2015-01-29

Studies on IT, Logistics, and the Structure of Production

Gong, Fengmei

UNIVERSITY OF CALGARY

Studies on IT, Logistics, and the Structure of Production

by

Fengmei Gong

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE

DEGREE OF DOCTOR OF PHILOSOPHY

HASKAYNE SCHOOL OF BUSINESS

CALGARY, ALBERTA

January, 2015

Abstract

Information Technology (IT) has changed how firms and industries run their businesses and how they organize production. This thesis examines the relationships between IT and three important aspects of production organization: the usage of logistics outsourcing, the interdependence with upstream suppliers for intermediate inputs, and the structure of production in an economy. The first essay examines whether the advent of the Internet coincided with a move to the market in one of the most connected industries in the economy: logistics. We find that the effects of IT on outsourced logistics have changed with the advent of the Internet. The second essay examines the impact of an industry's IT investment on its production interdependence with upstream suppliers, where we measure interdependence as direct backward linkage (DBL), and examines the relationship among DBL, total factor productivity (TFP), and value-added. We find that an industry's IT investment reduces its production interdependence with suppliers and leads to greater value-added. The third essay explores the relationship between IT and the structure of production in an economy. We take a unique perspective, network analysis, to generate a variety of measures of the structure of production which we categorize as connectivity among industries in an ego-centric network and concentration in an ego industry's supplying market. It is found that an industry's IT investment is associated with an increase in the connectivity within its supplying network and a decrease in concentration in the supplying market.

Acknowledgements

I would like to express my special appreciation and thanks to my advisor Dr. Barrie R. Nault. I would like to thank you for your guidance, encouragement, support and for being patient to allow me to grow as a better scholar. Your advice on my research and career has been priceless. I would also like to thank my committee members, Dr. Victoria L. Mitchell, Dr. Byungwan Koh, Dr. Deyi Xue, and Dr. Yuliang (Oliver) Yao for serving as my committee members. I thank the Robson Endowment and the Informatics Research Centre in the Haskayne School of Business at the University of Calgary and the Canadian Centre for Advanced Supply Chain Management and Logistics (C-ASL) for their support. I would like to thank Jeanette Burman for editing assistance.

I would like to give my special thanks to my family. Words cannot express how grateful I am to my mother-in law, father-in-law, my mother, and father for all of help and support you have given to me. I would also like to thank all of my friends who encouraged me to strive towards my goal. At the end I would like to express appreciation to my beloved husband Ming Zhou who took a lot of family responsibilities and my lovely son Charlie Zhou whose smiles always encouraged me to work hander.

Table of Contents

List of Figures and Illustrations

Chapter 1

INTRODUCTION

In order to improve production efficiency and create more business value, industries have been increasingly investing in information technology (IT). For example, IT investments in the manufacturing sector climbed steadily from $60.989 billion in 1987, and reached $185.179 billion in 2012 (measured in 2009 dollars), according to the U.S. Bureau of Labor Statistics (BLS). These IT investments have changed how firms and industries run their businesses and how they organize production.

Drawing on production theory, transaction costs analysis, growth accounting theory and by employing econometric analyses, this thesis examines the relationships between IT and three important aspects of production organization: the usage of logistics outsourcing, the interdependence with upstream suppliers for intermediate inputs, and the structure of production in an economy. It includes three essays, each focusing on the impact of IT on one aspect of production organization related to intermediate inputs. Fengmei Gong has done the majority of work and is the primary contributor to all three essays in this dissertation. Permissions have been given from the co-authors to include the three essays in this dissertation.

In the first essay, we collapse intermediate inputs into outsourced logistics and other intermediate inputs as well as examine the relationship between IT and outsourced logistics in a production function.[1] Logistics outsourcing, which refers to the use of external organizations to perform logistics activities that alternatively could be conducted by firms themselves, plays an important role in logistics management (Bowersox 1990, Lieb 1992).

[1]The original paper, titled "Information Technology, Logistics and Productivity" by Fengmei Gong, Barrie R. Nault, and Mohammad S. Rahman, was presented by Fengmei Gong at the 2011 INFORMS Annual Meeting, Charlotte, NC, U.S., November 14, 2011. The paper titled "An Internet-Enabled Move to the Market" was presented by Fengmei Gong at POMS 25th Annual Conference, Atlanta, GA, U.S., May 10, 2014

Over the past few decades, logistics outsourcing has grown substantially. For example, in U.S. manufacturing industries, one of the largest users of logistics outsourcing, the usage rate of logistics outsourcing remained stable in the early 1990s but increased sharply after the commercialization of the Internet. Based on transaction cost analysis, our study examines whether the advent of the Internet coincided with a move to the market in one of the most connected industries in the economy: logistics.

To show evidence of the move to the market for logistics activities, we take a novel approach using production theory—an approach that allows us to model the production of output in terms of capital, labor, and purchased goods and services. These latter goods and services are obtained by firms from the market rather than generated within the organization. Following our observation of an empirical regularity whereby after the advent of the Internet a substantially greater proportion of logistics was obtained from the market as intermediate inputs, that is, outsourced logistics, we examine the relationship between IT capital and outsourced logistics in a production model.

We create two datasets by employing the input-output (I-O) tables and industry productivity data, and estimate the relationship between an industry's own IT investment and outsourced logistics, and simultaneously estimate the relationship between the given industry's customers' IT investments and outsourced logistics before and after the commercialization of the Internet. We find that the relationship between IT investments and outsourced logistics changed after the commercialization of the Internet. As measured through our variable coefficient model, before the commercialization of the Internet, IT investments and outsourced logistics were substitutes, and they became complements after the Internet was widely used.

In the second essay, we examine the impact of an industry's IT investment on its production interdependence with upstream suppliers for intermediate inputs, where we measure interdependence as direct backward linkage (DBL), and examine the relationship among DBL, total factor productivity (TFP), and value-added.[2] Production interdependence be-

[2]The original paper, titled "IT Investments and the Structure of Production" by Fengmei Gong and

tween industries is demonstrated by inter-industry purchases and sales of goods and services (Santhanam and Patil 1972). As a widely used measure of production interdependence, DBL is defined as the amount of intermediate inputs required by a given industry to produce one unit of output (Miller and Blair 2009). IT has improved information sharing and coordination within and between industries. As a result, industries have become increasingly integrated with their suppliers' business processes such as purchasing and Just-in-time (JIT) production. However, whether industries in a supply chain have become more interdependent in production remains an open question.

Using estimation models that draw on production theory and growth accounting theory, we estimate the impact of IT on DBL and the impact of DBL on value-added through TFP based on two industry-level datasets in the U.S. The key result we find is that an industry's IT investment reduces its production interdependence with suppliers and leads to greater value-added. Specifically, an industry's IT increases the efficiencies of intermediate inputs from suppliers, resulting in a reduction in DBL; the reduction indicates a shift in production frontier towards more efficient production and leads to growth in TFP and greater value-added.

In the third essay, we examine the relationship between IT and the change in the structure of production.[3] In an economic network formed by industries that are connected in production processes, intermediate inputs flowing from upstream suppliers to an industry are transformed as products by the given industry and then flow downstream customer industries. The input-output relationships among industries in the economic network form the structure of production in the economy. We have seen the changes in the structure of production: some ties between industries are absent, some new ties are present, and some industries reallocate the amounts of inputs from supplier industries. Along with these changes, we have

Barrie R. Nault, was presented by Fengmei Gong at the 2012 INFORMS Annual Meeting, Phoenix, AZ, U.S., October 14, 2012.

[3]The original paper was titled as "The Impact of IT on the Structure of Production—A Network Analysis Perspective" by Fengmei Gong, Byungwan Koh, and Barrie R. Nault, and was presented by Fengmei Gong at the 2014 INFORMS Annual Meeting, San Francisco, CA, U.S., November 9, 2014.

observed a steadily increasing trend in IT investments of industries, and the deployment of IT has greatly reduced transaction costs in the economy. We are interested in whether the increase in IT investment is associated with the change in the structure of production in the economy.

To address this question, we employ economy-wide data in an exploratory study and take a unique perspective—network analysis, to generate a variety of measures of the structure of production which we categorize as connectivity among industries in an egocentric network and concentration in an ego-industry's supplying market. Next, we develop estimation models to examine the relationship between IT and those measures of the structure of production. We find that IT investment of an industry is associated with an increase in the connectivity within its supplying network and a decrease in concentration in the supplying market where it purchases intermediate inputs. It suggests that with an increase in IT, an industry is more likely to act as a hub or pivot that transmits goods and services along supply chains; and that the more IT an industry invests in, the less concentrated is the direct supplying market. We also find that IT may affect connectivity and concentration of manufacturing industries differently.

The remainder of the thesis is organized as follows: Chapter 2 presents the relationship between IT and outsourced logistics; Chapter 3 examines the impact of an industry's IT investment on its production interdependence with upstream suppliers; Chapter 4 explores whether the increase in IT investment is associated with the change in the structure of production in the economy; and Chapter 5 concludes the thesis with a brief summary.

Chapter 2

AN INTERNET-ENABLED MOVE TO THE MARKET

2.1 Introduction

Logistics is a process of planning, implementing, and controlling the effective and efficient flow of goods, services, and information between the point of origin and the point of consumption (Cooper et al. 1997, Rutner and Langley Jr. 2000). Its activities typically include inbound and outbound transportation, warehousing, inventory management, etc. It is an important part of the supply chain process because it is through logistics that raw materials flow into production and finished products are delivered to customers (Bowersox et al. 2010). Logistics are costly: in 2011, logistics expenditure was 8.5% of GDP, reaching about $1.28 trillion (Wilson 2012). Thus, by managing logistical processes efficiently and effectively, firms can achieve a competitive advantage.

Logistics outsourcing, which refers to the use of external organizations to perform logistics activities that alternatively could be conducted by firms themselves, plays an important role in logistics management (Bowersox 1990, Lieb 1992). For example, outsourced logistics expenditure accounted for on average 42% of the total logistics expenditure in 2011 (Langley and Capgemini 2012). Over the past few decades, logistics outsourcing has grown substantially. For example, in U.S. manufacturing industries—one of the largest users of logistics outsourcing, the usage rate of logistics outsourcing remained stable in the early 1990s but increased sharply after the commercialization of the Internet. In a survey of top-500 manufacturing firms in the U.S., 37% of firms used third-party logistics services (3PLs) in 1991 and 38% in 1994 (Lieb and Randall 1996), reaching 60% in 1995 and climbing to 80%

by 2004 (Lieb and Randall 1996, Ashenbaum et al. 2005). This climb from the mid-1990s to the mid-2000s was reflected in other percentage-using-3PLs surveys and was even more pronounced in estimated expenditures on 3PLs moving from under $31B in 1996 to almost $90B in 2004 (Ashenbaum et al. 2005).

Logistics outsourcing is the result of a make or buy decision made after assessing the production and transaction costs of internal or external provision. The substantial growth of logistics outsourcing starting in 1994–1995 suggests that the transaction costs of outsourcing logistics have fallen relative to internal or in-house logistics, and consequently the trend has been towards external provision. For example, as we can see from the Transportation Satellite Accounts (TSAs) data in Figure 2.1, for-hire or outsourced logistics accounts for around 80% of transportation costs for manufacturing industries in the U.S.[1] Although these transportation costs rose over 2% in real terms between 1992 and 1996, the for-hire component grew by 5% and the in-house component fell by over 8%, consistent with the percentage-using and estimated-expenditures-on 3PLs surveys quoted above.

Transaction costs in the context of logistics outsourcing depend on bounded rationality, opportunism, asset specificity, uncertainty, and frequency (Williamson 1981, 1985, 1989). Bounded rationality and opportunism are two assumptions about decision-makers' behaviors. *Bounded rationality* is the assumption that although decision-makers intend to act rationally, they are constrained by their information processing and communication capability (Williamson 1981). For example, when search costs for information about potential 3PL providers are high, the availability of this information is low and the outsourcers might mistakenly choose a 3PL provider charging high prices while providing low quality services. *Opportunism* is the behavioral assumption that decision makers seek their self-interest given the opportunity. For example, when it is costly for outsourcers to track and trace the shipment status and monitor the on-time delivery performance of carriers, the carriers may take

[1]Noting that, for-fire transportation covers about five modes of transportation, while in-house transportation here covers transportation by truck and bus, so the 80% is higher than the percentage of for-hire transportation on average, for example, 42% as we state above.

	1992		1996		(B-A)/A
	1992 nominal value	1997 real value (A)	1996 nominal value	1997 real value (B)	
Total transportation costs	102054	116747.45	116591	119266.13	0.022
For-hire	80248	91801.88	94275	96438.10	0.051
In-house	21806	24945.57	22316	22828.03	-0.085
For-hire/total transportation	0.786		0.809		
In-house/total transportation	0.214		0.191		
Total output	2,951,303		3,666,001		
Total transportation/output	0.035		0.032		
For-hire/output	0.027		0.026		
In-house/output	0.007		0.006		

Notes: (1) The data about transportation costs and output is from Table 5 and Table 2 of 1996 TSAs (Transportation Satellite Accounts), respectively. (2) For-hire transportation includes railroad and related services, passenger ground transportation, except transit; motor freight transportation and warehousing; water transportation; air transportation; and pipelines, freight forwarders, and related services. In-house transportation includes transportation by truck and bus provided by manufacturing industries for their own use. (3) The values are in millions of dollars at producers' prices. (4) Refer to U.S. Department of Transportation et al. (2011), nominal 1992 dollars and 1996 dollars are adjusted using the current series of consumer price index (CPI) published by the U.S. Bureau of Labor Statistics for all transportation. CPI for all transportation are 140.3, 156.9, and 160.5 in 1992, 1996, and 1997, respectively.

Figure 2.1: Transportation Costs of Manufacturing Industries in 1992 and 1996

advantage of the information asymmetry, such as disguising delay in delivery, and dishonestly reporting loss and damage. In order to monitor and measure carriers' performance, outsourcers would be subjected to substantial transaction costs.

Asset specificity, uncertainty, and frequency are three dimensions of transactions. *Asset specificity* is the degree to which an asset can be redeployed to alternative uses or users, and it can be measured as the difference in value between first-best use and second-best use (McGuigan et al. 2010, Williamson 1989). When asset specificity is high, the continuity of a transaction is important for both sides, resulting in high contractual and organizational safeguarding costs in order to avoid the hazards of opportunism (Williamson 1985). For example, when a 3PL provider sets up specific facilities and equipment and assign specific employees to work for an outsourcer, the premature termination of the relationship may cause a substantial loss for the provider because specialized assets cannot be redeployed without sacrificing productive value; and thus, the provider may require a long-term contract and the investment of the outsourcer in those transaction-specific assets to lock-in the relationship

and reduce risks. *Uncertainty* is the unpredictable changes in circumstances surrounding a transaction. A high level of uncertainty causes adaptation problems when decision-makers are limited by bounded rationality, resulting in costs of communicating new information, renegotiating agreements, and coordination to tackle new situations (Rindfleisch and Heide 1997). For example, at any given time, a manufacturer may have thousands of shipments in-transit, but limited visibility of inventory in-transit causes high uncertainty of product availability. The manufacturer has to communicate with the carriers to obtain the shipping status, and has to keep safety stock to buffer the impact of late delivery. *Frequency* refers to the occurrence of transactions. A high frequency of transactions gives buyers and sellers an opportunity to transfer tacit knowledge and build reputation and trust, resulting in a reduction in transaction costs. For instance, when a firm and a 3PL provider have experienced high volumes of transactions, they expect to behave in a trustworthy way in the future; and thus, the 3PL providers would commit to contract and the firm lowers the costs for monitoring and contractual adaptation.

IT has reduced both market transaction costs, that is, external transaction costs, and internal governance costs. We argue that the Internet reduces market transaction costs relatively more than internal governance costs. First, the Internet changes the behavioral characteristics of decision makers in transactions. The Internet loosens bounded rationality in the context of logistics. For instance, the Internet reduces buyer search costs for products and services (Bakos 1997, Brynjolfsson and Smith 2000), so that buyers of 3PL services can access information about price, capacity, and services at low cost. In addition, the Internet relieves decision makers' concerns about opportunism. For instance, it might be costly or even impossible to check shipment records through traditional EDI (electronic data interchange) for a manufacturer that has thousands of shipments delivered by a variety of carriers because of the low penetration rate of EDI in its trading community. However, through the Internet, a manufacturer can track and trace any shipments real-time and check

the records of all carriers' on-time delivery performance. Thus, opportunism is greatly reduced because of the visibility-in-transit enabled by the Internet.

Second, the Internet reduces transaction costs through the attributes of transactions: asset specificity, uncertainty, and frequency. The Internet reduces asset specificity of 3PL providers by facilitating an online logistics market that helps match logistics capacity and demand, and facilitating resource sharing among 3PL providers (e.g., Nault and Dexter 2006). For instance, online brokers are able to match carrier capacity and shipping demand, enabling logistics service providers to redeploy their human-specific and physical-specific assets over more users and uses to achieve economies of scale and scope. Information sharing among supply chain partners facilitated by Internet-based interorganizational information systems (IOSs) reduces the uncertainty of transactions by improving supply chain visibility and enabling firms to adaptively control logistics activities. For example, when a shipment is delayed and has not reached an expected waypoint, an event management system can notify shippers or customers regarding the exception, and firms can adapt to minimize or prevent potential problems. In addition, the Internet breaks down the financial and technological barriers that restricted the adoption of EDI, facilitating communication between firms and 3PL providers that do not have sufficient transaction volumes to justify the adoption of EDI. Thus, the Internet supports a range of transaction frequency at low communication costs, which greatly reduces the transaction costs of all transactions.

The logic of transaction cost analysis is that if transaction costs are high enough that they exceed the production cost advantages of the market, then firms prefer internal hierarchies, and when transaction costs are lower, firms favor the market (David and Han 2004, Grover and Malhotra 2003, Rindfleisch and Heide 1997, Williamson 1985). As the Internet reduces market transaction costs relatively more than internal governance costs, more firms prefer to outsource logistics and use more outsourced logistics services. By outsourcing logistics, firms avoid holding transportation and warehousing assets, which allows 3PL providers to

hold those assets because they have lower asset specificity as they can redeploy those assets to other customers. In addition, firms can focus on their core competencies to improve productivity and competitive advantages. At the same time, those firms can obtain additional efficiency by exploiting external logistics expertise of 3PL providers (Deepen 2007, Marasco 2008).

Previous studies about the impact of IT on organizational boundaries suggested that IT shaped the change in organizational boundaries. Some studies theoretically argued that IT led to the shift from hierarchies to markets by reducing transaction costs (Malone et al. 1987, Gurbaxani and Whang 1991, Clemons et al. 1993). Providing some indirect evidence of this, Brynjolfsson et al. (1994) found that IT was associated with smaller firm size which the authors suggested was a shift to externally provided services by reducing external coordination costs relative to internal. Hitt (1999) found that IT was associated with the decrease in vertical integration, which suggested that the increase in IT was related to the change in organizational structure towards less in-house making. Although these studies provide important initial insights regarding the impact of nascent IT on organizational boundaries, the fact is that the advent of the Internet radically enhanced IT capabilities which, in turn, fundamentally redefined the role of IT in organizations (Zhu and Kraemer 2002, Zhu 2004). Therefore, it is critical to juxtapose the impact of IT on organizational boundaries before and after the advent of the Internet. Using the reasoning underlying transaction costs, our study examines whether the advent of the Internet—with its universal connectivity and open standards—coincided with a move to the market in one of the most connected industries in the economy: logistics. The potential contribution of our work is direct empirical evidence that IT in its most interconnected instance—the Internet—favors a market form of the provision of goods and services.

To show evidence of this move to the market, we take a novel approach using production theory—an approach that allows us to model the production of output in terms of capital,

labor, and purchased goods and services where these latter goods and services are obtained by firms from the market rather than generated within the organization. Following our empirical regularity whereby after the advent of the Internet a substantially greater proportion of logistics was obtained from the market as intermediate inputs, that is, outsourced logistics, we measure the relationship between IT capital and outsourced logistics in a production model. If our transaction cost logic is correct, then the relationship between IT and outsourced logistics should change with the advent of the Internet.

We create two datasets by employing the input-output (I-O) tables and industry productivity data: one for Standard Industrial Classification (SIC)-defined manufacturing industries from 1987 to 1999, and the other for North American Industrial Classification System (NAICS)-defined manufacturing industries from 2000 to 2008. We estimate the relationship between an industry's own IT investment and outsourced logistics, and simultaneously estimate the relationship between the given industry's customers' IT investments and outsourced logistics, before and after the commercialization of the Internet. We find that the relationship between IT investments and outsourced logistics changed after the commercialization of the Internet. As measured through our variable coefficient model, before the commercialization of the Internet, IT investments and outsourced logistics were substitutes, and they became complements after the Internet was widely used. Together with the growth in the percent of firms using 3PLs, the growth of expenditures in 3PLs, the move from in-house logistics to for-hire logistics, and the fact that there were no substantial regulatory or transportation-related technological changes, this suggests that the Internet is responsible for the shift to the market in the context of logistics.

2.1.1 Prior Research on the Impact of IT on Organizational Boundaries

As noted earlier, prior studies suggested that IT affected organizational boundaries by reducing the costs of coordinating economic activities, and most adopted a transaction cost perspective. Malone et al. (1987) argued that IT would lead to a shift from hierarchies to

markets because IT reduced coordination costs. Gurbaxani and Whang (1991) classified coordination costs into internal and external coordination costs and argued that IT could affect either the horizontal or vertical dimension of firm size by reducing coordination costs. Clemons et al. (1993) argued that IT could reduce coordination costs without increasing the associated transaction risks, leading to more outsourcing but only from a few suppliers. Other studies empirically examined the impact of IT on firm size and vertical integration, providing some indirect evidence of the impact of IT on the shift from hierarchies to markets (note Brynjolfsson et al. (1994) and Hitt (1999) discussed above). In addition, Im et al. (2013) argued for an inverse relationship between IT and firm size, and found a sequential interaction between IT and firm size through transaction costs as a mediator. Their study supported transaction costs theory in terms of interpreting the relationship between IT and organizational changes.

Other studies conceptually analyzed the relationship between IT and logistics outsourcing. For example, Bowersox and Daugherty (1995) argued that information and communication capabilities would encourage firms to develop more external relationships, resulting in the growth in outsourcing and external alliances. Lewis and Talalayevsky (2000) argued that rapid development of IT supported more outsourcing of logistics activities as IT facilitates information sharing and communication between logistics service buyers and sellers, thereby reducing coordination costs.

In sum, prior studies suggest IT reduces transaction costs and leads to a shift from hierarchies to markets; however, there is still little empirical evidence. Our study provides direct empirical evidence that IT in its most revolutionary form—the Internet—enabled a shift to the market in the context of logistics.

The rest of the paper is organized as follows. Section 2 provides our conceptual and mathematical models. Section 3 presents the empirical estimation, including a description of the data, variables, econometric adjustments, and estimation results. Section 4 discusses

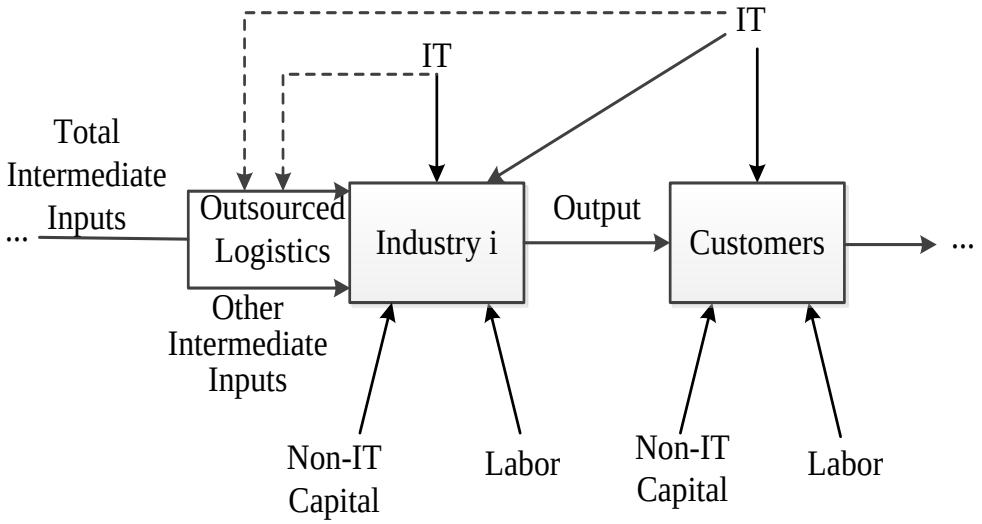

Figure 2.2: The Model of IT, Outsourced Logistics, and Output

our contributions, managerial implications, and research limitations.

2.2 Our Model of IT, Outsourced Logistics, and Output

Our conceptual model shown in Figure 2.2 is a model of productivity along a supply chain. In order to generate output that is supplied to customers, an industry uses non-IT capital, labor, IT capital, and purchases goods and services as intermediate inputs from the market. Intermediate inputs consist of purchased (outsourced) logistics, and other purchased goods and services. An industry benefits from the IT investments of customers because information sharing and coordination facilitated by customers' IT leads to more efficient production and operation (Cheng and Nault 2012).

IT is a multifaceted input, and as such IT not only has a direct impact on production—similar to other inputs, but also has indirect impacts on production by affecting other inputs (Mittal and Nault 2009). In particular, the Internet, which is characterized by global connectivity and open standards, helps remove incompatibility of legacy information systems within and between firms, greatly enhancing these systems' performance by allowing information sharing and coordination among broad trading partners (Zhu and Kraemer 2002, Zhu 2004). In our view, the Internet-enabled information sharing and coordination in supply chain lo-

gistics reduces external transaction costs and leads to a move to the market for provision of logistics. Outsourced logistics is typically more efficient than in-house logistics because 3PL providers have logistics-dedicated assets, economies of scale and scope, and managerial expertise. By outsourcing logistics, organizations can focus on their core competencies and exploit external logistics expertise, resulting in more efficient production.

From a supply chain perspective, customers' IT investments may affect an industry's output by affecting its in-bound logistics activities. The bullwhip effect, whereby small demand uncertainty at customer level causes greater demand fluctuations upstream, could trigger uncertainty in logistics activities which would be amplified further back along a supply chain (Lee et al. 1997). An increase in logistics uncertainty in a supply chain increases transaction costs along the supply chain. In particular, the external transaction costs associated with outsourcing in-bound logistics would increase. If transaction costs are high enough that they exceed operational advantages of the market, industries would prefer to use their own in-house logistics. However, on-time demand and inventory information sharing enabled by customers' investments in the Internet-based IOSs relieves information distortion in a supply chain, which in turn reduces demand uncertainty and logistics uncertainty (Lee et al. 2000, Cachon and Fisher 2000). As a result, transaction costs in a supply chain are reduced, allowing an industry to outsource more logistics. Consequently, an industry would increase productivity because of improved logistics as a result of greater logistics outsourcing. These effects are depicted in Figure 2.2.

We are interested in how IT, including an industry's own IT and customers' IT, affects its outsourced logistics and then consequently affects its output. We develop our mathematical form of the model based on the Cobb-Douglas production function,

$$Y = AK^{\alpha}L^{\beta}Z^{\gamma}M^{\phi},$$

where Y is the quantity of gross output; A is the technology change parameter; K, L, Z, M are the quantity of non-IT capital, labor, IT capital, and total intermediate inputs, respec-

tively. α, β, γ, and ϕ are the output elasticities of non-IT capital, labor, IT capital, and total intermediate inputs, respectively. The total intermediate inputs consist of outsourced logistics, W, and non-logistics intermediate inputs, X,

$$M = \{W, X\}.$$

Taking outsourced logistics and non-logistics intermediate inputs into the Cobb-Douglas production function and incorporating the impact of customers' IT on industry i's output, we specify the model with customers' IT investments as

$$Y_i = A_c K_i^{\alpha} L_i^{\beta} Z_i^{\gamma} X_i^{\theta} W_i^{\delta} C_i^{\varphi},$$

or in log form

$$y_i = a_c + \alpha k_i + \beta l_i + \gamma z_i + \theta x_i + \delta w_i + \varphi c_i, \tag{2.1}$$

where A_c is the corresponding technology change parameter, Z_i is industry i's own IT investment, C_i is the IT spillovers from industry i's customers, θ is the output elasticity of non-logistics intermediate inputs, and φ is the output elasticity of the customer-driven IT spillovers. The lower case variables are the log forms of the upper case variables. Next, to analyze how an industry's own IT investment and its customers' IT investments affect its outsourced logistics, we use a variable coefficient approach to specify the output elasticity of outsourced logistics, δ, as a linear function of industry i's own IT investment and IT investments from customers,

$$\delta(Z_i, C_i) = b + \mu Z_i + \nu C_i. \tag{2.2}$$

Taking (2.2) into (2.1), we have our estimation model corresponding to the conceptual model in Figure 2.2,

$$y_i = a_c + \alpha k_i + \beta l_i + \gamma z_i + \theta x_i + b w_i + \varphi c_i + \mu Z_i w_i + \nu C_i w_i. \tag{2.3}$$

Our estimation model is different from the variants of Cobb-Douglas (CD) production function in prior studies. Figure 2.3 compares our estimation model with those from prior

Studies	Level of Analysis	Main model	Main Contributions
Lichtenberg (1995)	Firm	$y = \alpha k + \gamma z + \beta_1 l_1 + \beta_0 l_0$	The excess returns to both IS capital and IS labor
Brynjolfsson and Hitt (1996)	Firm	$y = a + \alpha k + \gamma z + \beta_1 l_1 + \beta_2 l_2$	IS spending made significant contribution to output
Dewan and Kraemer (2000)	Country	$q = a + \alpha k + \gamma z + \beta l$	The different structure of returns from capital investment between developed and developing countries
Cheng and Nault (2007)	Industry	$y = a + \alpha k + \gamma z + \beta l + \theta m + \vartheta s$	The effect of supplier-driven IT spillovers on downstream output
Han et al. (2011)	Industry	$y = a + \alpha k + \gamma z + \beta l + \theta_0 m_0 + \theta_1 m_1$	The positive contribution of IT outsourcing to output and labor productivity
Cheng and Nault (2012)	Industry	$y = a + \alpha k + \gamma z + \beta l + \theta m + \varphi c$	The effect of customer-driven IT spillovers on upstream output
Current study	Industry	$y = a + \alpha k + \gamma z + \beta l + \theta x + \delta w + \varphi c$ $\delta(Z, C) = b + \mu Z + vC$	The Internet enables the move to the market form of the provision of goods and services

Notes: The models are log forms of Cobb-Douglas production function extensions. Each lowercase variable represents the log of the corresponding uppercase variable. On the left side of the above models, y is output, and q is annual GDP; On the right side of models, k is non-IT capital stock, z is IT capital stock, l_1 is IT staff labor, and l_0 is other labor, l_2 is other labor and expenses, l is total labor, m is total intermediate inputs, m_1 is IT services intermediate inputs, m_0 is non-IT services intermediate inputs, s is supplier-driven IT spillover index, c is customer-driven IT spillover index, x is non-Logistics intermediate inputs, and w is outsourced logistics. The controls for models are suppressed for brevity.

Figure 2.3: The Comparison of Current Study and the Prior

studies. Recognizing that conceptually diverse models can lead to similar estimation forms, the critical novelty of our model is the use of the variable coefficient approach to specify the relationship between IT and outsourced logistics and incorporating that relationship into an extended Cobb-Douglas production function. Through this we can focus on whether our specified relationship between IT and outsourced logistics has changed after the advent of the Internet.

