INFORMATION TECHNOLOGY FIRM PRODUCTIVITY: INVESTIGATING THE IMPACTS OF SOFTWARE AS A SERVICE MODEL AND WORKER MOBILITY

GE CHUNMIAN

(B.Eng.), South China University of Technology

A THESIS SUBMITTED

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF INFORMATION SYSTEMS

NATIONAL UNIVERSITY OF SINGAPORE

2014

DECLARATION

I hereby declare that the thesis is my original work and it has been written by me in its entirety. I have duly acknowledged all the sources of information which have been used in the thesis.

This thesis has also not been submitted for any degree in any university previously.

Ge Chunmian

GE CHUNMIAN

2 July 2014

ACKNOWLEDGEMENTS

The completion of this dissertation would not have been possible without the support from a group of people, to whom I am sincerely indebted.

First and foremost, I would like to express my sincere gratitude to my committee chair and advisor, Prof. Ke-Wei Huang. His research experience and extensive knowledge have greatly helped me develop my background in the field of economics of information systems. I appreciate all his contributions of time, ideas, and instructions to my research work, as well as his guidance and support in my professional career. Particularly, I would like to thank him for his patience, support and guidance during the tough times in my doctoral study. Prof. Ke-Wei Huang is a great advisor. It is my honor to be his first Ph.D. student.

Second, I am also very thankful to my thesis committee, Prof. Atreyi Kankanhalli and Prof. Seung Hyun Kim, who have devoted much time and effort to helping me improve the quality of this research. I appreciate their invaluable insights and comments to this research. More importantly, their rigorous way of designing and conducting scientific research has enlightened me and will continue benefiting me in my future academic career.

Third, I also would like to thank all the other professors in the department of information systems, who have encouraged and inspired me. In particular, I would like to express my sincere gratitude to Prof. Ivan Png, whose invaluable instructions have considerably enriched my research. Working with him is a great pleasure.

Fourth, I devote my special thanks to my fellow colleagues and friends in NUS for their continuous support and inspiration during my doctoral journey. Our friendship is a great fortune of my life.

Last but not least, I would like to thank my family for their unconditional support and selfless love, as well as their uplifting encouragement during my pursuit of Ph.D. To them I dedicate this dissertation.

TABLE OF CONTENTS

SUMMARY

Productivity analysis is an important tool in assessing a firm's ability of converting inputs into output. Previous information systems (IS) research often focused on examining the impacts of information technology (IT) capital investment on firm productivity. However, there could be other phenomena that significantly affect the productivity of IT firms. With this in mind, this dissertation includes two studies which investigate the productivity impacts of two important phenomena regarding the IT industry. Specifically, the two studies examine the impacts of Software as a Service (SaaS) model and worker mobility on firm productivity in the IT industry.

Study 1 examines the impacts of a new business model - SaaS, on the productivity of software firms. SaaS has been one of the fastest-growing delivery models in the software industry. The industry's trade press often considers economies of scale as the main benefit of SaaS firms because IT management and the associated resources are centralized at the SaaS vendors. However, centralized IT management also requires the associated cost of expanding the firm's IT infrastructure to serve more customers. Intuitively, it is not necessary that the former effect must dominate the latter. Using public firm-level data from Compustat, this study attempts to analyze the economies of scale of SaaS firms relative to their traditional counterparts by using the Stochastic Frontier Analysis (SFA) approach. Our empirical findings suggest SaaS firms have smaller economies of scale than traditional software firms. By

utilizing the technical efficiency score obtained from SFA, we further examine the effects of R&D expense and advertising expense on technical efficiency. The analysis suggests that it is the R&D expense, not the advertising expense, that could be the cause of smaller economies of scale at SaaS firms.

Study 2 examines the impacts of worker mobility on the productivity of IT firms. Leading software firms, such as Google, make considerable investments towards employee training and development. However, these investments not only enhance their productivity but could also benefit other IT firms which hire their employees. In other words, recipient IT firms are likely to experience productivity spillovers due to human resource (HR) investments by the leading firms. Prior studies have provided evidence of productivity spillover through labor mobility in the IT industry, but the sources of productivity spillover have not been differentiated. This paper contributes to the literature by examining the sources of labor productivity spillover in the IT industry. Using a novel structural approach with joint General Moment Method estimation and a large dataset on labor mobility derived from LinkedIn, we show that hiring one employee from leading software platform providers i.e., Google, Microsoft, Facebook, and Oracle, is associated with USD $1.14 million increase in value-added of the recipient IT firms. Further, we find that the spillover effect solely derives from leading firms' employees with work experience greater than 2 years or with post-graduate degrees, thus explicating possible sources of labor productivity spillover for IT firms.

The notable findings from this research enrich the productivity literature of information technology by empirically examining the productivity impacts of SaaS model and worker mobility. The findings also provide managerial implications for IT firm managers and government policy makers, which are further discussed.

Keywords: Software as a Service, productivity analysis, stochastic frontier analysis, R&D expense, human resource management, productivity spillover.

LIST OF TABLES

LIST OF FIGURES

1. GENERAL INTRODUCTION

Academic researchers have shown an enormous amount of interest in productivity research (e.g., Brynjolfsson and Hitt 1996; Griliches 1994; Solow 1957; Tambe and Hitt 2014b). Productivity research is important in that it quantifies the ability of a unit (such as a firm, an industry or a country) to convert inputs into output (Del Gatto et al. 2011). Based on the productivity analysis, implications and suggestions for productivity improvement could then be provided if we are able to further identify the source(s) that leads to different productivity level. In this regard, empirical studies of productivity analysis have strong implications for both corporate managers and government policy makers.

Entering the era of information technology (IT), information systems (IS) researchers were particularly interested in and focused on assessing the effect of IT capital investment on productivity (e.g., Dewan and Kraemer 2000; Hitt et al. 2002; Tambe and Hitt 2012b). Initially, researchers were not able to find positive and significant productivity contributions of IT investment (e.g., Loveman 1994; Morrison and Berndt 1991), given that computers had radically restructured the business processes in a lot of American corporations (Brynjolfsson and Hitt 1996). This puzzle is well known as the "productivity paradox", which was later addressed by the seminal work of Brynjolfsson and Hitt (1996). In their work, they used new firm-level data to assess several econometric models and found that IT investment had a substantial contribution to firm output. Therefore, they concluded that the "productivity paradox" disappeared after year 1991. After this seminal work, a lot of empirical evidence regarding the positive contribution of IT investment to productivity has been conducted at firm level (e.g., Tambe et al. 2012), industry level (e.g., Mittal and Nault 2009), and country level (e.g., Dewan and Kraemer 2000). By the time of this dissertation, IS researchers have already reached a consensus on the positive productivity contribution of IT investment (e.g., Hitt and Brynjolfsson 1996; Hitt et al. 2002; Kudyba and Diwan 2002; Park et al. 2007b; Tambe and Hitt 2012b).

However, in our modern economy, there are more phenomena in the IT industry that need examinations into their impacts on firm productivity issues beyond IT investment. *First*, there is a growing body of new business models in the IT industry. Some widely discussed ones include Software as a Service (SaaS), Platform as a Service (PaaS), and Infrastructure as a Service (IaaS), which jointly forms "cloud computing". A number of famous SaaS (cloud) services have also been provided and are accessible through Internet, such as cloud storage (e.g., Dropbox), enterprise customer relation management (CRM) service (e.g., Salesforce.com), human capital management service (e.g., SuccessFactors), marketing and public relations service (e.g., Vocus) and so on. New business models can represent a form of innovation and some of them will be much better adapted to customer needs and business environment than others (Teece 2010), leading to higher firm productivity. Armbrust et al. (2010)

described cloud computing as having "*the potential to transform a large part of the IT industry, making software even more attractive as a service and shaping the way IT hardware is designed and purchased*". Although researchers are aware of the importance of SaaS model in the IT industry, there is a lack of empirical studies regarding whether the new business model has substantially improved firm performance by improving firm productivity. For example, it is unclear to us whether the SaaS model indeed outperforms the traditional software delivery model in terms of productivity. In fact, many new business models in the IT industry remain largely unexplored regarding their impacts on firm productivity. Therefore, academic understanding on such issues is rather limited, as not much rigorous research has been done to empirically examine the impacts (Teece 2010). Given that SaaS has been one of the fastest-growing delivery models in the IT industry, it is salient to empirically examine the impacts of such business models on firm productivity in the IT industry. Findings regarding the productivity impacts of the SaaS model will provide useful information for firms which are considering entering this market. Further, identifying the source(s) of productivity efficiency (or inefficiency) could also provide implications for the existing SaaS firms to improve their productivity.

Second, talent war has been particularly prevalent in the IT industry, due to the intense competition for IT talents. The IT industry is widely perceived as human-capital intensive, as well as knowledge intensive (Boh et al. 2007;

Bresnahan et al. 2002). This industry is well known as having hired a large number of highly-educated knowledge workers (typically software and hardware engineers). The large demand of IT workers in the industry has lead to intense competition for talents. For instance, Apple poached former Google engineers to fix its map application (Wong 2012). Apple also invited a number of BlackBerry employees to a recruitment event (Pepitone 2013). The talent war within the IT industry is so intense that even Apple was reported to ask Google to stop poaching its employees (Whitney 2013). Given that the competition for IT talents is intense, it is interesting for us to know whether such recruitment does actually benefit the firms in terms of productivity improvement. If yes, what is the magnitude of such recruitment? While previous IS studies often focused on whether firm productivity could benefit from IT capital/hardware investment (e.g., Chang and Gurbaxani 2012a; Cheng and Nault 2007), it is natural for us to come out with the question: how is the productivity of IT firms affected by the labor side, specifically, the movement of their employees? Especially, from the perspective of productivity, which kind of recruited workers would benefit the recipient firms most? Answers to these questions will enrich the productivity literature and provide strong managerial implications for IT firms' recruitment practices.

Therefore, beyond the productivity impacts of IT investment, it is also critical to conduct productivity analysis to empirically examine the impacts of SaaS model, as well as the movement of IT firm workers, on firm productivity within the IT industry. With this in mind, this dissertation includes the following two studies.

The first study of this dissertation focuses on the economies of scale of SaaS firms by applying an advanced and sophisticated technique, namely stochastic frontier analysis. We find interesting results that are contrary to the popular claim of the media. The SaaS model is famous for the IT infrastructure being centralized and hosted by the vendors. While the media claims that economies of scale are one of the key properties of SaaS models due to such centralized and shared IT infrastructure (e.g., Desisto 2010), we do find empirical evidence that this may not be the case. We argue, theoretically, that the centralized IT infrastructure also brings significant variable cost to SaaS firms. By comparing the two models of software production – the traditional model vs. the SaaS model, we find that the economies of scale in SaaS are weaker than that in the traditional model. Indeed, SaaS firms exhibit diseconomies of scale. More theoretical arguments are also provided to elaborate the findings. Further, we find that the diseconomies of scale of SaaS firms may result from their over investing on Research and Development (R&D) activities.

The second study of this dissertation examines the productivity spillover from major software providers through worker mobility. We find significant productivity spillover effects from four major software providers to other IT firms through the movement of workers. Given that most existing studies focus on the productivity spillover of IT investment, it is important to examine whether there is also productivity spillover due to the labor side. Although some pioneering studies (Parrotta and Pozzoli 2012; Tambe and Hitt 2014b) have documented that there exists productivity spillover through worker mobility, this study adds to the literature by scrutinizing the *source* of labor productivity spillover in the IT industry. Identifying the source of productivity spillover is also critical and provides strong implications for human resource (HR) practices. By figuring out the best source to recruit employees, a recipient IT firm can improve its recruitment strategy. This study provides strong empirical evidence that IT firms should actively recruit employees from those major software providers – Google, Microsoft, Facebook and Oracle. In contrast, recruiting employees from many other firms would not lead to significant productivity spillover. Another point that we would like to highlight here is that this study employs the state-of-the-art productivity analysis techniques proposed by Wooldridge (2009). Applying this structural approach, we are able to establish the causality. Further, we show that the productivity spillover from the four major software providers is mostly contributed by their experienced and well-educated workers.

In conclusion, this dissertation empirically investigates the impacts of two important phenomena regarding the IT industry, i.e., SaaS model and worker mobility, on IT firm productivity. Through these investigations, this dissertation aims to complement and enrich the productivity literature of information technology. Our empirical findings suggest that, employing SaaS model does not enable SaaS firms to enjoy the economies of scale depicted by the media. Our results also show that recruiting workers from major software providers would significantly and greatly benefit the recipient firms in terms of productivity improvement. At the same time, this dissertation seeks to identify and suggest potential solutions to increase IT firm productivity. Our results suggest that the diseconomies of scale of SaaS firms may be alleviated by reducing their R&D investments. Meanwhile, IT firms can recruit well-educated workers with more than five years' experience in leading software providers to maximize the productivity spillover effects. More theoretical contributions and practical implications are provided in Section 2 and Section 3.

2. STUDY 1: EXPLORING THE IMPACTS OF SOFTWARE AS A SERVICE MODEL ON IT FIRM PRODUCTIVITY

2.1. Introduction

We have witnessed a sea of change in IT innovations for services management in the past decade (Rai and Sambamurthy 2006). One prominent innovation is the new software delivery model: Software as a Service (SaaS). In SaaS, software and the associated data are hosted centrally by the service providers, rather than being hosted in-house by the corporate clients. SaaS has been one of the fastest-growing segments in the industry since its inception, and is rapidly becoming an important consideration for enterprises of all types and sizes (Gartner 2009). The most successful SaaS vendor, Salesforce.com, grew its revenue from $176.4 million in 2004 to $1.66 billion in 2010. Consistent growth in the revenue of SaaS vendors suggests this new business model is not just a technology fad but rather an indication of where the software market is heading. Therefore, in this study, we will focus on studying the SaaS vendors.

A well-known property of the traditional software business is *economies of scale* (Schmidt and Schnitzer 2003). Economies of scale exist when the average production cost decreases as the number of units produced increases (Tirole 1988). Therefore, larger firms enjoy cost advantages with the presence of economies of scale. In the traditional software industry, the costs of replication and distribution are typically negligible after the significant cost incurred for the production of the "first copy" (Schmidt and Schnitzer 2003). This leads to significant economies of scale.[1] This zero variable cost property usually curb competition and yield oligopolies (Katz and Shapiro 1994), such as Microsoft, Oracle, and SAP.

In the SaaS era, economies of scale are also widely perceived as one of the key contributors to the fast adoption of the SaaS model, although the reasons behind this are different from those of the traditional software industry. For example, Gartner defines "sharing resources and economies of scale" as one of the four components of SaaS (Desisto 2010), since the IT infrastructure is centralized on the vendor side and shared among all customers, which in turn leads to economies of scale (e.g., Bonvanie 2007; ComputerWeekly 2009). The cost sharing and servers' load balancing are perceived as the main drivers of economies of scale in SaaS vendors. For example, according to Salesforce.com, its multi-tenant architecture leads to massive economies of scale to optimize computing resources across all customers (Salesforce.com 2011).

However, properties of the SaaS business models are frequently mentioned but rarely analyzed with rigor (Teece 2010). Although SaaS is recognized as a huge success of IT innovation, the widely accepted "economies of scale" have not been empirically tested in the literature. In theory, when the IT

[1] Adding to the supply-side economies of scale, a larger technology firm could outperform smaller competitors due to network effects as well.

infrastructure is centralized at the SaaS vendors, the associated costs are shifted from the customers to the vendors. Accordingly, to provide more units of "computing services" and to serve more corporate customers, SaaS vendors may incur a variety of costs related to the units of computing services provided. The most straightforward cost items include electricity bills and the data communication costs of delivering SaaS. SaaS vendors may incur semi-variable costs such as expanding their IT infrastructure in terms of installing more servers, procuring more storage devices, renting a larger space, as well as hiring more IT professionals. These IT infrastructure costs radically change the zero variable cost property in a unique way because infrastructure costs are neither purely fixed nor completely variable in nature (Huang and Sundararajan 2011).[2] In this way, there are two countervailing effects regarding the economies of scale of SaaS firms, motivating us to further examine the overall effect by applying advanced productivity analysis to this new and promising software business model.

