovation through information technology in smes. *Journal of Small Business Management* **46**(2) 203–218.

Dodgson, Mark, David Gann, Ammon Salter. 2006. The role of technology in the shift towards open innovation: The case of Procter & Gamble. *R&D Management* **36**(3) 333–346.

Dougherty, D., C. Hardy. 1996. Sustained product innovation in large, mature organizations: Overcoming innovation-to-organization problems. *Academy of Management Journal* **39**(5) 1120–1153.

Elliott, M. 2006. Integrating the drug discovery laboratory. *Scientific Computing* **May** 16–18.

Enkel, Ellen, Oliver Gassmann, Henry Chesbrough. 2009. Open R&D and open innovation: Exploring the phenomenon. *R&D Management* **39**(4) 311–316.

Fleming, L., O. Sorenson. 2003. Navigating the technology landscape of innovation. *Sloan Management Review* **44**(2) 15–23.

Gao, Guodong (Gordon), Lorin M. Hitt. 2011. IT and Product Variety: Evidence from Panel Data. *Management Science* Accepted for publication, 2011.

Garcia, R., R. Calantone. 2002. A critical look at technological innovation typology and innovativeness terminology: a literature review. *Journal of Product Innovation Management* **19** 110–132.

Gilbert, J., P. Henske, A. Singh. 2003. Rebuilding big pharma's business model. *In Vivo: The Business and Medicine Report* **21**(10). Article 2003800191.

Gordon, R. J. 2000. Does the "new economy" measure up to the great inventions of the past? *Journal of Economic Perspectives* **14**(4) 49–74.

Gordon, S., M. Tarafdar, R. Cook, R. Maksimoski, B. Rogowitz. 2008. Improving the front end of innovation with information technology. *Research Technology Management* **51**(3) 50–58.

Griliches, Zvi. 1979. Issues in assessing the contribution of research and development to productivity growth. *Bell Journal of Economics* **10**(1) 92–116.

Griliches, Zvi. 1981. Market Value, R&D, and Patents. *Economic Letters* **7** 183–187.

Griliches, Zvi. 1988. *Technology, Education and Productivity*. Basil Blackwell, New York, NY.

Griliches, Zvi. 1990. Patent statistics as economic indicators: A survey. *Journal of Economic Literature* **28**(4) 1661–1707.

Griliches, Zvi. 1992. The Search for R&D Spillovers. *Scandinavian Journal of Economics* **94** S29–S47. Supplement. Proceedings of a Symposium on Productivity Concepts and Measurement Problems: Welfare, Quality and Productivity in the Service Industries (1992).

Griliches, Zvi. 1994. Productivity, R&D, and the Data Constraint. *American Economic Review* **84**(1) 1–23.

Griliches, Zvi, Frank Lichtenberg. 1984. Interindustry technology flows and productivity growth: A reexamination. *Review of Economics and Statistics* **66**(2) 324–329.

Hall, Bronwyn H. 2000. Innovation and market value. R. Barrell, G. Mason, M. O'Mahoney, eds., *Productivity, Innovation and Economic Performance*. Cambridge University Press, Cambridge, England.

Hall, Bronwyn H. 2005. Measuring the Returns to R&D: The Depreciation Problem. *Annales D'Économie et de Statistique* **2005**(79-80) 341–381. Special Issue: Contributions in memory of Zvi Griliches.

Hall, Bronwyn H., Zvi Griliches, Jerry A. Hausman. 1986. Patents and R&D: Is There a Lag? *International Economic Review* **27**(2) 265–283.

Hall, Bronwyn H., Adam B. Jaffe, Manuel Trajtenberg. 2002. The NBER Patent-Citations Data File: Lessons, Insights, and Methodological Tools. A. Jaffe, M. Trajtenberg, eds., *Patents, Citations and Innovations*. The MIT Press, Cambridge, MA, 403–460.

Hall, Bronwyn H., Adam B. Jaffe, Manuel Trajtenberg. 2005. Market value and patent citations. *RAND Journal of Economics* **36**(1) 16–38.

Hall, Bronwyn H., J. Mairesse, P. Mohnen. 2009. Measuring Returns to R&D. NBER working paper W15622.

Hall, Bronwyn H., Jacques Mairesse. 1995. Exploring the relationship between R&D and productivity in French manufacturing firms. *Journal of Econometrics* **65** 263–293.

Hall, Bronwyn H., Rosemary Ziedonis. 2001. The patent paradox revisited: An empirical study of patenting in the u.s. semiconductor industry, 1979–1995. *RAND Journal of Economics* **32**(1) 101–128.