We are able to assess the impact of IT on outsourced logistics by examining the coefficients of the two interaction terms in (2.3). We use Edgeworth's definition of complementarity and substitutability between goods to interpret the estimates of the interaction terms: assuming the first derivatives of a production function $f(\cdot)$ with respect to r_i and r_j are positive, if

the cross partial derivative is positive, $\frac{\partial(\partial f(\cdot)/\partial r_i)}{\partial r_j} = \frac{\partial^2 f(\cdot)}{\partial r_i \partial r_j} > 0$, then the goods r_i and r_j are complements; if the cross partial derivative is negative, then the two goods are substitutes (Edgeworth 1925). The coefficient of the interaction term between own IT capital and outsource logistics, μ in (2.3), demonstrates how own IT investment affects the contribution of outsourced logistics to output, holding other factors fixed. The coefficient μ is defined as

$$\mu = \frac{\partial \delta(Z_i, C_i)}{\partial Z_i} = \frac{\partial \left(\frac{\partial Y_i}{\partial W_i} * \frac{W_i}{Y_i} \right)}{\partial Z_i} = \frac{W_i}{Y_i} * \frac{\partial^2 Y_i}{\partial W_i \partial Z_i}.$$

The magnitude of μ is determined by the factor share of outsourced logistics and a cross partial derivative of output with respect to own IT investment and outsourced logistics, and the sign of μ depends on the sign of the cross partial derivative. Therefore, when μ is positive, the cross partial of output with respect to own IT investment and outsourced logistics is positive, meaning an industry's own IT investment and outsourced logistics are complements; when μ is negative, the cross partial is negative, indicating an industry's own IT investment and outsourced logistics are substitutes. Similarly, the coefficient of the interaction term between customers' IT investments and an industry's outsourced logistics to its output, ν in (2.3), indicates the impact of customers' IT investments on the contribution of an industry's outsourced logistics. Customers' IT investment and an industry's outsourced logistics are complements when ν is positive, and substitutes when ν is negative.

Based on I-O use tables, we calculate the IT investments from the customers of industry i, C_i, as a weighted average of IT investments of downstream industries. We use the relative transaction values in the total transactions made by customers as weights. We formally specify C_i as

$$C_i = \sum_{j \neq i} \frac{V_{ij}}{\sum_{j \neq i} V_{ij}} Z_j. \tag{2.4}$$

where V_{ij} is the dollar transaction volume of industry i's sales to the customer industry j. Our constructions of customers' IT investments are consistent with those in Mun and Nadiri (2002) and Cheng and Nault (2012).

2.3 Empirical Estimation

2.3.1 Data and Variables

Our Measure of Outsourced Logistics Understanding what goes into our measure of logistics outsourcing is critical to interpreting our results. We measure outsourced logistics as the sum of for-hire transportation costs and warehousing and storage costs. Transportation and warehousing costs generally account for the largest proportion of total logistics expenditures—72.2% of the total in 2011 (Wilson 2012)—and more than half of those costs are dedicated to outsourcing (Langley and Capgemini 2012). Hence, these comprise a reasonable measure of outsourced logistics. We calculate transportation and warehousing costs based on I-O use tables whereby "Transportation costs are the freight charges paid to move the commodity from the producer to the intermediate user or the final user." (Streitwieser 2009, p.42). These costs consist of rail, truck, water, air, oil pipeline, and gas pipeline charges (Horowitz and Planting 2009, Streitwieser 2009).

In order to articulate what is included in our measure of outsourced logistics, we provide the scope of the measure of outsourced logistics in Figure 2.4. The measure of outsourced logistics includes purchased logistics for moving goods from producers to users between and within industries. Assume non-transportation industry i has two firms of F_A and F_B, and it purchases transportation and warehousing for moving intermediate inputs from other industries to F_A and F_B. L_A and L_B are the purchased logistics for moving goods from other industries to Firm A and Firm B, respectively. Both L_A and L_B are captured as the for-hire logistics of industry i in I-O accounts, even if producers in other industries pay for the logistics service. The measure of outsourced logistics for industry i also includes the purchased logistics used to move goods within industry i: L_{AB} is the freight charge and warehousing costs paid to logistics service providers for moving goods from Firm A to Firm B.

At the establishment level, assume F_B has two establishments E_1 and E_2, the outsourced

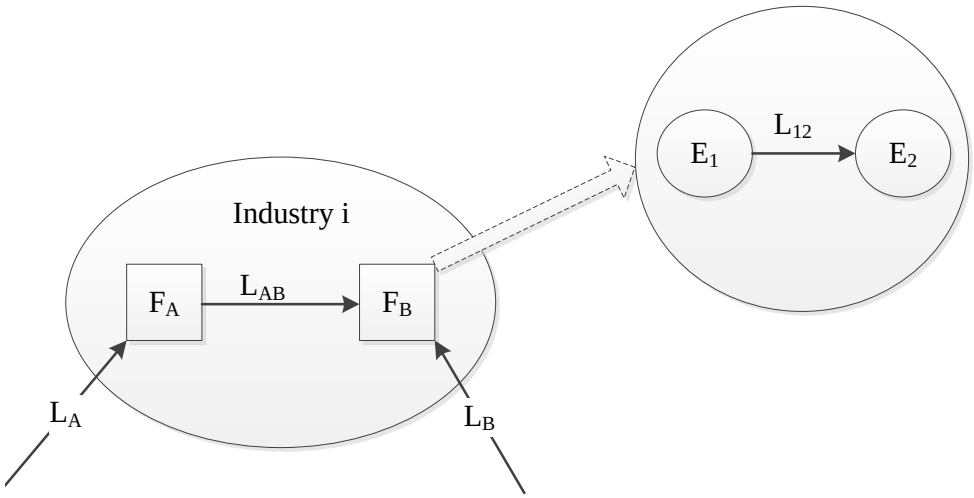

Figure 2.4: The Scope of Outsourced Logistics

logistics of industry i for moving goods from E_1 to E_2, L_{12}, captures purchased transportation and warehousing services from transportation industries. Note that, a proportion of in-house transportation is captured under outsourced transportation when it is provided by an establishment that is owned and operated by a non-transportation firm (U.S. Department of Transportation et al. 2011). For example, when one of establishments of F_B, E_3, whose core business is logistic services, moves goods from E_1 to E_2, this logistics activity is in-house logistics for F_B, while it is captured as for-hire transportation for industry i because E_3 is identified under transportation industries.

Although our measure of outsourced logistics may contain a proportion of in-house logistics due to the characteristics of IO accounts, the following evidence shows that the portion of in-house logistics is negligible. In order to figure out how much of in-house transportation costs was captured in our measure of for-hire transportation costs, we collected data about for-hire transportation costs from transportation satellite accounts (TSAs) provided by U.S. Department of Transportation. TSAs "only differ from the I-O accounts in that they separately measure in-house transportation" (U.S. Department of Transportation et al. 2011, p.14). Of the three available TSAs—1992, 1996, and 1997—only the 1997 TASs isolate freight from passenger transportation, and are thus comparable to our measure of outsourced

Transportation Costs	1997 I-O table	1997 TSAs	
		For-hire	In-house
Total transportation	96126.11 (A)	93927 (B)	36740
Air	15082.43	14904	3751
Rail	15308.45	12371	717
Water	968.94	2624	1969
Truck	59990.57	57427	30303
Pipeline	4775.72	6601	
(A-B)/B	0.023		

Notes: The data about for-hire and in-house transportation costs is from "Table 7. TSA's Use of Major Commodity Groups by Industry Sector: 1997" in 1997 TSAs (U.S. Department of Transportation et al. 2011); In-house transportation costs captured in 1997 TSAs cover the costs for moving an industry's intermediate inputs or output; For-hire transportation costs based on 1997 TSAs and 1997 I-O table are the costs for moving intermediate inputs from producers to users; Transportation costs are in millions of dollars at producers' prices.

Figure 2.5: Transportation Costs in 1997 Transportation Satellite Accounts

transportation costs which focuses on freight transportation.

Figure 2.5 shows the for-hire transportation costs based on 1997 I-O accounts and 1997 TSAs for manufacturing industries. Our measure of outsourced transportation costs from the I-O accounts is close to the corresponding for-hire transportation costs in 1997 TSAs: it is merely 2.3% larger than that in 1997 TSAs. As long as the methods of allocating transportation costs to industries in I-O accounts remain consistent across years, our measure of outsourced transportation costs consistently reflects purchases of transportation from the market. In addition, because outsourced transportation costs account for the largest proportion of the total outsourced logistics costs—on average 99.8% in Dataset I and 91.4% in Dataset II, our measure of outsourced logistics, consisting of outsourced transportation costs and warehousing and storage costs, represents purchases of logistics from the market.

Dataset I: 1987-1999 Dataset I consists of the multifactor productivity (MFP) dataset for 3-digit SIC manufacturing industries from 1987 to 1999 and the I-O tables from 1983 to 1999. [2] Both were requested from the Bureau of Labor Statistics (BLS). The data about output, intermediate inputs, labor, capitals is similar to that used by Cheng and Nault

[2]The Economic Census collects required data for developing the I-O tables at the level of establishment, where an establishment is defined as "a business or industrial unit at a single physical location that produces or distributes goods or that performs services". It classifies establishments into industries according to SIC or NAICS and aggregates establishment-level data to obtain industry-level data (Horowitz and Planting 2009).

(2007), Han et al. (2011b), and Cheng and Nault (2012). Figure A.1, in the Appendix, shows the data source and construction procedure of Dataset I. The MFP data contains 140 3-digit SIC manufacturing industries. It provides us with the series of output and the series of intermediate purchase compensation in millions of nominal dollars, and it also has the output deflator and the intermediate purchase price deflators with 1987 as the base year. Dividing the series of output and intermediate inputs by their corresponding deflators, we obtained the series for output Y and intermediate inputs M in millions of 1987 dollars, respectively. We also obtained the labor input L in millions of all employee hours from the MFP data.

We obtained information capital stock which was used as IT capital Z from BLS and calculated non-IT capital based on capital stock and information capital stock. The information capital stock are consisted of computers and related equipment, office equipment, communication, instruments, photocopy and related equipment; and all are in millions of 1987 dollars. Capital stock in millions of 1987 dollars consists of equipment, structures, inventories, land, and special tools. In order to obtain non-IT capital K, we summed the equipment and structure components and then subtracted the IT capital stock from the total.

Following the method of measuring outsourced logistics above, we obtained the data on outsourced logistics W for each manufacturing industry by summing the transportation commodities and warehousing and storage commodity in use tables: Railroad Transportation (SIC 40), Trucking and Courier Services, except Air (SIC 421,423), Water Transportation (SIC 44), Air Transportation (SIC 45), Pipelines, except Natural Gas (SIC 46), and Warehousing and Storage (SIC 422).[3] We calculated non-logistics intermediate inputs X by subtracting outsourced logistics from total intermediate inputs. Again, outsourced logistics and the non-logistics intermediate inputs are in 1987 dollars.

[3] The U.S. statistical system does not currently have a separate classification system for commodities. Each commodity is assigned the code of the industry in which the commodity is the primary product (Horowitz and Planting 2009).

Based on I-O tables, we calculated customers' IT investments. I-O tables contain manufacturing industries and non-manufacturing industries, and some of the rows/columns for manufacturing industries are the combination of more than one 3-digit SIC codes. We dropped all non-manufacturing industries in the I-O tables and aggregated the corresponding industries in the MFP data according to the manufacturing industries listed in I-O tables, which generated 98 manufacturing industries in the MFP data and the I-O tables. After matching MFP data and I-O data, we set the diagonals in the I-O tables to zero to isolate the transactions with other industries, and then we followed (2.4) to calculate the transaction weighted customers IT capital.

Based on 98 manufacturing industries, we dropped the industries with missing data and those without customer industries. Because of missing intermediate inputs data from 1997 to 1999, we dropped 6 industries: Logging (SIC 241), Newspapers (SIC 271), Periodicals (SIC 272), Books (SIC 273), Miscellaneous Publishing (SIC 274), and Greeting Cards (SIC 277). Next, we eliminated 7 industries which did not supply to others: Ordnance and Ammunition (SIC 348), Aerospace (SIC 372, 376), Ship and Boat Building and Repairing (SIC 373), Railroad Equipment (SIC 374), Toys and Sporting Goods (SIC 394), Footwear except Rubber and Plastic (SIC 313, 314), and Tobacco Products (SIC 21). In total, we have a balanced panel of 85 industries across 13 years.

Dataset II: 2000-2008 The second dataset is based on 3-digit 2002 NAICS codes and across 9 years from 2000 to 2008. It consists of the MFP data for 3-digit NAICS codes from 2000 to 2008 obtained from the BLS, gross domestic product (GDP) by industry obtained from the Bureau of Economic Analysis (BEA), and the I-O use tables from 2000 to 2008 that were also obtained from the BEA. Figure A.2, in the Appendix, shows the data source, construction procedure, and deflators used for each variable. We used chain-type quantity indexes as deflators to obtain the real values by multiplying the 2005 current-dollar value of the series by the corresponding chain-type quantity indexes and then dividing by 100.

We acquired Dataset II in December 2011. The gross output, Y, and the total intermediate inputs M were obtained from the GDP by industry accounts on the BEA website. Our measure of IT capital, non-IT capital, and labor were downloaded from the MFP databases of the BLS. IT capital stock includes computers, software, communication, and others, and we use the total as IT capital Z. The IT assets in Dataset II are different from those in Dataset I as software is included to reflect the importance of software in the current era. Capital includes equipment, structures, rental residential capital, inventories, and land. In order to calculate non-IT capital, K, we aggregated equipment and structures and then subtracted IT capital. Our measure of labor input, L, is in millions of hours.

We calculated outsourced logistics based on I-O use tables and by following the method of measuring outsource logistics above. We measure outsourced logistics as the purchases of the following commodities: Air Transportation (NAICS 481), Rail Transportation (NAICS 482), Water Transportation (NAICS 483), Truck Transportation (NAICS 484), Pipeline Transportation (NAICS 486), and Warehousing and Storage (NAICS 493). Subtracting outsourced logistics from total intermediate inputs, we obtained non-logistics intermediate inputs, X. We converted the nominal values of outsourced logistics and non-logistics intermediate inputs to chained 2005 U.S. dollars by using the chain-type quantity indexes for gross output of the warehousing and transportation sector and for intermediate inputs, respectively.

We followed the same method as that used in Dataset I to obtain customers' IT investments. In particular, I-O tables show transactions among manufacturing industries and transactions between manufacturing and non-manufacturing industries. To be consistent with Dataset I, we captured the transactions among manufacturing industries as weights for calculating customers' IT investments. In addition, to match the I-O tables with the MFP data, we combined codes 3361MV and 3364OT in I-O tables. Therefore, the final dataset is a balanced panel of 18 manufacturing industries across 9 years. Figure A.3, in the Appendix,

Variable	Obs	Mean	Std. dev.	Min	Max
Dataset I (1987-1999)					
Output (in millions of 1987 dollars)	1105	30294.40	46044.20	557.62	738130.80
Non-IT capital stock (in millions of 1987 dollars)	1105	20641.51	22836.99	461.80	135540.60
Labour (in millions of hours)	1105	414.45	342.97	12.20	2350.90
IT capital stock (in millions of 1987 dollars)	1105	1814.49	3165.67	30.30	27661.10
Total Intermediate inputs (in millions of 1987 dollars)	1105	16970.93	20999.34	313.18	202082.50
Non-Logistics intermediate inputs (in millions of 1987 dollars)	1105	16112.12	20043.91	307.17	193899.60
Outsourced logistics (in millions of 1987 dollars)	1105	858.81	1115.23	4.48	8211.68
Customers' IT investments (index)	1105	2969.07	2444.52	69.16	19508.81
Dataset II (2000-2008)					
Output (in millions of 2005 dollars)	162	253048.40	196900.70	20728	738084
Non-IT capital stock (in millions of 2005 dollars)	162	139626.80	95180.63	22269	336823
Labour (in millions of hours)	162	1644.82	989.78	239	4227
IT capital stock (in millions of 2005 dollars)	162	14693.33	16815.21	919	62336
Total Intermediate inputs (in millions of 2005 dollars)	162	170303.70	143106.3	7568.15	529658.60
Non-Logistics intermediate inputs (in millions of 2005 dollars)	162	164606.30	139135.90	7328.36	517997.90
Outsourced logistics (in millions of 2005 dollars)	162	5675.38	5093.44	614.01	24355.66
Customers' IT investments (index)	162	23694.58	7570.42	7200.11	39179.31

Notes: NAICS is a six-digit system and SIC is a 4-digit system. Dataset I is based on 3-digit SIC codes and dataset II is based on 3-digit NAICS codes. The 3-digit NAICS level (subsector) corresponds roughly to 2-digit SIC level (major group). Dataset II is at a higher aggregation level than dataset I, so the number of observations in dataset II is lower than that in dataset I.

Figure 2.6: Summary Statistics

describes the 18 manufacturing industries in 3-digit NAICS codes.

The summary statistics for both Datasets I and II are provided in Figure 2.6.

Econometric Adjustments The years covered in our two Datasets, 1987-1999 and 2000-2008, contain many changes in political and economic activities, such as the e-commerce boom in the late 1990s, the e-commerce collapse and the 9/11 terrorist attacks after 2000, and the financial crisis in 2008. These changes took place along with the variations in fiscal, monetary, and trade policies. Consequently, to control for any economy-wide shocks which could affect all industries, we add year-fixed effects in our estimation model. In addition, to better interpret the interaction terms and to reduce possible multicollinearity between the interaction effects and the main effects (Wooldridge 2009), we center the covariates, IT capital Z, logistics w, and customers' IT capital C, before constructing the interactions terms and without centering main terms in the estimation model.

Because both of our Datasets are cross-sectional time series, we test for the potential econometric problems: autocorrelation and heteroskedasticity (HE). First, we anticipate autocorrelation in error terms because the output of any industry is highly correlated with its output in the previous year under relatively smooth business cycles. Following the Wooldridge test for autocorrelation in panel data (Wooldridge 2002), we reject the null hypothesis of no first-order autocorrelation (AR1) at all reasonable levels of significance both in Dataset I (F (1, 84) = 73.60 for the full sample, 117.40 for the sample from 1987 to 1993, and 33.42 for the sample from 1994 to 1999) and Dataset II (F (1, 17) = 12.50). In addition, the autocorrelation could differ in magnitude for different industries if the magnitudes of the responses to the changes in business cycles differ across industries, so the AR1 process could be panel specific AR1 (PSAR1). We use the likelihood ratio test to check whether AR1 coefficients are common across the industries (Greene 2008). The null hypothesis that the regression with correction of AR1 is nested in the regression with the correction of PSAR1 is rejected at all levels of significance in Dataset I ($\chi^2(84) = 295.59$ for the full sample, $\chi^2(84) = 207.29$ for the sample from 1987 to 1993, and $\chi^2(84) = 224.77$ for the sample from 1994 to 1999), so we control for PSAR1 for Dataset I. In Dataset II, we fail to reject the null ($\chi^2(17) = 4.13$), so we adjust for AR1 in the estimations.

We also test for panel-level HE by using the likelihood ratio test (Greene 2008). It is reasonable to anticipate the panel-level HE, because the variances of the error terms for each industry are likely to fluctuate over time and the variances of the error terms could also differ across industries, resulting in panel-level HE. The null hypotheses of no panel-level HE is rejected at all levels of significance for the estimation model in Dataset I ($\chi^2(84) = 1624.82$ for the full sample, $\chi^2(84) = 572.89$ for the sample from 1987 to 1993, and $\chi^2(84) = 924.39$ for the sample from 1994 to 1999) and in Dataset II ($\chi^2(17) = 141.09$).

Consequently, after adding year fixed effects in the estimation model, we estimate our models by adjusting PSAR1 and panel-level HE for Dataset I, and AR1 and panel-level HE

for Dataset II. We use feasible generalized least squares to generate our estimates (Wooldridge 2002).

2.3.2 Estimation Results

We estimate the simple Cobb-Douglas production function and the extended Cobb-Douglas production function specified in (2.1) for two Datasets separately in order to compare our results with those from previous studies. Column 1 of Figure 2.7 and Column 1 of Figure 2.9 report the estimation results for the simple Cobb-Douglas production function for Dataset I and Dataset II, respectively. The estimation results for the extended Cobb-Douglas production function are shown in Column 2 of Figure 2.7 and Column 2 of Figure 2.9 for Dataset I and Dataset II, respectively. These results are similar across two Datasets and are consistent with those in previous studies. The consistency of our estimates with those of previous studies provides evidence for the validity of our Datasets.

Next, we estimate the model specified in (2.3) with corresponding econometric adjustments for Datasets I and II. In order to capture whether the effects of IT investments on outsourced logistics change over time, we split Dataset I into the pre-Internet era and the post-Internet era at 1993. According to the development history of the Internet, its commercialization began in 1994 when business and media started to notice the Internet (Zakon 2011). In April 1995, NSFNET (National Science Foundation Network) backbone was decommissioned, which marked that the last restriction on the use of the Internet to carry commercial traffic was removed (Leiner et al. 1997). Since 1995, the Internet has grown dramatically. Therefore, we split our first dataset into 1987–1993 and 1994–1999. We define the pre-Internet era as the 1987 to 1993 period in Dataset I, and the post-Internet era as the period from 1994 to 1999 in Dataset I and 2000 to 2008 in Dataset II. The main estimation results for the three time periods are shown in Figure 2.7, Figure 2.8, and Figure 2.9, respectively.

Variables	The Simple CD (1)	The Extended CD (2)	Full sample 1987-1999 (3)	Main Estimation 1987-1993 (4)	Robustness Tests: 2-digit SIC dummies (5)	Time span 1987-1994 (6)	Redefined IT 1 (7)	Redefined IT 2 (8)	Estimation using IVs (9)
Non-IT Capital	0.082*** (0.010)	0.037*** (0.010)	0.029*** (0.010)	0.062*** (0.009)	0.055*** (0.013)	0.060*** (0.010)	0.016* (0.009)	0.006 (0.008)	0.087*** (0.026)
Labour	0.237*** (0.008)	0.264*** (0.007)	0.272*** (0.007)	0.233*** (0.007)	0.215*** (0.010)	0.242*** (0.009)	0.245*** (0.007)	0.225*** (0.006)	0.198*** (0.024)
IT Capital	0.080*** (0.006)	0.083*** (0.005)	0.073*** (0.005)	0.088*** (0.004)	0.051*** (0.008)	0.099*** (0.004)	0.036*** (0.003)	0.048*** (0.004)	0.088*** (0.012)
Non-Logistics		0.539*** (0.009)	0.546*** (0.009)	0.581*** (0.007)	0.577*** (0.013)	0.553*** (0.007)	0.586*** (0.012)	0.614*** (0.007)	0.595*** (0.031)
Outsourced Logistics		0.104*** (0.006)	0.118*** (0.007)	0.052*** (0.005)	0.105*** (0.010)	0.060*** (0.005)	0.138*** (0.009)	0.118*** (0.006)	0.036** (0.017)
CustomerIT		0.025*** (0.005)	0.034*** (0.005)	0.012*** (0.004)	0.001 (0.007)	0.014*** (0.004)	0.027*** (0.005)	0.028*** (0.004)	0.028* (0.014)
OwnIT * Outsourced_ Logistics			4.24e-06** (1.72e-06)	-8.13e-06*** (1.70e-06)	-6.37e-06*** (1.06E-06)	-9.68e-06*** (1.98e-06)	-1.45e-05 (1.03e-05)	-4.98e-06 (5.41e-06)	-6.57e-06*** (2.17e-06)
CUSIT * Outsourced_ Logistics			6.19e-06*** (1.79e-06)	-8.57e-06*** (2.20e-06)	-1.06e-05*** (2.77E-06)	-7.83e-06*** (2.32e-06)	9.04e-07 (1.35e-05)	1.44e-06 (5.79e-06)	-1.53e-05** (6.69e-06)
Intermediate Inputs	0.639*** (0.010)								
Constant	1.275*** (0.0518)	1.653*** (0.064)	1.526*** (0.073)	1.597*** (0.063)	1.636*** (0.094)	1.682*** (0.068)	1.720*** (0.082)	1.691*** (0.060)	1.373*** (0.194)
Observations		1105		595	595	680	560	588	510

Notes: (1) OwnIT and CUSIT are an industry's IT investment and its customers' IT investments, respectively. OwnIT and CUSIT are in the levels and other variables are in natural logs. (2)We control for panel-level heteroskedasticity (HE) and panel specific autocorrelation (PSAR1) for dataset I. Details of the year fixed effects are suppressed for brevity. Standard errors are in parentheses following the estimated coefficients. (3) For redefined IT 1, we use the communication type of IT asset as IT capital (Z) in the estimation model. For redefined IT 2, we use the sum of two types of IT assets—communication and computers—as IT capital (Z) in the model. (4) Hansen/Sargan C-test statistics for the estimation using IVs (instrument variables) is 14.387 (p=0.072). Standard errors are in parentheses, *p<0.10, **p<0.05, ***p<0.01.

Figure 2.7: The Estimation Results for the Pre-Internet Era Based on Dataset I

The Effects of IT on Outsourced Logistics before the Advent of the Internet

The estimated coefficient of the interaction term "OwnIT * OutsourcedLogistics", μ, is negative and significant at 1% level from 1987 to 1993 (Column 4 of Figure 2.7). It suggests that an industry's own IT investment and outsourced logistics are substitutes during the pre-Internet era, meaning that increasing IT investment reduces outsourced logistics. For instance, installing an ERP system improves information timeliness, accuracy, and availability, facilitates information sharing across functions within firms, and upgrades internal connectivity. Thus, internal governance costs, such as information processing costs, monitoring costs, and opportunity costs due to poor information, etc., are reduced. Because

of the reduction in internal governance costs, the efficiency of existing business processes is improved. In particular, in-house logistics becomes more efficient, and firms prefer the internal provision of logistics.

The estimated coefficient of the interaction term "CUSIT*OutsourcedLogistics", ν, is negative and significant at 1% level from 1987–1993 (Column 4 of Figure 2.7). It suggests that customers' IT investments are substitutes for an industry's outsourced logistics, meaning that an increase in customers' IT investments reduces an industry's outsourced logistics. For example, the IT investments in EDI from an industry's customers electronically send demand information to a given industry, which in turn uses the data to forecast, analyze, and plan for orders, to optimize its production plan and procurement plan, and to better utilize its transportation capabilities and manage inventory. Thus, the given industry can more efficiently manage its transportation and warehousing across a supply chain from procurement to distribution, and this works towards lowering the costs of conducting logistics in-house.

The Effects of IT on Outsourced Logistics after the Advent of the Internet Column 1 of Figure 2.8 and Column 3 of Figure 2.9 show the estimation results for Dataset I from 1994 to 1999 and for Dataset II from 2000 to 2008, respectively.

The estimated coefficient of the interaction term "OwnIT * OutsourcedLogistics", μ, is positive and significant at 1% level. It suggests that an industry's own IT investment and outsourced logistics are complements after the advent of the Internet, which means that the greater the own IT investment, the more outsourced logistics. For instance, an industry's IT investment in Internet-based IOSs substantially facilitates information sharing and coordination among a focal industry, its suppliers, and 3PL providers, reducing external transaction costs associated with logistics. Because transaction costs are reduced, the focal industry outsources more logistics to 3PL providers that typically conduct logistics activities more efficiently.

Variables	Main Estimation 1994-1999	Robustness Tests				
		2-digit SIC dummies	Time span 1995-1999	Redefined IT 1	Redefined IT 2	Estimation using IV
	(1)	(2)	(3)	(4)	(5)	(6)
Non-IT Capital	0.058***	0.067***	0.045***	0.011	-0.020***	0.102***
	(0.009)	(0.017)	(0.008)	(0.010)	(0.007)	(0.039)
Labour	0.264***	0.194***	0.276***	0.204***	0.231***	0.206***
	(0.009)	(0.015)	(0.010)	(0.009)	(0.008)	(0.041)
IT Capital	0.085***	0.084***	0.084***	0.065***	0.072***	0.077***
	(0.006)	(0.015)	(0.006)	(0.004)	(0.004)	(0.019)
Non-Logistics	0.513***	0.509***	0.537***	0.575***	0.582***	0.566***
	(0.011)	(0.015)	(0.011)	(0.008)	(0.009)	(0.053)
Outsourced Logistics	0.116***	0.161***	0.111***	0.164***	0.162***	0.073**
	(0.008)	(0.009)	(0.008)	(0.008)	(0.007)	(0.031)
CustomerIT	0.053***	0.033***	0.065***	0.057***	0.049***	0.029
	(0.006)	(0.010)	(0.006)	(0.007)	(0.004)	(0.025)
OwnIT * Outsourced_ Logistics	5.63e-06***	9.80e-06***	7.63e-06***	5.89e-05***	2.71e-05***	7.05e-07
	(1.79e-06)	(1.53E-06)	(1.88e-06)	(1.64e-05)	(4.34e-06)	(7.60e-06)
CUSIT * Outsourced_ Logistics	8.67e-06***	4.29e-06**	1.18e-05***	8.64e-05***	1.24e-05***	7.01e-06
	(2.10e-06)	(2.17E-06)	(2.16e-06)	(1.35e-05)	(2.98e-06)	(7.88e-06)
Constant	1.355***	1.508***	1.145***	1.603***	1.579***	1.255**
	(0.082)	(0.093)	(0.088)	(0.010)	(0.074)	(0.361)
Observations	510	510	425	480	504	425

Notes: (1) OwnIT and CUSIT are an industry's IT investment and its customers' IT investments, respectively. OwnIT and CUSIT are in the levels and other variables are in natural logs. (2)We control for panel-level heteroskedasticity (HE) and panel specific autocorrelation (PSAR1) for dataset I. Details of the year fixed effects are suppressed for brevity. Standard errors are in parentheses following the estimated coefficients. (3) For redefined IT 1, we use the communication type of IT asset as IT capital (Z) in the estimation model. For redefined IT 2, we use the sum of two types of IT assets—communication and computers—as IT capital (Z) in the model. (4) Hansen/Sargan C-test statistics for the estimation using IVs (instrument variables) is 11.874 (p=0.157). Standard errors are in parentheses, *p<0.10, **p<0.05, ***p<0.01.

Figure 2.8: The Estimation Results for the Post-Internet Era Based on Dataset I

The estimated coefficient of the interaction term "CUSIT * OutsourcedLogistics", ν, is positive, and significant at 1% level for the dataset from 1994 to 1999 and at 5% level for the dataset from 2000 to 2008, respectively. It suggests that IT investments of a focal industry's customers and this focal industry's outsourced logistics are complements, meaning that an increase in customers' IT investments increases the focal industry's outsourced logistics. For example, an increase in customers' IT investments in the Internet-based IOSs enables demand and inventory information sharing. Because of the Internet's global connectivity, open standards, and low communication costs, information can be transferred swiftly to suppliers and 3PL providers, enabling the focal industry and its supply chain partners coordinate logistics at lower costs, leading to a greater demand for logistics outsourcing.