Therefore, this study's research objective is to investigate the firm-level economies of scale of SaaS firms. Specifically, we are interested in the following questions: (1) Do SaaS firms exhibit economies of scale? (2) Are the economies of scale of SaaS firms larger or smaller than those of traditional

[2] This infrastructure cost is not fixed because it depends on the number of buyers in a discontinuous fashion. Given a fixed number of servers, a SaaS firm can only serve a limited number of customers with satisfactory performance. In order to serve more customers, a SaaS firm needs to expand its "IT capacity." The infrastructure cost is also not purely variable like material costs because the marginal cost of providing one more unit of IT service is still close to zero.

software firms? (3) What are the major sources of (dis)economies of scale of SaaS firms? Answers to these questions are critical for the competition strategy of SaaS vendors. For example, if SaaS vendors exhibit economies of scale, then they should focus on developing a larger user base even at the cost of a loss in the early stage. But if not, then SaaS vendors have to focus more on differentiating and customizing their products.

We compiled an unbalanced panel dataset of 23 publicly listed SaaS firms and 480 publicly listed traditional software firms between 2002 and 2010. The firm-level measure of economies of scale is calculated by Stochastic Frontier Analysis (SFA), one of the most advanced methods in productivity analysis. One main benefit of SFA over linear regression, which is widely used in the existing literature, is that SFA produces a technical efficiency (TE) score for each firm in each year. Consistent with the productivity literature, we use capital and labor as two input variables of the production function. The output variable is the economic value added (Brynjolfsson and Hitt 1996). The results show that SaaS firms exhibit diseconomies of scale, and at the same time, smaller economies of scale than traditional software firms. In addition, the analysis shows that R&D contributes more to TE growth than advertising, while more efficient SaaS firms tend to spend less on R&D but not less on advertising expenses. These findings indicate that diseconomies of scale may result from the decreasing return in R&D investment.

2.2. Theory and Hypotheses

2.2.1. Software as a Service, Application Service Provider, and On-Demand Computing

SaaS is a relatively new software delivery business model. Compared with the traditional software delivery model, SaaS has three unique features. First, the SaaS model offers Web-based access to business software applications, while the traditional model requires the software to be installed on customers' own machines. Second, in the SaaS model, multiple customers access the same application based on the shared IT infrastructure provided, without having to make additional investments for hardware, installation, and maintenance (Choudhary 2007). Third, customers pay a small recurring subscription fee based on usage, rather than a large, one-time software license, as in the traditional model (see Table 1-1).

Table 1 - 1. Differences Between Two Software Delivery Models

	SaaS Delivery Model	Traditional Delivery Model
Installation	■ Vendors purchase the hardware. ■ Vendors install the software.	■ Customers purchase the hardware. ■ Customers install the software.
Maintenance	■ Customers do not need to have their own IT maintenance team.	■ Customers need to have their own IT maintenance team
License/Fee	■ Subscription-based usage ■ Customers pay small recurring fees.	■ Perpetual license ■ Customers pay large fees at one time.

Two frequently used jargons that are similar to SaaS are: Application Service Provider (ASP) and On-Demand Computing. Around year 2000, ASP and SaaS were totally equivalent concepts (SIIA 2001), but after 2005, minor differences between them started to emerge. SaaS vendors typically self-develop and deliver a new software application based on a powerful shared computing infrastructure. In contrast, ASP is more like a third-party distributor of existing solutions. ASP vendors obtain authorization from the software developers and release the software to the end users as a service, using subscription-based pricing plans. The underlying IT infrastructure of ASP is often dedicated rather than shared.

IS researchers have examined various issues of the ASP business model. Walsh (2003) provided an excellent overview of the technologies, economies, and strategies of ASP. Smith and Kumar (2004) developed a theory of ASP adoption from the customer's perspective. Currie and Parikh (2006) developed an integrative model to understand value creation in Web services from a provider's perspective. Susarla et al. (2003) empirically showed that expectations about ASP services had a significant impact on their performance evaluation. Cheng and Koehler (2003) derived an optimal pricing policy for ASP vendors. Ma and Seidmann (2007; 2008) studied the profitability of ASP pricing. Susarla and Barua (2011) studied the determinants of ASP survival.

Similarly, on-demand computing service (a.k.a. utility computing) is a popular synonym of SaaS. Some SaaS firms, such as Omniture Inc., use on-demand computing to describe their business model in their official annual reports. A few academic publications deal with on-demand computing or SaaS. Bhargava and Sundaresan (2004) studied various pricing mechanisms for on-demand computing with demand uncertainty. Choudhary (2007) contrasted

SaaS and perpetual licensing. Xin and Levina (2008) investigated the client-side determinants of SaaS model adoption. Fan et al. (2009) examined short- and long-term competition between SaaS and traditional software providers. Recently, Chen and Wu (2013) studied the impact of adopting on-demand services on market structure, firm profitability, and consumer welfare.

None of the above studies has examined the productivity of SaaS firms. However, there is a critical need to empirically measure the productivity of firms that adopt service innovations within service-oriented systems (Bardhan et al. 2010). The present study contributes to the literature by bridging this gap and providing more empirical evidence on the economic properties of SaaS firms.

2.2.2. Economies of Scale

Although economies of scale of software development (Banker et al. 1994) and maintenance (Banker and Slaughter 1997) have been investigated at the project level, economies of scale of SaaS firms have yet to be studied.

Trade magazine articles about SaaS generally cite economies of scale as one of the major benefits over the traditional software delivery model (Bonvanie 2007; ComputerWeekly 2009; Desisto 2010). For example, some of them state that "*the sheer economies of scale achieved by public cloud providers will inevitably mean they dominate in future*" (ComputerWeekly 2009).

Economies of scale, if exist, are important to both SaaS vendors and customers for the following reasons. First, strong economies of scale may lead to a winner-takes-all status in the equilibrium. With this in mind, executives of SaaS vendors should adjust their strategies to build a larger customer base as soon as possible, even at the cost of a loss in the early stage. Second, from the perspective of investors or shareholders of SaaS firms, a winner-takes-all situation makes their investments riskier, because the target firm could either turn out to be a winner like Microsoft eventually, or file bankruptcy in the near future. Last, clients of SaaS vendors should subscribe services from larger vendors even when the provided service is not the best fit for their business processes because smaller SaaS vendors will probably be forced out of the market, even if they offer better products. Therefore, it is critical to know whether economies of scale indeed exist in SaaS firms or not, given its important role in the strategy of both SaaS vendors and clients.

However, we posit that the wisdom of crowds in the trade press may not be scientifically correct. In fact, as we will discuss in the next two paragraphs, two countervailing effects exist in the SaaS model with regard to economies of scale. One increases economies of scale, while the other decreases it. Therefore, SaaS firms may indeed have diseconomies of scale, and smaller economies of scale than traditional software firms.

First, in the SaaS model, vendors provide applications based on a powerful server farm, a large data center, and a professional IT management team at their sites. The large fixed costs of the centralized IT infrastructure are indirectly shared among all customers (Viega 2009). This cost-sharing feature is the main source of economies of scale as mentioned in industry media articles. Furthermore, a shared IT infrastructure provides another source of economies of scale due to an increase in the utilization rate of computing resources resulting from load balancing. Studies have shown that the traditional software delivery model leads to overbuilding of IT assets: the utilization rate of the computing power of servers is around 10% to 35%, while that of desktop computers is only 5% (Carr 2005). In the SaaS delivery model, because multiple firms operate on the same infrastructure, the under-utilization of processing power and storage can be alleviated. In sum, the infrastructure cost sharing and CPU time-sharing features increase the economies of scale of SaaS vendors and buyers as a group.

Second, however, when the IT infrastructure and staff are centralized at SaaS vendors, all costs then shift from the customers to the vendors. After this shift, the cost structure of SaaS vendors may depend on the amount of computing services provided and the number of corporate customers served. If the SaaS vendors outsource hosting to third-party firms, such as Amazon Web Services, all infrastructure costs become "variable" because of the on-demand pricing (per-unit usage pricing) plans of the hosting vendors. If the SaaS vendors host the IT infrastructure, the "direct variable cost" includes electricity bills and the data communication costs of Internet connection. In theory, the infrastructure costs themselves are also “semi-variable” by nature. The centralized infrastructure imposes a capacity constraint on SaaS firms: there is a limit on the CPU processing power, memory, storage space, etc. Within the capacity constraint, the variable and marginal costs are zero. However, to serve more customers beyond the capacity limit, SaaS vendors have to install more servers, rent a larger space, and hire more IT workers (Campbell-Kelly 2009). The marginal cost of IT infrastructure is non-zero at the capacity limit when the vendor acquires additional capacity. Theoretically, this unique cost structure has been shown to be equivalent to a constant variable cost (Huang and Sundararajan 2011). In accounting, “cost of revenue” is closest to variable cost conceptually. Accounting practice also treats infrastructure costs as one type of variable cost. We observe that almost all of the SaaS firms in our sample mentioned “hosting our application suite” as one major component of the cost of revenue in the annual reports. SaaS vendors also include depreciation, amortization, and maintenance of infrastructure in “cost of revenue.” For example, the third-largest SaaS vendor, Constant Contact, explicitly stated in its annual report:

> *“... The expenses related to our hosted software applications are affected by the number of customers who subscribe to our products and the complexity and redundancy of our software applications and hosting infrastructure. We expect cost of revenue to increase in absolute dollars as we expect to increase our number of*

customers….”

Therefore, unlike traditional software firms, SaaS firms have significant costs that depend on the amount of computing services provided. Besides, the higher utilization rate achieved by the centralized infrastructure may also be offset by energy inefficiency due to unmanageably huge data centers. A *large* energy-inefficient data center is likely to attract negative attention from environmentalists and the media, damaging the corporate image.

“*A single data center can take more power than a medium-size town… However, on average, these data centers were using only 6 to 12 percent of the electricity powering their servers to perform computations.*” - The New York Times (TheNewYorkTimes 2012)[3]

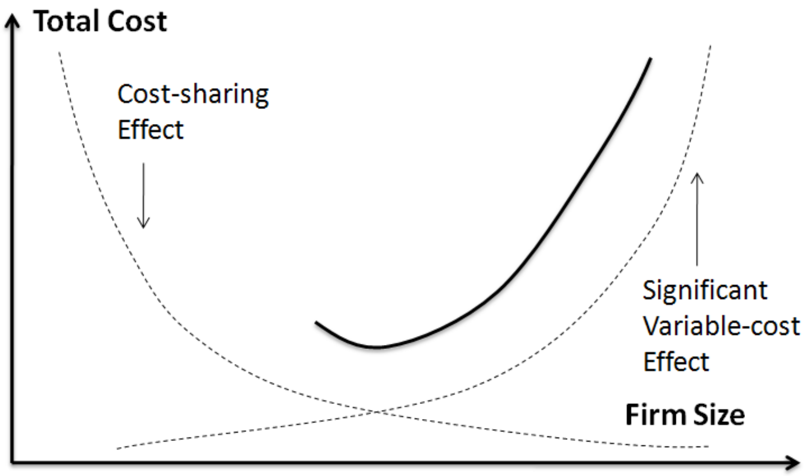


Figure 1 - 1. Economies of Scale From Two Effects

Considering these two effects together, we hypothesize that the economies of scale for SaaS firms will first increase before decreasing (as illustrated in

3 We do not intend to argue that all large data centers are inefficient but this report provides one piece of evidence that poorly managed large data centers could be inefficient. A large data center owned by famous companies is more likely to catch the attention of the media than small data centers. This is a potential cause of diseconomies of scale as well.

Figure 1-1). The reason is that, mathematically, the average cost from the first effect will converge to zero as long as the firm size keeps growing. Consequently, the total average cost of large firms will be dominated by the second effect (the increasing average cost). Even though all firms, theoretically, should operate at the most cost-efficient level if the firms have enough resources to operate at any production level (Varian and Repcheck 2010). In practice, SaaS firms may operate at a size larger than the efficient level because of the moral hazard and agency problems of executives (Varian and Repcheck 2010). The CEO's overconfidence has been documented in the literature to explain why firms kept the CEO overpaid to acquire smaller firms to accelerate company growth (Hayward and Hambrick 1997; Malmendier and Tate 2008). To sum up, we propose that SaaS firms may have the solid cost curve illustrated in Figure 1. We hypothesize that,

H1a. *The production function of publicly listed SaaS firms has diseconomies of scale.*

H1b. *The production function of publicly listed SaaS firms has smaller economies of scale than those of traditional software firms.*

2.2.3. Productivity Analysis

We expect that the productivity contribution (output elasticity) of capital in the SaaS firms' production function is larger than that of the traditional software firms' production function. This is because traditional software firms use capital or fixed assets mainly for R&D or administrative purposes, whereas a large part of the fixed assets (e.g. data centers) of SaaS firms are used to deliver SaaS services to their customers. If there are more customers, SaaS firms need to increase capital proportionally. Therefore, at SaaS firms, a large portion of the fixed assets costs may be directly accrued into the final service pricing formula. As a consequence, percentage increases in capital is highly correlated with the percentage changes in value-added. In contrast, fixed assets in traditional software firms are not that related to the output (value-added). Take computing equipments as an example of fixed assets. Computers at traditional software firms have lower utilization rate than those at SaaS firms as discussed because more computing equipments at SaaS firms are shared among multiple users or clients. Therefore, we hypothesize that,

H2.*The output elasticity of capital is higher in SaaS firms than in traditional software firms.*

Meanwhile, IT service firms need to employ both technical and client-facing personnel (Tambe and Hitt 2010; Tambe and Hitt 2012a). Prior studies have supported the existence of learning curves in software development and have highlighted its importance in the knowledge production processes (Boh et al. 2007). SaaS firms will not be able to leverage prior techniques and experience as much as traditional software firms do. In other words, SaaS firms enjoy less learning-by-doing effect. For example, the scaling problem of SaaS services, in terms of data center expansion and operation optimization, is a common challenge to the employees of SaaS firms (Holmes 2012). Lacking the required techniques and experiences would result in slow responses to service requests and customer dissatisfaction, which eventually leads to lower labor productivity and may require more R&D investment to get it fixed.

At the same time, since SaaS firms are responsible for the IT management of their customers, they will have to recruit more non-R&D IT professionals than their traditional competitors. Further, customer acceptance is relatively low for SaaS firms. Facing competition from non-SaaS incumbent competitors, many new SaaS vendors struggle to understand the key messages to put forth and how to articulate the value of their SaaS offering (Pring 2011). For example, Gartner observed that sales representatives from SaaS vendors have more difficulties in convincing customers about the reliability and security of SaaS services (Pring 2011). Moreover, many enterprises have substantial sunk cost in legacy software systems. Thereby, the switching costs are high, creating another barrier to migrate to SaaS products (Salesforce.com 2011). As a result, the sales team at SaaS firms needs to work harder to persuade and lure customers from their traditional counterparts, indicating that the marketing return on investment (ROI) of SaaS firms will be lower relative to traditional software firms. Therefore, we hypothesize that,

> **H3.** *The output elasticity of labor is lower in SaaS firms than in traditional software firms.*

Further, in research-intensive industries such as the SaaS industry, firms are forced by continual investment in R&D to introduce upgraded products for survival. R&D investment is broadly defined as investment in new knowledge that improves the production efficiency or the product quality. Extensive empirical studies have proven the benefits and necessity of R&D investment. For example, Lichtenberg and Siegel (1991) found that R&D investments pay off significantly by improving productivity. However, R&D investment is likely to have decreasing marginal return in the absence of real technical innovation (Knight 1944), meaning that high productivity firms may have less incentive to invest intensively because the return to further investment is low.