Hamermesh, Daniel, Sharon M. Oster. 2002. Tools or toys? The impact of high technology on scholarly productivity. *Economic Inquiry* **40**(4) 539–555.

Han, Kunsoo, Robert Kauffman, Barrie R. Nault. 2010. Returns to information technology outsourcing. *Information Systems Research* **Forthcoming**. Accepted for publication February 12, 2010.

Han, Shu, T. Ravichandran. 2006. Does IT impact firm innovativeness: An empirical examination of complementary and direct effects. *Proceedings of the Twelfth Americas Conference on Information Systems (AMCIS), Acapulco, Mexico*. 704–715.

Hatch, N. W., D. C. Mowery. 1998. Process innovation and learning by doing in semiconductor manufacturing. *Management Science* **44**(11) 1461–1477.

Henderson, Rebecca M., Iain M. Cockburn. 1994. Measuring competence? Exploring firm effects in pharmaceutical research. *Strategic Management Journal* **15**(S1) 63–84.

Hitt, Lorin M., Prasanna Tambe. 2006. Measuring spillovers from information technology investments. *Twenty-Seventh International Conference on Information Systems*.

Horn, P. M. 2005. The changing nature of innovation. *Research Technology Management* **48**(6) 28–33.

Howells, Jeremy. 1999. Research and technology outsourcing. *Technology Analysis & Strategic Management* **11**(1) 17–29.

Hu, Qing, Jing Quan. 2005. Evaluating The Impact Of IT Investments On Productivity: A Causal Analysis At Industry Level. *International Journal of Information Management* **25**(1).

Hulten, Charles R. 1992. Growth accounting when technical change is embodied in capital. *American Economic Review* **82**(4) 964–980.

Istook, C. 2000. Rapid prototyping in the textile & apparel industry: A pilot project. *Journal of Textile and Apparel Technology and Management* **1**(1).

Jorgenson, D. W. 2001. Information technology and the U.S. economy. *American Economic Review* **91**(1) 1–32.

Jorgenson, D. W., Z. Griliches. 1967. The explanation of productivity change. *Review of Economic Studies* **34**(3) 249–283.

Jorgenson, Dale W., Mun S. Ho, Kevin J. Stiroh. 2008. A Retrospective Look at the U.S. Productivity Growth Resurgence. *Journal of Economic Perspectives* **22**(1) 3–24.

Kleis, Landon, Paul D. Chwelos, Ronald V. Ramirez, Iain M. Cockburn. 2011. Information Technology and Intangible Output: The Impact of IT Investment on Innovation Productivity. *Information Systems Research* Published in Articles in Advance, April 8, 2011.

Kleis, Landon, Paul D. Chwelos, Ronald V. Ramirez, Kenneth L. Kraemer. 2003. Information technology and innovation: The impact of IT investment on intangible outputs. Workshop on Information Systems Economics (WISE), Seattle, WA.

Kohli, R., S. Devaraj. 2003. Measuring information technology payoff: A meta-analysis of structural variables in firm-level empirical research. *Information Systems Research* **14**(2) 127–145.

Konicki, S. 2002. Revving up. *InformationWeek* **891** 36.

Kortum, Samuel, Josh Lerner. 1998. Stronger protection or technological revolution: What is behind the recent surge in patenting? *Carnegie-Rochester Conference Series on Public Policy* **48** 247–304.

Kremp, E., J. Mairesse. 2004. Knowledge management, innovation and productivity: A firm level exploration based on french manufacturing cis3 data. *NBER Working Paper 10237* .

Kumar, K., H. G. van Dissel. 1996. Sustainable collaboration: Managing conflict and cooperation in interorganizational systems. *MIS Quarterly* **20**(3) 279–300.

Lee, H., B. Choi. 2003. Knowledge management enablers, processes, and organizational performance: An integrative view and empirical examination. *Journal of Management Information Systems* **20**(1) 179–228.

Lee, S. T., R. Gholami, T. Y. Tong. 2005. Time series analysis in the assessment of ict impact at the aggregate level - lessons and implications for the new economy. *Information & Management* **42**(7) 1009–1022.

Lichtenberg, Frank R., Donald Siegel. 1991. The Impact of R&D Investment on Productivity–New Evidence Using Linked R&D–LRD Data. *Economic Inquiry* **29**(2) 203–229.

Los, Bart, Bart Verspagen. 2002. R&D spillovers and productivity: Evidence from U.S. manufacturing microdata. *Empirical Economics* **25** 127–148.

Lunn, J. 1987. An empirical analysis of firm process and product patenting. *Applied Economics* **19** 743–751.