Variables	The Simple CD	The Extended CD	Main Estimation 2000-2008	Robustness Test				
				Sector dummies	Time span 2001-2008	Redefined IT 1	Redefined IT 2	Estimation using IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Non-IT Capital	0.196***	0.202**	0.164***	0.187***	0.188**	0.275***	0.185***	0.0029
	(0.044)	(0.083)	(0.062)	(0.062)	(0.074)	(0.050)	(0.049)	(0.040)
Labour	0.122***	0.209***	0.141***	0.112***	0.146***	0.129***	0.103***	0.045**
	(0.028)	(0.041)	(0.032)	(0.032)	(0.036)	(0.032)	(0.029)	(0.019)
IT Capital	0.062**	0.042	0.061*	0.047	0.054	-0.021	0.033	0.088**
	(0.0273)	(0.046)	(0.037)	(0.036)	(0.043)	(0.030)	(0.027)	(0.034)
Non-Logistics		0.521***	0.536***	0.550***	0.526***	0.572***	0.577***	0.745***
		(0.026)	(0.025)	(0.024)	(0.026)	(0.025)	(0.025)	(0.030)
Outsourced Logistics		0.068	0.107***	0.102***	0.104**	0.067**	0.103***	0.062***
		(0.051)	(0.038)	(0.036)	(0.047)	(0.033)	(0.0330)	(0.024)
CustomerIT		0.011	0.082**	0.090***	0.076**	0.055	0.074***	0.085***
		(0.031)	(0.033)	(0.031)	(0.037)	(0.035)	(0.029)	(0.026)
OwnIT * Outsourced_ Logistics			3.88e-06***	4.48e-06***	3.90e-06**	4.55e-05***	5.00e-06***	1.46e-06**
			(1.47e-06)	(1.40e-06)	(1.87e-06)	(1.70e-05)	(1.91e-06)	(7.43e-07)
CUSIT * Outsourced_ Logistics			2.35e-06**	2.38e-06**	2.24e-06*	9.84e-06	1.38e-06	-1.09e-06
			(1.10e-06)	(1.10e-06)	(1.20e-06)	(1.41e-05)	(1.51e-06)	(1.33e-06)
Intermediate Inputs	0.611***							
	(0.024)							
Constant	1.278***	1.138**	0.697	0.538	0.611	0.477	0.668*	0.919**
	(0.274)	(0.531)	(0.465)	(0.443)	(0.534)	(0.376)	(0.393)	(0.413)
Observations	162	162	162	162	144	162	162	144

Notes: (1) OwnIT and CUSIT are an industry's IT investment and its customers' IT investments, respectively. OwnIT and CUSIT are in the levels and other variables are in natural logs. (2)We control for panel-level heteroskedasticity (HE) and first-order autocorrelation (AR1) for dataset II. Details of the year fixed effects are suppressed for brevity. Standard errors are in parentheses following the estimated coefficients. (3) For redefined IT 1, we use the communication type of IT asset as IT capital (Z) in the estimation model. For redefined IT 2, we use the sum of three types of IT assets—communication, computers, and software—as IT capital (Z) in the model. (4) Hansen/Sargan C-test statistics for the estimation using IVs (instrument variables) is 13.371 (p=0.100). Standard errors are in parentheses, *p<0.10, **p<0.05, ***p<0.01.

Figure 2.9: The Estimation Results for the Post-Internet Era Based on Dataset II

2.3.3 Robustness Tests

Estimates with Sector-Fixed Effects In Datasets I and II, there might be unobserved heterogeneity related to different types of manufacturing such as raw materials and finished goods. To control for this kind of time-invariant industry-specific effects, we may use sector dummies or industry dummies. Estimating the models with industry-fixed effects uses a large degree of freedom because it needs to estimate different intercepts for different industries (Han et al. 2011b, Cheng and Nault 2012). Recognizing industries in a sector share relatively similar production processes, we use sector-fixed effects model.

In Dataset I, we group the industries by 2-digit subheadings, resulting in 19 sector dum-

mies. We continue to control for panel-level HE and PSAR1. The estimation results for the pre-Internet era with sector-fixed effects are reported in Column 5 of Figure 2.7, and the estimation results for the post-Internet era from 1994 to 1999 are reported in Column 2 of Figure 2.8. These results are similar to those in Column 4 of Figure 2.7 and Column 1 of Figure 2.8, respectively. In Dataset II, we classify the industries into a durable goods sector and a nondurable goods sector according to the BEA. The results of our regression with sector-fixed effects and controlling for panel-level HE and AR1 are presented in Column 4 of Figure 2.9, and they are similar to those in Column 3.

Estimates Using Different Time Spans We split the sample in Dataset I into 1987–1993 and 1994–1999 periods because the commercialization of the Internet started between 1993 and 1995. However, the widespread diffusion of the Internet across the economy may have taken time. To address this concern, we split Dataset I into 1987–1994 and 1995–1999 periods, and estimated our model controlling for year fixed effects, panel-level HE, and PSAR1. The results for the time period from 1987 to 1994 are reported in Column 6 of Figure 2.7, and the results for the time period from 1995 to 1999 are reported in Column 3 of Figure 2.8. The estimation results are consistent with the results based on the previous time split. Our main insights remain qualitatively unchanged even if we use the period 1987–1995 for the pre-Internet era and 1996–1999 for the post-Internet era.

For Dataset II, the e-commerce collapse in 2000 might be considered as an outlier, so we dropped year 2000. The estimation results based on the data from 2001 through 2008 are reported in Column 5 of Figure 2.9. The results are similar to those in Column 3.

Estimation with Redefined IT Capital IT capital stocks in both Datasets include a variety of IT assets. We use the total of those IT assets as IT capital in our model for the following reasons. First, it is reasonable to use total IT assets as the measure of IT capital because all types of IT assets are involved in external transactions and internal governance. Second, information systems implementations become increasing integrated. The same sys-

tems have functions for internal production management and external coordination, so it is hard to determine which IT investment is used for what purpose. For example, enterprise resource planning (ERP) systems really came from the prior MRP (materials requirements planning) I and MRP II implementations, and these systems dealt with specifically with material flows, inventory and logistics. Nowadays, ERP systems are organization-wide and integrate functions dealing with internal management—such as the functions of production management, accounting management, and human resource management—and external coordination, such as supply chain management and customer relationship management. Therefore, it is impossible to determine the IT investment for coordinating specific activities like logistics in integrated information systems implementations.

For a robustness check, we provide results using subsets of our measure of IT capital. We separately define IT capital as communication (Redefined IT 1), and as computers and communication (Redefined IT 2) for Dataset I. For Dataset II, we separately define IT capital as communication (Redefined IT 1), and computers, software, and communication (Redefined IT 2). The estimation results with redefined IT capital are shown in Columns 7 and 8 of Figure 2.7, Columns 4 and 5 of Figure 2.8, and Columns 6 and 7 of Figure 2.9. The estimated coefficients of interaction terms are positive and significant during the post-Internet era, which is similar to the main estimation results based on our broader measures of IT capital.

Estimates with Instrumental Variables In general, endogeneity can be caused by omitted variables, simultaneity, and measurement errors (Wooldridge 2002). It is possible that some omitted variables such as other productivity-related organizational initiatives may be correlated with the independent variables in our models, such as IT or outsourced logistics. In addition, there is potential for measurement errors in these variables due to aggregation because our variables are measured at the industry level. There might also be concerns about endogeneity caused by simultaneity. Simultaneity arises when explanatory variables

in our model are determined simultaneously along with the dependent variable output. For example, an increase in inputs increases output, and meanwhile, shocks in demand are likely to trigger temporary adjustments in output and inputs.

Our econometric adjustment with PSAR1, year-fixed effects, and time-invariant industry-specific effects already help relieve the endogeneity concerns. We further address such concerns by providing estimations using instrumental variables. Like previous related studies, we use one-year lags of the inputs as instruments (Stiroh 2002b, Han et al. 2011b, Cheng and Nault 2012, Han and Mithas 2013). If one or two main terms in an interaction term are endogenous, then the interaction term may be correlated with the error term, so we also use one-year lag of interaction terms as instruments.

We apply the two-stage generalized moment method (GMM) procedure and conduct the Hansen/Sargan C-test for the estimated model across two Datasets (Baum et al. 2003). The Hansen/Sargan C-test statistic cannot reject the exogeneity of suspect endogenous variables for the estimation based on the pre-Internet era from 1987 to 1993 ($\chi^2(8) = 14.387$, $p = 0.072$), the post-Internet era from 1994 to 1999 ($\chi^2(8) = 11.874$, $p = 0.157$), and the time period from 2000 to 2008 ($\chi^2(8) = 13.371$, $p = 0.100$). The results with the instruments are shown in Column 9 of Figure 2.7, Column 6 of Figure 2.8, and Column 8 of Figure 2.9. The estimation for the two interaction terms is negative and significant at 5% level in the pre-Internet era (Column 9 of Figure 2.7), which is similar as those in the main results (Column 5 of Figure 2.7). For the instrumental estimation results in the post-Internet era, the effects of own IT investment shown in Column 8 of Figure 2.9 are positive and significant at 5% level for the data from 2000 to 2008, which is similar to those in the main results (Column 3 of Figure 2.9). Regarding the estimations for the interaction effect of customers' IT investments in the post-Internet eras shown in Column 6 of Figure 2.8 and Column 8 of Figure 2.9, neither of them are significant, and this might be due to the efficiency of the instrumental variable estimation.

2.4 Conclusion

2.4.1 Contributions

Our study has three important contributions. First, we provide direct evidence that IT in its most interconnected instance—the Internet—enables a move to the market. Previous work theoretically argued that IT was associated with the shift from hierarchies to the market by reducing transaction costs, and other studies only provided indirect evidence in support of those arguments. In the context of logistics, we show that the increasing use of IT is associated with greater logistics outsourcing after the advent of the Internet. By focusing on logistics outsourcing, we can rule out the compounding factors that arise from offshoring in more general outsourcing.

Second, we show the impact of IT on output through logistics outsourcing. Previous studies about IT productivity converged to the contribution of IT to output (e.g., Lichtenberg 1995, Brynjolfsson and Hitt 1996, Dewan and Kraemer 2000, Cheng and Nault 2007, 2012, Han et al. 2011b). Our study shows the moderating role of IT in the contribution of logistics outsourcing to output. In particular, our study identifies a change in the relationship between IT and outsourced logistics. Before the advent of the Internet, an industry's own IT investment and outsourced logistics were substitutes, and they became complements after the Internet was widely used. We also find similar impacts of customers' IT investments on outsourced logistics.

Third, our study reconciles the theoretical arguments about the effects of IT on reducing coordination costs within and between firms. Previous studies argued that IT could reduce internal governance costs and external transaction costs, and IT would have different impacts on organizational change depending on which effect predominates. However, it was not known which effect dominated a priori. Our study suggests that the effect of IT on reducing internal governance costs dominates before the advent of the Internet, and the effect of IT on reducing external transaction costs dominates after the advent of the Internet.

2.4.2 Managerial Implications

One implication of our findings is that activities with the potential to be outsourced and that use the Internet for coordination are more likely to be outsourced in the Internet era. Our findings are based on the context of logistics, where coordination is a key function of management and Internet-based IOSs have been used to improve coordination between supply chain members. A current revolution, mobile computing, also affects coordination costs. Although it remains to be seen whether mobile computing has greater effects on internal or external coordination costs, our results are suggestive.

The other implication of our findings is that the increasing popularity of outsourcing has a technological and theoretical basis. The Internet with its universal connectivity and open standards has enabled the integration of IT infrastructure between firms and facilitated the information sharing and coordination, which in turn improves visibility and transparency of supply chains. In addition, the Internet enables online markets where firms can connect with potential providers of goods and services at low cost. As a result, the transaction costs involved in outsourcing can be reduced.

2.4.3 Limitations and Qualifications

It is worth noting two limitations in our study. First, we focus our attention on the impact of IT on outsourced logistics downstream in the supply chain. Our study examines the impact of IT, including a focal industry's own IT and customers' IT, on the focal industry's outsourced logistics before and after the advent of the Internet. Logistics outsourcing needs coordination among suppliers, 3PL providers, and customers. Thus, the IT capabilities of upstream supply chain members and 3PL providers also play an important role in improving coordination for moving goods from suppliers to customers.

Suppliers' IT investments not only improve the coordination of logistics activities, but also help increase downstream industry output through improved quality of goods and services

which are used as intermediate inputs for downstream industries (Cheng and Nault 2007). Consequently, with summary data on IT capital, it is not possible to differentiate the quality effect and the coordination effect from suppliers' IT. However, our study still sheds light on the impact of suppliers' IT on a focal industry's logistics outsourcing based on our unique data. In the I-O accounts, some supplying industries are also customer industries, so we can infer the impact of suppliers' IT from our results about the effects of customers' IT. Moreover, we avoid double-counting issue by the virtue of not estimating the impact of suppliers' IT in our model.

Second, for similar reasons, our industry-level data does not enable us to directly estimate the impact of IT on logistics processes to get insight about which specific transaction costs are reduced by IT since the advent of the Internet. Consequently, we can only study the summary effects of make or buy decisions, and infer that relative reductions in transaction costs have favored external provision. However, our industry level data allows us to investigate the economy-wide impact of IT on outsourced logistics, which increases the generalizability of our study.

Chapter 3

THE IMPACT OF IT ON PRODUCTION INTERDEPENDENCE

3.1 Introduction

In order to improve production efficiency and create more business value, industries have been increasingly investing in IT. According to the U.S. Bureau of Labor Statistics (BLS), IT investments in the manufacturing sector climbed steadily from $60.989 billion in 1987, reaching $185.179 billion in 2012 (measured in 2009 dollars). These IT investments have changed how firms and industries run their businesses and how they organize production. Information sharing and coordination captured by industries' IT investments enable them to integrate with upstream suppliers in their production processes such as purchasing and Just-in-time (JIT) production. A question left to be answered is the extent to which firms and industries are more or less dependent on upstream suppliers for purchases of goods and services to produce output; for example, the extent to which an industry needs fewer or more intermediate inputs from upstream supplier industries to produce one unit of output.

Production interdependence between industries is demonstrated by inter-industry purchases and sales of goods and services (Santhanam and Patil 1972). As an aspect of production structure based on input-output tables, production interdependence has drawn considerable attention. Two methods of measurement are widely used for measuring industry interdependence or linkage (Dietzenbacher 1992). One method, developed by Chenery and Watanabe (1958), focused on the direct linkage between industries, and used two indicators of interdependence among productive sectors. One indicator, called direct backward linkage (DBL), shows the extent of using purchases from upstream suppliers compared to the use

of labor and capital, and it is measured by the ratio of intermediate inputs to the value of total production. That is, DBL is the amount of intermediate inputs required by a given industry to produce one unit of output (Miller and Blair 2009). The other indicator, called direct forward linkage, is measured by the ratio of intermediate sales to total sales including final demand. The other method, developed by Rasmussen (1957), considered direct and indirect linkages of production. For instance, backward linkage measures how much output is produced by the whole system of industries in order to satisfy one unit increase in the final demand of an industry's products.

In order to understand the impact of IT on production interdependence, we study the impact of IT on the direct linkages between industries. In particular, we are interested in how an industry's IT investment affects its production interdependence with upstream suppliers. We capture this by examining the impact of IT on DBL. A high DBL for a given industry suggests that it depends heavily on suppliers for providing intermediate inputs, including materials, energy, and services. The magnitude of DBL of an industry could be affected by its IT investment through substitution for intermediate inputs and/or enhancement of the efficiency of intermediate inputs. For example, an industry's investment in Enterprise Resource Planning (ERP) allows cross-functional processes and real-time information sharing across departments. These enable better forecasting and scheduling for production, leading to more efficient usage of materials that are purchased from suppliers. In addition, an industry's IT investments in interorganizational information systems (IOSs), such as Supply Chain Management (SCM) systems and Collaborative Planning, Forecasting and Replenishment (CPFR), allow information sharing and coordination with upstream suppliers, and lead to a certain level of visibility of inventory in a supply chain as well as the removal of redundant inventory and logistics activities across tiers of production. Consequently, a given industry needs fewer intermediate inputs to produce a certain amount of output; this results in a reduction in its DBL.

A reduction in DBL represents a structural change in an industry's production frontier towards more efficient production, as it suggests that a given industry requires fewer intermediate inputs from upstream suppliers to produce a certain level of output. From growth accounting theory, technological change, also called total factor productivity (TFP), explains the variation in output that is not accounted for by production inputs and captures any shift in production frontier (Solow 1957), so the reduction in DBL is incorporated in the growth of TFP. From production theory, the growth in TFP leads to greater value-added. Therefore, IT acts as a substitute for, and increases the efficiencies of using, intermediate inputs from suppliers resulting in a reduction in DBL; the reduction in DBL represents a shift in production frontier towards more efficient production, and leads to growth in TFP and greater value-added.

Previous studies about the impact of IT on production factors showed that IT could both complement and substitute for labor and non-IT capital as well as augment labor and non-IT capital. Using a CES-Translog production function, Dewan and Min (1997) found that IT was a net substitute for both ordinary capital and labor. Chwelos et al. (2010) examined whether the substitution relationships between IT and other inputs have evolved over time. They confirmed the role of IT as a substitute for labor and found that IT capital was a complement of non-IT capital with the development of recent IT initiatives. In addition, Hitt and Snir (1999) examined the impact of IT on other inputs in different types of organizations and found that IT capital complemented non-IT capital in modern organizations and substituted non-IT capital in traditional organizations; in contrast, IT substituted for labor in all forms of organizations. Mittal and Nault (2009) measured the direct and indirect effects of IT and found that IT not only contributed to value-added directly as a production input, but also added value by augmenting non-IT capital and labor. Zhang et al. (2014) examined substitution of IT capital for other inputs by using the Morishima Elasticity of Substitution (MES) instead of the Allen elasticity of substitution (AES), and found the impact of the

change in IT capital prices on the quantity of other inputs is different from the impact of the change in prices of other inputs on the quantity of IT capital. However, few studies have examined the impact of IT on production interdependence, an aspect of production structure measured as the ratio of intermediate inputs to gross output.

Previous studies about IT and TFP have shown that the impact of IT on TFP arises from IT spillovers from supply chain partners, a firm's own IT investment, and IT-related complementary inputs. Cheng and Nault (2007) found that the IT spillovers from upstream supplier industries could benefit the productivity of downstream industries through improvements in output quality. Cheng and Nault (2012) examined the contribution of IT spillovers from the IT investments of downstream customer industries, and found that customers' IT investments have an impact on upstream industry output through information sharing and coordination. Han et al. (2011a) found that IT spillovers from suppliers could increase the TFP of a given industry; the effect of IT spillovers was moderated by the downstream industries' IT intensity and competitiveness. Tambe and Hitt (2014) studied the impact of IT investments of other firms from which a firm had hired IT labor, and found that IT investments of other firms could contribute to the given firm's productivity growth through IT labor flows. Some studies have investigated the impact of IT on TFP through a firm's own IT investment and associated organizational capital. Brynjolfsson and Hitt (2003) examined the impact of firms' computerization on productivity and output growth and found that the contribution of computerization was larger in the longer term than that in the short term; and they suggested that complementary inputs, such as organizational capital, might be associated with the greater growth in the long-term. However, little research has examined the impact of IT on TFP from the point-of-view of production structure—production interdependence.

The production interdependence of sectors or industries in an economy is represented by the purchases and sales of goods among different sectors (Santhanam and Patil 1972).

It has been used as a key factor in the international and intertemporal comparisons of production structure based on input-output (I-O) tables (Dietzenbacher 1992, Soofi 1992). In a seminal study about the production structure, Chenery and Watanabe (1958) compared the production interdependence across four countries and developed two measures of production interdependence that were used widely in the field: the ratio of purchases of inputs to total value of production of industry i, u_i, which measures the extent of indirect use of factors as compared to the direct use of labor and capital; and the ratio of intermediate demand to total demand for a given product of industry i, w_i. Many studies have used these measures. For example, Santhanam and Patil (1972) compared the production structure of India with that of selected advanced countries, and they maintained that u_i quantified the degree of industry i's dependence on upstream industries, and w_i quantified the dependence of downstream industries on the given industry. Hirschman (1958) defined backward and forward linkage effects as input-provision and output-utilization, respectively. These two concepts have been widely used by development economists to identify key sectors with above-average backward and forward linkages. Jones (1976) pointed out that linkage concepts were based on industrial interdependence, and concurred that u_i could be used to measure direct backward linkage (DBL). Miller and Blair (2009) reviewed the development of linkage measures and maintained the calculation of DBL defined by Chenery and Watanabe (1958).

We aim to add to the literature by examining whether and how an industry's IT investment affects its production interdependence with upstream suppliers, and how a change in an industry's production interdependence affects TFP and value-added. We use "production interdependence" not "production dependence" to describe the relationship between an industry and its suppliers because a supplier of an industry may also be supplied by the given industry in an economic network. We estimate the impact of IT on production interdependence, measured as DBL, as well as the impact of DBL on TFP and value-added. Our key result shows that IT can reduce an industry's production interdependence with upstream

suppliers by improving the efficiency associated with using intermediate inputs. In addition, the reduction of production interdependence is associated with an increase in TFP, which leads to greater value-added.

The remainder of the paper is organized as follows: section 2 develops our conceptual and mathematical models; section 3 presents the empirical estimation, including a description of the data, variables, econometric adjustments, and estimation results; and section 4 concludes the paper.

3.2 The Model of IT, DBL, and Value-added

3.2.1 A Conceptual Model

Under the production function framework, an industry invests IT capital, non-IT capital, and labor into production, and it also purchases materials, energy, and services, collectively called intermediate inputs, from upstream industries for production. The four inputs of production (IT capital, non-IT capital, labor, and intermediate inputs) are combined in a production function to model the production of gross output, while IT capital, non-IT capital, and labor create value-added for the given industry. The ratio of intermediate inputs to gross output, DBL, indicates an industry's production interdependence with upstream suppliers.

IT is a multi-faceted factor of production. We argue that an industry's IT capital not only contributes directly to its value-added as an input, but also contributes indirectly to its value-added by reshaping its production interdependence with upstream suppliers. IT applications bring about changes in business processes: automation, streamlining, and reengineering. Firstly, IT can computerize manual tasks related to processing intermediate inputs, so that the intermediate inputs can be utilized more efficiently and effectively. For example, computer-aided flexible manufacturing (CAM) uses software to control machine tools and related machinery, creating a faster production process and minimizing waste in intermediate inputs. Secondly, IT reduces the need for extra intermediate inputs due to bottlenecks

and redundant steps by facilitating streamlining in production. Streamlining simplifies or eliminates unnecessary steps in production, such as bottlenecks and redundancy, resulting in reduced intermediate inputs needed in business processes. Thirdly, IT can reengineer business processes within and between organizations, leading to a dramatic improvement in production efficiency. For example, Chrysler's Electronic Data Exchange (EDI) systems allowed information about planning, producing, and delivering of production requirements to be transmitted instantly within and between enterprises, facilitating the implementation of Just-in-time (JIT) (Mukhopadhyay et al. 1995). Consequently, the purchasing of goods and services to produce a certain level of output was reduced, resulting in reduced DBL.

A reduction in DBL shifts an industry's production frontier towards more efficient production: fewer intermediate inputs are needed to produce a certain level of output, while keeping other inputs constant. Technological change, or TFP, incorporates any shifts in the production function (Solow 1957) and can be classified as embodied and disembodied technological change; both are sources of economic growth (Jorgenson and Griliches 1967, Hulten 1992, Stiroh 2001). Embodied technological change refers to the technological change embodied in new investments that have been upgraded in terms of quality, such as new capital, new machinery and equipment, under the hypothesis that new capital is more productive than old capital. Disembodied technical change refers to generalized improvements in methods, processes and the enhancement of facilitating factors, such as communication, information and transportation networks. A reduction in DBL is brought about from improvements in methods and processes related to the usage of intermediate inputs and represented by a reduction in DBL. Thus, a reduction in DBL is captured as an increase in TFP associated with disembodied technical change, and an increase in TFP is associated with an increase in value-added.

Our conceptual model, which captures the impact of IT on DBL and value-added, is presented in Figure 3.1: an industry's IT capital directly contributes to its value-added with

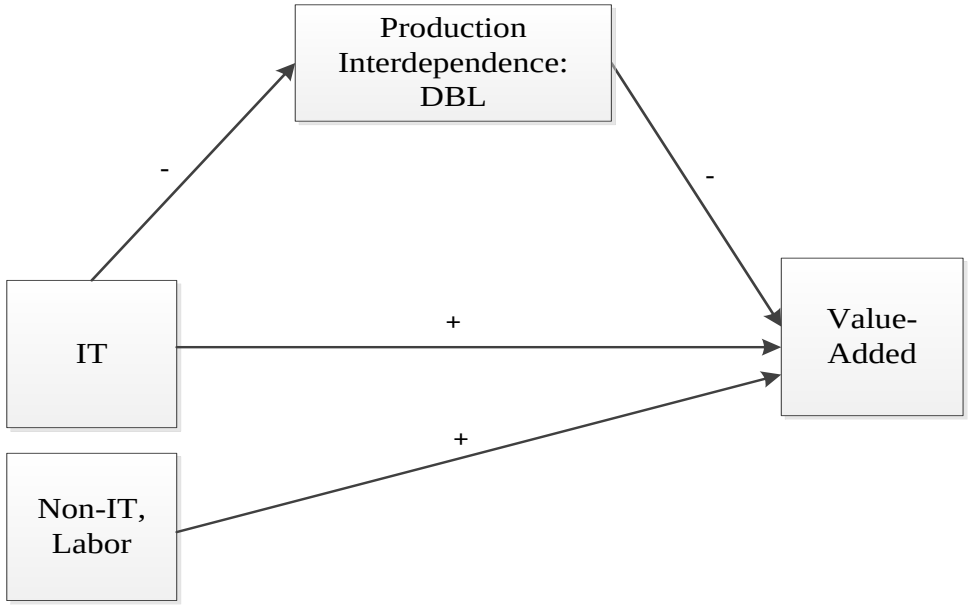

Figure 3.1: The Conceptual Model About IT, DBL, and Value Added

non-IT capital and labor inputs; IT capital also reduces the given industry's production interdependence with upstream suppliers, measured as DBL; the reduced DBL leads to greater value-added.

3.2.2 The Mathematical Models

Model 1 We develop our mathematical model based on the Cobb-Douglas production function. A standard Cobb-Douglas production function with gross output rather than value-added is

$$Y_{it} = A_i K_{it}^{\alpha} L_{it}^{\beta} Z_{it}^{\gamma} M_{it}^{\theta},$$

and the log form is

$$y_{it} = a_i + \alpha k_{it} + \beta l_{it} + \gamma z_{it} + \theta m_{it}.$$

Y_{it} is the quantity of gross output of industry i in year t, A_i is technological change or TFP of industry i, and $K_{it}, L_{it}, Z_{it}, M_{it}$ represent the quantities of non-IT capital, labor, IT capital, and intermediate inputs, respectively. $\alpha, \beta, \gamma, \theta$ are the output elasticities of non-IT capital, labor, IT capital, and intermediate inputs, respectively. Lower-case letters represent the log form of upper-case letters. The Cobb-Douglas form is theoretically supported as coming

from the accounting identity—output equals intermediate inputs plus the wage bill plus the total return on capital (Kundisch et al. 2014).

Cheng and Nault (2007) found that an industry's output is benefited from upstream supplying industries' IT spillovers through improvements in the quality of intermediate inputs sold by suppling industries. They argued that the measured deflator of intermediate inputs failed to capture quality improvement due to suppliers' IT investment, and that the quantity of intermediate inputs was over-deflated. In order to estimate the Cobb-Douglas function correctly, the real quantity of intermediate inputs should be used in the Cobb-Douglas model,

$$y_{it} = \bar{a}_i + \bar{\alpha} k_{it} + \bar{\beta} l_{it} + \bar{\gamma} z_{it} + \bar{\theta} \sum_{j \neq i} H_{jit} m^*_{jit},$$

where H_{jit} is the share of the intermediate inputs purchased from jth industry, V_{jit}, in the total value of intermediate inputs of industry i in time t, $\sum_{j \neq i} V_{jit}$; and m^*_{jit} is the log form of the real quantity of intermediate inputs of industry i purchased from industry j in time t.

m^*_{jit} can be rewritten as the sum of the log form of the quantity of intermediate inputs measured by us, m_{jit}, and the log form of mis-measurement due to the measured deflator without fully adjusting quality, ϵ_{jt}. The mis-measurement of an industry's intermediate inputs is proportional to supplier IT spillovers, $\epsilon_{jt} = \mu z_{jt}$, where μ is the constant of proportionality. The Cobb-Douglas production function incorporating the effect of suppliers' IT is specified as follows (Cheng and Nault 2007),

$$\begin{aligned} y_{it} &= \bar{a}_i + \bar{\alpha} k_{it} + \bar{\beta} l_{it} + \bar{\gamma} z_{it} + \bar{\theta} \sum_{j \neq i} H_{jit} m_{jit} + \bar{\theta} \sum_{j \neq i} H_{jit} \epsilon_{jt} \\ &= \bar{a}_i + \bar{\alpha} k_{it} + \bar{\beta} l_{it} + \bar{\gamma} z_{it} + \bar{\theta} m_{it} + \bar{\theta} \mu s_{it}, \end{aligned} \tag{3.1}$$

where $m_{it} = \sum_{j \neq i} H_{jit} m_{jit}$, and $s_{it} = \sum_{j \neq i} H_{jit} z_{jt}$ is IT spillovers from the supplying industry for industry i in time t.

We conjecture that an industry's own IT increases the efficiency of intermediate inputs, and that this effect in (3.1) is captured in $\bar{a}_i$—the log of TFP. The output elasticity of

intermediate inputs determines the percentage change in gross output in response to a one-percent change in intermediate inputs; it is a measure of the efficiency of intermediate inputs. We employ a variable coefficient approach to specify the impact of IT on the efficiency of intermediate inputs and assume a linear relationship between IT and the output elasticity of intermediate inputs. Use of the log form ensures consistency with other variables in terms of magnitude. We specify the relationship as

$$\bar{\theta}(z_{it}) = \omega + \nu z_{it}. \tag{3.2}$$

Bringing (3.2) into (3.1), we have

$$y_{it} = \tilde{a}_i + \tilde{\alpha} k_{it} + \tilde{\beta} l_{it} + \tilde{\gamma} z_{it} + \omega m_{it} + \omega \mu s_{it} + \nu z_{it} m_{it} + \nu \mu z_{it} s_{it}. \tag{3.3}$$

DBL is the amount of intermediate inputs used by an industry to produce one unit of output, and is calculated by dividing intermediate inputs by output (Miller and Blair 2009), and thus the log form of DBL, dbl, is written as

$$dbl = log \frac{M_{it}}{Y_{it}} = m_{it} - y_{it}. \tag{3.4}$$

Bringing (3.3) into (3.4), we have

$$\begin{aligned} dbl_{it} &= -\tilde{a}_i - \tilde{\alpha} k_{it} - \tilde{\beta} l_{it} - \tilde{\gamma} z_{it} + (1-\omega) m_{it} - \omega \mu s_{it} - \nu z_{it} m_{it} - \nu \mu z_{it} s_{it} \\ &= b_0 + b_1 k_{it} + b_2 l_{it} + b_3 z_{it} + b_4 m_{it} + b_5 s_{it} + b_6 z_{it} m_{it} + b_7 z_{it} s_{it} \end{aligned} \tag{3.5}$$

According to (3.5), own IT has two effects on DBL. The first, which can be assessed by b_3, is similar to the impact of labor and non-IT capital. Different production combinations can result in production of a certain amount of output: the more inputs in IT capital, non-IT capital, and labor an industry has, the fewer intermediate inputs the industry needs from suppliers. Thus, IT could be a substitute for intermediate inputs in production, implying that an increase in IT is associated with a decrease in intermediate inputs when producing a certain amount of output. Consequently, we expect a negative estimate of b_3. The second

effect of own IT is from the impact of an industry's own IT on the efficiency of intermediate inputs; this is assessed by b_6. If our conjecture that IT increases the efficiency associated with utilizing intermediate inputs is correct, an industry would need fewer intermediate inputs to produce a certain level of output. We therefore expect the estimate of the interaction term, b_6, to be negative.