Typically, most of the functionality provided by SaaS firms is similar to existing on-premise enterprise software. SaaS firms typically use R&D investments to migrate from old software architecture to the modern one with incremental changes such as adding new features, enhancing functionalities, and improving user-interface (Salesforce.com 2011), but not to conduct disruptive, breakthrough innovation. Therefore, the ROI of R&D may reach plateau earlier than their traditional counterparts. A larger proportion of the R&D expenses of SaaS firms are invested to overcome key technical issues for scaling of the provided service (Holmes 2012). After building up the required capabilities, efficient SaaS firms may no longer need as many R&D investments as before. Besides, SaaS firms provide all customers with services based on one version because of multi-tenancy (Choudhary 2007), while traditional software firms have to support and maintain dozens of old versions. As a result, leading traditional software firms may be required to undertake more R&D investments than SaaS firms for maintenance of legacy systems. Therefore, we hypothesize that,

H4. *Relative to traditional software companies, SaaS firms with higher productivity tend to spend less in R&D investments.*

2.3. Method and Model

2.3.1. Production Function

A production function describes the mathematical relationship between the input factors and the output of a firm, an industry, or an entire economy. Typically, the input factors consist of capital, labor and other tangible or intangible assets. Due to its mathematical properties, the Cobb-Douglas function is one of the most widely adopted production functions which satisfies all textbook assumptions. The most frequently used Cobb-Douglas production function with two inputs, capital (K) and labor (L), and one output (Y) is given by:

$$Y_{it} = A_t K_{it}^{\beta_K} L_{it}^{\beta_L}, \tag{1}$$

where Y_{it} denotes the output of the *i-th* firm at the *t-th* period. Here, A is a scale factor defined as Total Factor Productivity in the literature; K_{it} and L_{it} represent the capital input and labor input of the *i-th* firm at the *t-th* period. After taking the logarithms, it follows that:

$$\ln(Y_{it}) = \ln(A_t) + \beta_K \ln(K_{it}) + \beta_L \ln(L_{it}). \tag{2}$$

There are three useful properties. First, this functional form fits the linear regression estimation approach. Second, in this expression, β_K and β_L represent the output elasticities of capital and labor, respectively, which measure the percentage change in output after a one-percent increase in the corresponding input. For example, the output elasticity of capital, β_K, represents the percentage increase in output provided by a 1% increase in capital. Third, the sum of β_K and β_L can be defined as the economies of scale. To illustrate this, if sum of β_K and β_L equals X and all inputs in (1) are multiplied by N, the output will increase by N^X in (1). Consequently, output increases more than N times if and only if $X > 1$. In this study, we define economies of scale as follows: *A production function exhibits economies of scale if its output increases more than N times when all inputs increase N times.*

The literature on IT productivity has examined this logarithm expression by various regression methodologies. Since the seminal work of Brynjolfsson and Hitt (1996), Information Systems researchers have used similar approaches to study IT productivity (Aral et al. 2012b; Dewan and Ren 2007; Dewan and Ren 2011; Han et al. 2011a; Han et al. 2011b; Hitt et al. 2002; Kudyba and Diwan 2002; Park et al. 2007b; Tambe and Hitt 2012b).

Our study is unique in that we focus on contrasting the productivity of SaaS and non-SaaS software firms, whereas the majority of the literature focuses on the productivity of IT capital and IT labor, relative to non-IT capital and labor.

2.3.2. Stochastic Frontier Approach

In practice, given the same inputs, different firms may deliver different amounts of output. This deviation could result from random error (noise), or from the differences in production efficiency of the target firms. In economics, a production frontier is defined as the maximum output that can be achieved given certain inputs by the most efficient production technology at that scale of inputs. In other words, production frontier describes the production function of the most efficient firms of different sizes. Economists also use production functions to describe the mathematical properties of the production frontier.

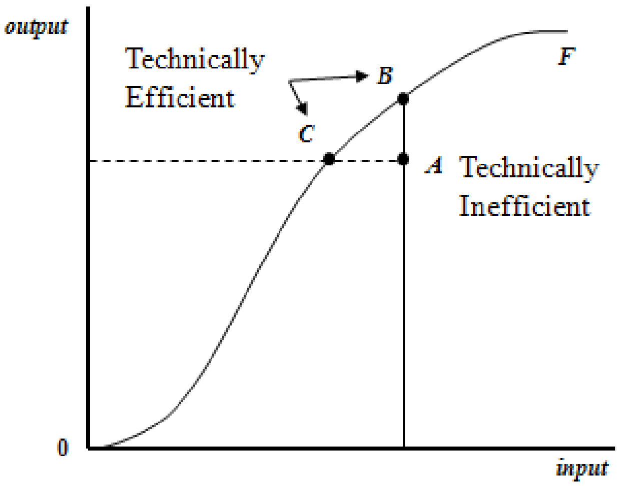


Figure 1 - 2. Production Frontier and Technical Efficiency

Designed to estimate the production frontier, SFA was developed independently by Aigner et al. (1977), and Meeusen and Van den Broeck (1977). If we ignore the random error, firms operate either on or beneath the frontier according to SFA (see Figure 1-2). They cannot operate above the frontier. Therefore, the shortage of output between the production frontier and a firm's actual output will be attributed to production inefficiency (line AB).

Formally, a Cobb-Douglas production frontier estimated by SFA is given as:

$$\ln(Y_{it}) = \beta_0 + \beta_K \ln(K_{it}) + \beta_L \ln(L_{it}) + V_{it} - U_{it}, \tag{3}$$

where V_{it} is a random variable that accounts for measurement error and other random factors. It is assumed to be normally distributed with variance σ_v and can be either positive or negative. Here, U_{it} is a *non-negative* random variable with variance σ_u. It is assumed to be independently distributed and represents production loss due to firm-specific technical inefficiency. Thus, it is not less than zero. More details about the definitions of parameters can be found in the works of Battese and Coelli (1995) and Greene (2011).

In empirical productivity analysis, there are three major approaches for estimating the production frontier: linear regression, Stochastic Frontier Analysis, and Data Envelopment Analysis (DEA). The most common approach used in the IS literature is linear regression. In this case, Equation (3) without U_{it} is estimated by various panel-data econometrics methods and all deviations from the Cobb-Douglas function are considered as “noise”. Therefore, one advantage of SFA over the linear regression approach is that it separates the deviations into random errors and the firm-specific inefficiency (U_{it}). With the consideration of U_{it}, SFA can provide more accurate estimation results regarding the production frontier. In contrast, DEA uses the convex hull of all data samples to estimate the production frontier. In other words, DEA assumes that all observed deviations from the production frontier are considered as firm-specific inefficiency. Interested readers can refer to the work of Banker et al. (1984) for details.

Using Equation (3), the Technical Efficiency score for the *i-th* firm at time *t* is:

$$TE_{it} = \frac{Y_{it}}{\max[Y_{it}]} = \exp(-U_{it}). \tag{4}$$

The TE score is an important measure of a firm's productivity performance. It gauges the percentage of output for the target firm divided by that of the most efficient firm. Therefore, TE is a percentage and is smaller when the target firm is less efficient.

There are few studies that use SFA in the IS literature, in sharp contrast to SFA's popularity in the economics literature. Lin and Shao (2000) use SFA to investigate the business value of IT at the firm level. Shao and Lin (2002) found strong statistical evidence to confirm that IT exerts a significant favorable impact on technical efficiency and, in turn, gives rise to the productivity growth. Li et al. (2010) used a stochastic frontier production function to measure the capability for each software firm in each time period by calculating their technical efficiency.

2.3.3. Stochastic Frontier Production Frontier

We follow the standard procedure in the literature to conduct a two-stage SFA. In the first stage, we apply SFA with the Cobb-Douglas production function to estimate the production frontier and the Technical Efficiency (TE) of each software firm in each year. In the second stage, we conduct regression analysis using TE as the dependent variable. The independent variables are R&D

expenses and advertising expenses (e.g., Li et al. 2010).

The equation for SFA is specified as follows:

$$\ln(Y_{it}) = \beta_0 + \beta_K \ln(K_{it}) + \beta_L \ln(L_{it}) + \sum_{t=2002}^{2009} \beta_t Year_t + V_{it} - U_{it}, \qquad (5)$$

where Y_{it} is the economic value added of firm i in year t; K_{it} and L_{it} are the capital and labor of firm i in year t; and $Year_t$ is a set of year dummies. Equation (5) is estimated for SaaS and traditional software firms separately. We also estimate (5) by OLS as robustness checks:

$$\ln(Y_{it}) = \beta_0 + \beta_K \ln(K_{it}) + \beta_L \ln(L_{it}) + \sum_{t=2002}^{2009} \beta_t Year_t + V_{it}. \qquad (6)$$

2.3.4. Technical Efficiency Model

In the second stage, potential causes such as R&D investment and advertising expense are included to regress upon the TE scores (Li et al. 2010). Formally, we estimate the following models,

Efficiency Model I:

$$TE_{it} = \delta_0 + \delta_1(RD_{it}) + \delta_2(AD_{it}) + \delta_3(FSIZE_{it}) + \sum_{t=2002}^{2009} \beta_t Year + \varepsilon_{it}, \qquad (7a)$$

Efficiency Model II:

$$TEG_{it} = \delta_0 + \delta_1(RD_{i,t-1}) + \delta_2(AD_{it}) + \delta_3(FSIZE_{it}) + \sum_{t=2002}^{2009} \beta_t Year + \varepsilon_{it}. \qquad (7b)$$

where TE_{it} is the TE score obtained from (4) and (5); TEG_{it} is the growth of TE score defined as ($TE_{i,t+1}$ - TE_{it}) / TE_{it}; $RD_{i,t}$ is the R&D investment of firm i in year t; AD_{it} is the advertising expense of firm i in year t; and $FSIZE_{it}$ is the firm size of firm i in year t, and ε_{it} is the error term. Firm size and year dummies are used as control variables in the literature and are included in our second-stage analysis (Knott and Posen 2005; Li et al. 2010).

The results of (7a) shed light on the strategies of efficiency leaders in the software industry. Specifically, it shows whether an efficient software firm spends more on R&D and advertising expenses in the same year. In contrast, (7b) examines: when a firm spends more on R&D or advertising, how does efficiency score change in the next year? We use $RD_{i,t-1}$ because R&D activities typically have impacts in the future years (Bardhan et al. 2013).

Sales and marketing expenses are included because they typically account for the single largest expense of software firms. For software firms, these costs are typically as high as 50% of the total revenue. There is ample evidence in the literature that suggests a positive relationship between advertising expense and firm performance (Altinkemer et al. 2011). Because the SaaS business model is new and the target market segment is small and medium firms (Demirkan et al. 2010), even large SaaS vendors may need to invest a significant amount of capital to educate potential corporate buyers, build brand awareness, and create new sales leads. For example, although Salesforce.com expected its revenue to rise in 2012, its overall profitability remains in the red due to increased sales and marketing costs. In its 2012 annual report, Salesforce.com stated: "*we expect marketing and sales costs, which were 52 percent of our total revenues for fiscal 2012 and 48 percent for the same period a year ago, to continue to represent a substantial portion of total* *revenues in the future as we seek to add and manage more paying subscribers, and build greater brand awareness"* (Salesforce.com 2011).

2.4. Data and Variables

The list of sample SaaS firms (full list in the appendix A1-1) was obtained from industry reports of the Software Equity Group[4], a consulting company. Traditional software firms are defined as publicly listed firms with NAICS code of 511210 (the software publisher) excluding SaaS firms. By this definition, we are forced to leave out some famous firms, such as IBM and Amazon, part of whose businesses are based on the SaaS model. In other words, our definition of SaaS is relatively strict: those firms are SaaS-only firms. Firms that have some SaaS products are still categorized as traditional software firms.

Financial data was obtained from Compustat. Samples with missing values in input or output variables were dropped. The data comprises an unbalanced panel of 23 SaaS firms and 480 traditional software firms from 2002 to 2010 with 135 and 2315 data points, respectively. The beginning year is 2002, because the first two pure-SaaS firms went public in 2002. The ending sample year is 2010, the most recent year with complete Compustat data.

---

4 http://www.softwareequity.com/research_annual_reports.aspx

2.4.1. Dependent Variable

The standard output measure used in the literature is economic value added, defined as the additional value of the final product over the cost of input materials used to produce it from the previous stage of production (Brynjolfsson and Hitt 1996; Dewan and Min 1997; Kudyba and Diwan 2002). We use the same definition from the literature (Kudyba and Diwan 2002): output, i.e., value added, is operationalized as the total annual sales minus the cost of goods sold (COGS) with total sales deflated by Producer Price Index (PPI) in the software industry and COGS deflated by PPI for intermediate goods.

2.4.2. Independent Variables

In the first stage of SFA, "capital" is operationalized as "total fixed assets," while "labor" is operationalized as "the number of employees." Both variables are standard input factors commonly used in the productivity literature.

In the second stage, we define R&D intensity as the R&D expense divided by total revenue. We define the advertising intensity as the advertising expense divided by total revenue. Firm size is operationalized as the natural logarithm of total asset. Following the literature (Dewan and Ren 2011; Lieberman et al. 1990), all variables are deflated to measure the real but not nominal values. A summary of the construction process and deflator for key variables is provided in Table 1-2. The deflated descriptive statistics are reported in Table 1-3. Please refer to Table A1-2 in the appendix for the full correlations of variables.

Table 1 - 2. Data Construction Procedure and Deflators[5]			
Variable	**Notation**	**Measurement Construction Process**	**Deflator**
Value Added (Output)	Y	Total sales (revt) minus cost of goods Sold (cogs), converted to constant 2002 dollars	Producer Price Index for software (NAICS code = 511210) (Bureau of Labor Statistics 2010)
Capital	K	Total assets (at) minus (total current assets (act) and intangible asset (intan)), converted to 2002 dollars	Producer Price Index for Intermediate Materials, Supplies and Components (Bureau of Labor Statistics 2010)
Labor	L	Total number of employees (emp)	N/A
R&D investments	RD	R&D expense (xrd) divided by total sales (sale)	Producer Price Index for Intermediate Materials, Supplies and Components (Bureau of Labor Statistics 2010)
Advertising	AD	Advertising expense (xad) divided by total sales (sale)	Producer Price Index for Intermediate Materials, Supplies and Components (Bureau of Labor Statistics 2010)
Firm size	FSIZE	Natural logarithm of total assets (at)	Producer Price Index for Intermediate Materials, Supplies and Components (Bureau of Labor Statistics 2010)

Table 1 - 3. Descriptive Statistics

	SaaS firms					Traditional software firms				
Variable	N	Mean	Std. Dev.	Min	Max	N	Mean	Std. Dev.	Min	Max
Value-Added (ln)	135	4.185	0.960	1.845	7.345	2315	3.683	2.131	-5.298	10.968
Capital (ln)	135	2.448	1.253	-0.372	6.921	2315	1.973	2.435	-7.269	9.764
Labor (ln)	135	-0.662	0.816	-2.847	1.669	2315	-1.102	1.812	-6.908	4.682
R&D	130	0.160	0.085	0.033	0.488	2190	0.302	1.372	0	39.102
Advertising	115	0.036	0.059	0.000	0.368	1401	0.033	0.177	-0.036	6.091
Firm Size	135	4.782	1.134	1.988	8.134	2315	4.3943	2.281	-4.304	11.472
TE*	135	0.715	0.164	0.323	0.959	2315	0.509	0.216	0.002	0.969

*TE are generated from SFA in Table 1-4

2.5. Analysis and Discussions

2.5.1. SFA for Economies of Scale

The estimation results of SFA in the first stage are summarized in Table 1-4.

The analysis is applied to the SaaS group and the traditional group separately,

[5] Names after variables in the parentheses are the variable names in Compustat.

since they may have different production functions, as well as different frontiers. However, in general, the Cobb-Douglas production function could be used for both groups. Main productivity coefficients (β_K and β_L) are significant at the 1% level for both SaaS and traditional software firms.