Madsen, J. B. 2008. Semi-endogenous versus Schumpeterian growth models: Testing the knowledge production function using international data. *Journal of Economic Growth* **13**(1) 1–26.

Mairesse, Jacques, Pierre Mohnen. 2005. The importance of R&D for innovation: a reassessment using French survey data. *Journal of Technology Transfer* **30**(1-2) 183–197.

Mairesse, Jacques, Mohamed Sassenou. 1991. R&D Productivity: A Survey of Econometric Studies at the Firm Level. NBER working paper W3666.

Majchrzak, Ann, Lynne P. Cooper, Olivia E. Neece. 2004. Knowledge reuse for innovation. *Management Science* **50**(2) 174–188.

Malhotra, Arvind, Ann Majchrzak, Robert Carman, Vern Lott. 2001. Radical Innovation without Collocation: A Case Study at Boeing-Rocketdyne. *MIS Quarterly* **25**(2) 229–249.

Mead, Charles Ian. 2007. R&D Depreciation Rates in the 2007 R&D Satellite Account. Tech. rep., Bureau of Economic Analysis.

Melville, Nigel, Vijay Gurbaxani, Kenneth Kraemer. 2007. The productivity impact of information technology across competitive regimes: The role of industry concentration and dynamism. *Decision Support Systems* **43**(1).

Melville, Nigel, Kenneth Kraemer, Vijay Gurbaxani. 2004. Information Technology And Organizational Performance: An Integrative Model Of IT Business Value. *MIS Quarterly* **28**(2).

Mithas, Sunil, Mayuram S. Krishnan, Claes Fornell. 2005a. Effect of information technology investments on customer satisfaction: Theory and evidence. Ross School of Business Working Paper No. 971. Available at SSRN: http://ssrn.com/abstract=901643.

Mithas, Sunil, Mayuram S. Krishnan, Claes Fornell. 2005b. Why do customer relationship management applications affect customer satisfaction? *Journal of Marketing* **69**(4) 201–209.

Mittal, Neeraj, Barrie R. Nault. 2009. Investments in information technology: Indirect effects and information technology intensity. *Information Systems Research* **20**(1) 140–154.

Mohammad, Hasan Qurban, Dawei Zhang, Zhuo June Cheng, Barrie R. Nault. 2009. On the Relationship of Information Technology with Other Inputs. Working Paper, University of Calgary.

Nadiri, M. Ishaq. 1993. Innovations and technical spillovers. NBER working paper W4423.

Nambisan, S. 2003. Information systems as a reference discipline for new product development. *MIS Quarterly* **27**(1) 1–18.

Narayanan, V., F. Douglas, D. Schirlin, G. Wess, D. Geising. 2004. Virtual communities as an organizational mechanism for embedding knowledge in drug discovery: The case of chemical biology platform. *Journal of Business Chemistry* **1**(2) 37–47.

Narver, J. C., S. F. Slater, D. L. MacLachlan. 2004. Responsive and proactive market orientation and new-product success. *Journal of Product Innovation Management* **21**(5) 334–347.

National Science Foundation. 2009. Research and Development in Industry: 2005. Detailed Statistical Tables NSF 10-319. URL `http://www.nsf.gov/statistics/nsf10319/`. Division of Science Resources Statistics.

Nellore, R., R. Balachandra. 2001. Factors influencing success in integrated product development (ipd) projects. *IEEE Transactions on Engineering Management* **48**(2) 164–174.

Nelson, Richard R. 1982. The Role of Knowledge in R&D Efficiency. *Quarterly Journal of Economics* **97** 453–470.

Nerkar, Atul, Srikanth Paruchuri. 2005. Evolution of R&D capabilities: The role of knowledge networks within a firm. *Management Science* **51**(5) 771–785.

Oliver, J. R. 2003. Accounting and tax treatment of R&D: An update. *The CPA Journal* **73**(7) 46–49.

Owen-Smith, J., W. W. Powell. 2003. The expanding role of university patenting in the life sciences: Assessing the importance of experience and connectivity. *Research* **32**(9) 16951711.

Pakes, A., Z. Griliches. 1984. Patents and R&D at the firm level: A first look. Z. Griliches, ed., *R&D, Patents, and Productivity*. University of Chicago Press, Chicago, 55–72.

Paulraj, A., A. Lado, I. Chen. 2008. Inter-organizational communication as a relational competency: Antecedents and performance outcomes in collaborative buyersupplier relationships. *Journal of Operations Management* **26**(1) 45–64.

Pisano, Gary P. 1990. The R&D Boundaries Of The Firm: An Empirical Analysis. *Administrative Science Quarterly* **35**(1) 153–176.