Suppliers' IT spillovers also have two effects on DBL in (3.5). The first effect of suppliers' IT on DBL can be assessed by b_5. Because suppliers' IT increases the quality of intermediate inputs, the given industry may need less to produce a certain amount of output. Therefore, we expect that the estimate of b_5 is negative. The second effect of suppliers' IT spillovers is through their relationship with an industry's own IT: when there is a fit between own IT and suppliers' IT spillovers, for example, when both are above their own average levels, there would be better information sharing and coordination in a supply chain; and thus the given industry needs less intermediate inputs to produce a certain level of output. Consequently, we expect a negative estimate of b_7.

Model 2 A standard production function with value-added as the dependent variable is

$$VA_{it} = \hat{A}_{it} K_{it}^{\hat{\alpha}} L_{it}^{\hat{\beta}} Z_{it}^{\hat{\gamma}}.$$

We conjecture that DBL is incorporated in TFP. We specify TFP in the equation above, $\hat{A}_{it}$, as a function of DBL_{it},

$$\hat{A}(DBL_{it}) = C_i e^{\lambda DBL_{it}}, \tag{3.6}$$

where C_i is an industry-specific productivity level effect, which captures embodied technological change and disembodied technological change that is not captured by DBL. The specification of (3.6) is consistent with the formulation of the function of technological change parameter in the literature (Solow 1957, Scherer 1982, Stiroh 2002a, Han et al. 2011a). The function $\hat{A}(DBL_{it})$ represents technological change related to the efficiency of intermediate inputs, and we can incorporate DBL into the production function with value-added as the

dependent variable:

$$VA_{it} = C_i e^{\lambda DBL_{it}} K_{it}^{\hat{\alpha}} L_{it}^{\hat{\beta}} Z_{it}^{\hat{\gamma}}.$$

The log form of the equation above is

$$va_{it} = c_i + \lambda DBL_{it} + \hat{\alpha} k_{it} + \hat{\beta} l_{it} + \hat{\gamma} z_{it}. \tag{3.7}$$

We are interested in the estimate of λ, which allows us to assess the impact of DBL on TFP in (3.6). When the efficiency of intermediate inputs is increased, an industry can create more value added with the same level of gross output and keeping other inputs constant, so we expect a negative estimate of λ.

3.3 Empirical Estimation

3.3.1 Data and Variables

Dataset I: 1987-1999 We used two datasets in our analyses. These datasets cover different time periods and have different levels of aggregation.

Dataset I is an industry-level dataset based on the 3-digit Standard Industrial Classification (SIC) level. It has three sources from the U.S. Bureau of Labor Statistics (BLS): the multifactor productivity (MFP) dataset for 140 manufacturing industries from 1987 to 1999, the information capital stock and capital stock for 140 manufacturing industries from 1987 to 1999, and input-output (I-O) use tables for 98 manufacturing industries and 78 other industries with 3-digit SIC codes from 1983 to 2000. Some rows/columns of I-O tables, which have specific industry numbers in, are the combination of more than one 3-digit SIC industries. In order to match three sources of data, we dropped non-manufacturing industries in I-O tables, and aggregated data in the MFP dataset, IT capital stock, and non-IT capital stock according to the correspondence between industry numbers and SIC codes listed in I-O tables; this generated data for 98 manufacturing industries from 1987 to 1999.

We obtained the data on output, intermediate inputs, and labor from the MFP dataset, which contains the data on the series of output and intermediate purchase compensation in millions of nominal dollars and has the output deflators and the intermediate purchase price deflators with 1987 as the base year. Dividing the series of output and intermediate inputs by their corresponding deflators and multiplying by 100, we obtained the series for output Y and intermediate inputs M in millions of 1987 dollars, respectively. Subtracting the real values of intermediate inputs from gross output, we obtained the time series of value-added, VA. We calculated direct backward linkage, DBL, for each industry by dividing M by Y. We also obtained the labor input, L, in millions of all employee hours from the MFP data.

The information capital stock consists of computers and related equipment, office equipment, communication, instruments, photocopy and related equipment; all values are in millions of 1987 dollars. We used the total information capital stock as IT capital, Z. Capital stock (in millions of 1987 dollars) consists of equipment, structures, inventories, land, and special tools. In order to obtain non-IT capital, K, we summed the equipment and structure components then subtracted the total IT capital stock from the sum.

Based on the transaction volumes among 98 manufacturing industries in I-O tables, we calculated the index of suppliers' IT spillovers. We set the diagonals in the I-O tables to zero to isolate the transactions with other industries, and then calculated it as a weighted average of the log form of IT investments of upstream industries, $s_{it} = \sum_{j \neq i} \frac{V_{ji}}{\sum_{j \neq i} V_{ji}} z_{jt}$. Our constructions of suppliers' IT investments are consistent with those in Mun and Nadiri (2002) and Cheng and Nault (2007).

Based on 98 manufacturing industries, we dropped the industries with missing data and those without customer industries. Because of missing data about intermediate inputs from 1997 to 1999, we dropped 6 industries: Logging (SIC 241), Newspapers (SIC 271), Periodicals (SIC 272), Books (SIC 273), Miscellaneous Publishing (SIC 274), and Greeting Cards (SIC 277). Next, we eliminated 7 industries which did not supply to others: Ordnance and

Ammunition (SIC 348), Aerospace (SIC 372, 376), Ship and Boat Building and Repairing (SIC 373), Railroad Equipment (SIC 374), Toys and Sporting Goods (SIC 394), Footwear except Rubber and Plastic (SIC 313, 314), and Tobacco Products (SIC 21). In total, we have a balanced panel of 85 manufacturing industries across 13 years, which is similar to that used by Cheng and Nault (2007), Han et al. (2011b), Cheng and Nault (2012) and Gong et al. (2013).

Dataset II: 1998-2008 The second dataset is based on the 3-digit 2002 North American Industry Classification Systems (NAICS) codes and covers 11 years from 1998 to 2008. It consists of the MFP data for the 3-digit NAICS codes obtained from the BLS, gross domestic product (GDP) by industry obtained from the Bureau of Economic Analysis (BEA), and the I-O use tables from 1998 to 2008 that were also obtained from the BEA. The GDP industry accounts and I-O use tables from the BEA contain data for 61 industries, and the MFP databases from the BLS provide us the data on IT capital, non-IT capital, and labor for 59 industries. In order to match the data from the BEA and BLS, based on 61 industries, we dropped farms industry and combined the industries of Motor Vehicles, Bodies and Trailers, and Parts (NAICS 3361MV) and the industry of Other Transportation Equipment (NAICS 3364OT) as the industry of Transportation Equipment (NAICS 336); this generated 59 industries based on 3-digit NAICS codes with matched data. The data details for our variables are in the following.

We obtained data on value added, gross output, and intermediate inputs from 1998 to 2008 through the GDP by industry accounts on the BEA website. The data on value-added is in millions of chained 2005 dollars, and we used this as variable VA. From the same GDP by industry accounts, we obtained the nominal values of gross output and intermediate inputs, along with the corresponding deflators. Consistent with the BEA's methodology of converting nominal values to real ones, we used chain-type quantity indices as deflators to obtain the real values of gross output Y and total intermediate inputs M by multiplying

the 2005 current-dollar value of the series of gross output and intermediate inputs by the corresponding chain-type quantity indices, then dividing by 100, respectively. We calculated DBL for each industry by dividing M by Y.

Obtained from the MFP databases, IT capital stock, which is in millions of 2005 dollars, includes computers, software, communication, and others; we used the total as IT capital, Z. The IT assets in Dataset II are different from those in Dataset I as software is included to reflect the importance of software in the more recent period. Capital stock, which is in millions of 2005 dollars, includes equipment, structures, rental residential capital, inventories, and land. In order to calculate non-IT capital, K, we aggregated equipment and structures and then subtracted IT capital. Our measure of labor input, L, is in millions of hours.

We followed the same method as that used in dataset I to obtain suppliers' IT spillovers. In particular, we calculated the index of suppliers' IT spillovers based on the transaction volumes among 59 industries in I-O tables. After we obtained the data for 59 industries, we dropped two industries which do not supply to other industries: Hospitals and Nursing and Residential Care Facilities (NAICS 622HO) and Social Assistance (NAICS 624). Therefore, we have a panel dataset with 57 3-digit NAICS industries from 1998 to 2008.

Summary statistics for both Datasets are provided in Figure 3.2.

3.3.2 Econometric Adjustments

Fixed Effects and Econometric Adjustments Related to DBL The years contained in our two datasets, 1987-1999 and 1998-2008, cover many changes in political and economic activities, such as the e-commerce boom in the late 1990s, the e-commerce collapse, the 9/11 terrorist attacks in 2001, and the financial crisis in 2008. These changes took place alongside the variations in fiscal, monetary, and trade policies. Consequently, to control for any economy-wide shocks that could affect all industries, we added year-fixed effects in our estimation model.

Industries differ in production processes, production organization, etc., which drive the

Variable	Obs.	Mean	Std. Dev.	Min	Max
Dataset I (1987–1999)					
Gross Output (in millions of 1987 dollars)	1105	30294.4	46044.2	557.619	738130.8
Value-added (in millions of 1987 dollars)	1105	13323.47	35347.11	143.0986	679890.2
Non-IT Capital Stock (in millions of 1987 dollars)	1105	20641.51	22836.99	461.8	135540.6
Labor (in millions of hours)	1105	414.4487	342.9677	12.2	2350.9
Own IT Capital Stock (in millions of 1987 dollars)	1105	1814.489	3165.669	30.3	27661.1
Intermediate Inputs (in millions of 1987 dollars)	1105	16970.93	20999.34	313.1791	202082.5
Normalized Direct Backward Linkage	1105	1	.2123983	.128721	1.682773
Suppliers' IT Index	1105	7.207371	.8298654	5.4153	9.6363
Dataset II (1998–2008)					
Gross Output (in millions of 2005 dollars)	627	323781.5	353473.8	20728	2187837
Value-added (in millions of 2005 dollars)	627	174969.7	233002.6	5000	1545182
Non-IT Capital Stock (in millions of 2005 dollars)	627	237772	319179.4	9381	1876281
Labor (in millions of hours)	627	3183.657	4584.261	84	25810
Own IT Capital Stock (in millions of 2005 dollars)	627	29313.18	49115.24	468	344497
Intermediate Inputs (in millions of 2005 dollars)	627	149867.8	144569	6020.251	728609.1
Normalized Direct Backward Linkage	627	1	.3176625	.3551718	1.970463
Suppliers' IT Index	627	10.09531	.3908323	8.9868	11.1546

Notes: NAICS is a six-digit system and SIC is a 4-digit system. Dataset I is based on 3-digit SIC codes and Dataset II is based on 3-digit NAICS codes. The 3-digit NAICS level (subsector) corresponds roughly to the 2-digit SIC level (major group). Dataset II uses a higher aggregation level than Dataset I, so the number of observations in Dataset II is lower than that in Dataset I.

Figure 3.2: Summary Statistics

differences in their requirements for goods and services from suppliers, so each industry's DBL may be driven by specific industry characteristics. In order to control for this type of industry-specific effect, we normalized DBL by the simple average of all direct backward linkage,

$$NDBL_{jt} = \frac{DBL_{jt}}{\frac{1}{n}\sum_{j=1}^{n} DBL_{jt}},$$

where $NDBL_{jt}$ is the normalized DBL for industry j in year t (Miller and Blair 2009). In addition, there might be unobserved heterogeneity related to different types of manufacturing production processes, and this unobserved heterogeneity may be correlated to an industry's DBL. Our adjustment of normalized DBL partially addresses the effect of unobserved het-

erogeneity. Additionally, we may use sector dummies or industry dummies to control for time-invariant industry-specific effects. In Dataset I, industries from the groups with identical 2-digit subheadings share relatively similar manufacturing production processes, so we use sector-fixed effects in our estimation in order to preserve degrees of freedom. Dataset II includes both manufacturing and non-manufacturing industries, which differ in production processes and organization. We use manufacturing and non-manufacturing dummies to control for the effect due to unobserved heterogeneity in manufacturing and non-manufacturing processes for similar reasons.

To better interpret the interaction term and reduce possible multicollinearity between the interaction effect and the main effects (Wooldridge 2009, Ozer-Balli and Sørenson 2010), we center IT capital z, suppliers' IT index s, and intermediate inputs m when constructing the interactions terms, $z * m$ and $z * s$, without centering the main terms in the estimation model.

Autocorrelation and Heteroskedasticity Because both of our datasets are cross-sectional time-series, we test for the potential econometric problems: autocorrelation and heteroskedasticity (HE). First, we anticipate autocorrelation in error terms because the DBL of any industry is highly correlated with its output in the previous year under relatively smooth business cycles. Following the Wooldridge test for autocorrelation in panel data (Wooldridge 2002), we reject the null hypothesis of no first-order autocorrelation (AR1) at all reasonable levels of significance both in Dataset I (F(1, 84) = 37.81 for (3.5), and F(1, 84) = 10.19 for (3.7), and Dataset II (F (1, 56) = 86.86 for the (3.5), F (1, 56) = 159.96 for (3.7)). In addition, autocorrelation could differ in magnitude for different industries if the magnitude of response to changes in business cycles differ across industries, so the AR1 process could be panel-specific AR1 (PSAR1). We use the likelihood ratio test to check whether AR1 coefficients are common across the industries (Greene 2008). The null hypothesis—that the regression with the correction of AR1 is nested in the regression with the correction

of PSAR1—is rejected at all levels of significance in Dataset I ($\chi^2(84) = 241.17$ for (3.5), $\chi^2(84) = 376.16$ for (3.7)), so we control for PSAR1 for Dataset I. In Dataset II, we reject the null hypothesis ($\chi^2(56) = 130.59$ for (3.5), $\chi^2(56) = 353.06$ for (3.7)), so we adjust for PSAR1 for Dataset II in the estimations.

	Dataset I 1987–1999			Dataset II 1998–2008		
Variable	Log of Value-added (1)	Log of Normalized DBL (2)	Log of Value-added (3)	Log of Value-added (4)	Log of Normalized DBL (5)	Log of Value-added (6)
Log of Non-IT Capital	0.239*** (0.0406)	-0.0955*** (0.0157)	0.353*** (0.0184)	0.304*** (0.0162)	-0.102*** (0.0122)	0.318*** (0.0165)
Log of Labor	0.651*** (0.0353)	-0.244*** (0.0155)	0.567*** (0.0139)	0.643*** (0.0203)	-0.237*** (0.0112)	0.544*** (0.0155)
Log of OwnIT	0.190*** (0.0279)	-0.0747*** (0.0122)	0.130*** (0.00820)	0.0639*** (0.0171)	-0.0689*** (0.00854)	0.0844*** (0.0126)
Log of Intermediate Inputs		0.367*** (0.0132)			0.409*** (0.0155)	
Log Supplier IT		-0.0422*** (0.0113)			-0.111*** (0.0237)	
Log of OwnIT * Log of Intermediate Inputs		0.00166 (0.00423)			-0.0248*** (0.00667)	
Log of OwnIT * Log of SupplierIT		-0.0218*** (0.00493)			-0.00423 (0.00857)	
Year Dummies	***	***	***	***	***	***
Sector Dummies	***	***	***	-0.389*** (0.0308)	0.157*** (0.0204)	0.0305 (0.0277)
Normalized DBL			-1.623*** (0.0317)			-1.000*** (0.0195)
Constant	1.557*** (0.190)	-0.290*** (0.104)	3.110*** (0.123)	2.594*** (0.199)	-0.138 (0.242)	3.791*** (0.163)
Observations	1,105	1,105	1,105	627	627	627
The number of industries	85	85	85	57	57	57

Notes: We control for panel-level heteroskedasticity (HE) and panel-specific autocorrelation (PSAR1) for Datasets I and II. We use 2-digit SIC dummies and manufacturing versus non-manufacturing sector dummies for Datasets I and II, respectively. Details of the sector dummies and year-fixed effects are suppressed for brevity. Standard errors are in parentheses following the estimated coefficients. *p<0.10, **p<0.05, ***p<0.01.

Figure 3.3: The Estimation Results for Two Datasets

We also test for panel-level HE using the likelihood ratio test (Greene 2008). It is reasonable to anticipate panel-level HE, because the variances of the error terms for each industry are likely to fluctuate over time and the variances of the error terms could also differ across industries, resulting in panel-level HE. The null hypothesis of no panel-level HE is rejected

at all levels of significance for the estimation model in Dataset I ($\chi^2(84) = 1585.21$ for (3.5), $\chi^2(84) = 84$ for (3.7)) and in Dataset II ($\chi^2(56) = 605.17$ for (3.5), $\chi^2(56) = 776.10$ for (3.7)).

Consequently, after adding year-fixed effects, normalizing DBL, controlling for sector-fixed effects, and centering the interaction term in the estimation models, we estimate our models by adjusting for PSAR1 and panel-level HE for both Datasets I and II.

For each regression, we use feasible generalized least squares to generate our estimates (Wooldridge 2002). We address endogeneity issues which may bias our estimators later.

3.3.3 Estimation Results

We estimate the simple Cobb-Douglas production function with value-added as the dependent variable for Datasets I and II in order to compare our results with those from previous studies. Columns 1 and 4 of Figure 3.3 report the estimation results for the simple Cobb-Douglas production function for Datasets I and II, respectively. These results are consistent with those in previous studies (e.g., Lichtenberg 1995, Brynjolfsson and Hitt 1996, Dewan and Kraemer 2000, Cheng and Nault 2007, 2012, Han et al. 2011b, Gong et al. 2013). The consistency of our estimates with those of previous studies provides evidence for the validity of our datasets.

The Impact of Own IT on DBL We estimate two effects of own IT on DBL. The direct impact of IT on DBL is assessed by b_3 in (3.5), which is negative and significant at the 1% level (Columns 2 and 5 of Figure 3.3 for the full sample of Datasets I and II, respectively.) It suggests that a 1% increase in an industry's IT investment is associated with a 0.07 % decrease in its DBL for both datasets, keeping other factors constant. In order to produce a certain amount of output, an industry can use different combinations of inputs. When IT input increases, keeping non-IT capital and labor constant, the intermediate inputs required from upstream suppliers are reduced, and thus DBL decreases. This direct effect of IT on

DBL suggests substitution between IT and intermediate inputs.

Variable	1987–1993			1994–1999		
	Log of Value-added (1)	Log of Normalized DBL (2)	Log of Value-added (3)	Log of Value-added (4)	Log of Normalized DBL (5)	Log of Value-added (6)
Log of Non-IT Capital	0.275***	-0.106***	0.243***	0.194***	-0.127***	0.334***
	(0.0281)	(0.00954)	(0.0273)	(0.0441)	(0.0168)	(0.0304)
Log of Labor	0.576***	-0.208***	0.509***	0.547***	-0.196***	0.505***
	(0.0234)	(0.00736)	(0.0203)	(0.0322)	(0.0166)	(0.0218)
Log of OwnIT	0.176***	-0.0576***	0.196***	0.288***	-0.152***	0.195***
	(0.0200)	(0.00756)	(0.0179)	(0.0358)	(0.0157)	(0.0202)
Log of Intermediate Inputs		0.351***			0.434***	
		(0.0103)			(0.0149)	
Log Supplier IT		-0.0321***			-0.0681***	
		(0.00800)			(0.0149)	
Log of OwnIT * Log of Intermediate Inputs		0.00671**			-0.0153**	
		(0.00298)			(0.00711)	
Log of OwnIT * Log of SupplierIT		-0.00885**			-0.0243***	
		(0.00346)			(0.00656)	
Year Dummies	***	***	***	***	***	***
Sector Dummies	***	***	***	***	***	***
Normalized DBL			-1.452***			-1.904***
			(0.0372)			(0.0417)
Constant	1.796***	-0.456***	3.947***	2.029***	-0.270*	3.519***
	(0.147)	(0.0653)	(0.139)	(0.188)	(0.144)	(0.164)
Observations	595	595	595	510	510	510
The number of industries	85	85	85	85	85	85

Notes: We control for panel-level heteroskedasticity (HE) and panel-specific autocorrelation (PSAR1) for two periods in Dataset I. Details of the sector dummies and year-fixed effects are suppressed for brevity. Standard errors are in parentheses following the estimated coefficients. *p<0.10, **p<0.05, ***p<0.01.

Figure 3.4: The Estimation Results for 1987–1993 and 1994–1999

Separate from the substitution effects, IT also has an impact on DBL through the effect of intermediate inputs on DBL. We access this impact of own IT on DBL by the coefficient of the interaction term "Own IT * Intermediate Inputs", β_6. The estimate of this interaction term in (3.5) is not significant for the full sample of Dataset I, but it is negative (-0.025) and significant at the 1% level for Dataset II (Column 5, Figure 3.3). The negative and significant estimate of the interaction term suggests that IT negatively moderates the effect

of intermediate inputs on DBL: when IT investment is high, the effect of intermediate inputs on DBL is reduced compared with when IT investment is low. In particular, a 1% increase in IT decreases the impact of intermediate inputs on DBL by 2.5% for Dataset II.

From (3.2), we can calculate the impact of IT on the marginal product of intermediate inputs. For an industry with the mean level of IT investment and the mean level of the factor share of intermediate inputs, an increase in IT capital by 293.132 dollars (1% of IT capital) is associated with an increase in its marginal product of intermediate inputs by 0.304 dollars (which is calculated based on the formula, $\frac{\nu * log_e(1\% * Mean(IT))}{\frac{Mean(IntermediateInputs)}{Mean(GrossOutput)}} = \frac{0.0248 * log_e 293.132}{149867.8/323781.5}$) for Dataset II, keeping other factors fixed.

The Impact of Own IT in the Pre- and Post-Internet Eras The advent of the Internet radically enhanced IT capabilities which, in turn, fundamentally redefined the role of IT in organizations (Zhu and Kraemer 2002, Zhu 2004). As a result, industries have dramatically changed the way they organize their production; for example, the Internet enables industries to outsource more activities to external providers (Gong et al. 2013). Implementation of IOSs has reduced the number of suppliers used by firms (Dedrick et al. 2008). It is possible that the effect of IT on production interdependence may differ between the pre- and post-Internet eras.

We examine the effect of IT on DBL before and after the advent of the Internet. According to the development history of the Internet, its commercialization began in 1994 when it started to become noticed by business and media (Zakon 2011). In April 1995, National Science Foundation Network (NSFNET) backbone was decommissioned, which marked regional NSFNET networks' migration of their connections to commercial network providers (Leiner et al. 1997). Since 1995, the Internet has grown dramatically. Therefore, we split our first Dataset into 1987–1993 (the pre-Internet era) and 1994–1999 (the post-Internet era). We then test the effect of IT on DBL with the same econometric adjustments as those for the full sample of Dataset I. The estimation results based on the 1987–1993 and 1994–1999

time-periods are shown in Figure 3.4.

As seen, the main effect of own IT on DBL is negative and significant, consistent with the main effect estimated based on full sample from 1987 to 1999; while the impact of IT on DBL through the effect of intermediate inputs on DBL has changed after the advent of the Internet: before the advent the Internet, the estimate of the interaction term "Own IT * Intermediate Inputs" in (3.5) is positive and significant (Column 2, Figure 3.4), and it is negative and significant for the sample after the advent of the Internet (Column 5, Figure 3.4). Although firms had spent a lot of IT assets to satisfy functional requirements before the advent of the Internet, those IT assets were not integrated to facilitate efficient usage of intermediate inputs. Consequently, firms use more intermediate inputs than what is needed in efficient production to produce one unit of output. After the advent of the Internet, because of global connectivity, open standards, and low communication costs, IT assets in firms have been integrated to improve the efficiency of intermediate inputs, resulting in reduced DBL.

The Impact of Suppliers' IT Spillovers We estimate the impact of suppliers' IT spillovers on DBL. The main effect of suppliers' IT on DBL is negative and significant at the 1% level of significance (Columns 2 and 5, Figure 3.3; Columns 2 and 5, Figure 3.4). The results imply that suppliers' IT spillovers through improvements in the quality of intermediate inputs reduce an industry's production interdependence with suppliers—the given industry needs less intermediate inputs to produce a certain level of output.

In addition, suppliers' IT spillovers have an impact on DBL through their relationship with an industry's own IT. Note that we center the items in the interaction term "Own IT * Suppliers' IT". When an industry's IT investment is above the average of all industries' IT investments and when the level of suppliers' IT spillovers is above the average of suppliers' IT spillovers, there is a fit between the given industry's own IT investment and suppliers' IT spillovers, resulting in better information sharing and coordination. The negative and

significant estimation result of the "Own IT * Suppliers' IT" (Column 2 of Figure 3.3, and Columns 2 and 5 of Figure 3.4) suggests that when there is such a fit, the given industry needs less intermediate inputs to produce a certain level of output, keeping other factors constant.

In sum, the results about the impact of own IT and suppliers' IT on DBL suggest that after controlling for the impact of suppliers' IT spillovers on DBL, an industry's own IT can reduce its production interdependence by improving the efficiency of intermediate inputs. For example, use of ERP systems (e.g., in inventory management, material requirement planning, material management, production planning, and shipping) reduces redundant inventories at various stages of production, and thus an industry needs fewer intermediate inputs to produce a certain level of output.

The Impact of DBL on TFP The coefficient of DBL, λ, explains the relationship between DBL and TFP, which is defined in (3.6) and estimated in (3.7). The negative and significant estimate of DBL, λ (Columns 3 and 6 of Figure 3.3, and Columns 3 and 6 of Figure 3.4), suggests that a reduction in DBL is associated with an increase in TFP. For example, in the full sample of Dataset I, a one-unit decrease in the normalized DBL is associated with a 162.3% increase in TFP, keeping other factors constant.

The decrease in DBL suggests a shift in an industry's production frontier towards more efficient production in the sense of needing less intermediate inputs to produce a certain level of output; this is captured by TFP which explains the variation in value-added that is not explained by the change in IT capital, non-IT capital, and labor.

The Indirect Effects of IT on Value-added through DBL When examining the indirect effect of IT on value-added, we focus on the impact of IT on value-added through DBL and TFP. As shown in Columns 3 and 6 of Figure 3.3, the estimate of normalized DBL is negative and significant at the 1% level. This suggests that a one-unit decrease in normalized DBL is associated with a 162.3% increase in value-added for Dataset I and a

100% increase in value-added for Dataset II. We find that own IT reduces DBL and that the reduction in DBL is associated with an increase in TFP. Therefore, IT reduces DBL and leads to the growth of TFP, which in turn contributes to greater value-added.

Variable	1987–1999		1987–1993		1994–1999		1998–2008	
	Log of Normalized DBL (1)	Log of Value-added (2)	Log of Normalized DBL (3)	Log of Value-added (4)	Log of Normalized DBL (5)	Log of Value-added (6)	Log of Normalized DBL (7)	Log of Value-added (8)
Log of Non-IT Capital	-0.074*** (0.0278)	0.336*** (0.0401)	-0.090*** (0.0243)	0.312*** (0.0355)	-0.122*** (0.0353)	0.338*** (0.0514)	-0.060 (0.0368)	0.230*** (0.0674)
Log of Labor	-0.227*** (0.0341)	0.531*** (0.0394)	-0.211*** (0.0253)	0.570*** (0.0350)	-0.203*** (0.0436)	0.549*** (0.0495)	-0.147*** (0.0355)	0.524*** (0.0688)
Log of OwnIT	-0.081*** (0.0186)	0.141*** (0.0207)	-0.081*** (0.0149)	0.143*** (0.0179)	-0.085*** (0.0262)	0.144*** (0.0274)	-0.010 (0.0248)	0.188*** (0.0515)
Log of Intermediate Inputs	0.380*** (0.0342)		0.366*** (0.0298)		0.405*** (0.0496)		0.204*** (0.0520)	
Log Supplier IT	-0.020 (0.0189)		-0.010 (0.0158)		-0.004 (0.0223)		-0.179** (0.0809)	
Log of OwnIT * Log of Intermediate Inputs	0.004 (0.0150)		0.001 (0.0108)		0.001 (0.0201)		-0.021 (0.0167)	
Log of OwnIT * Log of SupplierIT	-0.007 (0.0170)		0.0005 (0.0115)		-0.008 (0.0276)		-0.058** (0.0269)	
Year Dummies	***	***	***	***	***	***	***	***
Sector Dummies							0.246*** (0.0573)	
Normalized DBL		-1.078*** (0.143)		-1.009*** (0.128)		-1.317*** (0.169)		-0.623*** (0.221)
Constant	-0.849*** (0.197)	2.877*** (0.280)	-0.754*** (0.163)	2.715*** (0.256)	-0.878*** (0.243)	2.974*** (0.354)	1.332 (0.8228)	3.713*** (0.826)
Observations	935	1,020	425	510	340	425	570	570
R-squared	0.449	0.901	0.694	0.920	0.300	0.896	0.615	0.812
Hansen/Sargan C-test Statistics	Chi-sq(4) = 5.492 P = 0.241	Chi-sq(1) = 2.110 P = 0.146	Chi-sq(4) = 5.624 P = 0.229	Chi-sq(1) = 0.270 P = 0.603	Chi-sq(4) = 2.50 P = 0.643	Chi-sq(1) = 0.378 P = 0.539	Chi-sq(4) = 9.411 P = 0.052	Chi-sq(1) = 0.149 P = 0.700

Notes: Details of the year-fixed effects are suppressed for brevity. Standard errors are in parentheses following the estimated coefficients. *p<0.10, **p<0.05, ***p<0.01.

Figure 3.5: The Endogeneity Tests for Suspected Endogenous Variables

3.3.4 Robustness Tests

Estimates with Instrumental Variables In general, endogeneity can be caused by omitted variables, measurement errors, and simultaneity (Wooldridge 2002). It is possible that some omitted variables, such as other productivity-related organizational initiatives, may be correlated with the independent variables in our models. In addition, there is potential for measurement errors in variables relating to inputs and output due to aggregation, sampling, etc., such that the measure of DBL is likely subjected to measurement errors as well. There might also be concerns about endogeneity caused by simultaneity. Simultaneity arises when explanatory variables in our models are determined simultaneously with the dependent variables. For example, a decrease in DBL is associated with an increase in value-added through TFP, and the increase in value-added may reduce an industry's direct requirement for intermediate inputs from suppliers, leading to a reduction in DBL.

Our econometric adjustments with PSAR1, normalization of DBL, year-fixed effects, and time-invariant industry-specific effects help relieve these endogeneity concerns. We further address such concerns by providing estimates using instrumental variables (IVs). Firstly, we use lags of the variables as IVs, similar to previous related studies (Stiroh 2002a, Han et al. 2011b, Cheng and Nault 2012, Han and Mithas 2013). Specifically, for the estimation of (3.5), we suspect that the variables we are interested, IT capital, suppliers' IT, intermediate inputs, and the interaction term "Own IT * Intermediate Inputs", might be endogenous, so we use one-year and two-year lags of those variables as IVs for the full sample and time-split samples of Dataset I. Regarding the estimation of (3.5) for Dataset II, we follow the same procedure of estimation as for Dataset I, use one-year lags of the those four variables as instruments because we have fewer years of data. For the estimation of (3.7), we are interested in whether NDBL is endogenous; we use one year lag of value added and sector dummies as instruments for both datasets since they are highly correlated with NDBL.