Table 1 - 4. Estimation Results of Stochastic Frontier Analysis[6]

Independent Variable	SaaS firms			Traditional software firms		
	Coefficient	Std. Err.	P-value	Coefficient	Std. Err.	P-value
β_K	0.217***	0.051	0.000	0.151***	0.014	0.000
β_L	0.672***	0.073	0.000	0.813***	0.019	0.000

Note: *** p<0.01, ** p<0.05, * p<0.1 All intercept estimates are omitted for brevity.

Based on Table 1-4, we can calculate the economies of scale of the two groups, defined as the sum of the beta coefficients of the two input factors. Consistent with H1, the sum of two beta coefficients of SaaS firms is smaller than one (0.889 versus 1), and smaller than that of the traditional software firms (0.889 versus 0.964). From the formal tests (Wald tests of joint significance) (Morley 2006; Temple 2001) reported in Table 1-5, we conclude that SaaS firms exhibit diseconomies of scale and smaller economies of scale than traditional software firms. The two null hypotheses are listed in Table 1-5. As a consequence, *Hypothesis 1a and 1b are both supported*.

Table 1 - 5. Wald Test for Economies of Scale in SaaS firms

Null hypothesis (H_0)	χ^2	**P-value**	**Conclusion**
(SaaS) $\beta_K + \beta_L >= 1$	7.01	0.0081	H_0 is rejected at 1% confidence level.
(SaaS) $\beta_K + \beta_L >= 0.964$	3.20	0.0734	H_0 is rejected at 10% confidence level.

6 Year dummies are included in the estimation but not reported in the Table for brevity. All the following estimations in this study include year dummies. The year dummies of both groups in Table 4 have been increasing in recent years, indicating that the total factor productivity has been growing in the software industry.

As robustness checks, we also estimate the Cobb-Douglas production function by four commonly used linear regressions: (1) fixed-effect panel regression (FE), (2) random-effect panel regression (RE), (3) panel linear regression with panel-corrected standard errors (PCSE), and (4) panel linear regression with AR1 errors (AR1). Results are in Table 1-6. The sum of these two input factors is consistently smaller in SaaS firms than in traditional software firms in all cases except the fixed-effect model. However, β_K of SaaS firms in the fixed-effect model is not significant and hence is not comparable.[7] Therefore, H1 is supported in our robustness checks.

Table 1 - 6. Estimation Results of Fixed-Effect Linear Regressions

	SaaS Firms				Traditional Software Firms			
	(1) FE	(2) RE	(3) PCSE	(4) AR1	(1) FE	(2) RE	(3) PCSE	(4) AR1
β_K	0.061	0.252***	0.399***	0.289***	0.083***	0.139***	0.217***	0.205***
	(0.058)	(0.051)	(0.047)	(0.053)	(0.016)	(0.015)	(0.022)	(0.028)
β_L	0.750***	0.611***	0.452***	0.566***	0.611***	0.809***	0.830***	0.821***
	(0.114)	(0.083)	(0.078)	(0.102)	(0.027)	(0.022)	(0.031)	(0.041)
$\beta_K + \beta_L$	0.811	0.863	0.851	0.855	0.694	0.948	1.047	1.027
Overall R^2	0.822	0.879	0.890	0.799	0.877	0.883	0.885	0.803

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1 All intercept estimates are omitted for brevity.

Considering that R&D activity is important in the production of software firms, we also included R&D investment (measured as absolute amount, differs from Table 2) as an input of production function as a robustness of our baseline analysis. Results are reported in Table 1-7. The sum of β_K, β_L and β_{RD} is smaller for SaaS firms, relative to traditional firms (except the FE model, in which β_K is not significant). In Table 1-6 and Table 1-7, β_K of SaaS firms is

[7] Indeed, Ackerberg et al. (2007) pointed out that fixed effect estimations usually resulted in unreasonably low estimate of capital coefficient for firm production function.

consistently larger than that of traditional firms, while β_L of SaaS firms is consistently smaller. Further, the contribution of R&D in SaaS firms is consistently smaller except the FE model.

Table 1 - 7. Robustness Check with R&D Included as An Input

	SaaS Firms				Traditional Software Firms			
	(1) RE	(2) FE	(3) PCSE	(4) AR1	(1) RE	(2) FE	(3) PCSE	(4) AR1
β_K	0.184***	0.079	0.329***	0.214***	0.093***	0.059***	0.124***	0.130***
	(0.048)	(0.054)	(0.045)	(0.043)	(0.015)	(0.016)	(0.019)	(0.024)
β_L	0.453***	0.375**	0.418***	0.450***	0.599***	0.440***	0.605***	0.633***
	(0.085)	(0.147)	(0.071)	(0.085)	(0.025)	(0.031)	(0.032)	(0.044)
β_{RD}	0.311***	0.398***	0.157***	0.279***	0.319***	0.247***	0.349***	0.310***
	(0.065)	(0.095)	(0.050)	(0.056)	(0.021)	(0.025)	(0.020)	(0.027)
$\beta_K + \beta_L + \beta_{RD}$	0.948	0.852	0.904	0.943	1.011	0.746	1.078	1.073
Overall R^2	0.889	0.844	0.902	0.869	0.912	0.909	0.913	0.912

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1 All intercept estimates are omitted for brevity.

2.5.2. Output Elasticities of Input Factors

This section examines H2 and H3. Recall that β_K and β_L indicate the marginal contribution to productivity of the input factors. Specifically, they measure the percentage change in the output once the input is increased by one percent, i.e., the output elasticity of that input factor. Results of the Wald Test are reported in Table 1-8.

Table 1 - 8. Wald Test for Comparing Elasticities

Null hypothesis (H_0)	χ^2	**P-value**	**Conclusion**
(SaaS) $\beta_K <= 0.151$	1.69	0.1934	H2 is not supported
(SaaS) $\beta_L >= 0.813$	3.72	0.0537	H3 is supported

Capital Productivity: SaaS firms have an insignificantly larger coefficient than that of traditional software firms (p-value is 0.1934). *Hypothesis 2 is not supported*. Rigorously speaking, this test is inconclusive due to the small sample size of SaaS firms. However, this result[8] is consistent with the conjecture that capital in traditional software firms contributes less to its output, because fixed assets in traditional software firms, such as computers or buildings, are used to support R&D and back office operations. In contrast, a significant proportion of the fixed assets of SaaS firms are directly used for delivering SaaS services (e.g., data centers). The productivity difference is also economically significant[9]: the productivity of the capital at SaaS firms is 1.44 times (0.217 divided by 0.151) larger than that of non-SaaS firms. For every 1% increase in capital, there is a 0.217% and 0.151% increase in economic value added for SaaS firms and traditional software firms, respectively. In dollar amount, every $1 investment in capital contributes to $1.207 and $0.835 (i.e. marginal products) value-added in SaaS firms and non-SaaS firms, respectively.[10]

Labor Productivity: Our results suggest that the output elasticity for employees in SaaS firms is much lower than in traditional software firms. Specifically, a 1% increase in employees leads to a 0.672% and 0.813%

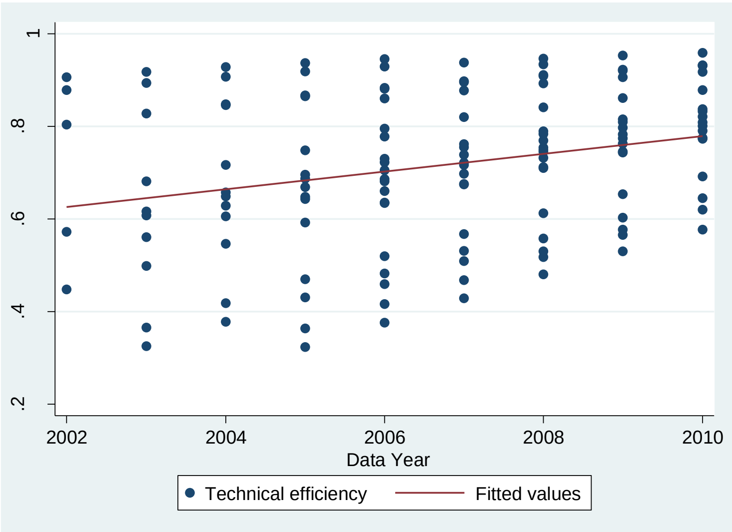
[8] Although the result in Table 1-8 is only significant at a 20% significance level, Table 1-7 does show that the coefficient of capital of SaaS firms is significantly larger than that of traditional software firms. The robustness checks with R&D as an input actually support H2.

[9] A coefficient is economically significant when it has a significant influence on the amount of the dependent variable. A statistically significant but economically insignificant coefficient does not really influence the amount of the dependent variable.

[10] The calculation is derived by the following procedure for an average SaaS firm. In Table 3, the average value added is exp(4.081) and the average capital is exp(2.365). So 1% increase in capital is equivalent to 0.01 * exp(2.365) increase in capital, which produces 0.22% * exp(4.081) value added. Dividing 0.22% * exp(4.081) by 0.01 * exp(2.365) leads to 1.207. The number for non-SaaS can be derived by the same procedure.

increase in value added for SaaS firms and traditional software firms, respectively. This difference is both statistically and economically significant. Labor productivity at SaaS firms is only 82.66% of that of non-SaaS firms (0.672% versus 0.813%). In dollar amount, one additional employee leads to $77,131 and $ 97,317 (i.e., marginal products) in value added of SaaS firms and traditional software firms, respectively. In this case, the labor productivity of SaaS firms is 79.26% of that of non-SaaS firms. *Hypothesis 3 is supported*.

2.5.3. Do efficient firms invest more in R&D or advertising?



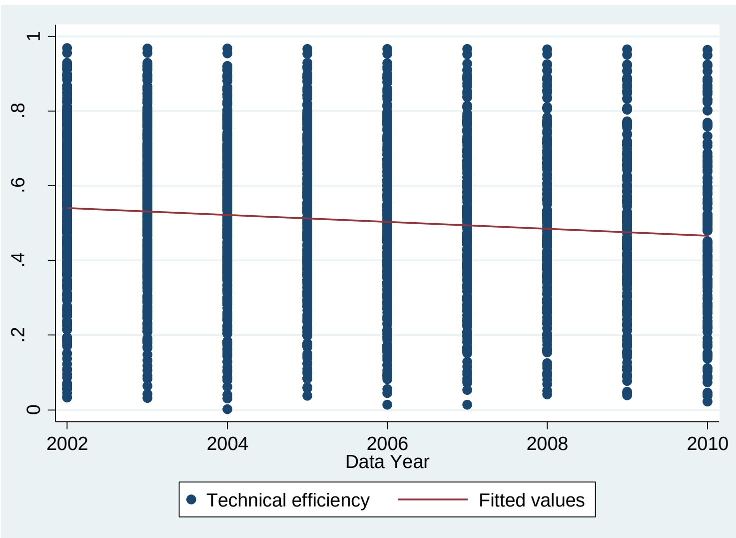


Figure 1 - 3. TE over years of SaaS firms (top) and Traditional Software firms (bottom)

Figure 1-3 depicts the scatter plots of TE scores over years for SaaS and traditional software firms. The average TE scores of SaaS firms have been increasing over time, and the differences in performance among SaaS firms have been shrinking. In contrast, the average TE scores of non-SaaS firms have been decreasing recently. A possible explanation for this observation is that the surge of interests in SaaS and cloud computing has shifted market share from traditional firms to SaaS firms. In Sections 2.5.3 and 2.5.4, we will examine the correlation among TE score, R&D expenses, and advertising expenses.

Table 1 - 9. Estimation Results of Efficiency Model I

	SaaS Firms					Traditional Software Firms				
	(1) FE	(2) FE-R	(3) RE	(4) RE	(5) RE	(1) FE	(2) FE-R	(3) RE	(4) RE	(5) RE
δ_1	-0.430***	-0.430**	-0.417***	-0.475***		0.006***	0.006***	0.006***	0.003***	
(R&D)	(0.085)	(0.187)	(0.089)	(0.076)		(0.001)	(0.002)	(0.002)	(0.001)	
δ_2	-0.115	-0.115	-0.008		-0.412**	0.021***	0.021	0.021***		0.018***
(Advertising)	(0.215)	(0.266)	(0.211)		(0.208)	(0.007)	(0.014)	(0.007)		(0.007)
δ_3	-0.004	-0.004	0.004	-0.003	0.005	0.001***	0.001**	0.001***	0.001***	0.001***
(Firm Size)	(0.007)	(0.010)	(0.008)	(0.006)	(0.007)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
N	110	110	110	130	115	1349	1349	1349	2190	1401
Within R^2 [11]	0.881	0.881	0.880	0.889	0.844	0.909	0.909	0.909	0.926	0.910
Overall R^2	0.065	0.065	0.097	0.093	0.058	0.039	0.039	0.041	0.036	0.039

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1 All intercept estimates are omitted for brevity.

We estimate equation (7a) by five models. The results are reported in Table 1-9. Model 1 is estimated using fixed effect panel regression (FE); Model 2 is estimated by fixed effect with robust standard errors (FE-R); and

[11] Within R^2 is the proportion of variability that is explained by the model within the group of each firm.

Model 3 is estimated by the random effect model (RE). Based on Model 3, Model 4 is estimated without advertising expense and Model 5 is estimated without R&D investment.

SaaS Firms: The results in Table 1-9 suggest that R&D intensity is negatively correlated with TE, whereas advertising expense is uncorrelated with TE in the same year. In other words, more efficient SaaS firms are associated with spending less on R&D investment while keeping the advertising investment at a consistent level. Even with a small sample of SaaS firms, we can show that this result is significant at the 1% level. The absolute value is also quite large: a 4.17% increase in TE score is correlated with a 10% decrease in R&D intensity.

At the same time, it is also possible that the negative correlation is caused by laggard firms being associated with spending more on R&D. To rule out this possibility, we run the same regressions on two subsamples of SaaS firms: (1) laggard firms in around bottom 30% in terms of TE scores and (2) leading firms in around top 30% in terms of TE scores. The results are shown in Table 1-10. The coefficient of R&D in Model (1) is not significant, indicating that laggard firms are not significantly correlated with less R&D. In contrast, the coefficient of R&D in Model (2) is significant and negative, showing that leading firms are correlated with significant less R&D. Therefore, the negative correlation of R&D is not caused by laggard firms but leading firms. Therefore, *Hypothesis 4 is supported*.

Table 1 - 10. Laggard and Leading SaaS Firms

	(1) Laggard	(2) Leading
δ_1	-0.024	-0.152**
(R&D)	(0.041)	(0.061)
N	37	37

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1 All intercept estimates are omitted for brevity.

Changes in R&D expense typically result from changes in R&D headcounts, stock-based compensation to R&D staffs, or third-party consulting fees. It is intuitive that a firm with high TE may have better performance and bestows larger stock-based compensation to its R&D staffs, which contradicts our finding. Therefore, the observed negative correlation between TE and R&D suggests that leading SaaS firms either recruit fewer R&D employees or spend less on third-party consulting fees.

Interestingly, advertising expense does not exhibit a similar pattern (Table 1-9), which provides one falsification test. In other words, results from the advertising expense suggest that the negative correlation between R&D and TE does not appear in all types of expenses.

Traditional Software Firms: Results in Table 1-9 show that efficient traditional software firms spend more on both R&D and advertising. This finding is statistically significant but with very small absolute values. It provides a good robustness check that our finding of SaaS firms is not a property of all software firms.

Overall, the results support our previous findings about the diseconomies of scale of SaaS firms. Leading traditional firms spend more on both R&D and advertising, whereas leading SaaS firms spend significantly less on R&D. Since R&D and advertising are the two most important expenses, reducing investment in R&D implies that SaaS firms expect a lower return for developing new products relative to their traditional counterparts. This is consistent with the decreasing return in R&D. Our results seem to suggest that one potential cause of the diseconomies of scale of SaaS firms could be the decreasing return in R&D.