Quinn, J. B., J. J. Baruch, K. A. Zien. 1997. *The Innovation Explosion : using intellect and software to revolutionize growth strategies*. Free Press, New York.

Ramirez, Ronald V., Landon Kleis. 2010. The firm innovation process: How critical is information technology? Working paper, University of Colorado Denver, Denver, CO.

Rice, R. E. 1994. Relating electronic mail use and network structure to R&D work networks and performance. *Journal of Management Information Systems* **11**(1) 9–29.

Ringle, Christian Marc, Sven Wende, Alexander Will. 2005. Smartpls release 2.0 (beta). University of Hamburg, Hamburg, Germany. URL `http://www.smartpls.de`.

Roberts, P. W. 1999. Product innovation, productmarket competition and persistent profitability in the u.s. pharmaceutical industry. *Strategic Management Journal* **20** 655670.

Rothwell, R. 1994. Towards the fifth generation innovation process. *International Marketing Review* **11**(1) 7–31.

Ryssel, R., T. Ritter, H. Gemunden. 2004. The impact of information technology deployment on trust, commitment and value creation in business relationships. *Journal of Business and Industrial Marketing* **19**(3) 197–207.

Sambamurthy, V., A. Bharadwaj, V. Grover. 2003. Shaping agility through digital options: Reconceptualizing the role of information technology in contemporary firms. *MIS Quarterly* **27**(2) 237–263.

Sangiovanni-Vincentelli, A. L. 2003. Electronic-system design in the automobile industry. *IEEE Micro* **21**(3) 8–18.

Scherer, Frederic M. 1982. Inter-industry technology flows and productivity growth. *Review of Economics and Statistics* **64**(4) 627–634.

Schilling, M., C. Hill. 1998. Managing the new product development process: Strategic imperatives. *Academy of Management Executive* **12**(3) Schilling, M., C. Hill. 1998. Managing the new product development process: Strategic imperatives. Acad. Management Executive 12(3) 6781.

Scotchmer, S. 2004. *Innovation and incentives. MIT Press, Cambridge.*. MIT Press, Cambridge.

Scott, J. 2000. Facilitating interorganizational learning with information technology. *Journal of Management Information Systems* **17**(2) 81–113.

Siegel, Donald. 1994. Errors in output deflators revisited: Unit values and the producer price index. *Economic Inquiry* **32**(1) 11–32.

Siegel, Donald. 1997. The impact of computers on manufacturing productivity growth: A multiple-indicators, multiple-causes approach. *Review of Economics and Statistics* **79**(1) 68–78.

Stiroh, Kevin J. 2002. Are ICT spillovers driving the new economy? *Review of Income and Wealth* **48**(1) 33–57.

Sudarsan, Rachuri, Steven J. Fenves, Ram D. Sriram. 2005. A product information modeling framework for product lifecycle management. *Computer-Aided Design* **37**(13) 1399–1411.

Swink, M. 2006. Building collaborative innovation capability. *Research Technology Management* **49**(2) 37–47.

Tanriverdi, Huseyin. 2005. Information technology relatedness, knowledge management capability, and performance of multibusiness firms. *MIS Quarterly* **29**(2) 311–324. Special Issue on Information Technologies and Knowledge Management.

Tatikonda, M. V., S. R. Rosenthal. 2000. Successful execution of product development projects: Balancing firmness and flexibility in the innovation process. *Journal of Operations Management* **18**(4) 401–425.

Teresko, J. 2004. P&G's secret: Innovating innovation. *Industry Week* **253**(12) 26.

Thomke, S. H. 2006. Capturing the real value of innovation tools. *MIT Sloan Management Review* **47**(2) 24–32.

Thomke, Stefan H. 1998. Managing experimentation in the design of new products. *Management Science* **44**(6) 743–762.

Thompson, Peter, Melanie Fox-Kean. 2005. Patent citations and the geography of knowledge spillovers: A reassessment. *American Economic Review* **95**(1) 450–460.

Trajtenberg, M. 1990. A penny for your quotes: Patent citations and the value of innovations. *RAND Journal of Economics* **21**(1) 172–187.

USPTO, United States Patent and Trademark Office. 2006. General information concerning patents. URL `http://www.uspto.gov/web/offices/pac/doc/general/index.html`.

Weill, Peter. 1992. The relationship between investment in information technology and firm performance: A study of the valve manufacturing sector. *Information Systems Research* **3**(4) p307 – 333.

Zahay, Debra, Abbie Griffin, Elisa Fredericks. 2004. Sources, uses, and forms of data in the new product development process. *Industrial Marketing Management* **33**(7) 657–666.