Secondly, we employ a two-step generalized method of moments (GMM) procedure, con-

trolling for arbitrary heteroskedasticity and within-panel autocorrelation, and conduct the Hansen/Sargan C-test for the estimated model across two datasets (Baum et al. 2003). The Hansen/Sargan C-test statistics reported in Figure 3.5 cannot reject the exogeneity of suspect endogenous variables for the estimation for two datasets ($\chi^2(4) = 5.492, p = 0.241$ for (3.5), and $\chi^2(1) = 2.110, p = 0.146$ for (3.7) for the full sample of Dataset I, and $\chi^2(4) = 9.411, p = 0.052$ for (3.5), and $\chi^2(1) = 0.149, p = 0.700$ for (3.7) for Dataset II).

Next, we use "two-step efficient GMM in the presence of arbitrary heteroskedasticity and autocorrelation" to estimate our models with instrumental variables as "the advantages of GMM over IV are clear: if heteroskedasticity is present, the GMM estimator is more efficient than the simple IV estimator, whereas if heteroskedasticity is not present, the GMM estimator is no worse asymptotically than the IV estimator" (Baum et al. 2003, pp.11). As the presence of heteroskeasticity is suggested by our tests for both datasets, and there is panel-specific autocorrelation, we use "two-step efficient GMM in the presence of arbitrary heteroskedasticity and autocorrelation". We estimate (3.5) for Dataset I without sector dummies because the singleton dummy variables cause the estimated covariance matrix of moment conditions to be less than full rank, and thus it can not generate optimal weighting matrix for GMM estimation.

The estimated results are presented in Figure 3.6, Columns 1 to 6 for different samples of Dataset I and Columns 7 and 8 for Dataset II. As can be seen, the estimates of IT capital (Columns 1, 3, and 5) and the estimates of normalized DBL (Columns 2, 4, 6 and 8) are negative and significant, which is similar to the main results. Regarding the estimates for the interaction term "Own IT * Intermediate Inputs" shown in Columns 1, 3, 5, and 7 of Figure 3.6, neither is significant, but the one for Dataset II is still negative. The negative sign is consistent with the estimation result of this interaction term reported in Column 5 of Figure 3.3. The non-significance of the estimate of the interaction term might be due to the loss of efficiency of the estimation with instrumental variables. When turning to the efficient

GMM estimation, there is a cost of loss of efficiency for the sake of consistency (Baum et al. 2003, pp. 19-20), which may cause large asymptotic variance of the GMM estimator, so it is reasonable that the GMM estimators of variables we are interested are not significant.

Variable	1987–1999		1987–1993		1994–1999		1998–2008	
	Log of Normalized DBL (1)	Log of Value-added (2)	Log of Normalized DBL (3)	Log of Value-added (4)	Log of Normalized DBL (5)	Log of Value-added (6)	Log of Normalized DBL (7)	Log of Value-added (8)
Log of Non-IT Capital	-0.085***	0.319***	-0.080***	0.304***	-0.112***	0.337***	-0.038	0.225***
	(0.0292)	(0.0403)	(0.0262)	(0.0357)	(0.0368)	(0.0524)	(0.0405)	(0.0677)
Log of Labor	-0.214***	0.563***	-0.241***	0.591***	-0.206***	0.554***	-0.148***	0.530***
	(0.0350)	(0.0433)	(0.0288)	(0.0404)	(0.0467)	(0.0517)	(0.0375)	(0.0703)
Log of OwnIT	-0.077***	0.146***	-0.088***	0.146***	-0.073***	0.142***	0.004	0.192***
	(0.0189)	(0.0208)	(0.0155)	(0.0178)	(0.0282)	(0.0272)	(0.0268)	(0.0533)
Log of Intermediate Inputs	0.375***		0.377***		0.395***		0.163***	
	(0.0347)		(0.0304)		(0.0502)		(0.0585)	
Log Supplier IT	-0.020		-0.015		-0.014		-0.166**	
	(0.0195)		(0.0161)		(0.0256)		(0.0848)	
Log of OwnIT * Log of Intermediate Inputs	0.011		0.003		0.003		-0.010	
	(0.0154)		(0.0110)		(0.0210)		(0.0181)	
Log of OwnIT * Log of SupplierIT	-0.011		-0.003		-0.021		-0.047	
	(0.0188)		(0.0119)		(0.0370)		(0.0315)	
Year Dummies	***	***	***	***	***	***	***	***
Sector Dummies							0.307***	
							(0.0618)	
Normalized DBL		-0.872***		-0.915***		-1.274***		-0.555*
		(0.184)		(0.154)		(0.214)		(0.301)
Constant	-0.808***	2.603***	-0.709***	2.556***	-0.857***	2.921***	1.230	3.613***
	(0.201)	(0.320)	(0.174)	(0.310)	(0.262)	(0.381)	(0.932)	(0.873)
Observations	935	1,020	425	510	340	425	570	570
R-squared	0.441	0.895	0.701	0.919	0.301	0.894	0.627	0.810

Notes: Details of the year-fixed effects are suppressed for brevity. Standard errors are in parentheses following the estimated coefficients. *p<0.10, **p<0.05, ***p<0.01.

Figure 3.6: The Estimation Results Using IVs

Estimates with redefined IT variable We use total assets of IT capital in our model as the impact of IT on the utilization of intermediate inputs is through automation, streamlining, and reengineering, which requires the support of a variety of information systems. In addition, as the implementations of information systems have become increasing integrated, it is hard to tell which one is for what purpose. For example, ERP systems are organization-wide and integrate functions dealing with internal management—such as the functions of

production management, accounting management, and human resource management—and external coordination, such as supply chain management and customer relationship management. Therefore, it is reasonable to use total IT assets as the IT variable.

Variable	1987–1999		1987–1993		1994–1999		1998–2008		
	Log of Normalized DBL (1)	Log of Value-added (2)	Log of Normalized DBL (3)	Log of Value-added (4)	Log of Normalized DBL (5)	Log of Value-added (6)	Log of Normalized DBL (7)	Log of Value-added (8)	Log of Value-added (9)
Log of Non-IT Capital	-0.124***	0.484***	-0.114***	0.403***	-0.187***	0.483***	0.003	-0.038**	-0.026
	(0.0157)	(0.0213)	(0.0107)	(0.0162)	(0.0191)	(0.0190)	(0.0095)	(0.0162)	(0.0201)
Log of Labor	-0.236***	0.494***	-0.223***	0.539***	-0.206***	0.481***	-0.309***	0.680***	0.749***
	(0.0146)	(0.0232)	(0.0092)	(0.0167)	(0.0150)	(0.0188)	(0.0118)	(0.0126)	(0.0142)
Log of OwnIT	-0.029***	0.076***	-0.030***	0.061***	-0.068***	0.089***	-0.051***	0.109***	0.081***
	(0.0081)	(0.0076)	(0.0047)	(0.0070)	(0.0098)	(0.0059)	(0.0090)	(0.0124)	(0.0149)
Log of Intermediate Inputs	0.370***		0.347***		0.419***		0.363***		
	(0.0128)		(0.0101)		(0.0167)		(0.0148)		
Log Supplier IT	-0.033***		-0.047***		-0.089***		-0.072***		
	(0.0105)		(0.0087)		(0.0133)		(0.0267)		
Log of OwnIT * Log of Intermediate Inputs	0.006		0.005		-0.032***		0.013***		
	(0.0055)		(0.0033)		(0.0079)		(0.0040)		
Log of OwnIT * Log of SupplierIT	-0.015***		-0.006*		-0.019***		-0.014		
	(0.0048)		(0.0032)		(0.0062)		(0.0100)		
Year Dummies	***	***	***	***	***	***	***	***	***
Sector Dummies	***	***	***	***	***	***	-0.268***	-0.082***	-0.456***
							(0.0193)	(0.0271)	(0.0332)
Normalized DBL		-1.703***		-1.284***		-1.698***		-0.916***	
		(0.0365)		(0.0378)		(0.0357)		(0.0218)	
Constant	-0.462***	2.786***	-0.363***	2.911***	0.039	2.837***	-0.937***	6.839***	5.626***
	(0.0892)	(0.153)	(0.0679)	(0.128)	(0.113)	(0.132)	(0.276)	(0.181)	(0.209)
Observations	1,040	1,040	560	560	480	480	591	591	591
The Number of Industries	80	80	80	80	80	80	57	57	57

Notes: Own IT is redefined as a subset of total of IT — communication. Non-IT capital is calculated as follows: Structure + Equipment - Communication. Details of the year-fixed effects are suppressed for brevity. Standard errors are in parentheses following the estimated coefficients. *p<0.10, **p<0.05, ***p<0.01.

Figure 3.7: The Estimation Results Using Communication as IT Variable

Variable	1987–1999		1987–1993		1994–1999		1998–2008		
	Log of Normalized DBL (1)	Log of Value-added (2)	Log of Normalized DBL (3)	Log of Value-added (4)	Log of Normalized DBL (5)	Log of Value-added (6)	Log of Normalized DBL (7)	Log of Value-added (8)	Log of Value-added (9)
Log of Non-IT Capital	-0.127*** (0.0122)	0.454*** (0.0192)	-0.126*** (0.010)	0.377*** (0.0143)	-0.135*** (0.0105)	0.435*** (0.0212)	0.022** (0.0096)	-0.022 (0.0183)	-0.007 (0.0237)
Log of Labor	-0.246*** (0.0129)	0.522*** (0.0201)	-0.234*** (0.0097)	0.549*** (0.0143)	-0.246*** (0.0151)	0.524*** (0.0171)	-0.282*** (0.0113)	0.659*** (0.0132)	0.753*** (0.0152)
Log of OwnIT	-0.032*** (0.0088)	0.040*** (0.0092)	-0.021*** (0.0052)	0.058*** (0.0072)	-0.088*** (0.0103)	0.071*** (0.0109)	-0.057*** (0.0095)	0.089*** (0.0130)	0.083*** (0.0188)
Log of Intermediate Inputs	0.373*** (0.0120)		0.359*** (0.0104)		0.424*** (0.0139)		0.335*** (0.0145)		
Log Supplier IT	-0.037*** (0.0104)		-0.037*** (0.0082)		-0.084*** (0.0139)		-0.083*** (0.0277)		
Log of OwnIT * Log of Intermediate Inputs	-0.005 (0.0038)		0.006** (0.0030)		-0.027*** (0.0056)		0.013*** (0.0044)		
Log of OwnIT * Log of SupplierIT	-0.016*** (0.0040)		-0.015*** (0.0031)		-0.016*** (0.0052)		0.0001 (0.0098)		
Year Dummies	***	***	***	***	***	***	***	***	***
Sector Dummies	***	***	***	***	***	***	0.258*** (0.0202)	-0.0100 (0.0245)	-0.433*** (0.0322)
Normalized DBL		-1.707*** (0.0342)		-1.336*** (0.0403)		-1.646*** (0.0379)		-0.902*** (0.0221)	
Constant	-0.335*** (0.0868)	3.084*** (0.140)	-0.372*** (0.0636)	3.148*** (0.120)	-0.123 (0.0992)	3.040*** (0.134)	-0.828*** (0.291)	6.812*** (0.197)	5.301*** (0.220)
Observations	1,105	1,105	595	595	510	510	591	591	591
The Number of Industries	85	85	85	85	85	85	57	57	57

Notes: Redefined IT2: own IT is redefined as the sum of communication and computers in Dataset I, and it is redefined as the sum of communication, software, and computers in Dataset II. Non-IT capital is calculated as follows: Structure + Equipment - Redefined IT2. Details of the year-fixed effects are suppressed for brevity. Standard errors are in parentheses following the estimated coefficients. *p<0.10, **p<0.05, ***p<0.01.

Figure 3.8: The Estimation Results Using Redefined IT2 as IT Variable

For a robustness check, we consider different types of IT assets as the IT variable in our model. Instead of using total IT assets, we use a subset of IT assets as the IT variable in (3.5) and (3.7). Specifically, we use the IT asset of communication as the first redefined IT for two Datasets, and use the sum of communication and computers and the sum of

communication, software, and computers as the second type of redefined IT capital (redefined IT2) in Datasets I and II, respectively. In addition, we adjust non-IT capital by subtracting redefined IT capital from the sum of structure and equipment.

We follow the same procedure to estimate models, the estimation results shown in Figure 3.7 and Figure 3.8 are similar to those in the main results. As shown in Columns 1, 3, and 5 of both figures, the estimates of the interaction term "OwnIT * IntermediateInputs" are negative and significant after 1993, suggesting that IT reduces the impact of intermediate inputs on DBL. Regarding the estimation results for Dataset II, the estimate of non-IT capital is abnormal, which is negative and/or not significant in Columns 8 and 9 of Figure 3.7 and Figure 3.8. Theoretically, the contribution of non-IT capital is positive in production theory, and many studies based on production theory have empirically shown the positive impact of non-IT capital (e.g., Brynjolfsson and Hitt 1996, Cheng and Nault 2007, Han et al. 2011b, Cheng and Nault 2012). The negative and/or not significant estimate of non-IT capital suggests that using a subset of IT capital as IT variable is not appropriate for the models based on production theory for Dataset II. A possible reason for a negative or not significant estimate of non-IT capital is that a partial of natural IT capital is counted as non-IT capital, resulting in high correlation between IT and non-IT capitals, and thus leads to a non-significant estimate of non-IT capital. As the basic production model does not work well for Dataset II when using redefined IT variable, the estimate of the interaction term "Own IT * Intermediate Inputs" based on Dataset II is meaningless.

3.4 Conclusion

Our study makes three contributions. Firstly, we find that an industry's IT investment can reduce its production interdependence with upstream suppliers by improving the efficiency of intermediate inputs. Our results suggest that an industry's own IT enhances the efficiency of intermediate inputs by improving existing business processes or creating new business pro-

cesses, and facilitating information sharing and coordination within and between industries. Thus, a given industry needs fewer intermediate inputs from suppliers to produce a certain level of output. As a result, the industry reduces its production interdependence with its suppliers for providing resources to generate economic values.

Secondly, we contribute to the literature on the impact of IT on TFP. Previous studies on the relationship between IT and TFP examined the impact of IT spillovers from upstream suppliers and downstream customers, as well as the impact of organizational capital, which is related to the long-term contribution of IT capital to outcomes. Our study shows that the impact of an industry's IT on TFP also comes from reshaping its production structure by reducing its production interdependence with suppliers, as measured by DBL. The information sharing and coordination captured by an industry's IT investment substitute for intermediate inputs from suppliers and improve the efficiency of intermediate inputs. Consequently, IT reduces DBL, and the reduction of DBL is captured by the growth of TFP.

Thirdly, we have identified an indirect effect of IT on value-added through DBL. Previous studies showed that IT investment can directly contribute to value-added as a production input, and that IT also has an impact on production by augmenting other production factors, such as non-IT capital and labor. Our study shows that IT has an indirect effect on value-added by reducing its production interdependence with suppliers, DBL. Specifically, an industry's IT investment reduces its direct requirements for intermediate inputs from upstream suppliers to produce a certain level of output, and this shift of production frontier is captured by TFP, leading to greater value-added.

3.4.1 Managerial Implications

One implication of our study is that industries' IT investments help reduce their interdependence with upstream suppliers in terms of purchases of goods and services for production. Investment in IT by an industry facilitates information sharing and coordination within and between industries, allows integration of business processes among its supply chain part-

ners, and enables its synchronization with upstream suppliers in terms of product design, production scheduling, and inventory replenishment. Interestingly, our study suggests that although an industry's IT investment increases the integration between itself and its suppliers in terms of business processes, its IT investment reduces its production interdependence with upstream suppliers by substituting for goods and services purchased from suppliers and enhancing the efficiency of their use. As a result, the industry needs fewer goods and services to produce a certain level of output. Consequently, IT both integrates and reduces interdependence of production in the supply chain.

Our industry-level study also sheds lights on the operation and management of firms. Manufacturing firms need to consider different types of costs of goods sold, including costs of materials used in production, the direct labor costs used to produce the product, and manufacturing overhead, etc. Our results suggest that firms' IT investments help reduce their dependence on upstream suppliers and create more value-added; thus firms should preferentially direct their IT investments towards improving the efficiency of purchased goods and services in order to reduce their production interdependence with suppliers. In addition, firms may realize improved negotiating power over their suppliers. For example, manufacturing firms may be able to negotiate the prices of goods and services purchased from suppliers because the alternative is to spend money on IT to reduce the requirements of goods and services.

From a supply chain perspective, our study implies one means to create more value-added in a supply chain. If each firm in a supply chain invests in IT to improve its efficiency of goods and services purchased from upstream suppliers, the whole supply chain becomes more efficient. As a result, more value-added is created by the supply chain.

3.4.2 Limitations and Future Research

It is worth noting a limitation of our study. Our industry-level data does not allow a direct estimate of the impact of IT on business processes and efficiency, while firm-level data would

enable us to investigate the intangible benefits of IT on production. However, firm-level data that covers the breadth of the economy that our industry-level data covers is not available. Our industry-level data allows us to investigate the economy-wide impact of IT on value-added through DBL, which increases the generalizability of our study.

In future research, we can examine the impact of IT on other aspects of production structure. For example, we may examine the impact of IT on direct and indirect backward linkage, which indicates the change in the output of an economic production system in response to one unit increase in the final demand for an industry's product. We may also examine how an industry's IT investment affects its direct forward linkage, which represents the increased total amounts of goods and services required by downstream industries to directly utilize output from one unit of primary input into an industry.

Chapter 4

THE IMPACT OF IT ON THE STRUCTURE OF PRODUCTION

4.1 Introduction

In our economy, industries are connected in production processes. An industry purchases goods and services from a variety of supplier industries, produces products, and sells those products to downstream customer industries. Each industry is involved in an economic network, functioning as a node, and has ties with upstream suppliers and downstream customers through goods and services flows. Intermediate inputs flowing from upstream suppliers to an industry are transformed as products by the given industry and then flow downstream customer industries. Those input-output relationships among industries in the economic network form the structure of production in the economy.

Over the past few decades, we have seen the changes in the structure of production. Firstly, there is a change in the connectivity in an economic network: some ties between industries are absent and some new ties are present. Secondly, some industries reallocate the amounts of inputs from supplier industries, resulting in a different distribution of the amounts of inputs. As organizations in industries are information processing entities (Galbraith 1974), one might reasonably suspect that the changes in the structure of production for industries, which are aggregations of organizations, are related to the massive deployment of IT.

We have observed a steadily increasing trend in IT investment, and the deployment of IT has greatly reduced transaction costs in the economy. Powerful information processing capabilities brought about from IT investment relieves decision makers' bounded rationality, reducing costs in transactions due to limited information processing and communication

(Gurbaxani and Whang 1991, Clemons et al. 1993). And with the increase in IT investment, industries have digitalized production processes (McAfee and Brynjolfsson 2008). Industries can check, track, and monitor the production processes in real time; they can trace back to historical production data; and they can forecast production activities more accurately, resulting in a reduction in uncertainty and opportunism in transactions. The digitalization of production processes also allows process innovations to be widely propagated throughout in the economy, greatly enhancing the scope of the impact of IT (McAfee and Brynjolfsson 2008). Finally, the applications of interorganizational information systems (IOSs), such as Enterprise Resource Planning (ERP), Supply Chain Management (SCM), and Customer Relationship Management (CRM), facilitate information sharing and coordination within and between industries. In turn, there is a reduction in coordination costs in terms of communicating environmental and behavioral changes, renegotiating, and rescheduling production along supply chains. Therefore, transaction costs in production processes along supply chains are substantially reduced. Consequently, it is less costly for industries to reshape their supply chains by streamlining and reengineering their production processes, resulting in the change in the structure of production in the economy.

Previous research has examined the relationship between IT and organizational structure, such as the impact of IT on firm size, the number of suppliers, and vertical integration, etc. Brynjolfsson et al. (1994) found that IT was associated with smaller firm size based on transaction costs theory. Im et al. (2013) argued for an inverse relationship between IT and firm size, and found a sequential interaction between IT and firm size through transaction costs as a mediator. Bakos and Brynjolfsson (1993) applied the theory of incomplete contracts to determine the optimal number of suppliers, and found that buyers could maximize profits by limiting the number of suppliers, giving suppliers more incentives to invest in non-contractibles such as innovation, responsiveness, and information sharing. Also based on the theory of incomplete contracts, Banker et al. (2006) found that IT increased contract com-

pleteness, resulting in an increase in the costs of coordinating with suppliers and monitoring contract terms, thus the number of suppliers was reduced. Dedrick et al. (2008) examined the relationship between IT usage (particularly electronic procurement) and the number of suppliers, and found that the use of electronic procurement was associated with purchasing from more suppliers for custom goods but from fewer suppliers for standard goods. Wang and Seidmann (1995) analyzed how a supplier's adoption of Electronic Data Interchange (EDI) affected its competitive position in a simple two-level hierarchical market, and found that the adoption of EDI could increase its own production profits and generate negative externalities on the other suppliers in terms of reducing their production profits. Based on transaction costs theory, Hitt (1999) examined the relationship between IT and the firm structure, measured as vertical integration and diversification, and found a strong negative relationship between IT and firm structure in both causal directions. In sum, previous studies focus on the impact of IT at the firm level; however, few studies have examined the impact of IT on the structure of production in the economy. This aspect is important in order to learn more about the broader impact of IT.

The principal aim of this study is to empirically examine the relationship between IT and changes in the structure of production in the economy. To address this question, we obtained economy-wide data on investment in IT in the U.S. and employ a network analysis approach to generate different measures of the structure of production in terms of connectivity and concentration. Our results show that IT investment of an industry is associated with an increase in connectivity within its supplying network and a decrease in concentration in the supplying market where it purchases intermediate inputs. We also find that IT may affect connectivity and concentration of manufacturing industries differently.

The remainder of the paper is organized as follows: section 2 describes the measures of the structure of production; section 3 introduces our data and methodology, including a description of data and variables, factor analysis, estimation models, and econometric

adjustments; section 4 presents estimation results, robustness tests, and granger causality analysis; and section 5 concludes the paper.

4.2 Our Measures of the Structure of Production

According to network theory, a network has two important components: nodes, representing individual actors in the network and ties, representing the relationship between actors (Hanneman and Riddle 2005). An input-output table shows how industries purchase intermediate inputs from upstream suppliers and sell output to downstream customers, and it can be presented as a directed (in/out) network, as shown in Figure 4.1 as an example. The ego-industry purchases goods and services from upstream suppliers, alter industries A, B, and C, with transaction volumes of 10, 20, and 70, respectively. Suppliers purchases goods and services from suppliers' suppliers upstream. The products of the ego-industry are sold to downstream customers, alter industries D, E, and F, with transaction volumes of 20, 20, and 60, respectively. Customer industries sell their products to customers' customers downstream.

In Figure 4.1, the first-tier in-neighborhood consists of the ego-industry's direct suppliers, and the first-tier out-neighborhood consists of its direct customers. The ego-industry and its first-tier in or out neighborhood form an egocentric network. The strengths of ties in this scenario are the transaction volumes among industries in the network. We capture an industry's structure of production by using two categories of measures: connectivity and concentration. Figure 4.2 provides a relationship map of the measures of the structure of production.

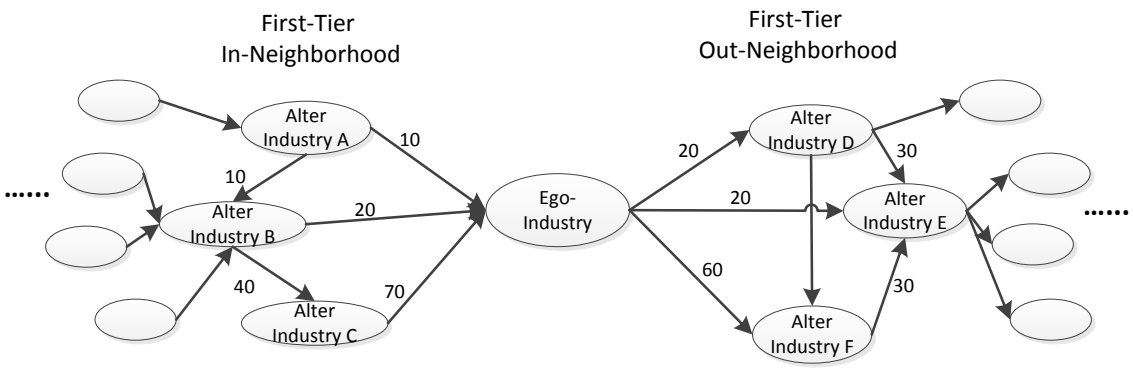

Figure 4.1: An Example of an Ego-Centric Network

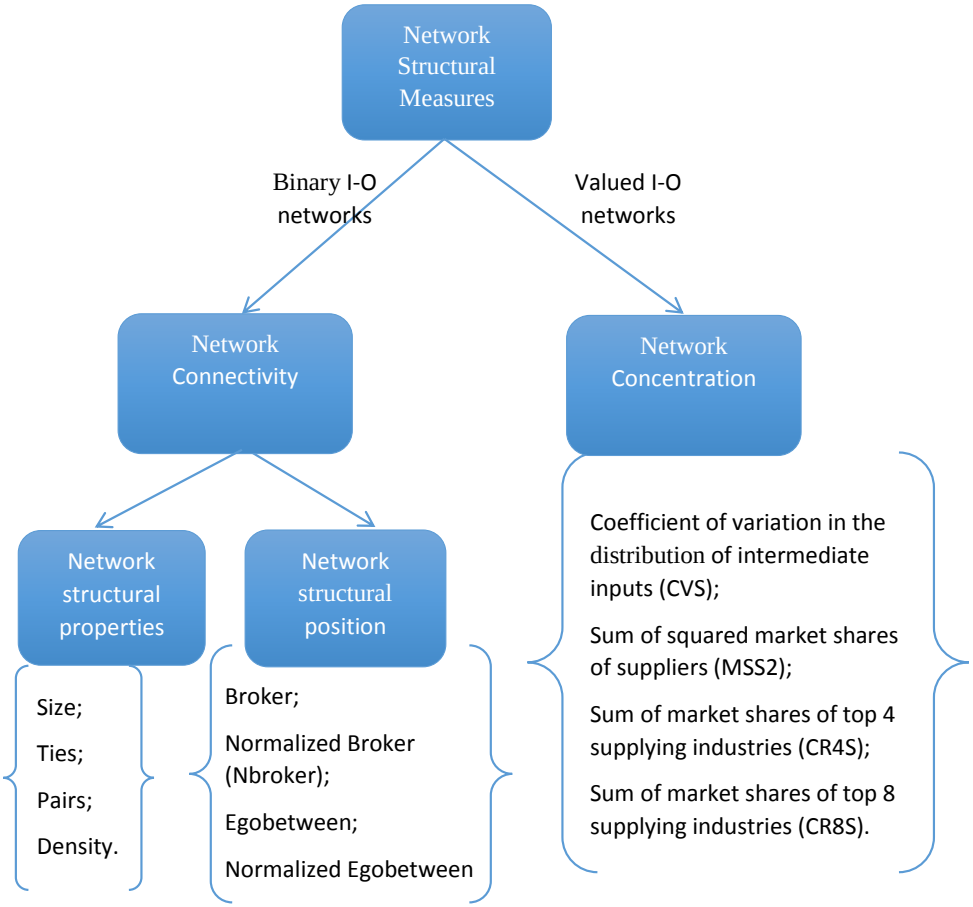

Figure 4.2: The Relationship Map of Network Measures

Basic Measures	Definition	Interpretation and Calculation
Size	The number of actors (alters) that ego is directly connected to	How many direct suppliers does an industry have; for example, 113FF has 47 direct suppliers in 2002.
Ties	The number of directed links among all alters in the ego network (not counting ties involving ego)	For example, in the network of 113FF as ego, there are 2011 directed ties between all of its suppliers in 2002.
Pairs	The number of (potential maximum) possible directed ties in each ego network	The number of directed pairs among an industry's suppliers; for example, in the network of 113FF as ego, it has 2162 ordered pairs, which is permutation p (47, 2).
Density	The number of ties divided by the number of pairs, times 100	What percentage of all possible ties in each ego network is actually present? eg. The density of the network with l13FF as ego is 2011/2162*100=93.02.
Broker	The number of times ego lies on the shortest path between two alters (i.e., the number of pairs of alters that are not directly connected)	The idea of brokerage is that ego is the "go-between" for pairs of other actors. $Broker = (Pairs - Ties)/2$
Normalized Broker (Nbroker)	Broker divided by number of pairs	It assesses the extent to which ego's role is that of broker. $Nbroker = Broker/Pairs$
Ego Betweenness (Egobetween)	The sum of ego's proportion of times ego lies on the shortest path between each part of alters	$Egobetween = \sum_{j<k} \frac{g_{jk}(n_i)}{g_{jk}}$, where g_{jk} is the total number of geodesics linking the two nodes j and k, and $g_{jk}(n_i)$ is the number of those geodesics that containn_i. Ego is "between" two other actors if ego lies on the shortest directed path from one to the other. For alters connected to each other, the contribution to between of that pair is 0, for alters connected to each other only through ego, the contribution is 1, for alters connected through ego and one or more other alters, the contribution is 1/k, where k is the number of nodes which connects that pair of alters.
Normalized Ego Betweenness (Negobetween)	Ego Betweenness normalized by a function of the number of nodes in the ego network	$Negobetween = \frac{Egobetween}{[(g-1)(g-2)/2]} * 100$, where g is the number of nodes in an ego network.

Notes: The definitions of basic measures are adopted from Borgatti et al. (2002). UCINET 6.0 uses the formula, $Negobetween = \frac{Egobetween}{[(g-1)(g-2)]} * 100$, to generate data.

Figure 4.3: The Connectivity Measures for an Egocentric Network

4.2.1 The Measures of Connectivity

Connectivity measures are about the presence of ties between industries, generated based on binary I-O networks. We refer to basic measures of the structure of an egocentric network (Hanneman and Riddle 2005, Kim et al. 2011) to generate connectivity measures, including

Size, Ties, Density, Broker, Normalized Broker (Nbroker), Ego Betweenness (Egobetween), and Normalized Ego Betweenness (Negobetween). The measures titled Size, Ties, Pairs, and Density give us a big picture about the connections within each ego-industry's one-step supplier neighborhood; the metrics titled Broker, Normalized Broker, Ego Betweenness, and Normalized Ego Betweenness indicate how an ego-industry embeds in its supplying network. Figure 4.3 shows the definition, calculation, and interpretation of connectivity measures.

4.2.2 The Measures of Concentration

The second category of measures of the structure of production focus on the concentration in the supplying market where an industry purchases inputs. They are about the strengths of ties in an egocentric network, calculated from a valued I-O network. We generate four metrics of concentration by borrowing the concepts of classical measures of market concentration: coefficient of variation (CVS), the sum of squared market shares of suppliers ($MSS2$), the sum of market shares of four largest suppliers ($CR4S$), and the sum of market shares of eight largest suppliers ($CR8S$).

The Coefficient of Variation An industry purchases different amounts of intermediate inputs from a variety of supplying industries and the distribution of those inputs may be disperse or concentrated. We may capture the dispersion of the distribution of inputs by using the concept of the coefficient of variation—defined as the ratio of standard deviation σ to the mean μ, $CV = \frac{\sigma}{\mu}$ (Abdi 2010).

Based on input-output tables, we can calculate the coefficient of variations in the distribution of inputs, CVS, for each industry. We can compare the dispersion in intermediate inputs across industries and over different time periods. The larger a CVS, the more concentrated is the supplying market.

The Concentration Measure based on HHI Herfindahl-Hirschman index (HHI) is commonly accepted as a measure of market concentration and it is calculated by summing

the squared market share of each firm in a market where market shares are used as weights (Church and Ware 2000). We borrow the concept of HHI to generate a metric that is used to examine whether inputs are concentrated from a few large suppliers.