2.5.4. How do R&D or advertising expenses affect technical efficiency?

Equation (7b) is estimated by the same five models in Section 2.5.3. There are two differences: the dependent variable is the growth of TE score and the R&D expense is from the previous year. We investigate how R&D and advertising contribute to the productivity growth. The results are summarized in Table 11.

SaaS Firms: Table 1-11 shows that R&D investment is positively correlated with TE growth, whereas advertising expense does not have significant impacts. Interestingly, these two coefficients have similar absolute values. In other words, our results imply that the absolute values of "*productivity return on investment* (ROI)" for R&D and adverting are similar, while the ROI for R&D is less volatile than for advertising. This finding is also economically significant: a 1% increase in R&D intensity is associated with a 0.123% increase in TE for the following year. This shows that R&D investment does pay off in terms of improving TE and confirms that R&D

indeed exhibited decreasing return. The reason is that efficient SaaS firms should be equipped with more resources to spend on R&D. If they do not spend more R&D, it is highly possible that R&D exhibited decreasing return.

Table 1 - 11. Estimation Results of Efficiency Model II

	SaaS Firms					Traditional Software Firms				
	(1) FE	(2) FE-R	(3) RE	(4) RE	(5) RE	(1) FE	(2) FE-R	(3) RE	(4) RE	(5) RE
δ_1	0.125***	0.125**	0.123***	0.140***		0.0005***	0.0005	0.0005***	0.0001	
(R&D 1-yr lag)	(0.028)	(0.046)	(0.028)	(0.025)		(0.0001)	(0.0003)	(0.0001)	(0.0001)	
δ_2	0.143*	0.143	0.0812		0.203***	0.0003	0.0003	0.0003		0.0008
(Advertising)	(0.081)	(0.084)	(0.072)		(0.005)	(0.0006)	(0.0012)	(0.0006)		(0.0007)
δ_3	0.0024	0.0024	-0.0002	0.0006	-0.0013	0.0001***	0.0001	0.0001***	0.0001***	0.0000
(Firm Size)	(0.0028)	(0.0031)	(0.0028)	(0.0025)	(0.0028)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
N	92	92	92	108	96	1079	1079	1079	1706	1118
Within R^2	0.783	0.783	0.778	0.792	0.721	0.695	0.695	0.694	0.706	0.661
Overall R^2	0.034	0.034	0.073	0.101	0.026	0.056	0.056	0.064	0.054	0.028

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1 All intercept estimates are omitted for brevity.

Traditional Software Firms: The coefficients in our results are statistically significant yet economically insignificant. It again provides a falsification check that our findings are applicable on SaaS firms, but not all software firms.

2.5.5. Robustness Check on Two Industries That Produce Non-Information Goods

Considering that software firms produce information goods, it is critical to know whether decreasing return of R&D also influence scale economies of the industries that produce non-information goods, We tested two industries: “electronic computers industry (ECI)” (NAICS 334111) and “computer storage devices (CSD)” (NAICS 334112) between 2002 and 2010. The sample includes 296 observations for 57 firms. We applied the same method by first running SFA and next regressing R&D intensity on the generated technical efficiency scores. The results of SFA are reported in Table 1-12. The economies of scale of ECI and CSD are significantly larger than that of SaaS firms by Wald Test. The results for equation (7a) are given in Table 1-13. R&D is not significantly related to technical efficiency, indicating that our finding on SaaS firms, does not apply to these two industries that produce non-information goods. This, again as a falsification check, supports our findings.

Table 1 - 12. Stochastic Frontier Analysis Results for ECI and CSD

Independent Variable	Coefficient	Std. Err.	P-value
β_K	0.141***	0.033	0.000
β_L	0.840***	0.057	0.000

*** p<0.01, ** p<0.05, * p<0.1. All intercept estimates are omitted for brevity.

Table 1 - 13. Second Stage Estimation Results of ECI and CSD

	(1) RE	(2) FE
δ_1	0.004	0.005
(R&D)	(0.007)	(0.007)

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1 All intercept estimates are omitted for brevity.

2.6. Conclusion

This study is the first attempt to use SFA to examine the economies of scale of SaaS firms as well as to contrast the productivity differences between SaaS firms and traditional software firms. Because SFA takes both inefficiency and random noise into account, it is generally considered as a better approach for productivity analysis. We first investigate the economies of scale of SaaS firms and then examine the potential sources of (dis)economies of scale.

Our results demonstrate diseconomies of scale in SaaS firms. This finding results from the fact that SaaS firms simultaneously sell a software application and offer the associated IT infrastructure management service for corporate clients. Software application is well accepted as a typical example enjoying both supply- and demand-side economies of scale. However, the IT infrastructure management service does not have zero variable costs, which reduces the supply-side economies of scale. As a result, the production function of SaaS firms has smaller economies of scale than that of traditional software firms. Our productivity analysis suggests that the input factor of capital contributes more to the output of SaaS firm than to that of traditional software firms. At the same time, labor contributes significantly less to the output of SaaS firms, which may be due to the difficulty in scaling the centralized IT infrastructure. Our results also indicate that the diseconomies of scale may result from the decreasing return in R&D investment of SaaS firms. More practical implications are discussed in the "General Conclusion" of the thesis.

There are several limitations to this study. First, we only have publicly available, firm-level data from Compustat. We are not able to examine small SaaS firms which are not publicly listed. Besides, we do not have detailed product-level numbers to calculate the exact productivity of SaaS production at the product level. Particularly, some firms provide both SaaS and non-SaaS software products. These firms are categorized as traditional software firms in this study. With product-level data, we may be able to examine the productivity of SaaS versus non-SaaS software production at a granular level. Furthermore, our sample size of SaaS firms is much smaller than that of the non-SaaS firms because the SaaS market is still in a nascent and morphing phase. As the market becomes more mature, researchers may observe more interesting results by conducting similar analysis on this dataset. Finally, since we only have firm-level financial data, we focus on production from the viewpoint of SaaS firms. We cannot analyze IT service productivity by considering the corporate buyers of SaaS vendors along with the vendors themselves. In other words, we underestimate the benefits of productivity improvement by the SaaS delivery model. Some benefits of the SaaS model may be realized only on the buyers due to competition, as it was in the traditional software industry documented in the literature (Hitt and Brynjolfsson 1996). Therefore, our analysis does not suggest that SaaS is not a beneficial technology and service innovation. Our findings only suggest that SaaS firms do not exhibit economies of scale in their operational performance.

Since SaaS is an emerging phenomenon, there exist several important future research directions. First, IT human capital and labor productivity at SaaS firms have not been thoroughly studied in the literature. Our study has identified significantly lower labor contribution at SaaS firms. Therefore, it would be interesting to further explore the antecedents of this finding at granular level. Similarly, it could be fruitful to analyze the productivity contribution of different types of employees (marketing versus R&D, IT versus non-IT). Unfortunately, we do not have access to this information.

Future research could also examine other benefits of the SaaS model, such as customization. On the client side, corporate buyers may require more customization as the SaaS market matures. However, the customization of service provided by SaaS firms may call for individualized design, whereas efficient production may call for more process standardization. This trade-off remains unexplored in the productivity literature. The impact of the centralized risks on the valuation of SaaS firms is another important issue. Since SaaS firms centralize the IT infrastructure and IT management, risks such as disruption or security problems also become centralized. For example, Salesforce.com had several outage events in the past, leaving thousands of businesses without access to their applications. These risks are not captured in the productivity analysis framework. Finally, SFA itself is a growing area in economics. Applying more advanced SFA tools to this dataset is another research direction.

3. STUDY 2: EXPLORING THE IMPACTS OF WORKER MOBILITY ON IT FIRM PRODUCTIVITY

3.1. Introduction

Leading (high-value) software firms make considerable investments towards employee training and human capital development. For instance, Oracle delivered 3.7 million training hours to its employees in FY2012,[12] while Microsoft offers around 2,000 employee training programs taught by instructors from leading educational institutions.[13] Google is known for its rigorous recruitment procedures and innovative human resource (HR) practices, such as allowing employees 20% time to devote to projects of their interest, that have often been cited as reasons for its success (Walker 2012). Another leader, Facebook, has a well-known Bootcamp program for software training and cultural indoctrination of new recruits where one of the activities involves pushing out live software updates to the company's billion plus users within a few days of joining.[14] In addition to the benefits from training, work experience acquired at these leading software firms can be valuable for employees. One way this occurs is through highly-qualified mentors that such firms provide to help entry-level engineers sharpen their technical skills and

---

12 http://www.oracle.com/us/corporate/citizenship/workforce/employee-development/index.html

13 http://careers.microsoft.com/careers/en/us/careerdevelopment.aspx

14 http://www.thestar.com/business/2012/04/18/a_look_inside_facebooks_bootcamp_for_new_employees.html

cultivate innovative ideas. In this study, we posit that the HR investments of these leading software firms not only enhance their own productivity but also benefit other IT firms (termed as recipients here) that recruit employees from these firms.

Our study builds on prior research which suggests that the flow of personnel across organizational boundaries is critical to organizational knowledge transfer (Argote and Ingram 2000), since it brings tacit knowledge and skills that add to the firm's knowledge stock (Madsen et al. 2003). For example, the acquisition of external IT labor often infuses the recipient firm with new technical know-how gained through their on-the-job training at the previous employer (Tambe and Hitt 2014a). Thus, recruitment could lead to spillovers, which occur when the recipient firm enjoys productivity gains due to the investments of the source firm. This is particularly relevant for the IT industry with its high human-capital and knowledge intensiveness (Boh et al. 2007; Bresnahan et al. 2002).

However, although recruitment has been recognized as a cost-effective way of acquiring knowledge from a source firm for a recipient (Rosenkopf and Almeida 2003), productivity spillover through labor mobility remains under-explored in the literature. Existing information systems (IS) studies have mainly examined productivity spillover from IT investment through other means such as IT-service and supplier-driven spillovers (e.g., Chang and Gurbaxani 2012a; Cheng and Nault 2007). Further, prior studies typically investigated the spillover impacts of IT investments across all industries (e.g., Chang and Gurbaxani 2012a; Cheng and Nault 2007). Yet, the spillover effects could be particularly salient for software firms, where around 80% of the productivity contribution comes from labor inputs (Huang and Wang 2009). Thus, productivity analysis on other sectors e.g., manufacturing companies, may not be generalizable to software firms. The lack of studies on labor productivity spillover could be attributed to the unavailability of relevant data till recently. As an exception, Tambe and Hitt (2014a) examined spillover through labor mobility due to the IT investment of the source firms. However, factors beyond IT investment, such as training and development of human capital, may contribute to the spillover effect through labor mobility. The objective of this study, therefore, is to model and empirically examine the sources of firm-level labor productivity spillover effects in the IT industry.

In this regard, the pioneering study by Tambe and Hitt (2014a) showed the existence of productivity spillover effects at the aggregate level from IT labor mobility. Our study adds to this literature by exploring which type of source firms and which type of IT workers may contribute greater spillover effects to the recipient IT firms. Particularly, it is important to understand how the spillover effects depend on the nature of the source firms and the type of human capital (e.g., educational level and experience) such that firms can draw implications for their HR practices. Here, we posit that leading software platform providers are the main sources of productivity spillovers through labor mobility in the IT industry. The IT industry, particularly the software industry, is different from other industries in terms of its "platform economics" (Gawer 2009). A few software giants control key technology platforms which serve as foundations upon which other firms build complementary products. In other words, many smaller software firms must develop their products based on the Application Programming Interfaces (APIs) provided by the technology platform owners and may benefit from hiring employees of these platform owners. Thus we propose that employees from the leading software platform providers may generate stronger spillover effects for the recipient than from other firms.

Even among the employees of these leading firms, we believe that those with higher education levels would generate greater spillover effects. Higher education e.g., in the form of postgraduate degrees, increases the ability of workers to acquire and employ specific knowledge (Hatch and Dyer 2004). In the IT industry which has considerable demand for highly educated workers[15], greater gains should be obtainable from more educated workers from leading firms who could have acquired specific knowledge more readily at the source firm and would then be able to apply it at the recipient firm. Additionally, longer work experience at the leading firms can allow their employees to avail greater benefits from their HR investments who should then be able to transfer the advantages to recipient firms that they move to. Thus, we explore the

[15] http://www9.georgetown.edu/grad/gppi/hpi/cew/pdfs/fullreport.pdf

effects of both educational level and work experience of workers on the productivity spillover due to labor mobility from the leading software firms.

After screening through all software firms, we chose Google, Microsoft, Oracle, and Facebook as the primary source firms to examine productivity spillover. These leading software firms were chosen because their market capitalization exceeds USD 100 billion in market value and importantly, each of them controls at least one major software platform. Throughout this study, these four firms are referred to as “leading software firms” or “leading firms”. To investigate the aforementioned spillover effects, we developed programs to scrape more than three million profiles from LinkedIn.com, the largest professional social networking site and online resume database in the world. We were able to construct a firm-level dataset with the 4 leading firms mentioned above as sources and 242 publicly listed U.S.-based IT recipient firms from 2002 to 2012. With the dataset, we conducted productivity analysis on the sample firms with four inputs i.e., labor, capital, and two unique variables that represent the HR spillover effects (via labor mobility) from the four leading firms and the other sample firms, respectively.

We find significant, robust productivity spillover effects resulting from hiring from the leading firms. Specifically, we show that hiring one employee from leading software platform providers is associated with USD $1.14 million increase in value-added of the recipient IT firms. In sharp contrast, hiring from non-leading firms on average does not produce significant spillover effects. Our results also show that among the employees recruited from the leading firms, those with longer work experience or higher education degree contribute to the spillover effect. Specifically, hiring workers with less than two years' experience or less than a Master's degree from leading software firms does not produce significant productivity spillover effects. Thus, our study shows that the spillover effect from labor mobility in the IT industry concentrates on well-educated and experienced employees from leading software platform providers. In this manner, this paper contributes to the literature by examining the sources of labor productivity spillover in the IT industry.

Additionally, the study uses a new publicly available dataset and contributes to IS research by applying novel econometrics models for productivity analysis. In the last decade or so, there has been a surge of interest in applying the structural approach in economics after the seminal paper by Berry et al. (1995). The counterpart method in productivity analysis is pioneered by Olley and Pakes (1996) (OP for short). OP solves the simultaneity issue that firm managers may adjust inputs after observing productivity shocks that are unobservable to researchers. If this issue exists, the widely used OLS estimators are shown to be biased, which was ignored by researchers till the introduction of the OP method. In the past decade, the OP method has been generalized in the economics literature (details will be discussed in Section 4). The state-of-the-art method along this line of studies is proposed by Wooldridge (2009). Wooldridge (2009) estimates a more general empirical model by joint Generalized Moment Methods (GMM) to simplify the complex procedure used in OP and the subsequent papers. To the best of our knowledge, our paper is an initial effort in the IS literature to employ Wooldridge (2009)'s method to conduct productivity analysis.

At the same time, our empirical study provides a rigorous analysis of "the value of employees" at inter-firm level. Such findings can be used in practice to guide HR executives in measuring the value of IT talent and to decide the appropriate salary and compensation package to recruit or retain extraordinary IT personnel such as those from the leading firms in our study. These findings could contribute to the emerging "People Analytics" approach in HR that has been pioneered by firms such as Google. For instance, Google executives have calculated the performance differential between exceptional and average technologists which could be as much as 300 times higher while the difference of salaries is much smaller (Sullivan 2013).

Our findings also have public policy implications. Specifically, governments should encourage free labor mobility because this can improve recipient firms' productivity and therefore the IT industry's total productivity. There is also an opportunity for governments to subsidize training efforts of leading firms such that the entire IT industry can benefit from these investments. From employees' viewpoint, our analysis reveals the surprisingly high "value-add" of work experience at major software platform providers, which can be a relevant but possibly less known factor for IT workers to consider when they search for and choose their jobs.