Considering each industry as a firm in the calculation of HHI, and all supplier industries together as a supplying market, we can obtain the market share of each supplier industry in the supplying market. We can then sum the squared market shares of the supplier industries to obtain the total of squared market shares of supplier industries—$(MSS2)_i$. The calculation is mathematically specified as follows:

$$(MSS2)_i = \sum_{j=1}^{n} \left(\frac{m_{ji}}{\sum_{j=1, j \neq i}^{n} m_{ji}} \right)^2,$$

where n is the number of suppliers of industry i, m_{ji} is the dollar values of intermediate inputs purchased from industry j by industry i. The larger $(MSS2)_i$, the more concentrated is the supplying market of industry i.

The Concentration Measures Based on Concentration Ratios To understand the distribution of the inputs from suppliers, we also borrow other classical measures of concentration in a market—four firms concentration ratio, $CR4$, and eight firms concentration ratio, $CR8$ (Church and Ware 2000).

Similarly, we consider all supplier industries together as a supplying market and each supplier industry as a firm in this supplying market. We obtain a concentration measure, $CR4S$, by summing the market shares of the top four largest supplier industries, and obtain $CR8S$ by summing the market shares of top eight largest supplier industries. For example, the calculation of $CR4S$ for the supplying market of industry i, $CR4S_i$, is mathematically specified as

$$CR4S_i = \sum_{j=1}^{4} S_j = \sum_{j=1}^{4} \frac{m_{ji}}{\sum_{j=1, j \neq i}^{n} m_{ji}},$$

where S_j is the market share of the jth largest supplier, m_{ji} is the dollar values of intermediate inputs purchased from j^{th} largest suppliers of industry i, and n is the total number of

suppliers of industry i. The larger $CR4S$ or $CR8S$, the more concentrated is the supplying market.

4.3 Data and Methodology

4.3.1 Dataset I: 1987-1999

We used two Datasets in our analyses. These Datasets cover different time periods and have different levels of aggregation.

Dataset I is an industry-level dataset based on the 3-digit Standard Industrial Classification (SIC) level. It has three sources from the U.S. Bureau of Labor Statistics (BLS): the multifactor productivity (MFP) dataset for 140 manufacturing industries from 1987 to 1999; the information capital and capital stock for 140 manufacturing industries from 1987 to 1999; and the input-output (I-O) use tables for 98 manufacturing industries and 78 other industries with 3-digit SIC codes from 1983 to 2000. Some rows/columns of I-O tables which have specific industry numbers in I-O use tables are the combination of more than one 3-digit SIC industries in MFP data. In order to match the two sources of data, we dropped non-manufacturing industries in I-O use tables and aggregated data in the MFP dataset, IT capital stock, and non-IT capital stock according to the correspondence between industry numbers and SIC codes listed in I-O tables, generating data for 98 manufacturing industries from 1987 to 1999.

The information capital stock consists of computers and related equipment, office equipment, communication, instruments, photocopy and related equipment; all values are in millions of 1987 dollars. We used the total information capital stock as IT capital, Z. Capital stock (in millions of 1987 dollars) consists of equipment, structures, inventories, land, and special tools. In order to obtain non-IT capital, K, we summed the equipment and structure components and then subtracted the total IT capital stock from the sum.

We obtained the data on output and labor from the MFP dataset, which contains the

data on the series of output in millions of nominal dollars and has the output deflators with 1987 as the base year. Dividing the series of output by the corresponding deflators and multiplying by 100, we obtained the series for output Y in millions of 1987 dollars, respectively. We also obtained the labor input L in millions of all employee hours from the MFP data.

To obtain the variables about connectivity measures, we employed a popular software product in network analysis—UCINET 6 (Borgatti et al. 2002). During the procedure of calculation, each I-O table with original transaction volumes in cells, a 98*98 matrix, is used as an input network, and UCINET 6 automatically dichotomizes the matrix at zero transaction volume as a cut-point, turning valued data into binary data. We chose "in-neighborhood" as the type of ego-neighborhood as we are interested in an ego-industry's first-tier suppliers. The calculation generated the measures we are interested in: Size, Ties, Pairs, Density, Broker, Normalized Broker, Ego Betweenness, and Normalized Ego Betweenness.

We directly calculated concentration measures from the 98*98 I-O use tables. We obtained the mean and standard deviation of transaction volumes for each column of an I-O use table. Dividing the standard deviation of transaction volumes in the column by the mean, we obtained the measure of coefficient of variation, CVS. To determine the other three concentration measures in the supplying market, we computed the relative proportion of intermediate inputs purchased from each supplier to the total volumes of intermediate inputs, and used the proportion as the market share of a supplier in the supplying market for an industry. Next, we followed the corresponding mathematical formulas specified above to calculate $MSS2$, $CR4S$, and $CR8S$.

Based on 98 manufacturing industries, we dropped the industries with missing data and those without customer industries. Because of missing intermediate inputs data from 1997 to 1999, we dropped 6 industries: Logging (SIC 241), Newspapers (SIC 271), Periodicals (SIC 272), Books (SIC 273), Miscellaneous Publishing (SIC 274), and Greeting Cards (SIC

277). Next, we eliminated 7 industries which did not supply to others: Ordnance and Ammunition (SIC 348), Aerospace (SIC 372, 376), Ship and Boat Building and Repairing (SIC 373), Railroad Equipment (SIC 374), Toys and Sporting Goods (SIC 394), Footwear except Rubber and Plastic (SIC 313, 314), and Tobacco Products (SIC 21). In total, we have a balanced panel of 85 industries across 13 years, which is similar to that used by Cheng and Nault (2007), Han et al. (2011b), Cheng and Nault (2012) and Gong et al. (2013).

4.3.2 Dataset II: 2000-2009

The second dataset is based on the 3-digit 2002 North American Industry Classification Systems (NAICS) codes and covers 10 years from 2000 to 2009. It consists of the MFP data for the 3-digit NAICS codes obtained from the BLS, gross domestic product (GDP) by industry obtained from the Bureau of Economic Analysis (BEA) as well as the I-O use tables from 2000 to 2009 that were also obtained from the BEA. The GDP industry accounts and I-O use tables from the BEA contain data for 61 industries. The MFP databases from the BLS provide us the data on IT capital, non-IT capital, and labor for 59 industries. In order to match the data from the BEA and BLS, we dropped farms industry and combined the industry of Motor Vehicles, Bodies and Trailers, and Parts (NAICS 3361MV) and the industry of Other Transportation Equipment (NAICS 3364OT) as the industry of Transportation Equipment (NAICS 336); this generated 59 industries based on 3-digit NAICS codes with matched data. The data details for our variables are as follows.

Obtained from the MFP databases, IT capital stock, in millions of 2005 dollars, includes computers, software, communication, and others. We used the total as IT capital, Z. The IT assets in Dataset II are different from those in Dataset I as software is included to reflect the importance of software in the more recent period. Capital stock, in millions of 2005 dollars, includes equipment, structures, rental residential capital, inventories, and land. In order to calculate non-IT capital, K, we aggregated equipment and structures and then subtracted IT capital. Our measure of labor input, L, is in millions of hours.

Variable	Obs	Mean	Std.dev.	Min	Max
Dataset I (1987-1999)					
Gross Output	1,105	30,294	46,044	557.6	738,131
IT Capital Stock	1,105	1,814	3,166	30.30	27,661
Labor	1,105	414.4	343.0	12.20	2,351
Non-IT Capital	1,105	20,642	22,837	461.8	135,541
Broker	1,105	121.0	87.49	5.500	449.5
Nbroker	1,105	0.0666	0.0387	0.00384	0.186
Egobetween	1,105	21.73	45.92	0	357.6
Negobetween	1,105	0.564	1.102	0	8.870
CVS	1,105	4.409	1.696	2.031	9.439
MSS2	1,105	0.236	0.173	0.0519	0.910
CR4S	1,105	0.682	0.175	0.343	0.976
CR8S	1,105	0.838	0.117	0.514	0.994
Dataset II (2000-2009)					
Gross Output	580	336,423	361,112	18,597	2187837
IT Capital Stock	580	32,390	52,902	548	358,847
Labor	580	3,188	4,540	84	25,810
Non-IT Capital	580	241,216	325,607	11,266	1.925e+06
Broker	580	88.67	44.64	4	231.5
Nbroker	580	0.0730	0.0260	0.0114	0.150
Egobetween	580	7.126	7.770	0	62.17
Negobetween	580	0.287	0.264	0	2.172
CVS	580	0.130	0.136	0.0457	0.886
MSS2	580	2.168	0.882	1.126	6.713
CR4S	580	0.526	0.121	0.311	0.990
CR8S	580	0.709	0.0909	0.501	0.995

Figure 4.4: Summary Statistics

	Gross Output	IT Capital Stock	Labor	Non-IT Capital	Broker	Nbroker	Egobetween	Negobetween	CVS	MSS2	CR4S	CR8S
Gross Output	1.000											
IT Capital Stock	0.562	1.000										
Labor	0.560	0.315	1.000									
Non-IT Capital	0.538	0.554	0.472	1.000								
Broker	0.360	0.364	0.528	0.353	1.000							
Nbroker	0.321	0.344	0.450	0.322	0.909	1.000						
Egobetween	0.218	0.122	0.263	0.221	0.564	0.519	1.000					
Negobetween	0.203	0.116	0.235	0.213	0.520	0.484	0.995	1.000				
CVS	0.043	0.005	0.011	-0.089	-0.344	-0.345	-0.215	-0.200	1.000			
MSS2	0.067	0.016	0.037	-0.070	-0.302	-0.311	-0.187	-0.176	0.983	1.000		
CR4S	-0.030	-0.043	-0.063	-0.148	-0.406	-0.392	-0.240	-0.216	0.919	0.849	1.000	
CR8S	-0.083	-0.075	-0.166	-0.168	-0.481	-0.443	-0.248	-0.214	0.836	0.752	0.958	1.000

Figure 4.5: Correlation Matrix for Dataset I

	Gross Output	IT Capital Stock	Labor	Non-IT Capital	Broker	Nbroker	Egobetween	Negobetween	CVS	MSS2	CR4S	CR8S
Gross Output	1.000											
IT Capital Stock	0.370	1.000										
Labor	0.627	0.268	1.000									
Non-IT Capital	0.519	0.331	0.264	1.000								
Broker	0.276	0.222	0.445	-0.018	1.000							
Nbroker	0.271	0.215	0.431	-0.008	0.986	1.000						
Egobetween	0.223	0.159	0.460	0.038	0.761	0.769	1.000					
Negobetween	0.218	0.141	0.449	0.054	0.709	0.737	0.990	1.000				
CVS	-0.032	-0.084	-0.229	0.103	-0.208	-0.211	-0.149	-0.146	1.000			
MSS2	-0.048	-0.097	-0.206	0.058	-0.256	-0.276	-0.160	-0.163	0.961	1.000		
CR4S	-0.075	-0.083	-0.292	0.099	-0.340	-0.330	-0.238	-0.224	0.939	0.875	1.000	
CR8S	-0.066	-0.051	-0.358	0.146	-0.460	-0.433	-0.330	-0.297	0.833	0.766	0.933	1.000

Figure 4.6: Correlation Matrix for Dataset II

We obtained data on gross output through the GDP by industry accounts on the BEA website. We obtained the nominal values of gross output and the corresponding deflators. Consistent with the BEA's methodology of converting nominal values to real ones, we used chain-type quantity indices as deflators to obtain the real values of gross output Y by multiplying the 2005 current-dollar value of the series of gross output and intermediate inputs by the corresponding chain-type quantity indices, then dividing by 100.

By using each I-O table with original transaction volumes in cells, a 59*59 matrix, as an input network and following the same procedure to calculate network measures based on 1987 to 1999 dataset, we obtained connectivity and concentration measures.

After we obtained the data for 59 industries, we dropped one industry with negative transaction volume in 2003, titled as Forestry, fishing, and related activities (NAICS 113FF). Therefore, we have a panel dataset with 58 3-digit NAICS industries from 2000 to 2009.

Summary statistics for both Datasets are provided in Figure 4.4 and the correlation matrices are shown in Figure 4.5 and Figure 4.6, respectively.

4.3.3 Factor Analysis

We have a variety of metrics for connectivity and concentration measures and each captures an aspect of network structure. We conduct principal component analysis (PCA) for all metrics in interest in order to reduce the number of variables in data and extract as much variance as possible by using few components (factors). A principal component is a linear combination of those variables in interest, the loading patterns of variables indicate what variables load on which component.

In the results of PCA for both Datasets, the eigenvalues of the first two components are larger than 1 and the eigenvalue of the third component is close to 1. If the eigenvalue of a component is larger than 1, the component explains a relative larger variance in data. We conduct PCA with a specification of 2 components and 3 components for both Datasets separately. In order to obtain a relatively simple structure—variables should have high loadings on few (one) factors and factors should ideally have only low and high values, we apply orthogonal rotations to accomplish this. We use a rule of thumb—a standardized loading with a magnitude above 0.3 is considered substantive (Kline 2005).

As seen in Figure 4.7 and Figure 4.8, connectivity measures are loaded on one component and concentration measures are loaded on the other component. The first two principal components explain a partial of total variances of individual components, 74.85% for Dataset I and 85.57% for Dataset II. They do not contain all information in the data, and therefore some of the variances in the variables are unexplained. These unexplained variances equal the sums of squares of the loadings in the deleted components, weighted by the associated eigenvalues. The average unexplained variance is equal to the overall unexplained variance of 25.15% ($1-0.7485$) for Dataset I and 14.43% for Dataset II. In addition, the loading plots seen in Figure 4.9 and Figure 4.10 show the same patterns. Regarding the loading pattern of variables for three components, Figure 4.11 and Figure 4.12 show that concentration measures are still loaded on one component and connectivity metrics are loaded on the

other two components. The loading patterns of variables for both 2 and 3 components provide evidence that it is reasonable for us to classify our measures into connectivity and concentration.

Although we may use the principal components as dependent variables (DVs) in our models, we choose to use specific metrics for two reasons. Firstly, there are loss of variances when we collapse the measures into composites, which affects the significance of inference. Secondly, it is hard to interpret the impact of IT on a composite, which is mathematically a combination of all measures without a specific economic meaning.

Principal components/correlation — Number of obs = 1105
Number of comp. = 2
Trace = 10
Rotation: orthogonal varimax (Kaiser off) — Rho = 0.7485

Component	Variance	Difference	Proportion	Cumulative
Comp1	3.79263	.100419	0.3793	0.3793
Comp2	3.69221	.	0.3692	0.7485

Rotated components

Variable	Comp1	Comp2	Unexplained
Size	0.4185	0.0376	.3725
Ties	0.4024	0.0348	.4188
Broker	0.4530	-0.0543	.1454
Nbroker	0.3560	-0.0922	.4005
Egobetween	0.4087	0.0330	.3983
Negobetween	0.3972	0.0410	.4388
CVS	0.0237	0.5173	.0423
MSS2	0.0420	0.5037	.1128
CR4S	-0.0153	0.5004	.0542
CR8S	-0.0606	0.4598	.1315

Figure 4.7: The Factor Analysis with Two Principal Components for Dataset I

Principal components/correlation

Number of obs = 580
Number of comp. = 2
Trace = 10
Rho = 0.8557

Rotation: orthogonal varimax (Kaiser off)

Component	Variance	Difference	Proportion	Cumulative
Comp1	4.74085	.924364	0.4741	0.4741
Comp2	3.81649	.	0.3816	0.8557

Rotated components

Variable	Comp1	Comp2	Unexplained
Size	0.3548	-0.1270	.1973
Ties	0.3383	-0.1329	.2463
Broker	0.4496	0.0114	.0576
Nbroker	0.4509	0.0155	.0577
Egobetween	0.4221	0.0832	.2414
Negobetween	0.4067	0.0856	.2994
CVS	0.0616	0.5198	.05319
MSS2	0.0288	0.4944	.1087
CR4S	-0.0085	0.4955	.0494
CR8S	-0.0778	0.4376	.1317

Figure 4.8: The Factor Analysis with Two Principal Components for Dataset II

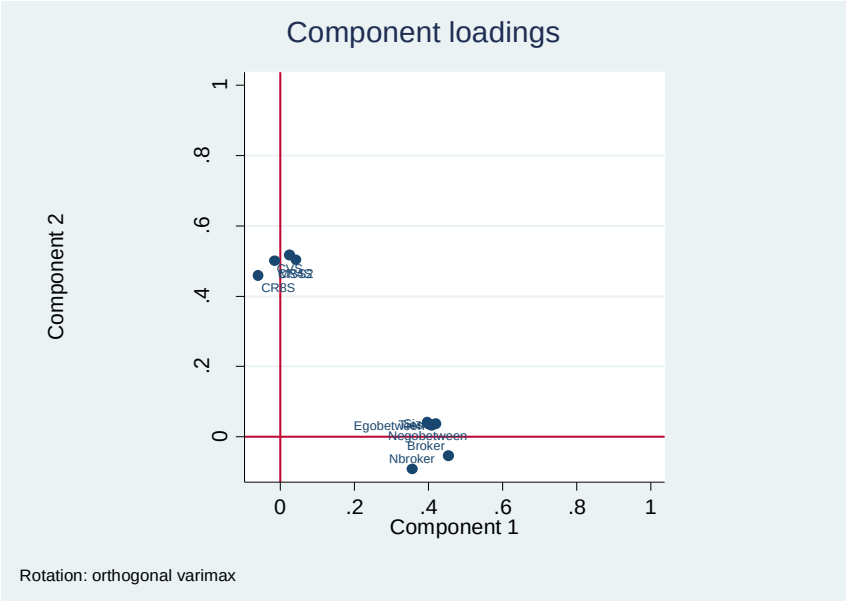

Figure 4.9: The Loading Plot of Network Measures for Dataset I

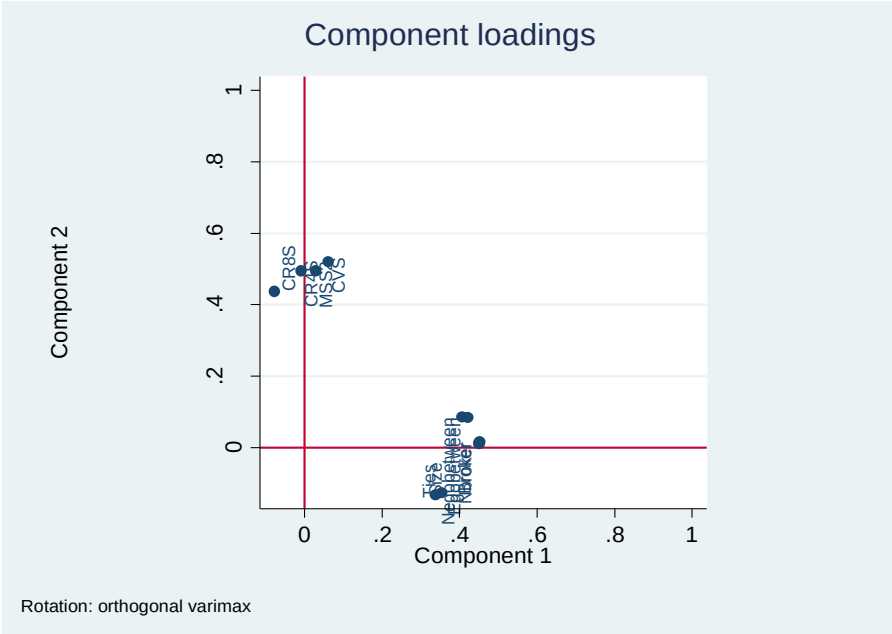

Figure 4.10: The Loading Plot of Network Measures for Dataset II

Principal components/correlation — Number of obs = 1105
Number of comp. = 3
Trace = 10
Rotation: orthogonal varimax (Kaiser off) — Rho = 0.8920

Component	Variance	Difference	Proportion	Cumulative
Comp1	3.78899	1.21429	0.3789	0.3789
Comp2	2.57469	.0179142	0.2575	0.6364
Comp3	2.55678	.	0.2557	0.8920

Rotated components

Variable	Comp1	Comp2	Comp3	Unexplained
Size	0.0212	0.6333	-0.0436	.04666
Ties	0.0190	0.6446	-0.0780	.04728
Broker	-0.0709	0.3755	0.2614	.1366
Nbroker	-0.1048	0.1532	0.3463	.3731
Egobetween	0.0192	-0.0484	0.6314	.06349
Negobetween	0.0277	-0.0687	0.6359	.07929
CVS	0.5161	0.0467	0.0132	.04159
MSS2	0.5018	0.0656	0.0194	.1115
CR4S	0.5007	-0.0066	0.0108	.0539
CR8S	0.4618	-0.0734	0.0119	.1261

Figure 4.11: The Factor Analysis with Three Principal Components for Dataset I

Principal components/correlation — Number of obs = 580
Number of comp. = 3
Trace = 10
Rotation: orthogonal varimax (Kaiser off) — Rho = 0.9548

Component	Variance	Difference	Proportion	Cumulative
Comp1	3.58064	.0333358	0.3581	0.3581
Comp2	3.5473	1.12698	0.3547	0.7128
Comp3	2.42032	.	0.2420	0.9548

Rotated components

Variable	Comp1	Comp2	Comp3	Unexplained
Size	-0.0231	0.5589	-0.0918	.01969
Ties	-0.0190	0.5802	-0.1407	.02366
Broker	0.0554	0.4193	0.2027	.04
Nbroker	0.0517	0.3958	0.2312	.04854
Egobetween	-0.0033	-0.0102	0.6460	.01106
Negobetween	-0.0130	-0.0578	0.6767	.0132
CVS	0.5459	0.0702	-0.0163	.02432
MSS2	0.5020	-0.0053	0.0214	.1015
CR4S	0.5034	-0.0278	-0.0094	.0407
CR8S	0.4356	-0.0969	-0.0325	.129

Figure 4.12: The Factor Analysis with Three Principal Components for Dataset II

4.3.4 The Models of IT and the Structure of Production

IT and the Structure of Production An industry's IT investment may affect the position of this ego-industry within its in-neighborhood network. According to transaction costs theory, IT can reduce uncertainty in transactions, relieve opportunism, and reduce internal and external transaction costs (Malone et al. 1987, Gurbaxani and Whang 1991, Clemons et al. 1993). With a reduction in transaction costs, the ego-industry is able to reshape its connection with upstream suppliers: streamlining and reengineering its production processes along supply chains. As a result, the ego-industry is more likely to be in a position to broker other industries in the network, allowing more industries to go-between it, and thus, we expect a positive effect of IT on connectivity. In addition, the ego-industry can reallocate the amounts of inputs from suppliers. With improved information sharing and coordination brought about from IT investment, the ego-industry is able to allocate inputs relatively evenly in order to reduce the bargaining power of a small number of suppliers who provide a large proportion of inputs. Thus we expect that the impact of IT on concentration measures

is negative.

The Models of IT and the Structure of Production We develop two types of models to estimate the impact of IT on connectivity and concentration. The four metrics of connectivity—Broker, Normalized Broker, Egobetween, and Normalized Egobetween—measure the embeddedness of the ego-industry in a directed egocentric network and indicate how "central" or "powerful" an ego-industry is within their own neighborhood. We are more interested in how an industry's IT affects its role in the network, so we focus on Broker, Egobetween, and their normalizations in terms of connectivity. In Model 1, we use the logarithm transformation of IT as an independent variable in order to stabilize the variance in IT variable for inference efficiency. On the left side of equations, we use four metrics of connectivity and four metrics of concentration, which are together titled as $NetworkMeasures_{it}$ in Model 1.

In Model 2, we use logarithm transformations of Broker and Egobetween, which are two variables at the levels for two reasons: one is to stabilize variances in these two variables, the other is to capture the relationships between IT and dependent variables that are non-linear in the levels, but linear in logs.

We control for year-fixed effects and industry-fixed effects by adding year dummies and industry dummies in models. The years contained in two Datasets cover changes in political and economic activities, which could affect every industry and result in changes in connections between industries, so we add year-fixed effects in our estimation models. In addition, industries differ in production processes which may drive their requirements for goods and services from suppliers, leading to a specific distribution of inputs. We add industry dummies in models in order to control for this type of industry-specific effects.

Model 1

$$NetworkMeasures_{it} = \alpha_1 + \beta_1 log_e(IT)_{it} + YearDummies + IndustryDummies + \epsilon_{it}; \quad (4.1)$$

Model 2

$$log_e(Broker_{it}) = \alpha_2 + \beta_2 log_e(IT)_{it} + YearDummies + IndustryDummies + \epsilon_{it}; \quad (4.2)$$

$$log_e(Egobetween_{it}) = \alpha_3 + \beta_3 log_e(IT)_{it} + YearDummies + IndustryDummies + \epsilon_{it}. \quad (4.3)$$

4.3.5 Econometric Adjustments

Both of our Datasets are cross-sectional time-series, so we test for autocorrelation and heteroskedasticity (HE). We anticipate autocorrelation in error terms because the connectivity and concentration measures of an industry are highly correlated with those measures in the previous year. Following the Wooldridge test for autocorrelation in panel data (Wooldridge 2002), we reject the null hypothesis of no first-order autocorrelation (AR1) at all reasonable levels of significance for both Datasets (eg. $F(1, 57) = 44.71$ when Broker as the DV in Model 1 for Dataset II, $F(1, 57) = 19.13$ when CVS as the DV in Model 1 for Dataset II). In addition, we test if the magnitude of autocorrelation differs for different industries, which causes panel-specific AR1 (PSAR1). We use the likelihood ratio test to check whether AR1 coefficients are common across industries (Greene 2008). If the null hypothesis—that the regression with the correlation of AR1 is nested in the regression with the correction of PSAR1—is rejected, it suggests controlling for PSAR1. As suggested by the likelihood ratio tests, we control for PSAR1 for all estimation models for Dataset I, except the ones with Broker and Nbroker as independent variables. For Dataset II, when Broker, Nbroker, Egobetween, Negobetween, and $MSS2$ are DVs, we control for AR1, and we control for PSAR1 in the estimation models with other DVs.

We also test for panel-level HE using likelihood ratio test (Greene 2008). We anticipate panel-level HE because the variances of the error terms for each industry are likely to fluctuate over time and also differ across industries. The null hypothesis of no panel-level HE is rejected at all levels of significance for all the estimation models in Dataset I (eg.$\chi^2(84) = 406.34$ for Broker as the DV in Model 1, and $\chi^2(84) = 2227.55$ for CVS as

the DV in Model 1) and in II ($\chi^2(57) = 888.57$ for Broker as the DV in Model 1, and $\chi^2(57) = 744.30$ for CVS as the DV in Model 1), so we control for panel-level HE for all estimation models.

For each regression, we use feasible generalized least squares to generate estimates (Wooldridge 2002).

4.4 Estimation Results

4.4.1 The Impact of IT on Connectivity and Concentration

We estimate the impact of IT on four measures of connectivity separately. The estimation results based on Dataset II are provided in Figure 4.13. As seen in columns 5 and 6, the estimates of $log_e(IT)$ are positive and significant. It suggests that, keeping other factors constant, a 1% increase in an industry's IT investment is associated with a 0.016% increase in the number of times the given industry being a broker and 1% increase in an industry's IT is associated with a 0.047% increase in the number of the given industry's proportion of times as an ego lies on the shortest path between each part of alters. It implies that when an industry increases its IT investment, its importance as a role of brokerage is improved.

The estimation results of the impact of IT on concentration for Dataset II are shown in Figure 4.14. The estimates of $log_e(IT)$ in Columns 1 through 4 are significant and negative. For example, the estimate of $log_e(IT)$ in Column 1 suggests that a 1% increase in IT is associated with a 0.057 unit decrease in CVS, keeping other factors constant. It implies that when an industry increases its IT investment, its supplying market becomes less concentrated.

The estimation results based on Dataset I are shown in Figure 4.15 and Figure 4.16. The estimates of $log_e(IT)$, shown in Column 1 of Figure 4.16, is consistent with the results based on Dataset II, when CVS is the DV. Interestingly, the estimates of $log_e(IT)$ with connectivity measures as DVs, shown in Columns 3, 4, and 6 of Figure 4.15, are negative and significant,

opposite to the sign of the corresponding estimate for Dataset II. As Dataset II covers almost all industries in the U.S. economy, while Dataset I focuses on manufacturing industries, the different signs of IT estimates suggest that IT may affect connectivity of manufacturing industries differently. Alternatively, the different signs of IT estimates may be due to the different aggregation levels of two Datasets. Dataset II is based on 3-digit NAICS codes roughly corresponding to 2-digit SIC level, which is at a relatively higher aggregated level.

Considering the complexity of the relationships between IT and network measures, we conduct estimations using subsamples of two Datasets and test robustness of results in the following sections. A summary of estimation results regarding the effect of IT on connectivity and concentration is provided in Figure B.1.

Variables	(1) Broker	(2) Nbroker	(3) Egobetween	(4) Negobetween	(5) Lbroker	(6) Legobetween
Log(IT)	0.734	0.000563	0.103	0.00330	0.0155*	0.0467***
	(0.538)	(0.000379)	(0.135)	(0.00486)	(0.00820)	(0.0163)
Constant	49.01***	0.0530***	5.903***	0.326***	3.876***	1.502***
	(5.206)	(0.00366)	(1.234)	(0.0442)	(0.0815)	(0.146)
Observations	580	580	580	580	580	550
Number of Industries	58	58	58	58	58	55
Industry FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES

Notes: When Legobetween is the dependent variable, it controls for PSAR1 and PH; otherwise, controls for AR1 and PH. * P<.10, ** P<0.05, ***P<0.01, standard errors in parentheses. The same in the below.

Figure 4.13: The Estimation Results of Connectivity Measures Based on Dataset II

Variables	(1) CVS	(2) MSS2	(3) CR4S	(4) CR8S
Log(IT)	-0.0568*	-0.00688*	-0.0174***	-0.00936***
	(0.0292)	(0.00385)	(0.00355)	(0.00319)
Constant	2.430***	0.158***	0.657***	0.805***
	(0.311)	(0.0378)	(0.0398)	(0.0313)
Observations	580	580	580	580
Number of Industries	58	58	58	58
Industry FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES

Notes: When CVS, CR4S, and CR8S are dependent variables, it controls for PSAR1 and PH, respectively; for MSS2, controls for AR1 and PH.

Figure 4.14: The Estimation Results of Concentration Measures Based on Dataset II

Variables	(1) Broker	(2) Nbroker	(3) Egobetween	(4) Negobetween	(5) Lbroker	(6) Legobetween
Log(IT)	0.162	4.30e-05	-0.0785***	-0.00302***	0.00285	-0.0258***
	(0.386)	(0.000193)	(0.0257)	(0.000731)	(0.00430)	(0.00422)
Constant	69.31***	0.0437***	10.35***	0.328***	4.228***	2.443***
	(2.247)	(0.00113)	(0.150)	(0.00428)	(0.0254)	(0.0248)
Observations	1,105	1,105	1,105	1,105	1,105	1,066
Number of Industries	85	85	85	85	85	82
Industry FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES

Notes: Controlling for PH for all models; when Broker and Nbroker as dependent variables, it controls for AR1; otherwise, controls for PSAR1.

Figure 4.15: The Estimation Results of Connectivity Measures Based on Dataset I

Variables	(1) CVS	(2) MSS2	(3) CR4S	(4) CR8S
Log(IT)	-0.0473**	-0.000660	-0.00203	-0.00113
	(0.0204)	(0.00184)	(0.00236)	(0.00136)
Constant	9.825***	0.932***	1.021***	0.987***
	(0.136)	(0.0158)	(0.0250)	(0.00917)
Observations	1,105	1,105	1,105	1,105
Number of Industries	85	85	85	85
Industry FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES

Notes: Controlling for PSAR1 and PH.