3.2. Literature and Research Background

Productivity is a measure of the efficiency of the production process (Coelli et al. 2005). Extant studies have proposed various econometrics methods to measure productivity at the firm-level. These methods have been applied to study the contribution of IT investments to firm productivity since the seminal study by Brynjolfsson and Hitt (1996). This paper seeks to examine the productivity spillover effects of hiring workers from leading software firms for the recipient IT firms. For this purpose, we review literature on productivity spillover effects from IT investments and labor mobility.

3.2.1. Two types of Productivity Spillovers

Productivity spillover is defined as the productivity gain of the recipient firm that results from the investment of the source firms. Two types of productivity spillovers, i.e., rent spillovers and knowledge spillovers, are differentiated in the literature (Chang and Gurbaxani 2012a). *Rent spillovers* are external benefits embodied in goods or services when they are purchased from other firms at a price lower than their full quality-adjusted price (Chang and Gurbaxani 2012a). For example, PC manufactures are able to produce faster computers through their IT investments. However, the manufacturers may not enjoy the benefits from the gain in product quality due to competition (Cheng and Nault 2007). Rather, buyers of PCs may benefit from the IT investments of PC manufacturers. Similarly, only the buyers, not the supplier, may benefit from the reduced transaction costs due to the implementation of a new supply chain management system when the supplier does not have the pricing power to charge customers for the improved service (Chang and Gurbaxani 2012a).

Knowledge spillovers occur when the accumulated knowledge at one firm can be transmitted to other firms due to its public good characteristics (Griliches 1979). For example, firms are increasingly outsourcing their information systems development to external service providers. Under such circumstances, client firms may benefit from the accumulated knowledge, such as business process innovation knowledge, possessed by the IT service providers (Chang and Gurbaxani 2012b). Relevant to our study, knowledge spillovers could also be induced by the mobility of IT workers. Firms can benefit from the knowledge acquired by incoming IT workers through their training at their previous employer (Tambe and Hitt 2014b).

3.2.2. IT Investment and Rent Spillovers

IS research in this area began with studying how IT investments may enhance a firm's productivity (Brynjolfsson and Hitt 1996). Building on this literature, researchers then began to examine how one firm's IT investment may affect the productivity of other firms, i.e., IT spillover effects. Empirical studies on IT productivity spillover have been conducted at three levels, i.e., country, industry, and firm.

At the country level, Park et al. (2007a) argued that IT investment in a country had a positive influence on the productivity of its import partner country when IT products are traded across borders, after controlling for openness, innovative capacity, and IT infrastructure. Using data collected from 39 developing and developed countries from 1992 to 2000, Park et al. (2007a) also showed that such country-level IT spillover occurs only when the source country is an IT-intensive or hi-tech export country. *At the industry level*, Cheng and Nault (2007) estimated the effects of upstream IT investments on downstream productivity and found substantial supplier-driven IT spillovers. These spillovers result from the fact that suppliers cannot fully capture the productivity improvement from their IT investments. The authors highlighted that a proper supplier output deflator would be needed to account for these IT spillovers in order to calculate the real input for the production of the downstream industries. Han et al. (2011a) also studied inter-industry IT spillover and found that IT investments made by supplier industries increased the productivity of downstream industries. Further, they examined the moderating effects of IT intensity and competitiveness of the downstream industry. They concluded that industries that are more IT intensive and competitive will benefit more from the IT spillovers. *At the firm level*, Chang and Gurbaxani (2012a) examined IT spillovers over a long-term horizon. Specifically, they found that firms with high IT intensity received greater and sustained spillover benefits from the IT service industry. In contrast, they found that the impact of IT spillover did not persist in low IT intensity firms regardless of the source.

The other stream of research on productivity spillover relates to knowledge spillovers of which *spillovers through labor mobility* are directly relevant to our study.

3.2.3. Labor Mobility and Knowledge Spillovers

The mobility of employees is an important mechanism for knowledge flows between firms. Both anecdotal and empirical evidence suggest that organizations have come to rely more on the acquisition of human assets from other organizations, as opposed to internal development and promotion, to meet their demands for talent (Somaya et al. 2008). This has brought about interest in studying the performance impacts of labor mobility. However, empirical studies on spillover due to labor mobility have been hampered by the lack of availability of relevant data, which has been alleviated very recently.

An exception is the Integrated Database for Labor Market Research provided by Statistics Denmark since 1980.[16] This dataset provides the complete data set on each individual employed in the recorded population of Danish firms and is suitable for studying labor mobility. Using this data set, Parrotta and Pozzoli (2012) examined the impact of recruiting technicians and

---

[16] http://www3.druid.dk/wp/20100016.pdf

highly educated workers on a firm's value-added in five industries.[17] Similar to Tambe and Hitt (2014a), Parrotta and Pozzoli (2012) show the existence of spillover effect from incoming labor mobility. However, their work does not examine which type of firms and which type of workers contribute more to the observed spillover effect. Moreover, the Danish dataset and the findings derived from it may not be generalizable to the IT industry in the US that forms the focus of our study. Denmark has a population of 5.6 million (about 55 times smaller than the US) and the major industries in Denmark include agricultural and dairy, transport, and pharmaceuticals (CIA 2013). Relevant to our study, there is lack of large software or e-commerce providers headquartered in Denmark and few IT companies in the country. On the other hand, the US is a leader in IT innovation and hosts 8 of the 14 largest IT companies globally (ranked by revenue in 2012).

Recent, alternative data sources that can provide information about labor mobility are online resume databases, such as Monster.com or LinkedIn.com. Tambe and Hitt (2014a) made use of such a database to analyze hundred-thousands of US-based IT workers' employment history. They found that by hiring IT workers, firms obtained significant productivity benefits from the IT investments of the source firms. Their results showed that 1% increase in the external IT employment pool would increase the productivity of recipient firms by around 0.018%. In our study, we use another data source -

[17] The five industries are: (1) manufacturing, (2) construction, (3) wholesale and retail trade, (4) transport and (5) financial and business activities.

LinkedIn, which is similar to their dataset in terms of the data variables available for empirical analysis.

Our paper differs from these related studies (Parrotta and Pozzoli 2012; Tambe and Hitt 2014a) in several significant ways. First, conceptually, the basis of productivity spillover in Tambe and Hitt (2014a) is IT investment and labor mobility is the carrier of the spillover effect. In our paper, the basis of productivity spillover includes all types of HR related investment of major software platform providers. Second, both existing studies are conducted across all industries whereas our study focuses on the IT industry. Third, by not using the IT investment variable in our study, we could avoid the data limitation faced by Tambe and Hitt (2014a) where IT investment data was available from 1987-1994 and IT employment data was used as an alternative measure from 1987- 2006. In the fast-changing IT industry earlier period results may not be generalizable to recent years because Google and Facebook both gained dominance after 2005. Fourth, and most importantly, even if their results could be applicable to the IT industry in recent years, their study does not examine the sources of spillover effect, which is the focus of our study. Our findings show that the spillover effect in the IT industry is heavily concentrated on well-educated, experienced employees from a few (in this case, four) major software platform providers. Last, Tambe and Hitt (2014a) use Arellano-Bond estimators (AB) as their main method with Levinsohn-Petrin (LP) estimator mentioned as a robustness check. We also contribute to the literature by using a more advanced econometrics model developed by Wooldridge (2009), which is a generalized version of LP.

3.3. Conceptual Model and Hypotheses

A conceptual model of IT firms' production with spillover effect is shown in Figure 2-1. The production inputs and output are illustrated in this figure. IT Firm *i* generates output using (1) capital, (2) labor, (3) spillovers by incoming HR flow from leading software firms, and (4) spillovers by incoming HR flow from other firms. We incorporate these two spillover terms as inputs, since knowledge is considered a critical input in production (Grant 1996). This setup is used in Parrotta and Pozzoli (2012) and is conceptually similar to Tambe and Hitt (2014a).

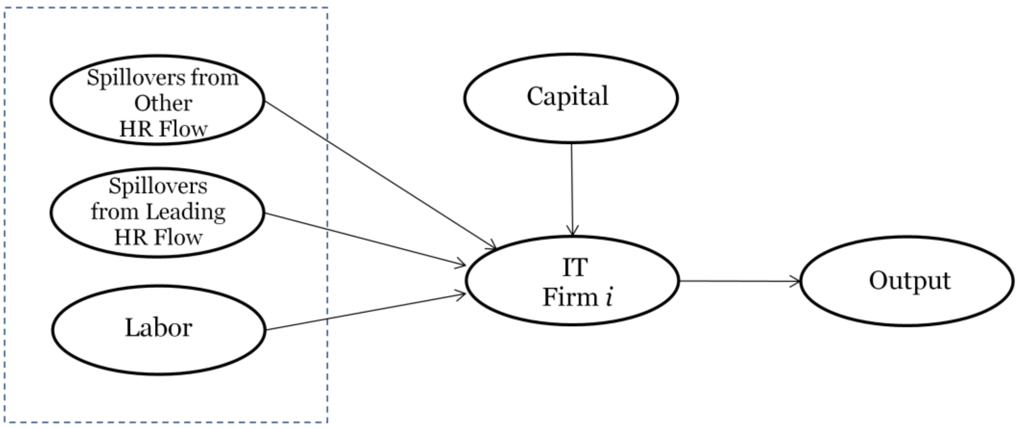


Figure 2 - 1. The Productivity Spillovers from Leading Software firms to the Recipient Recipient Firm

3.3.1. HR Flow and Productivity Spillover

In practice, few organizations internally create all the knowledge required for technological development (Song et al. 2003). Instead, they need to acquire the knowledge from external sources. A significant portion of the knowledge that organizations seek is embedded in individuals (Song et al. 2003). Thus, labor mobility is a key channel for such purpose (Tambe et al. 2012) as workers are able to transfer tacit as well as explicit knowledge when they move and adapt their knowledge to the new context (Argote and Ingram 2000). Such recruitment is referred to as a “learning-by-hiring” strategy (Song et al. 2003).

In the IT industry, leading software firms invest heavily in their human resources through employee hiring, training, and other human capital development activities, making them desirable sources for labor mobility. Taking Google as an example, only 4,000 to 6,000 people are hired from over one million resumes received every year (BusinessInsider 2011). In the early years of the firm, the two founders and the CEO themselves went through all the resumes of applicants (Oreskovic 2013). The rigorous hiring practices indicate Google’s HR investment in selecting talent. At the same time, the company also invests in various training programs for its elite employees (Walker 2012). Further, Google encourages its employees to work on their own ideas and innovate (He 2013), which has helped the firm develop several core offerings such as Gmail, Google News and Ad Sense (Kantrowitz 2013).

There are several ways through which labor mobility could lead to productivity spillovers from leading software firms to recipient firms in the IT industry. First, hiring ex-employees of leading firms can be a cost effective way for recipient IT firms to learn and adopt the new technologies and systems developed or used in leading software firms. Leading firms are usually at the frontier of new technology creation and adoption (Palomeras and Melero 2010). Therefore, when workers move from a leading software firm to the recipient firm, their technical know-how of these new technologies could be transferred to the new employer. The transfer of technical know-how through employees is more crucial for new technologies, since hands-on implementation experience is rare to find and becomes critical for their utilization (Tambe and Hitt 2014a). For example, hiring workers who are familiar with developing APIs at leading software firms (e.g., Microsoft) could help recipient IT firms better utilize these APIs (e.g., Windows API), to develop software and thereby increase the productivity of the recipient IT firms. Windows APIs are used by thousands of software firms to develop products for Windows OS, while Microsoft's Visual Basic and Oracle's Java[18] suite serve as the foundation for a large majority of software products.

Second, hiring workers from leading software firms may increase innovation activities in the recipient IT firm. Studies on technology management and innovation have found that labor mobility could fuel rapid innovation, with employees taking ideas gained in one organization and applying them to the next organization (Casper and Murray 2005). This is also observed in studies of labor mobility, where regions with high mobility saw

[18] Java is the most common required programming class for computing majors in universities partly because Java is a frequent job requirement in programming job recruitment advertisements.

firms citing each other's patents more often (Almeida and Kogut 1999). Particularly, *leading* organizations in high-tech industries are often the sources of a great deal of innovative ideas (Palomeras and Melero 2010). For example, Google is known for its people-management practices such as Google Café and Google Moderator that spur continuous innovation (Cope 2012). Thus, employees from leading software firms steeped in innovation activities are likely to continue to generate innovations that benefit recipient firms when they move.

Third, newly recruited workers from leading software firms may raise the productivity of the existing workers in the recipient IT firms. This can occur when the new employees transfer valuable knowledge acquired at the source firm to existing workers of the recipient firm. Indeed, Moretti (2004a) found that low-skilled workers benefited more from an increase in high-skilled workers in the firm than high-skilled workers. Similarly, Henderson (2007) noted that high-skilled workers raise the productivity of low-skilled workers. Particularly, workers accumulate both tacit and explicit knowledge through the observation of, or informal conversations with their colleagues (Nonaka 1994; Rosen 1972). Therefore, the lower skilled workers at the recipient IT firms could learn practices e.g., coding and debugging techniques, from their colleagues who previously worked for leading software firms. Consequently, the acquisition of HR flows from leading software firms could increase the productivity of the low skilled workers at the recipient IT firm which in turn increases the firm's productivity.

The above paragraphs discuss several ways in which labor mobility could have a positive impact on recipient firm's productivity. Further, the impact should be stronger for HR flow from leading software firms where the employees recruited are likely to have better knowhow of relevant new technologies and systems, spur innovation, and transfer their knowledge in part to lower-skilled workers, as compared to HR flow from other firms. Therefore, we hypothesize

H1a. *Spillovers resulting from the HR flow from leading software firms have a positive effect on the production output of the recipient IT firm.*

H1b. *Spillovers resulting from the HR flow from leading software firms are greater than those resulting from the HR flow from other firms.*

Moreover, prior research has classified knowledge as articulable (explicit) or tacit and suggested that professionals gain articulable knowledge through formal education, while tacit knowledge is acquired through on-the-job learning (Hitt et al. 2001). From a human capital perspective, the quality of the knowledge and skills embedded in HR inflows can thus be reflected in the educational level and work experience of the incoming employees. Accordingly, this perspective suggests that higher levels of education and work experience raise the productivity of workers (Becker 1962). Thus, in order to better understand the spillover effects due to different human capital quality of HR inflows, we further examine the impact of educational level and work experience of the inflow on recipient firm productivity.

3.3.2. Educational Level

Higher educational levels have been found to be correlated with greater extent of cognitive processing and problem-solving ability (e.g., Kimberly and Evanisko 1981). Indeed, the level of education is considered a proxy for employees' cognitive skills and motivational need for achievement (Hatch and Dyer 2004). A number of empirical studies have found that workers' education level impacts firm performance. Specific to the IT industry, Banker et al. (2008) found that the average education level of employees was positively associated with firm performance. Moretti (2004b) reported that better educated workers would make other workers more productive, which indicates potential synergy gains from acquiring such workers. At the same time, hiring well-educated employees increases the likelihood that they will combine and exchange their ideas to form new knowledge to benefit the firm (Smith et al. 2005). Further, higher levels of education (e.g., Masters and PhD) could *increase the ability of workers to acquire and employ specific knowledge* (Hatch and Dyer 2004). As a result, we expect that the education level of the HR inflow from leading software firms will be positively related to the productivity spillover at the recipient firm.

H2. *In general, the spillover effects from better educated HR flows from leading software firms are greater than those from less educated HR flows from leading software firms.*

3.3.3. Work Experience

A specific feature of an IT career is that after graduation, most professionals must continue to learn and gain knowledge through learning by doing (Pisano 1994). In the IT industry, technical competencies of workers are essentially skill-based competencies which are acquired through on-the-job experience (Ang et al. 2002). Thus, IT workers acquire tacit technical expertise through their work experience. Further, with the increasing standardization of hardware, software, and methodologies in the IT industry, workers' experience and tacit knowledge acquired at one firm could be applied to other firms (Mithas and Krishnan 2008). This is particularly true for leading software providers in our study whose platforms serve as foundations upon which other IT firms build complementary products. Additionally, on-the-job training is an investment that expedites the flow of tacit as well as articulable knowledge into the stock of human capital (Hatch and Dyer 2004). Thus, employees from leading software firms with greater work experience should be able to transfer their technical skills more readily to the recipient IT firms.

Last, more experienced workers are expected to have greater technical expertise and relevant knowledge to facilitate knowledge exchange and combination processes at the recipient firms (Smith et al. 2005). This is because experienced employees have more knowledge to recall and apply, and less bias in recalling and applying their knowledge as compared to novices (Lord and Maher 1990). Accordingly, employees from leading software firms with greater work experience should be more capable of transferring technical knowledge to the recipient firm and also more likely to generate innovative ideas or facilitate innovation activities in the recipient firm, as compared to less experienced workers. Consequently, the inflow of experienced workers should be more beneficial. Therefore, we hypothesize,

H3. *The spillover effects from HR flows from leading software firms are stronger if the work experience embedded in the HR flows from leading software firms is longer.*

3.4. Empirical Models

A production function describes the mathematical relationship between the input factors and the output of a firm, an industry, or a country. Typically, the input factors consist of capital, labor, and other tangible or intangible assets. Researchers have spent considerable efforts in improving methods for estimating the functional relationship between inputs and outputs. We refer the reader to Ackerberg et al. (2007) and Del Gatto et al. (2011) for recent surveys about using linear regression methods for productivity analysis. The present study utilizes a novel estimation method developed by Wooldridge (2009). To justify our use of this method, we will describe the evolution of econometrics methods for productivity analysis here.

In the literature, when researchers investigate the impacts of spillover effects from labor inflows, they consider three major inputs as the independent variables: capital (*K*), labor (*L*), and *one* spillover term (Parrotta and Pozzoli 2012; Tambe and Hitt 2014a). Since the focus of our study is the difference of spillover effects from leading software firms and other firms, we split the spillover term into *two* terms in our regression. Formally, we specify a Cobb-Douglas production function with one output (*Y*) and four inputs: capital (*K*), labor (*L*), spillover effect by labor inflow from leading software firms (*LHR*), and spillover effect by labor inflow from other firms (*nonLHR*).

$$Y_{it} = A_t K_{it}^{\beta_K} L_{it}^{\beta_L} LHR_{it}^{\beta_{LHR}} nonLHR_{it}^{\beta_{nonLHR}}, \tag{1}$$

where Y_{it} denotes the output of firm *i* in the period *t*. In the literature, A_t is defined as Total Factor Productivity (e.g., Mas-Collel et al. 1995). K_{it} and L_{it} represent the capital input and labor input of firm *i* in year *t*. LHR_{it} captures the productivity spillover effect from the leading firms to firm *i* in year *t*. It is measured by the number of leading software firms' former employees recruited by firm *i* in year *t*. Similarly, $nonLHR_{it}$ captures the productivity spillover effect from other firms to firm *i* in year *t*. After taking logarithms, it follows that:

$$\begin{aligned} \ln(Y_{it}) = \ln(A_t) + \beta_K \ln(K_{it}) + \beta_L \ln(L_{it}) + \beta_{LHR} \ln(LHR_{it}) \\ + \beta_{nonLHR} \ln(nonLHR_{it}) + e_{it}, \end{aligned} \tag{2}$$

where β_K, β_L, β_{LHR} and β_{nonLHR} represent the output elasticity of each input variable, respectively. Output elasticity measures the percentage change in the output from a one-percent increase in the corresponding input, while e_{it} is the error term. Specifically, β_{LHR} represents the productivity spillover effect of HR flows from leading software firms. If we consider only one spillover term in eq.(2), eq.(2) becomes the baseline regression in Parrotta and Pozzoli (2012) and is similar to Tambe and Hitt (2014a).

In the IS literature, Eq.(2) is usually estimated by panel regression models e.g., fixed effects or random effects, or extensions of panel regressions (e.g., Cheng and Nault 2012; Dewan and Kraemer 2000). However, panel regression methods do not solve the critical underlying econometrics issues identified in the economics literature. First, firm-level analysis may suffer from simultaneity issues. For example, when the manager of the focal firm observes factors (unobservable to researchers in the dataset) that positively affect the firm's productivity, the manager may adjust (typically increase) input factors accordingly. In our study context, employees may be attracted to firms that have better growth prospects or to firms that released positive growth information unobservable to the researchers in the dataset. Second, in productivity analysis, the capital variable and the spillover variable are unobservable by nature. The commonly used proxy variables for capital may suffer from potential measurement error, which is aggravated by fixed effect estimation (Ackerberg et al. 2007). Both issues can be handled by the newer methods discussed in the following paragraphs.

The earlier solution for the simultaneity issue is the Arellano and Bond estimator (Arellano and Bond 1991), which was later improved upon in

Blundell and Bond (1998). This approach has been employed by scholars to address the identification issue of the production function in IS research (e.g., Aral et al. 2012a; Tambe and Hitt 2012b). The idea of this method is to estimate the first-differenced eq. (2) on both sides of this regression equation. The estimation method is by applying GMM with lagged values of dependent variable and independent variables as the instrumental variables. The benefit of this approach is that the simultaneity issue can be alleviated; especially those effects that can be captured by the lagged values of the dependent variables.

The more recent solution of the simultaneity issue is proposed in the seminal paper by Olley and Pakes (1996) (OP for short). OP presents a two-step estimation method for the following regression function with underlying structural equation modeling setup.

$$\begin{aligned}\ln(Y_{it}) = \ln(A_t) + \beta_K \ln(K_{it}) + \beta_L \ln(L_{it}) + \beta_{LHR} \ln(LHR_{it}) \\ + \beta_{nonLHR} \ln(nonLHR_{it}) + v_{it} + e_{it},\end{aligned} \quad (3)$$

where $v_{it} + e_{it}$ is the only difference between eq. (2) and eq. (3). The unobservable productivity shock $v_{it} = I^{-1}(I_{it}, K_{it}) = g(I_{it}, K_{it})$ in eq. (3) explicitly models the simultaneity issue. In OP, v_{it} is allowed to be correlated with K_{it} and L_{it} and by assumption, the managers will maximize firm profit by adjusting I_{it} (investment) after observing the productivity shock, conditional on the value of K_{it}. A theoretical contribution of OP is to show by dynamic programming that v_{it} can be expressed as a monotone function of capital (K) and the "proxy variable" – investment (I_{it}). The introduction of I_{it} is the main novel solution to the simultaneity problem. In the first stage of estimation, semi-parametric methods are used to estimate the coefficients on L_{it}. In the second stage, the parameters on K_{it} and I_{it} can be identified by iterated optimization procedure. Other details can be found in Del Gatto et al. (2011).

Compared with other estimation methods, OP has several advantages. OP is more general cross-sectionally because v_{it} can be any function of the proxy variable and the capital. Along the time-series dimension, OP is more general than Arellano and Bond because OP also leaves the flexibility to the researchers to specify the stochastic process of v_{it}. OP also has a stronger theoretical foundation in production theory (Del Gatto et al. 2011, p.984) whereas Arellano and Bond estimator is a general-purpose estimator for dynamic panels.

Building on OP, Levinsohn and Petrin (2003) (LP for short) proposed a modification of OP by replacing I_{it} by the *intermediate input* (m_{it}) because the investment variable is very lumpy in real-world datasets. Ackerberg et al. (2006) (ACF for short) pointed out that both OP and LP estimators (especially LP) are not identified when the labor input is also included into $v_{it} = g(m_{it}, K_{it}, L_{it})$, a setup that is arguably more reasonable. ACF resolves the potential lack of identification by devising an enhanced two-step approach in which the first stage does not identify any coefficients of the input variables. However, ACF is difficult to implement due to its complexity.

Wooldridge (2009), our main estimation approach, simplified ACF by using joint-GMM estimation. Consistent with Wooldridge (2009), we estimate the following two equations by joint-GMM.

$$\begin{aligned}\ln(Y_{it}) &= \ln(A_t) + \beta_K \ln(K_{it}) + \beta_L \ln(L_{it}) + \beta_{LHR} \ln(LHR_{it}) \\ &\quad + \beta_{nonLHR} \ln(nonLHR_{it}) + g(\ln(m_{it}), \ln(K_{it}), \ln(L_{it})) + e_{it}.\end{aligned} \tag{4}$$

$$\begin{aligned}\ln(Y_{it}) &= \ln(A_t) + \beta_K \ln(K_{it}) + \beta_L \ln(L_{it}) + \beta_{LHR} \ln(LHR_{it}) \\ &\quad + \beta_{nonLHR} \ln(nonLHR_{it}) + f(g(\ln(m_{i,t-1}), \ln(K_{i,t-1}), \ln(L_{i,t-1}))) + u_{it}.\end{aligned} \tag{5}$$

where $v_{it} = f(v_{i,t-1}) + a_{it}$. $f()$ captures the time-series movement of the unobservable shock v_{it} and a_{it} is the error of this stochastic process. Following the notation in Wooldridge (2009), we define $u_{it} = a_{it} + e_{it}$, which is a noise term uncorrelated with any covariates. The moment conditions used to estimate coefficients of (4) and (5) are given by (6) and (7), respectively. Please refer to the appendix (see A2-1) for the additional mathematical details.

$$E(e_{it} \mid L_{it}, K_{it,} m_{it}, L_{i,t-1}, K_{i,t-1,} m_{i,t-1}, ..., L_{i1}, K_{i1,} m_{i1}) = 0. \tag{6}$$

$$E(u_{it} \mid K_{it,} L_{i,t-1}, K_{i,t-1,} m_{i,t-1}, ..., L_{i1}, K_{i1,} m_{i1}) = 0. \tag{7}$$

The advantages of Wooldridge (2009) over OP/LP/ACF include: (1) researchers can use existing Stata commands to implement the procedure without the need to develop customized programs in econometrics software; (2) it enables more efficient estimation than the two-step approach used by OP/LP/ACF because these methods ignore the contemporaneous correlation in the error terms of eq.(4) and eq.(5), and do not account for serial correlation or heteroskedasticity in the errors (Wooldridge 2009); (3) it provides insights about how labor may be affected by unobservable productivity shocks, while OP and LP are subject to identification problem as mentioned earlier in this section. The last issue is critical to our analysis because labor inputs are the focus of our study.

To test H2 and H3, we further divide LHR_{it}, and $nonLHR_{it}$ in (2) into several variables. For example, when testing H2, we further divide LHR_{it} and $nonLHR_{it}$ into four variables by the new recruits' highest degree (Bachelors/Masters/PhD/Other). Similarly, when testing H3, we also categorize LHR_{it} and $nonLHR_{it}$ into two ranges: less than 5 years, and 5 years and above.

3.5. Data and Variables

3.5.1. Sample

In this study, we utilized two datasets. First, firm-level accounting variables (Vale Added, Capital) and the official number of employees (Labor) are collected from the Compustat database (e.g., Yang et al. 2012). Second, the number of newly recruited employees from leading software firms (LHR) and other firms (nonLHR) is calculated from the public profiles scraped from LinkedIn.com by computer programs. The work experience (HighExp, LowExp) and education level (Bachelors, Masters, PhD, Other) information are also scraped from LinkedIn.com. Most users on LinkedIn.com post a public profile, which is very similar to a detailed resume that includes a brief bio, educational background, employment history, patents, skills, as well as other job related information. From the employment history, we could extract and aggregate the individual profiles to different employers. As a result, we could compile a firm-year panel data with firm-level financial variables from Compustat and the worker inflows information aggregated from the individual profiles scraped from LinkedIn.com.

As mentioned earlier, we selected Google, Microsoft, Facebook, and Oracle as the leading software firms based in the US that serve as sources of productivity spillover in our study. In this study, we define software industries by NAICS code 511210 (software publishers) and 519130 (Internet Publishing and Broadcasting and Web Search Portals).[19] In terms of value, these four are the only software companies with market capitalization greater than 100 billion US dollars in 2013. Further, all four firms' flagship products cover important technology platforms of O/S (Microsoft), search engine and maps (Google), social networking (Facebook), Java and database (Oracle), and enterprise software (Microsoft SharePoint and Oracle). Employees from these four companies are likely to be familiar with the APIs of these platforms. When these employees move to other software vendors that offer products built on the APIs, their work experience, training, and networking at the former employer may become important sources of productivity spillover for software R&D.

We collected more than 40 million LinkedIn profiles in the US. Next, we scrutinized the recipient firms of the outflows of the four leading software

[19] According to Wharton Research Data Services, the NAICS code for Microsoft and Oracle is 511210, while the NAICS code for Google and Facebook is 519130.

firms and selected firms in the top seven IT industry categories as our sample firms.[20] There are 8,802 individuals who left one of the four leading firms and subsequently joined one of the publicly listed IT firms in the top seven IT industry categories.[21] The remaining industries are either irrelevant to the software business, or have a very small number (less than 1% of the total outflow) of inflows from the four leading firms. The sample period for this study is from 2002 to 2012, where 2002 is chosen as the start year to avoid the influence of the Internet bubble. As a result, there are 2,968,155 LinkedIn profiles collected in the recipient firms of these seven industry categories during the sample period.

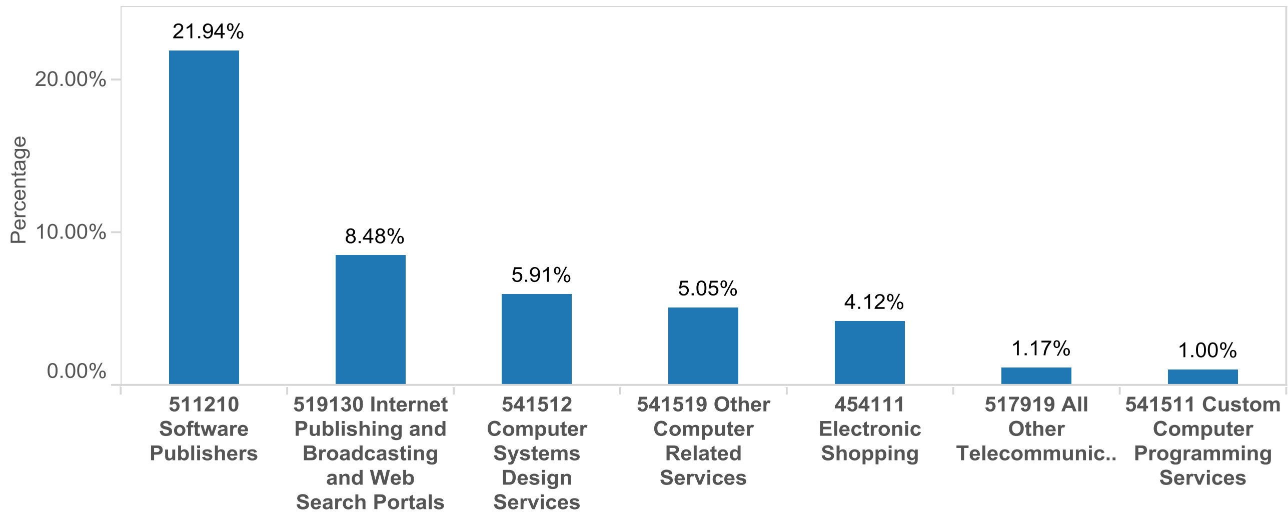


Figure 2 - 2. Recipient Industries – Top Seven IT Industries among Publicly Listed Firms.