Figure 4.16: The Estimation Results of Concentration Measures Based on Dataset I

4.4.2 Results on Subsamples of the Data

Different Time Periods of Dataset I The advent of the Internet greatly enhanced IT capabilities of organizations, resulting in reduced transaction costs and improved coordination among organizations. It is possible that the effect of IT on connectivity and concentration might be different for the pre- and post-Internet eras.

We examine the effect of IT on connectivity and concentration before and after the advent of the Internet. We split Dataset I into 1987-1993 (the pre-Internet era) and 1994-1999 (the post-Internet era) as the Internet has grown substantially as of 1995 (Zakon 2011, Leiner et al. 1997).

The estimation results about the impact of IT on connectivity for the pre- and post-Internet eras are shown in Figure 4.17 and Figure 4.18. As seen in Columns 3 through 6

of Figure 4.17 and Column 6 of Figure 4.18, the estimates of $log_e(IT)$ are negative and significant, which is consistent with the estimates for the full samples of Dataset I.

The estimation results about the effect of IT on concentration for two eras are provided in Figure 4.19 and Figure 4.20. The estimates of $log_e(IT)$ are positive and significant at the pre-Internet era, and become negative and significant at the post-Internet era. It suggests that the impact of IT on concentration has evolved with the advent of the Internet.

	(1)	(2)	(3)	(4)	(5)	(6)
Variables	Broker	Nbroker	Egobetween	Negobetween	Lbroker	Legobetween
Log(IT)	0.0537	-3.20e-05	-0.0839***	-0.00453***	-0.0309***	-0.0370***
	(0.542)	(0.000254)	(0.0254)	(0.000686)	(0.00929)	(0.00926)
Constant	69.71***	0.0440***	10.35***	0.336***	4.421***	2.502***
	(3.108)	(0.00146)	(0.146)	(0.00394)	(0.0534)	(0.0530)
Observations	595	595	595	595	595	574
Number of Industries	85	85	85	85	85	82
Industry FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES

Notes: Controlling for PH for all models; when Broker and Nbroker are dependent variables, it controls for AR1; otherwise, controls for PSAR1.

Figure 4.17: The Estimation Results of Connectivity Measures Based on 1987–1993 Dataset

	(1)	(2)	(3)	(4)	(5)	(6)
Variables	Broker	Nbroker	Egobetween	Negobetween	Lbroker	Legobetween
Log(IT)	-0.118	-0.000339	-0.0279	-0.000841	0.000413	-0.0177**
	(0.957)	(0.000488)	(0.0335)	(0.00107)	(0.00990)	(0.00703)
Constant	69.00***	0.0450***	10.07***	0.315***	4.233***	2.405***
	(5.973)	(0.00305)	(0.212)	(0.00678)	(0.0622)	(0.0442)
Observations	510	510	510	510	510	492
Number of Industries	85	85	85	85	85	82
Industry FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES

Notes: Controlling for PH for all models; when Broker and Nbroker are dependent variables, it controls for AR1; otherwise, controls for PSAR1.

Figure 4.18: The Estimation Results of Connectivity Measures Based on 1994–1999 Dataset

Variables	(1) CVS	(2) MSS2	(3) CR4S	(4) CR8S
Log(IT)	0.00553	5.48e-05	0.00934***	0.00451***
	(0.0181)	(0.00156)	(0.00252)	(0.00112)
Constant	9.349***	0.900***	0.919***	0.956***
	(0.104)	(0.00933)	(0.0146)	(0.00656)
Observations	595	595	595	595
Number of Industries	85	85	85	85
Industry FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES

Notes: Controlling for PSAR1 and PH.

Figure 4.19: The Estimation Results of Concentration Measures Based on 1987-1993 Dataset

Variables	(1) CVS	(2) MSS2	(3) CR4S	(4) CR8S
Log(IT)	-0.110***	-0.00307	-0.0155***	-0.0116***
	(0.0276)	(0.00245)	(0.00336)	(0.00145)
Constant	9.929***	0.893***	1.065***	1.052***
	(0.219)	(0.0285)	(0.0224)	(0.00954)
Observations	510	510	510	510
Number of Industries	85	85	85	85
Industry FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES

Notes: Controlling for PSAR1 and PH.

Figure 4.20: The Estimation Results of Concentration Measures Based on 1994-1999 Dataset

Manufacturing Industries in Dataset II In order to examine the impact of IT on the structure of production for manufacturing industries in a relative recent period, we capture a subnetwork that is composed of all 18 manufacturing industries and their ties. We apply the same methodology to generate measures of network structure and estimate models.

The results of the sample of manufacturing industries in Dataset II are shown in Figure 4.21 and Figure 4.22. Compared to the estimation results for the full sample of Dataset II, the estimates of IT variable for the sample of manufacturing industries have opposite signs in terms of the effects on connectivity and concentration. This suggests that the impact of IT on the structure of production for the whole economy differs from that for manufacturing industries. Compared to the estimation results based on Dataset I which is composed of manufacturing industries at a different aggregation level, the estimates of IT variable are consistent regarding the effect on connectivity as seen in Figure 4.17, Figure 4.18, and

Figure 4.21; while the estimates of IT variable in terms of the effect on concentration are inconsistent as seen in Figure 4.19, Figure 4.20, and Figure 4.22. A possible reason for the inconsistence is that the aggregation differences between two datasets might affect concentration more than connectivity.

Variables	(1) Broker	(2) Nbroker	(3) Egobetween	(4) Negobetween	(5) Lbroker	(6) Legobetween
Log(IT)	-0.490	-0.00980***	-0.277**	-0.195***	-0.0501	-0.226**
	(0.308)	(0.00303)	(0.141)	(0.0613)	(0.0587)	(0.0879)
Constant	14.89***	0.167***	4.578***	2.508***	2.767***	2.606***
	(2.353)	(0.0231)	(1.075)	(0.467)	(0.448)	(0.667)
Observations	180	180	180	180	180	180
Number of Industries	18	18	18	18	18	18
Industry FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES

Notes: Controlling for PH and AR1 for all models.

Figure 4.21: The Estimation Results of Connectivity Measures Based on the Sample of Manufacturing Industries in Dataset II

Variables	(1) CVS	(2) MSS2	(3) CR4S	(4) CR8S
Log(IT)	0.116	0.00632	0.0342**	0.0104***
	(0.0947)	(0.0219)	(0.0141)	(0.00300)
Constant	0.146	0.0611	0.316***	0.755***
	(0.718)	(0.166)	(0.107)	(0.0230)
Observations	180	180	180	180
Number of Industries	18	18	18	18
Industry FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES

Notes: Controlling for PH for all models; controlling for AR1 when CVS, MSS2, and CR4S are dependent variables, and controlling for PSAR1 when CR8S is the dependent variable.

Figure 4.22: The Estimation Results of Concentration Measures Based on the Sample of Manufacturing Industries in Dataset II

4.4.3 Robustness Tests

Estimation Results Using Instrument Variables (IVs) In general, endogeneity can be caused by omitted variables, measurement errors, and simultaneity (Wooldridge 2002). It is possible that some omitted variables, such as other productivity-related organizational initiatives, may be correlated with the independent variables in our models. In addition,

there is potential for measurement errors in variables relating to IT capital. There may also be concerns about endogeneity caused by simultaneity. Simultaneity arises when explanatory variables in our models are determined simultaneously with the dependent variables. For example, an increase in the connectivity in a network of an ego-industry may lead to more IT investment of the given industry.

Our econometric adjustments with year-fixed effects and time-invariant industry-specific effects help relieve these endogeneity concerns. We further address such concerns by providing estimates using instrumental variables (IVs). Firstly, we use lags of the variables as IVs, similar to previous related studies (Stiroh 2002a, Han et al. 2011b, Cheng and Nault 2012, Han and Mithas 2013). Specifically, we use one-year and two-year lags of IT variable as the instruments. We then employ a two-step generalized method of moments (GMM) procedure, controlling for arbitrary heteroskedasticity, and conduct the Hansen/Sargan C-test for the estimated model across two Datasets (Baum et al. 2003). The Hansen/Sargan C-test statistics, shown at the bottom of Figure 4.23 through Figure 4.26, cannot reject the exogeneity of suspect endogenous variables for the estimation for two Datasets.

Next, we use "two-step efficient GMM in the presence of arbitrary heteroskedasticity" to estimate our models with instrumental variables as "the advantages of GMM over IV are clear: if heteroskedasticity is present, the GMM estimator is more efficient than the simple IV estimator, whereas if heteroskedasticity is not present, the GMM estimator is no worse asymptotically than the IV estimator" (Baum et al. 2003, pp.11).

The estimated results are presented in Figure 4.23 through Figure 4.26. As seen, the estimates of IT variable using instrumental variables are similar to the main results for Dataset I. The estimates of IT variable are not significant for Dataset II, this might be due to the loss of efficiency of the estimation with instrumental variables. When turning to the efficient GMM estimation, there would be a cost of loss of efficiency for the sake of consistency (Baum et al. 2003, pp. 19-20) which may cause large asymptotic variance of

the GMM estimator, so it is reasonable that the GMM estimators of IT variable we are not significant.

The estimation results using IVs for subsamples of two Datasets are provided in Figure B.2 through Figure B.7 in Appendix B.

Variables	(1) Broker	(2) Nbroker	(3) Egobetween	(4) Negobetwee n	(5) Lbroker	(6) Legobetween
Log(IT)	1.227	-3.66e-05	0.0395	-0.00109	0.0131	0.0327
	(1.775)	(0.00116)	(0.146)	(0.00502)	(0.0279)	(0.0441)
Constant	43.60***	0.0580***	6.442***	0.368***	3.872***	1.594***
	(16.78)	(0.0110)	(1.374)	(0.0473)	(0.265)	(0.412)
Observations	464	464	464	464	464	440
R-squared	0.990	0.986	0.996	0.996	0.987	0.992
Industry FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Hansen/Sargan	Chi-sq(1) =	Chi-sq(1) =	Chi-sq(1) =	Chi-sq(1) =	Chi-sq(1) =	Chi-sq(1) =
C-test Statistics	0.339,	0.936,	0.072,	0.001,	0.001,	5.108 ,
	P=0.560	P=0.333	P=0.788	P=0.976	P=0.976	P=0.024

Notes: It use one year and two year lags of IT variable as instrumental variables, controlling for industry dummies (by industry) and year dummies. It uses 2-step GMM controlling for arbitrary heteroskedasticity.

Figure 4.23: The Estimation Results of Connectivity Measures Using IVs for Dataset II

Variables	(1) Broker	(2) Nbroker	(3) Egobetween	(4) Negobetwee n	(5) Lbroker	(6) Legobetween
Log(IT)	0.259	-0.000135	-0.135	-0.00632	-0.00418	-0.0420*
	(0.744)	(0.000342)	(0.157)	(0.00479)	(0.0164)	(0.0223)
Constant	65.92***	0.0434***	10.74***	0.351***	4.250***	2.568***
	(5.079)	(0.00235)	(1.067)	(0.0325)	(0.112)	(0.152)
Observations	935	935	935	935	935	902
R-squared	0.999	0.999	1.000	1.000	0.995	0.999
Industry FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Hansen/Sargan	Chi-sq(1) =	Chi-sq(1) =	Chi-sq(1) =	Chi-sq(1) =	Chi-sq(1) =	Chi-sq(1) =
C-test Statistics	1.080,	1.536,	0.562,	0.616,	1.490,	0.219 ,
	P=0.299	P=0.215	P=0.454	P=0.434	P=0.222	P=0.640

Figure 4.24: The Estimation Results of Connectivity Measures Using IVs for Dataset I

Variables	(1) CVS	(2) MSS2	(3) CR4S	(4) CR8S
Log(IT)	-0.0820	-0.0110	-0.00950	-0.00416
	(0.0626)	(0.00872)	(0.00876)	(0.00573)
Constant	2.744***	0.203**	0.607***	0.778***
	(0.588)	(0.0817)	(0.0817)	(0.0538)
Observations	464	464	464	464
R-squared	0.963	0.976	0.963	0.972
Industry FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Hansen/Sargan C-test Statistics	Chi-sq(1) = 0.016, P=0.900	Chi-sq(1) = 0.259, P=0.611	Chi-sq(1) = 0.008, P=0.927	Chi-sq(1) = 0.888, P=0.346

Figure 4.25: The Estimation Results of Concentration Measures Using IVs for Dataset II

Variables	(1) CVS	(2) MSS2	(3) CR4S	(4) CR8S
Log(IT)	-0.140	-0.0122	-0.0164*	-0.00846
	(0.104)	(0.0107)	(0.00911)	(0.00573)
Constant	10.31***	0.973***	1.083***	1.040***
	(0.715)	(0.0734)	(0.0623)	(0.0390)
Observations	935	935	935	935
R-squared	0.981	0.978	0.987	0.989
Industry FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Hansen/Sargan C-test Statistics	Chi-sq(1) = 0.551, P=0.45	Chi-sq(1) = 1.956, P=0.162	Chi-sq(1) = 0.454, P=0.501	Chi-sq(1) = 0.262, P=0.609

Figure 4.26: The Estimation Results of Concentration Measures Using IVs for Dataset I

Estimation Results with Alternative Independent Variables We focus on the impact of IT on the structure of production measured as connectivity and concentration, while there are alternative explanations about changes in the structure of production. Both of our connectivity and concentration measures are based on the ties between supplier industries and an ego-industry, where intermediate inputs are flows of ties, so two types of measures of the structure of production are related to intermediate inputs. From production theory, intermediate inputs, IT capital, non-IT capital, and labor are production factors in a production function and there are substitution and complementation relationships among them (Dewan and Min 1997, Mittal and Nault 2009, Chwelos et al. 2010, Zhang et al. 2014). It is possible that labor and non-IT capital may be associated with changes in the structure of

production through IT capital. In order to address this concern, we add the log transformations of non-IT and labor in Models 1 and 2. The corresponding estimation results based on two Datasets are shown in Figure 4.27 through Figure 4.30. As seen, these estimation results are consistent with main results.

In addition, there might be concerns that the impact of IT on the structure of production is driven by the scale of industries. Although our controlling for industry-fixed effects in our models has at least partially addressed this concern, we address it further by using a normalized IT variable in models, where IT capital is normalized by gross output. The estimation results are shown in Figure 4.31 through Figure 4.34. The estimates of IT variable in terms of the effect on connectivity for Dataset I are consistent with main results. The estimates of IT variable in terms of the effect on concentration are positive and significant, with different signs from main results for both Datasets. As the normalized IT has a different economic meaning, IT intensity, the results suggest that an industry's IT intensity may have different effects from its IT investment: the more intensive investment in IT, the more concentrated is its supplying market.

	(1)	(2)	(3)	(4)	(5)	(6)
Variables	Broker	Nbroker	Egobetween	Negobetween	Lbroker	Legobetween
log(IT)	-0.144	-0.000245	0.0393	0.00131	0.0107	0.0185
	(0.693)	(0.000511)	(0.206)	(0.00739)	(0.0107)	(0.0260)
log(Labor)	2.820***	0.00240***	0.200	0.00788	0.0363***	0.0350
	(1.055)	(0.000776)	(0.252)	(0.00924)	(0.0130)	(0.0308)
log(Non-IT)	0.485	0.000640	0.0245	-0.00190	-0.0168	0.0144
	(1.967)	(0.00129)	(0.578)	(0.0205)	(0.0218)	(0.0538)
Constant	34.16	0.0377**	4.999	0.326	3.946***	1.357**
	(22.70)	(0.0147)	(6.713)	(0.238)	(0.244)	(0.595)
Observations	580	580	580	580	580	550
Number of Industries	58	58	58	58	58	55
Industry FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Econometrics Adjustment	AR1+HE	AR1+HE	AR1+HE	AR1+HE	AR1+HE	PSAR1+HE

Figure 4.27: The Estimation Results of Connectivity Measures Controlling for the Effects of Labor and Non-IT Capital for Dataset II

Variables	(1) Broker	(2) Nbroker	(3) Egobetween	(4) Negobetween	(5) Lbroker	(6) Legobetween
log(IT)	0.247	-2.91e-05	-0.112***	-0.00382***	-0.000854	-0.0353***
	(0.475)	(0.000247)	(0.0334)	(0.000967)	(0.00646)	(0.00512)
log(Labor)	-0.149	0.000112	0.0664	0.00156	0.00415	0.0280***
	(0.728)	(0.000366)	(0.0428)	(0.00133)	(0.00796)	(0.00823)
log(Non-IT)	-0.170	0.000111	0.0332	0.00161	0.00734	0.0118
	(0.963)	(0.000464)	(0.0556)	(0.00177)	(0.0111)	(0.00989)
Constant	71.37***	0.0424***	9.818***	0.308***	4.153***	2.214***
	(7.683)	(0.00365)	(0.448)	(0.0144)	(0.0846)	(0.0820)
Observations	1,105	1,105	1,105	1,105	1,105	1,066
Number of Industries	85	85	85	85	85	82
Industry FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Econometrics Adjustment	AR1+HE	AR1+HE	PSAR1+HE	PSAR1+HE	PSAR1+HE	PSAR1+HE

Figure 4.28: The Estimation Results of Connectivity Measures Controlling for the Effects of Labor and Non-IT Capital for Dataset I

Variables	(1) CVS	(2) MSS2	(3) CR4S	(4) CR8S
log(IT)	-0.0959**	-0.00587	-0.0158***	-0.0120***
	(0.0451)	(0.00541)	(0.00530)	(0.00403)
Log(Labor)	0.131***	0.0109*	0.0492***	0.0363***
	(0.0505)	(0.00583)	(0.0108)	(0.00779)
log(Non-IT)	-0.249***	-0.0248***	-0.0464***	-0.0226**
	(0.0834)	(0.00937)	(0.0154)	(0.0111)
Constant	5.484***	0.428***	1.007***	0.935***
	(0.917)	(0.0995)	(0.185)	(0.132)
Observations	580	580	580	580
Number of Industries	58	58	58	58
Industry FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Econometrics Adjustment	PSAR1+HE	PSAR1+HE	PSAR1+HE	PSAR1+HE

Figure 4.29: The Estimation Results of Concentration Measures Controlling for the Effects of Labor and Non-IT Capital for Dataset II

Variables	(1) CVS	(2) MSS2	(3) CR4S	(4) CR8S
log(IT)	0.00206	-1.87e-06	-0.00562**	-0.00340**
	(0.0181)	(0.00153)	(0.00286)	(0.00153)
log(Labor)	-0.144***	-0.0108***	-0.00957**	-0.0101***
	(0.0420)	(0.00358)	(0.00484)	(0.00250)
log(Non-IT)	0.0810	0.00481	0.0148*	0.0145***
	(0.0597)	(0.00530)	(0.00799)	(0.00432)
Constant	9.696***	0.949***	0.917***	0.920***
	(0.503)	(0.0443)	(0.0645)	(0.0350)
Observations	1,105	1,105	1,105	1,105
Number of Industries	85	85	85	85
Industry FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Econometrics Adjustment	PSAR1+HE	PSAR1+HE	PSAR1+HE	PSAR1+HE

Figure 4.30: The Estimation Results of Concentration Measures Controlling for the Effects of Labor and Non-IT Capital for Dataset I

Variables	(1) Broker	(2) Nbroker	(3) Egobetween	(4) Negobetween	(5) Lbroker	(6) Legobetween
IT/Gross_Output	-3.173	-0.000592	-0.0207	0.00370	-0.0333	-0.0155
	(1.995)	(0.00175)	(0.516)	(0.0201)	(0.0375)	(0.0739)
Constant	55.59***	0.0580***	6.819***	0.356***	4.013***	1.921***
	(2.114)	(0.00146)	(0.218)	(0.00801)	(0.0391)	(0.0183)
Observations	580	580	580	580	580	550
Number of Industries	58	58	58	58	58	55
Industry FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Econometrics Adjustment	AR1+HE	AR1+HE	AR1+HE	AR1+HE	AR1+HE	PSAR1+HE

Figure 4.31: The Estimation Results of Connectivity Measures with Normalized IT for Dataset II

Variables	(1) Broker	(2) Nbroker	(3) Egobetween	(4) Negobetween	(5) Lbroker	(6) Legobetween
IT/Gross_Output	-0.907	-0.00152	-0.745*	-0.0298**	-0.00125	-0.211***
	(3.147)	(0.00154)	(0.391)	(0.0134)	(0.0356)	(0.0553)
Constant	70.26***	0.0440***	9.908***	0.311***	4.244***	2.299***
	(0.391)	(0.000239)	(0.0264)	(0.000820)	(0.00625)	(0.00532)
Observations	1,105	1,105	1,105	1,105	1,105	1,066
Number of Industries	85	85	85	85	85	82
Industry FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Econometrics Adjustment	AR1+HE	AR1+HE	PSAR1+HE	PSAR1+HE	PSAR1+HE	PSAR1+HE

Figure 4.32: The Estimation Results of Connectivity Measures with Normalized IT for Dataset I

Variables	(1) CVS	(2) MSS2	(3) CR4S	(4) CR8S
IT/Gross_Output	0.301*	0.00600	0.134***	0.0747***
	(0.154)	(0.0160)	(0.0282)	(0.0207)
Constant	1.923***	0.0967***	0.501***	0.721***
	(0.158)	(0.0151)	(0.0224)	(0.0122)
Observations	580	580	580	580
Number of Industries	58	58	58	58
Industry FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Econometrics Adjustment	PSAR1+HE	AR1+HE	PSAR1+HE	PSAR1+HE

Figure 4.33: The Estimation Results of Concentration Measures with Normalized IT for Dataset II

Variables	(1) CVS	(2) MSS2	(3) CR4S	(4) CR8S
IT/Gross_Output	0.268*	0.000591	0.00992	0.133***
	(0.154)	(0.0215)	(0.0289)	(0.0140)
Constant	9.560***	0.887***	0.971***	0.978***
	(0.0735)	(0.00983)	(0.00648)	(0.00244)
Observations	1,105	1,105	1,105	1,105
Number of Industries	85	85	85	85
Industry FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Econometrics Adjustment	PSAR1+HE	PSAR1+HE	PSAR1+HE	PSAR1+HE

Figure 4.34: The Estimation Results of Concentration Measures with Normalized IT for Dataset I

4.4.4 Granger Causality Analysis

We argue that IT leads to changes in the production of structure in the economy, one may argue that the relationship between IT and the structure of production is in the reverse direction: if an ego-industry is in a more connected network, there is a need to increase IT investment in order to support its coordination with a variety of suppliers; if the supplying market is more competitive, which is less concentrated, suppliers are more likely to increase their IT investment to obtain competitive advantages, which in turn impose the ego-industry to invest more in IT.

In order to address these concerns, we conduct Granger causality analysis, which allows us to understand whether there exists a causality relationship between IT and the structure of production, and if so, it is in one or bidirectional directions. The basic of Granger

causality test is that: build a autoregressive model for Y, and add lags of variable X to the autoregressive model; if there is a significant improvement in predicting Y after adding lags of X, then X Granger-causes Y; otherwise, there is no evidence of Granger causality from X to Y (Dutta 2001).

We follow the three steps of Granger test in Dutta (2001) and also employed the approach of Granger tests for panel data in Hurlin and Venet (2001) and Dietrich (2009). The first step is to take a unit root test to make sure our data is stationary. The unit root test is used to examine if a time-serial variable is non-stationary by using an autoregressive model. We conduct an augmented Dickey-Fuller unit root test, Levin-Lin-Chu panel unit root test (Levin et al. 2002), as this test is good for small T (time series) and larger N (industries). The null hypothesis of the augmented Dickey-Fuller test for panel data is that the unit root exists in a time series. If this null hypothesis is not rejected, it suggests that a time series is not stationary. The non-stationarity of time series may cause spurious regressions, where the relationship between two variables could be driven by trends in variables. Our unit root tests suggest that all variables are stationary as the tests were rejected for Dataset II (eg. Levin-Lin-Chu test statistics for IT variable Z, $t-star = -6.27, p < 0.001$; $t-star = -20.00, p < 0.001$ for Broker; $t-star = -13.84, p < 0.001$ for Egobetween, $t-star = -9.72, p < 0.001$ for CVS; and $t-star = -8.60, p < 0.001$ for $MSS2$). The augmented Dickey-Fuller test suggests similar results to Dataset I. Therefore, the results of augmented Dickey-Fuller test provide evidence that our time-series variables are stationary.

In the second step, we decide the optimal number of lags. When testing homogeneous non-causality test (HNC), the Wald statistic exists if and only if $T > 5 + 2m$, where T is time period and m is the number of lags (Dumitrescu and Hurlin 2012). In our Datasets, $T = 13$ for Dataset I and $T = 10$ for Dataset II, so we use two lags in the estimation models.

In the third step, we conduct homogeneous non-causality test (HNC) for panel data, referring to (Hurlin and Venet 2001, Dietrich 2009). The null hypothesis is: $\beta_i^{(m)} = 0, \forall i \in$

$[1, N], \forall m \in [1, p]$, testing whether or not the regression slope coefficients associated with variable X, β, for all individual i and all lag m are nulls. If we fail to reject the null, the variable X is not Granger-causing Y in all the N industries of the samples, then the test procedure stops. If it is rejected, it suggests that it exists overall homogeneous causality (HC) or causality for some individuals. We use "xtpcse" command in Stata, where parameters are estimated by ordinal least square (OLS), but with panel-corrected standard error (PCSE) and controlling for industry-fixed effect by adding industry dummies as there is heterogeneity in industries and the HNC test is suggested to control for fixed effect Hurlin and Venet (2001). We use panel-specific AR1 autocorrelation structure. By default, in "xtpcse" the disturbances are assumed to be heteroskedastic and that the disturbances are contemporaneously correlated across the panels.

The results of Granger causality tests shown in Figure 4.35 and Figure 4.36 suggest that there is relatively stronger evidence that IT Granger-causes connectivity and concentration for Dataset II. Some of test statistics for Dataset I are significant, this rules out the alternative that there is no Granger causality between IT and concentration and connectivity measures.

		Dataset I (1987-1999)		Dataset II (2000-2009)	
Models	Variables	IT → Connectivity	Connectivity → IT	IT → Connectivity	Connectivity → IT
Model 1 (log(IT))	Broker	3.05	2.41	10.23 ***	2.70
	Nbroker	2.02	2.62	7.79 **	3.70
	Egobetween	1.35	1.05	6.45**	0.49
	Negobetween	2.59	1.39	4.36 (p=0.11)	0.53
Model 1 (log(IT), log(Labor), and log(Non-IT))	Broker	2.19	2.40	0.28	1.93
	Nbroker	1.93	2.67	0.05	3.08
	Egobetween	0.12	1.34	1.92	0.25
	Negobetween	0.22	1.87	2.35	0.27
Model 1 (IT/Gross_output)	Broker	1.76	0.06	8.66**	2.38
	Nbroker	0.83	0.05	8.15**	3.06
	Egobetween	5.92*	0.04	6.82**	1.94
	Negobetween	5.07*	0.20	4.31(p=0.12)	2.13
Model 2 (log(IT))	Log(broker)	1.64	3.26	4.32(p=0.12)	5.61*
	Log(egobetween)	0.37	5.12*	1.38	4.89*

Notes: The chi-square statistics are in cells and the degree of freedom is 2. * p<0.10, ** P<0.05, *** p<0.01.

Figure 4.35: The Test of Homogeneous Non-Causality for Connectivity Measures

		Dataset I (1987-1999)		Dataset II (2000-2009)	
Models	Variables	IT → Concentration	Concentration → IT	IT → Concentration	Concentration → IT
Model 1 (log(IT))	CVS	0.02	0.46	5.80*	10.19***
	MSS2	0.01	1.51	8.01**	10.02***
	CR4S	1.56	0.31	2.45	5.50*
	CR8S	3.28	1.20	5.76*	6.83**
Model 1 (log(IT), log(Labor), and log(Non-IT))	CVS	0.78	5.57*	8.25**	11.26***
	MSS2	1.69	4.73*	8.60**	9.67***
	CR4S	0.99	4.04	6.16**	6.74**
	CR8S	0.69	4.33	8.56**	8.59**
Model 1 (IT/Gross_output)	CVS	6.75**	7.06**	6.51**	0.71
	MSS2	5.05*	7.21**	6.00**	0.45
	CR4S	8.42**	5.10*	3.51	3.43
	CR8S	1.98	1.74	6.69**	1.64

Notes: The chi-square statistics are in cells and the degree of freedom is 2. * p<0.10, ** P<0.05, *** p<0.01.

Figure 4.36: The Test of Homogeneous Non-Causality for Concentration Measures

4.5 Conclusion

4.5.1 Summary

In an economy, industries are connected in terms of production processes, forming an economic network. In order to measure the structure of production in the economy, we employ network analysis to generate a variety of measures, which we categorize as connectivity among industries in an egocentric network and concentration in an ego-industry's supplying market.

There is substantial evidence of a relationship between IT investment and the structure of production in the economy in terms of connectivity and concentration. The overall relationship is robust to a variety of measures of connectivity and concentration. We find that with an increase in IT, an industry is more likely to act as a hub or pivot that transmits goods and services along supply chains in an economy with broad connections among manufacturing and non-manufacturing industries. An industry's investment in IT improves the digitalization of production processes and reduces uncertainty and opportunism in transactions. This enables new business processes and facilitates the given industry's role of "go-between" other

industries. However, our findings based on the economy-wide data differ from findings based on the sample of manufacturing industries, suggesting that IT may affect the connectivity amongst manufacturing industries differently.

We also find that IT is associated with the change in concentration in an industry's first-tier supplier neighborhood. The more IT an industry invests in, the less concentrated is the direct supplying market. A possible interpretation is that an industry's IT investment improves its coordination and information sharing with upstream suppliers, enabling relatively even distribution in the amount of inputs.

Our study provides evidence that IT has reshaped the production processes along supply chains, resulting in the realignment of input-output relationships among industries in the economy.

4.5.2 Limitations and Future Work

As this is an exploratory study on the relationship between IT and the structure of production, it is worth noting three limitations. Firstly, other factors (omitted variables) may have played a role in the relationship between IT and the structure of production. We have provided different models to examine the relationship, conducted robustness tests using instrumental variables and considered alternative explanations about the change in the structure of production, such as the impact of labor, non-IT capital, and the ratio of IT to gross output. However, we cannot rule out more complex relationships.

Secondly, our Granger causality analysis is inconclusive. We have conducted the Granger causality analysis to examine the directions of causal links in the relationship between IT and the structure of production. We find that the evidence for causality from IT to the structure of production is relatively stronger than that for causality in the opposite direction. As Granger tests provide evidence for causality in both directions, there may be a more complex relationship between IT and the structure of production.

Thirdly, our interpretations for the estimation results focus on the results based on

Dataset II which covers relatively broader industries in the U.S. economy. We also find that IT may affect the structure of production differently for economic networks among manufacturing industries. A refined interpretation of the different effects of IT for different samples will enrich our study.

In future research, we will refine our models, address more alternative explanations, refine the Granger causality analysis, and improve our interpretations for all estimation results. The relationship between IT and the structure of production is complex, our study represents an initial step in analyzing this relationship. However, we believe that our results have the potential to provide significant insight into the nature of the relationship between IT and the structure of production after refined work is done.

Chapter 5

CONCLUSION

In this thesis, we examine the relationship between IT and three aspects of organizing production: outsourcing logistics activities, purchases of intermediate inputs from upstream supplies, and the structure of production in an economy by drawing on a variety of theories and employing different methodologies. Each of these has the potential to contribute the field about the impact of IT on production organization.

In the first essay, we find that the effects of IT on outsourced logistics have changed with the advent of the Internet. Before the advent of the Internet, an industry's own IT investment and outsourced logistics were substitutes, implying that IT reduced internal governance costs relatively more than external transaction costs. After the advent of the Internet, IT and outsourced logistics became complements. This suggests that because of the unique characteristics of the Internet as an enabler, IT reduced external transaction costs relatively more than internal governance costs, thus industries favored the market form of the provision of logistics. We also find similar impacts of customers' IT investments on a focal industry's outsourced logistics.

The first essay makes three contributions. Firstly, we provide direct evidence that IT in its most interconnected instance—the Internet—enables a move to the market. Previous work theoretically argued that IT was associated with the shift from hierarchies to the market by reducing transaction costs, while other studies only provided indirect evidence in support of those arguments. Our study shows that in the context of logistics, the increasing use of IT is associated with greater logistics outsourcing after the advent of the Internet. Secondly, we show the impact of IT on output through logistics outsourcing. Previous studies about IT productivity converged to the contribution of IT to output. Our study shows the moderating

role of IT in the contribution of logistics outsourcing to output. Thirdly, our study reconciles the theoretical arguments about the effects of IT on reducing coordination costs within and between firms. Our study suggests that the effect of IT on reducing internal governance costs dominated before the advent of the Internet, and the effect of IT on reducing external transaction costs dominated after the advent of the Internet. In previous studies, it was not known which effect dominated a priori.

Our findings in the first essay have two managerial implications. One implication is that activities with the potential to be outsourced and that use the Internet for coordination are more likely to be outsourced in the Internet era. The other implication is that the increasing popularity of outsourcing has a technological and theoretical basis.

In the second essay, we find that an industry's IT investment can reduce its production interdependence with upstream suppliers by improving the efficiency of intermediate inputs. It has three contributions. Firstly, previous studies about the impact of IT on production factors showed that IT could both complement and substitute for labor and non-IT capital as well as augment labor and non-IT capital, but few studies have examined the impact of IT on production interdependence. We show that IT can not only substitute intermediate inputs, but also affect the efficiency of using intermediate inputs, resulting in a decrease in its production interdependence with suppliers, as measured by DBL. Secondly, it contributes to the literature on the impact of IT on TFP. Previously, little research has examined the impact of IT on TFP from the point-of-view of production structure—production interdependence. We find the effect of IT on TFP through DBL. Thirdly, we have identified an indirect effect of IT on value-added through DBL. Our study shows that IT reduces DBL, leading to greater TFP and value-added.

The second essay also provides important managerial implications. At the industry level, it implies that industries' IT investments help reduce their interdependence with upstream suppliers in terms of purchases of goods and services for production. At the firm level,

our results suggest that firms' IT investments help reduce their dependence on upstream suppliers and create more value-added. Thus firms should preferentially direct their IT investments towards improving the efficiency of purchased goods and services in order to reduce their production interdependence with suppliers. From a supply chain perspective, the whole supply chain becomes more efficient and thus more value-added is created by the supply chain if each firm in a supply chain invests in IT to improve its efficiency of goods and services purchased from upstream suppliers.

The third essay explores the relationship between IT and the structure of production, and it represents an initial step in analyzing this complex relationship. Creatively, we develop two categories of measures about the structure of production by employing network analysis: connectivity and concentration, each has a variety of metrics. We find that IT investment of an industry is associated with an increase in the connectivity within its supplying network and a decrease in concentration in the supplying market, where it purchases intermediate inputs. It suggests that with an increase in IT, an industry is more likely to act as a hub or pivot that transmits goods and services along supply chains and that the more IT an industry invests in, the less concentrated is the direct supplying market. We also find that IT may affect connectivity and concentration of manufacturing industries differently. Previous studies about the relationship between IT and organizational structure focus on the impact of IT at the firm level. However, few studies have examined the impact of IT on the structure of production in the economy as we do.

Bibliography

Abdi, H. 2010. Coefficient of variation. *Encyclopedia of Research Design. SAGE Publications, Inc., Thousand Oaks, CA* 169–171.

Ashenbaum, B., A. Maltz, E. Rabinovich. 2005. Studies of trends in third-party logistics usage: What can we conclude? *Transportation Journal* **44**(3) 39–50.

Bakos, J. Y. 1997. Reducing buyer search costs: Implications for electronic marketplaces. *Management Science* **43**(12) 1676–1692.

Bakos, J. Y., E. Brynjolfsson. 1993. Information technology, incentives, and the optimal number of suppliers. *Journal of Management Information Systems* 37–53.

Banker, R. D., J. Kalvenes, R. A. Patterson. 2006. Research note-information technology, contract completeness, and buyer-supplier relationships. *Information Systems Research* **17**(2) 180–193.

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

Borgatti, S. P., M. G. Everett, L. C. Freeman. 2002. Ucinet 6 for windows: Software for social network analysis. *Harvard, MA: Analytic Technologies* .

Bowersox, D. J. 1990. The strategic benefits of logistics alliances. *Harvard Business Review* **68**(July-August) 36–45.

Bowersox, D. J., P.J. Daugherty. 1995. Logistics paradigms: The impact of information technology. *Journal of Business Logistics* **16**(1) 65–80.

Bowersox, D.J., D.J. Closs, M.B. Cooper. 2010. *Supply Chain Logistics Management*. 3rd ed. McGraw-Hill, New York.

Brynjolfsson, E., L. Hitt. 1996. Paradox lost? Firm-level evidence on the returns to information systems spending. *Management Science* **42**(4) 541–558.

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

Brynjolfsson, E., T. W. Malone, V. Gurbaxani, A. Kambil. 1994. Does information technology lead to smaller firms? *Management Science* **40**(12) 1628–1644.

Brynjolfsson, E., M. D. Smith. 2000. Frictionless commerce? A comparison of Internet and conventional retailers. *Management Science* **46**(4) 563–585.

Cachon, G. P., M. Fisher. 2000. Supply chain inventory management and the value of shared information. *Management Science* **46**(8) 1032–1048.

Chenery, Hollis B., T. Watanabe. 1958. International comparisons of the structure of production. *Econometrica* **26**(4) 487–521.

Cheng, Z., B. R. Nault. 2007. Industry level supplier-driven IT spillovers. *Management Science* **53**(8) 1199–1216.

Cheng, Z., B. R. Nault. 2012. Relative industry concentration and customer-driven IT spillovers. *Information Systems Research* **23**(2) 340–355.

Church, J. R., R. Ware. 2000. *Industrial organization: a strategic approach*. McGraw-Hill, New York.

Chwelos, P., R. Ramirez, K. L. Kraemer, N. P. 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.

Clemons, E. K., S. P. Reddi, M. C. Row. 1993. The impact of information technology on the organization of economic activity: The "move to the middle" hypothesis. *Journal of Management Information Systems* **10**(2) 9–35.

Cooper, M. C., D. M. Lambert, J. D. Pagh. 1997. Supply chain management: More than a new name for logistics. *International Journal of Logistics Management* **8**(1) 1–14.

David, R. J., S. K. Han. 2004. A systematic assessment of the empirical support for transaction cost economics. *Strategic Management Journal* **25**(1) 39–58.

Dedrick, J., S.X. Xu, K.X. Zhu. 2008. How does information technology shape supply-chain structure? Evidence on the number of suppliers. *Journal of Management Information Systems* **25**(2) 41–72.

Deepen, J. M. 2007. *Logistics Outsourcing Relationships: Measurement, Antecedents, and Effects of Logistics Outsourcing Performance*. Physica-Verlag HD.

Dewan, S., K. L. Kraemer. 2000. Information technology and productivity: Evidence from country-level data. *Management Science* **46**(4) 548–562.

Dewan, S., C. Min. 1997. The substitution of information technology for other factors of production: A firm level analysis. *Management Science* **43**(12) 1660–1675.

Dietrich, A. 2009. Does growth cause structural change, or is it the other way round? a dynamic panel data analyses for seven OECD countries. Tech. rep., Jena economic research papers.

Dietzenbacher, E. 1992. The measurement of interindustry linkages: Key sectors in the Netherlands. *Economic Modelling* **9**(4) 419–437.

Dumitrescu, E. I., C. Hurlin. 2012. Testing for granger non-causality in heterogeneous panels. *Economic Modelling* **29**(4) 1450–1460.

Dutta, A. 2001. Telecommunications and economic activity: An analysis of granger causality. *Journal of Management Information Systems* **17**(4) 71–96.

Edgeworth, F. Y. 1925. The pure theory of monopoly. *Papers Relating to Political Economy*, vol. 1. Macmillan & Co., London, 111–142.

Galbraith, J. R. 1974. Organization design: An information processing view. *Interfaces* **4**(3) 28–36.

Gong, F., B. R. Nault, R. Rahman. 2013. An Internet-Enabled move to the market. University of Calgary, working paper.

Greene, W. H. 2008. *Econometric Analysis*. 6th ed. Pearson Prentice Hall, Upper Saddle River, N.J.

Grover, V., M. K. Malhotra. 2003. Transaction cost framework in operations and supply chain management research: Theory and measurement. *Journal of Operations Management* **21**(4) 457–473.

Gurbaxani, V., S. Whang. 1991. The impact of information systems on organizations and markets. *Communications of the ACM* **34**(1) 59–73.

Han, K., Chang Y. B., J. Hahn. 2011a. Information technology spillover and productivity: The

role of information technology intensity and competition. *Journal of Management Information Systems* **28**(1) 115–146.

Han, K., R. J. Kauffman, B. R. Nault. 2011b. Returns to information technology outsourcing. *Information Systems Research* **22**(4) 824–840.

Han, K., S. Mithas. 2013. Information technology outsourcing and non-it operating costs: An empirical investigation. *MIS Quarterly* **37**(1) 315–331.

Hanneman, R. A., M. Riddle. 2005. *Introduction to social network methods*. University of California Riverside.

Hirschman, A. O. 1958. *The Strategy of Economic Development*. Yale University Press, New Haven.

Hitt, L. M. 1999. Information technology and firm boundaries: Evidence from panel data. *Information Systems Research* **10**(2) 134–149.

Hitt, L.M., E. Snir. 1999. The role of information technology in modern production: Complement or substitute to other inputs? University of Pennsylvania, working paper.

Horowitz, K., M. Planting. 2009. Concepts and methods of the U.S. input-output accounts. Tech. rep., the U.S. Bureau of Economic Analysis of the U.S. Department of Commerce.

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

Hurlin, C., B. Venet. 2001. Granger causality tests in panel data models with fixed coefficients. *Cahier de Recherche EURISCO, September, Université Paris IX Dauphine* .

Im, K. S., V. Grover, J. T. C. Teng. 2013. Research note-do large firms become smaller by using information technology? *Information Systems Research* **24**(2) 470–491.

Jones, L. P. 1976. The measurement of Hirschmanian linkages. *The Quarterly Journal of Economics* **90**(2) 323–333.

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

Kim, Y., T. Y. Choi, T. Yan, K. Dooley. 2011. Structural investigation of supply networks: A social network analysis approach. *Journal of Operations Management* **29**(3) 194–211.

Kline, T. 2005. *Psychological testing: A practical approach to design and evaluation*. Sage Publications.

Kundisch, D., N. Mittal, B. R. Nault. 2014. Using income accounting as the theoretical basis for measuring it productivity. Working paper.

Langley, C. J. Jr., Capgemini. 2012. 2012 third party logistics: Results and findings of the 16th annual study. Tech. rep., Capgemini, Penn State University, The Panalpina Group, Heidrick & Struggles, and eyefortransport.

Lee, H. L., V. Padmanabhan, S. Whang. 1997. Information distortion in a supply chain: The bullwhip effect. *Management Science* **43**(4) 546–558.

Lee, H. L., K. C. So, C. S. Tang. 2000. The value of information sharing in a two-level supply chain. *Management Science* **46**(5) 626–643.

Leiner, B. M., V. G. Cerf, D. D. Clark, R. E. Kahn, L. Kleinrock, D. C. Lynch, J. Postel, L. G. Roberts, S. S. Wolff. 1997. The past and future history of the Internet. *Communications of the ACM* **40**(2) 102–108.

Levin, A., C. F. Lin, C. S. James Chu. 2002. Unit root tests in panel data: asymptotic and finite-sample properties. *Journal of econometrics* **108**(1) 1–24.

Lewis, I., A. Talalayevsky. 2000. Third-party logistics: Leveraging information technology. *Journal of Business Logistics* **21**(2) 173–185.

Lichtenberg, F.R. 1995. The output contributions of computer equipment and personnel: A firm-level analysis. *Economics of Innovation and New Technology* **3** 201–217.

Lieb, R. 1992. The use of third-party logistics services by large American manufacturers. *Journal of Business Logistics* **13**(2) 29–42.

Lieb, R. C., H. L. Randall. 1996. A comparison of the use of third-party logistics services by large american manufacturers, 1991, 1994, and 1995. *Journal of Business Logistics* **17**(1) 305–320.

Malone, T. W., J. Yates, R. I. Benjamin. 1987. Electronic markets and electronic hierarchies. *Communications of the ACM* **30**(6) 484–497.

Marasco, A. 2008. Third-party logistics: A literature review. *International Journal of Production Economics* **113**(1) 127–147.

McAfee, A., E. Brynjolfsson. 2008. Investing in the it that makes a competitive difference. *Harvard Business Review* **86**(7/8) 98.

McGuigan, J. R., R. C. Moyer, F. H. deB Harris. 2010. *Managerial economics: Application, Strategy, and Tactics*. 12th ed. South-Western, Cengage Learning.

Miller, R. E., P. D. Blair. 2009. *Input-output Analysis: Foundations and Extensions*. Cambridge University Press, New York. 2nd ed.

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

Mukhopadhyay, T., S. Kekre, S. Kalathur. 1995. Business value of information technology: A study of electronic data interchange. *MIS Quarterly* **19**(2) 137–156.

Mun, S., M. I. Nadiri. 2002. Information technology externalities: Empirical evidence from 42 U.S. industries. Working Paper 9272, National Bureau of Economic Research.

Nault, B. R., A. S. Dexter. 2006. Agent-intermediated electronic markets in international freight transportation. *Decision Support Systems* **41**(4) 787–802.

Ozer-Balli, H., B. Sørenson. 2010. Interaction effects in econometrics. CEPR Discussion Papers 7929, C.E.P.R. Discussion Papers.

Rasmussen, P.N. 1957. *Studies in Inter-sectoral Relations*. Amsterdam, North-Holland.

Rindfleisch, A., J. B. Heide. 1997. Transaction cost analysis: Past, present, and future applications. *The Journal of Marketing* **61**(4) 30–54.

Rutner, S. M., C. J. Langley Jr. 2000. Logistics value: Definition, process and measurement. *International Journal of Logistics Management* **11**(2) 73–82.

Santhanam, K. V., R. H. Patil. 1972. A study of the production structure of the Indian economy: An international comparison. *Econometrica* **40**(1) 159–176.

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

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

Soofi, A. 1992. Industry linkages, indices of variation and structure of production: An international comparison. *Economic Systems Research* **4**(4) 349–376.

Stiroh, K. 2001. What drives productivity growth? *Economic Policy Review* **7**(1) 37–59.

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

Stiroh, K. J. 2002b. Information technology and the U.S. productivity revival: What do the industry data say? *American Economic Review* **92**(5) 1559–1576.

Streitwieser, M.L. 2009. A primer on BEA's industry accounts. *Survey of Current Business,* **June** 40–52.

Tambe, P., L. Hitt. 2014. Job hopping, information technology spillovers, and productivity growth. *Management Science* **60**(2) 338–355.

U.S. Department of Transportation, Research and Innovative Technology Administration, Bureau of Transportation Statistics. 2011. *Transportation Satellite Accounts: A Look at Transportation's Role in the Economy*. Washington, DC.

Wang, E. T. G., A. Seidmann. 1995. Electronic data interchange: Competitive externalities and strategic implementation policies. *Management Science* **41**(3) 401–418.

Williamson, O. E. 1981. The economics of organization: The transaction cost approach. *American Journal of Sociology* **87**(3) 548–577.

Williamson, O. E. 1985. *The Economic Institutions of Capitalism*. Free press, New York.

Williamson, O.E. 1989. *Handbook of Industrial Organization*, vol. 1. Elsevier Science Publishers B.V.

Wilson, R. 2012. 23rd state of logistics report: The long and winding recovery. CSCMP annual state of logistics report presentation, June 13, 2012.

Wooldridge, J. M. 2002. *Econometric Analysis of Cross Section and Panel Data*. MIT Press, Cambridge, MA.

Wooldridge, J. M. 2009. *Introductory Econometrics: A Modern Approach*. 4th ed. South-Western, Cengage Learning.

Zakon, R. H. 2011. Hobbes' Internet timeline 10.2. http://www.zakon.org/robert/internet/timeline.

Zhang, D., Z. Cheng, H. Q. Mohammad, B. R. Nault. 2014. Information technology substitution revisited. *Information Systems Research* Forthcoming.

Zhu, K. 2004. The complementarity of information technology infrastructure and e-commerce capability: A resource-based assessment of their business value. *Journal of Management Information Systems* **21**(1) 167–202.

Zhu, K., K. L. Kraemer. 2002. E-commerce metrics for net-enhanced organizations: Assessing the value of e-commerce to firm performance in the manufacturing sector. *Information Systems Research* **13**(3) 275–295.

Appendix A

First Appendix

Variable	Source	Construction Procedure	Deflator
Gross Output (Y)	MFP dataset of the BLS	Gross output by industry converted to 1987 U.S. dollars.	Gross output price deflator
IT Capital (Z)	Requested from the BLS	Productive information capital stock by asset type (Direct aggregate-millions of 1987 dollars).	None
Non-IT Capital (K)	Calculated by using capital stock and information capital stock from the BLS	Productive capital stock of equipment and structure, excluding information capital stock (Direct aggregate-millions of 1987 dollars).	None
Labour (L)	MFP dataset of the BLS	Hours of all persons (in millions).	None
Total Intermediate Inputs (M)	MFP dataset of the BLS	The total intermediate inputs converted to 1987 U.S. dollars.	Intermediate inputs price deflator
Non-Logistics Intermediate Inputs (X)	Calculated by using intermediate inputs and the input-output tables from the BLS	An industry's total intermediate inputs subtract purchased logistics inputs. The input-output tables are in 1987 dollars.	None
Outsourced Logistics (W)	The input-output tables from the BLS	The aggregation of an industry's purchases of the following commodities: SIC40; SIC 421, 423; SIC 422; SIC 44; SIC 45; SIC 46. The input-output tables are in 1987 dollars.	None
Customers' IT Capital (C)	Calculated by using the input-output tables and information capital	$C_i = \sum_{j \neq i} \frac{V_{ij}}{\sum_{j \neq i} V_{ij}} Z_j$, using transactions between industry i and its customers as the weights and the IT capital of each customer as base to get the aggregation of customers' IT capital stock.	None

Figure A.1: Data Sources and Construction Procedure for Dataset I

Variable	Source	Construction Procedure	Deflator
Gross Output (Y)	BEA Annual Industry Account	Gross output by industry converted to 2005 U.S. dollars.	Chain-type quantity index for output by industry from BEA
IT Capital (Z)	MFP dataset from the BLS	Productive information capital stock by asset type (Direct aggregate-millions of 2005 Dollars).	None
Non-IT Capital (K)	Calculated by using capital stock and information capital stock from the BLS	Productive capital stock of equipment and structure, excluding information capital stock (Direct aggregate-millions of 2005 Dollars).	None
Labour (L)	Requested from the labour productivity and costs section of the BLS	Hours of all persons (in millions)	None
Total Intermediate Inputs (M)	BEA Annual Industry Account	Intermediate inputs by industry converted to 2005 U.S. dollars.	Chain-type quantity indexes for intermediate inputs (IIQI)
Non-Logistics Intermediate Inputs (X)	Calculated by using total intermediate inputs and the input-output tables from the BEA	An industry's total intermediate inputs from the use tables, excluding purchased logistics inputs. Converted to 2005 dollars.	Chain-type quantity indexes for intermediate inputs (IIQI)
Outsourced Logistics (W)	The input-output tables from the BEA	The aggregation of an industry's purchases of the following commodities: NAICS 481, 482, 483, 484, 486, and 493. Converted to 2005 U.S. dollars.	Chain-type quantity indexes for gross output (GOQI) of the sector of warehousing and transportation
Customers' IT Capital (C)	Calculated by using the input-output tables and information capital	$C_i = \sum_{j \neq i} \frac{V_{ij}}{\sum_{j \neq i} V_{ij}} Z_j$, using the transactions between a given industry i and its customers as the weights to get the aggregation of customers' IT capital stock.	None

Figure A.2: Data Sources and Construction Procedure for Dataset II

Industry Number	2002 NAICS Industry Code	Industry Title
8	321	Wood products
9	327	Nonmetallic mineral products
10	331	Primary metals
11	332	Fabricated metal products
12	333	Machinery
13	334	Computer and electronic products
14	335	Electrical equipment, appliances, and components
15	336	Transportation equipment
16	337	Furniture and related products
17	339	Miscellaneous manufacturing
18	311, 312 (311FT)	Food and beverage and tobacco products
19	313, 314 (313TT)	Textile mills and textile product mills
20	315, 316 (315AL)	Apparel and leather and allied products
21	322	Paper products
22	323	Printing and related support activities
23	324	Petroleum and coal products
24	325	Chemical products
25	326	Plastics and rubber products

Notes: Industry number is the sequence in which the industries appear in the input-output tables. The codes in parentheses are the ones used in the input-output tables, if different from the NAICS codes.

Figure A.3: Dataset II (2000-2008): 3-digit NAICS Industry Description

Appendix B

Second Appendix

Variable	2000-2009 Full Sample	2000-2009 Manufacturing	1987-1999	1987-1993	1994-1999
Connectivity	+	-	-	-	-
Concentration	-	+	-	+	-

Notes: 1987-1999 industries are all in manufacturing sectors.

The Summary of Results Using IVs (ivreg2)

Variable	2000-2009 Full Sample	2000-2009 Manufacturing	1987-1999	1987-1993	1994-1999
Connectivity	NS	-	-	-	NS
Concentration	NS	+	-	-	-

Notes: 1987-1999 industries are all in manufacturing sectors; 2-step GMM, and one-year and two-year lags of IT as IVs. "NS" means estimation results are not significant. The same in the below.

The Summary of Results Considering Labor and Non-IT Capital

Variable	2000-2009 Full Sample	2000-2009 Manufacturing	1987-1999
Connectivity	NS	-	-
Concentration	-	+	-

The Summary of Results Using Normalized IT

Variable	2000-2009 Full Sample	2000-2009 Manufacturing	1987-1999
Connectivity	NS	-	-
Concentration	+	-	+

Figure B.1: Summary of Estimation Results

Variables	(1) Broker	(2) Nbroker	(3) Egobetween	(4) Negobetween	(5) Lbroker	(6) Legobetween
Log(IT)	-2.152*	-0.0239***	-0.839***	-0.440***	-0.298*	-0.625***
	(1.208)	(0.00752)	(0.281)	(0.0928)	(0.155)	(0.167)
Constant	28.44***	0.285***	9.239***	4.569***	4.769***	5.907***
	(9.571)	(0.0596)	(2.223)	(0.733)	(1.233)	(1.317)
Observations	144	144	144	144	144	144
R-squared	0.932	0.926	0.922	0.913	0.943	0.959
Industry FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Hansen/Sargan C-test Statistics	Chi-sq(1) = 1.094, P=0.296	Chi-sq(1) = 1.118, P=0.290	Chi-sq(1) = 1.579, P=0.209	Chi-sq(1) = 1.748, P=0.186	Chi-sq(1) = 0.445, P=0.505	Chi-sq(1) = 1.129 , P=0.288

Notes: It uses one year and two year lags of IT variable as IVs, controls for industry dummies (by industry) and year dummies. It uses 2-step GMM controlling for arbitrary heteroskedasticity. Eg. xi: ivreg2 broker (lz=l.lz l2.lz) i.year i.nbr, gmm2s robust

Figure B.2: The Estimation Results of Connectivity Measures Using IVs for the Sample of Manufacturing Industries in Dataset II

Variables	(1) CVS	(2) MSS2	(3) CR4S	(4) CR8S
Log(IT)	0.266**	0.0900**	0.0335**	0.0121***
	(0.126)	(0.0378)	(0.0165)	(0.00447)
Constant	-1.018	-0.597**	0.327**	0.746***
	(1.000)	(0.301)	(0.130)	(0.0351)
Observations	144	144	144	144
R-squared	0.975	0.969	0.962	0.975
Industry FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Hansen/Sargan C-test Statistics	Chi-sq(1) = 3.799, P=0.051	Chi-sq(1) = 4.516, P=0.034	Chi-sq(1) = 1.780, P=0.182	Chi-sq(1) = 0.003, P=0.959

Figure B.3: The Estimation Results of Concentration Measures Using IVs for the Sample of Manufacturing Industries in Dataset II

VARIABLES	(1) Broker	(2) Nbroker	(3) Egobetween	(4) Negobetween	(5) Lbroker	(6) Legobetween
Log(IT)	-3.196*	-0.00204**	0.0800	-0.00947	-0.0504	-0.0303
	(1.913)	(0.000915)	(0.255)	(0.00687)	(0.0787)	(0.108)
Constant	88.68***	0.0558***	9.340***	0.366***	4.548***	2.473***
	(11.80)	(0.00565)	(1.560)	(0.0420)	(0.485)	(0.666)
Observations	425	425	425	425	425	410
R-squared	1.000	1.000	1.000	1.000	0.997	0.999
Industry FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Hansen/Sargan C-test Statistics	Chi-sq(1) = 5.557, P=0.02	Chi-sq(1) = 7.311, P=0.007	Chi-sq(1) = 0.002, P=0.968	Chi-sq(1) = 0.203, P=0.652	Chi-sq(1) = 4.962, P=0.026	Chi-sq(1) = 0.950 , P=0.330

Figure B.4: The Estimation Results of Connectivity Measures Using IVs for 1987–1993 Dataset

VARIABLES	(1) Broker	(2) Nbroker	(3) Egobetween	(4) Negobetween	(5) Lbroker	(6) Legobetween
Log(IT)	1.337	0.000376	0.405	0.0108	0.0102	0.0295
	(3.343)	(0.00161)	(0.307)	(0.00893)	(0.0593)	(0.0660)
Constant	59.59***	0.0405***	7.144***	0.237***	4.172***	2.097***
	(22.76)	(0.0110)	(2.071)	(0.0599)	(0.405)	(0.450)
Observations	340	340	340	340	340	328
R-squared	0.999	0.998	1.000	1.000	0.991	0.998
Industry FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Hansen/Sargan C-test Statistics	Chi-sq(1) = 0.073, P=0.787	Chi-sq(1) = 0.359, P=0.548	Chi-sq(1) = 0.013, P=0.909	Chi-sq(1) = 0.062, P=0.803	Chi-sq(1) = 0.530, P=0.466	Chi-sq(1) = 0.794, P=0.373

Figure B.5: The Estimation Results of Connectivity Measures Using IVs for 1994–1999 Dataset

VARIABLES	(1) CVS	(2) MSS2	(3) CR4S	(4) CR8S
Log(IT)	-0.278***	-0.0336***	-0.00814	0.00426
	(0.0998)	(0.0104)	(0.00928)	(0.00599)
Constant	11.07***	1.102***	1.024***	0.958***
	(0.618)	(0.0643)	(0.0571)	(0.0368)
Observations	425	425	425	425
R-squared	0.997	0.996	0.998	0.998
Industry FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Hansen/Sargan C-test Statistics	Chi-sq(1) = 1.208, P=0.272	Chi-sq(1) = 0.959, P=0.328	Chi-sq(1) = 0.919, P=0.338	Chi-sq(1) = 0.121, P=0.728

Figure B.6: The Estimation Results of Concentration Measures Using IVs for 1987–1993 Dataset

VARIABLES	(1) CVS	(2) MSS2	(3) CR4S	(4) CR8S
Log(IT)	0.0638	0.0107	-0.0103	-0.0117**
	(0.0926)	(0.00976)	(0.00949)	(0.00552)
Constant	8.803***	0.799***	1.036***	1.058***
	(0.633)	(0.0672)	(0.0646)	(0.0375)
Observations	340	340	340	340
R-squared	0.997	0.997	0.997	0.998
Industry FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Hansen/Sargan C-test Statistics	Chi-sq(1) = 2.610, P=0.106	Chi-sq(1) = 2.570, P=0.109	Chi-sq(1) = 3.586, P=0.058	Chi-sq(1) = 0.846, P= 0.358

Figure B.7: The Estimation Results of Concentration Measures Using IVs for 1994–1999 Dataset

Copyright Authorization

I, Barrie R. Nault, as a co-author of the paper “An Internet-Enabled Move to the Market ", give permission to Fengmei Gong to include the work in a thesis titled “Studies on IT, Logistics, and the Structure of Production ". I am aware that Fengmei Gong will sign a Thesis Non-Exclusive License that authorizes Library and Archives Canada to reproduce, communicate to the public on the Internet, loan, distribute or sell copies of this thesis.

_____

Name

_____ _____

Signature Date

Copyright Authorization

I, Barrie R. Nault, as a co-author of the paper “The Impact of IT on Production Interdependence", give permission to Fengmei Gong to include the work in a thesis titled “Studies on IT, Logistics, and the Structure of Production ". I am aware that Fengmei Gong will sign a Thesis Non-Exclusive License that authorizes Library and Archives Canada to reproduce, communicate to the public on the Internet, loan, distribute or sell copies of this thesis.

_____
Name

_____ _____
Signature Date

Copyright Authorization

I, Barrie R. Nault, as a co-author of the paper “The Impact of IT on the Structure of Production", give permission to Fengmei Gong to include the work in a thesis titled “Studies on IT, Logistics, and the Structure of Production ". I am aware that Fengmei Gong will sign a Thesis Non-Exclusive License that authorizes Library and Archives Canada to reproduce, communicate to the public on the Internet, loan, distribute or sell copies of this thesis.

_____
Name

_____ _____
Signature Date

Copyright Authorization

I, Mohammad S. Rahman, as a co-author of the paper “An Internet-Enabled Move to the Market ", give permission to Fengmei Gong to include the work in a thesis titled “Studies on IT, Logistics, and the Structure of Production ". I am aware that Fengmei Gong will sign a Thesis Non-Exclusive License that authorizes Library and Archives Canada to reproduce, communicate to the public on the Internet, loan, distribute or sell copies of this thesis.

_____
Name

_____ _____
Signature Date

Copyright Authorization

I, Zhuo (June) Cheng, as a co-author of the paper “The Impact of IT on Production Interdependence", give permission to Fengmei Gong to include the work in a thesis titled “Studies on IT, Logistics, and the Structure of Production ". I am aware that Fengmei Gong will sign a Thesis Non-Exclusive License that authorizes Library and Archives Canada to reproduce, communicate to the public on the Internet, loan, distribute or sell copies of this thesis.

_____

Name

_____ _____

Signature Date

Copyright Authorization

I, Byungwan Koh, as a co-author of the paper “The Impact of IT on the Structure of Production", give permission to Fengmei Gong to include the work in a thesis titled “Studies on IT, Logistics, and the Structure of Production ". I am aware that Fengmei Gong will sign a Thesis Non-Exclusive License that authorizes Library and Archives Canada to reproduce, communicate to the public on the Internet, loan, distribute or sell copies of this thesis.

_____

Name

_____ _____

Signature Date