[20] The NAICS codes for these 7 industries are: 454111, 511210, 517919, 519130, 541511, 541512, and 541519. The business of companies in these industries is related to software technologies. The sampling ratio is provided in Figure 2-2.

[21] In our dataset, around 30% of the employees who left the four leading firms would subsequently join a publicly listed firm that has accounting information in Compustat for productivity analysis. Of these, 47.66% of the outflows were to the IT firms in our sample.

The average sampling rate of the four leading firms is 42.6%. The sampling rate is calculated by the number of LinkedIn profiles divided by the official number of employees from Compustat. For recipient firms in the seven industry categories, we exclude firms with average sampling rate lower than 10% during the sample period. We also exclude firms that do not have any ex-employees of the four leading firms. Including these firms may lead to overestimated spillover effect because these firms' employees typically have worse performance. Therefore, the spillover effect (i.e., the additional productivity contribution of workers from leading software firms relative to other workers in the recipient firm) will be overestimated. In other words, even if those firms are included, our significant results will become even more significant. In our final sample, there are 242 firms and the average sampling rate for the recipient firms is 31%.

Although there exists potential sampling bias of LinkedIn users relative to the population of all IT workers, researchers seem to embrace the benefits from the high sampling rate at the inevitable costs of relatively mild sampling issues that could exist in all online resume databases. A number of recent studies utilized similar datasets from online resume repository or LinkedIn for their analysis (e.g., Agrawal and Tambe 2013; Arora and Nandkumar 2012; Garg and Telang 2012; Ge et al. 2014; Guillory and Hancock 2012; Milanov and Shepherd 2013; Tambe 2012; Tambe 2014). Particularly, Tambe (2014) found no systematic bias between workers who self-report their skills in their

LinkedIn profiles and workers who do not. This implies that LinkedIn users may not be systematically biased to those who are more willing to look for a job, because intuitively such workers will also be more likely to report their skills. The sampling bias is also alleviated because of our high sampling rate, which is much higher than studies that make use of surveys or cases. Also, our productivity analysis is based on the historical job turnover information aggregated to the firm-level number of employees. Indeed, Ge et al. (2014) found that LinkedIn profiles are much more accurate than patent data, which is usually used for labor mobility studies. There is no obvious reason to argue that the job *history* in our sample is not a representative sample of the population job history. Besides, the correlation between the recipient firms' official numbers of employees and their numbers of employees from LinkedIn is as high as 92%, suggesting that our sample closely tracked the firms' labor inputs. Last, our hypotheses are framed on a "relative basis". In H1b, H2 and H3, we compare two or more terms of spillover effects and the potential sampling bias effects (if any) could be canceled after we take the difference of two regression coefficients.

3.5.2. Dependent Variable

The standard output measure used in the literature is economic value added, which is defined as the additional value of the final product over the cost of input materials used to produce it from the previous stage of production (Brynjolfsson and Hitt 1996; Dewan and Min 1997; Kudyba and Diwan 2002).

We borrow the definition from the literature (Brynjolfsson and Hitt 1996): output Y (equivalently, value-added) is operationalized as the total annual sales (revt) minus the cost of goods sold (COGS) with total sales deflated by Producer Price Index (PPI) in the software industry and COGS deflated by PPI for intermediate goods. The benchmarking year for deflator is 2000.

3.5.3. Independent Variables

The independent variables in this study are capital, labor, and spillovers resulting from HR inflows. Both capital and labor are standard inputs in the productivity literature. The two spillover terms are measured using the natural logarithms of the estimated number of recruited workers from leading software firms or other firms in each year, respectively, as used in Tambe and Hitt (2014a) and Parrotta and Pozzoli (2012). The number of recruited workers from leading software firms is estimated by the number of new employees from leading software firms observed on LinkedIn divided by the sampling ratio of the leading firms in the year. Similarly, the number of recruited workers from other firms is estimated by the number of new employees from other firms observed on LinkedIn divided by the sampling ratio of the recipient firm in the year. The definitions of independent and control variables are given in Table 2-1.

The descriptive statistics of variables are displayed in Table 2-2. Please refer to Table A2-2 in the appendix for the full correlation table of variables.

Table 2 - 1. Data Construction Procedure and Deflators[22]

Variable	Notation	Measure
Value Added	*Y*	Total annual sales (revt) minus cost of goods sold (cogs), converted to 2000 dollars.
Capital	*K*	Total assets (at) minus (total current assets (act) and intangible assets (intan)), converted to 2000 dollars.
Labor	*L*	Total number of employees (emp).
Spillovers from HR Flow of Leading software firms	*LHR*	Estimated total number of former leading software firms' employees that join the firm: Firm's new employees from leading software firms observed on LinkedIn divided by the sampling ratio of the leading firms in the year.
Spillovers from HR Flow of Other Firms	*nonLHR*	Estimated total number of former other firms' employees that join the firm: Firm's new employees from other firms observed on LinkedIn divided by the sampling ratio of the recipient firm in the year.
Other Degree Holders	*Other*	Estimated total number workers who do not have a Bachelors/Masters/PhD degree.
Bachelors Degree Holders	*Bachelor*	Estimated total number of Bachelors degree holders.
Masters Degree Holders	*Master*	Estimated total number of Masters degree holders.
PhD Degree Holders	*PhD*	Estimated total number of PhD degree holders.
High Experienced Workers	*HighExp*	Estimated total number of workers who have five years' work experience or more.
Low Experienced Workers	*LowExp*	Estimated total number of workers who have less than five years' work experience.
Workers Having Gap in Resume	*WithGap*	Estimated total number of workers who have three months' gap in resume or more.
Workers Not Having Gap in Resume	*WithoutGap*	Estimated total number of workers who do not have gap or have less than three months' gap in resume.
Year Dummies		Year dummies from year 2002 to 2012

Note: The deflator for *Y* and *K* is Producer Price Index (PPI) for software and PPI for Intermediate Materials, Supplies and Components respectively. Terms in parentheses are the variables names in Compustat. *WithGap* and *WithoutGap* are two variables that we will use for robustness checks in Section 3.6.3.

[22] Names after variables in the parentheses are the variable names in Compustat.

Table 2 - 2. Descriptive Statistics

Variable	Observa-tion	Mean	Std. Dev.	Min	Max
Sampling ratio	1923	0.31	0.12	0.06	0.93
Y (log)	1923	5.05	1.47	0.90	10.93
K (log)	1923	3.78	1.94	-1.94	10.88
L	1923	5.68	30.10	0.03	434.25
LHR	1923	4.44	18.62	0.00	419.00
Bachelor - LHR	1923	1.92	8.63	0.00	219.00
Master - LHR	1923	1.17	5.26	0.00	117.00
Phd – LHR	1923	0.09	0.61	0.00	16.00
Other - LHR	1923	1.26	4.69	0.00	82.00
HighExp - LHR	1923	3.44	14.45	0.00	332.00
LowExp - LHR	1923	1.00	4.64	0.00	87.00
WithoutGap - LHR	1923	3.05	14.44	0.00	338.00
WithGap - LHR	1923	1.39	4.78	0.00	81.00
WithoutGap - LHR (6 months)	1923	3.26	15.17	0.00	359.00
WithGap - LHR (6 months)	1923	1.19	4.09	0.00	60.00
nonLHR	1923	143.66	391.19	0.00	6029.00
Bachelor - nonLHR	1923	60.63	161.40	0.00	2779.00
Master - nonLHR	1923	29.21	89.60	0.00	1699.00
Phd - nonLHR	1923	2.29	8.67	0.00	141.00
Other - nonLHR	1923	47.27	137.79	0.00	1823.00
HighExp - nonLHR	1915	104.49	281.47	0.00	4356.00
LowExp - nonLHR	1915	39.78	115.28	0.00	1673.00
WithoutGap - nonLHR	1923	84.78	231.31	0.00	3834.00
WithGap - nonLHR	1923	58.89	163.64	0.00	2223.00
WithoutGap - nonLHR (6 months)	1923	91.20	248.10	0.00	4130.00
WithGap - nonLHR (6 months)	1923	52.47	147.61	0.00	2057.00

3.6. Results and Discussion

In this paper, we examined the productivity spillover effects from 4 leading software firms to 242 U.S. publicly listed IT firms through labor mobility from 2002 to 2012. The estimation results are shown in Table 2-3 using the following six models: (1) Ordinary least square (OLS); (2) Random effect (RE); (3) Fixed effect (FE); (4) OP estimator; (5) LP estimator; (6) Wooldridge estimator. Results are reported with robust standard errors. Models (1) - (3) are benchmarking models widely used in the IS literature. Models (4) and (5) have been widely applied in the recent productivity analysis literature to solve the simultaneity issue. Model (6) is our baseline model and is more general than (4) and (5). Among all six models, model (6) is the only one that tackles the endogeneity issue of labor input directly (as discussed in the section of empirical models), and therefore it is the baseline estimator used in this study.

Table 2 - 3. Estimation Results of Productivity Spillover Effect

Models	(1) OLS	(2) RE	(3) FE	(4) OP	(5) LP	(6) Wooldridge
Dependent Variable	Value Added	Value Added	Value Added	Value Added	Value Added	Value Added
Capital (K)	0.291***	0.141***	0.115***	0.193***	0.174***	0.228***
	(0.012)	(0.020)	(0.022)	(0.035)	(0.037)	(0.070)
Labor (L)	0.519***	0.710***	0.722***	0.503***	0.503***	0.711***
	(0.026)	(0.038)	(0.046)	(0.045)	(0.057)	(0.147)
Leading HR	0.035***	0.007***	0.005**	0.030***	0.030***	0.020**
Flow (LHR)	(0.004)	(0.002)	(0.002)	(0.007)	(0.008)	(0.010)
Other HR Flow	0.017	0.004	0.003	-0.014	-0.014	-0.015
(nonLHR)	(0.022)	(0.011)	(0.011)	(0.033)	(0.035)	(0.031)
Observations	1,923	1,923	1,923	1,918	1,918	1,539
R^2	0.880	0.864	0.859			
No. of Firms	242	242	242	242	242	226

Note: *** p<0.01, ** p<0.05, * p<0.1; Robust standard errors in parentheses; Year dummies are included but omitted for brevity.

Results of Model (1) - (6) consistently show that the productivity contribution of the firm’s labor and capital are positive and significant. Our variable of interest, the spillover effect from leading software firms (*LHR*), is significant and positive. At the same time, the spillover effect from other firms (*nonLHR*) is not significant in any of the models. Further, the coefficients of *nonLHR* are all smaller than those of *LHR* in each of the six models, indicating that the elasticity of *nonLHR* is smaller than that of *LHR*.

In general, FE or RE analysis may suffer from simultaneity issues e.g., outperforming firms may recruit or attract better employees. Our analysis is robust to this simultaneity issue for two reasons. First, we report results using the state-of-the-art econometrics models (OP/LP/Wooldridge) tailored to solve the simultaneity issues. Second, the insignificance of *nonLHR* also rules out the possibility that the positive coefficient of *LHR* is caused by spurious effect of the recipient firm, such as opening a new data center which could enhance firm productivity and increase hiring of talent simultaneously. If that is the case, we should observe a positive and significant coefficient of *nonLHR* as well and the values of coefficients of *LHR* and *nonLHR* should be similar. Our analysis confirms that the employees recruited from four leading firms indeed contribute to stronger spillover effects than other source firms.

Overall, the results of Model (1) - (6) support the proposed spillover effect. Therefore, *Hypothesis 1 is supported*. From Model (6), we see that 1% increase in HR flows from leading software firms would increase the recipient IT firm's output by 0.02%. The marginal effects of hiring one more worker form leading software firms are given in Table 2-4. Here, the marginal effect represents the value-added created by an additional worker from leading software firms, and is shown in dollars. The marginal productivity spillover per employee from leading firms is also larger than the spillover from other firms.

Table 2 - 4. Marginal Effects of Productivity Spillover Effect

Marginal Effect (Dollar amount per additional worker)			
Model 1 (OLS) - *LHR*	1,994,221***	Model 1 (OLS) - *nonLHR*	15,760
Model 2 (RE) - *LHR*	398,383***	Model 2 (RE) - *nonLHR*	3,708
Model 3 (FE) - *LHR*	284,535**	Model 3 (FE) - *nonLHR*	2,781
Model 4 (OP) - *LHR*	1,708,979***	Model 4 (OP) - *nonLHR*	-12,978
Model 5 (LP) - *LHR*	1,708,979***	Model 5 (LP) - *nonLHR*	-12,978
Model 6 (Wooldridge) - *LHR*	1,138,848**	Model 6 (Wooldridge) - *nonLHR*	-13,905

Note: The calculation of marginal effects is based on results in Table 2-3. We use the mean value of each variable to calculate the marginal effects. T-test shows that the marginal effect of LHR (left portion) is significantly larger than that of nonLHR (right portion).

3.6.1. Educational Level

To test H2, we divided the HR flows from leading software firms and the other firms into 4 categories (Bachelors, Masters, PhD, and Other[23]). The estimation results are reported in Table 2-5. The key finding is that all coefficients of the leading firms are significantly positive whereas almost all coefficients of non-leading i.e., other, firms are not significantly positive.

[23] "Other" category includes two types of profiles. First, those with degrees lower than Bachelors degree, such as high school and associate degree. Second, unidentifiable records are also categorized under this type. The original LinkedIn data about degree information is in unstructured text format. LinkedIn users may report certifications, degrees with typos, or degrees in foreign languages that are difficult to extract.

Table 2 - 5. Estimation Results of Productivity Spillover Effect by Educational Levels

Models	(1) OLS	(2) RE	(3) FE	(4) Wooldridge
Dependent Variable	Value Added	Value Added	Value Added	Value Added
Capital (K)	0.283***	0.138***	0.113***	0.203***
	(0.013)	(0.020)	(0.022)	(0.032)
Labor (L)	0.503***	0.702***	0.717***	0.688***
	(0.023)	(0.036)	(0.044)	(0.056)
Bachelor_LHR	0.019***	0.000	-0.001	0.006*
	(0.005)	(0.003)	(0.003)	(0.004)
Master_LHR	0.023***	0.012***	0.010***	0.017***
	(0.006)	(0.003)	(0.003)	(0.004)
Phd_LHR	0.048***	0.011***	0.009**	0.015**
	(0.011)	(0.004)	(0.004)	(0.007)
Other_LHR	0.013**	0.006**	0.005*	0.005
	(0.005)	(0.003)	(0.003)	(0.004)
Bachelor_nonLHR	-0.022*	-0.001	0.000	-0.019**
	(0.012)	(0.006)	(0.006)	(0.009)
Master_nonLHR	0.015	0.004	0.004	0.007
	(0.009)	(0.004)	(0.004)	(0.006)
Phd_nonLHR	0.005	0.004	0.004	0.004
	(0.004)	(0.003)	(0.003)	(0.004)
Other_nonLHR	0.008**	-0.001	-0.002	0.004
	(0.003)	(0.003)	(0.003)	(0.004)
Observations	1,919	1,919	1,919	1,538
R^2	0.884	0.868	0.862	
No. of Firms	241	241	241	225

Note: *** p<0.01, ** p<0.05, * p<0.1; Robust standard errors in parentheses; Year dummies are included but omitted for brevity.

For HR flows from the leading software firms, in Models (1) - (4), the coefficient of the Bachelors degree holders is the lowest and least statistically significant among all cases. The coefficients for PhD degree holders and Masters degree holders are positive and significant. In Model (1), the elasticity of PhD degree holders is larger than that of Masters degree holders, which is larger than that of Bachelors degree holders. In Models (2) - (4), the elasticity of PhD degree holders is slightly smaller than that of Masters degree holders. However, the marginal effect of hiring one PhD is still much larger than that of hiring a Masters degree holder because the number of Master holders is much larger than the number of PhD holders. Therefore, recruiting one more PhD degree holder still leads to higher marginal effect relative to recruiting one more Masters degree holder. Table 2-6 reports the marginal effects of hiring one additional worker of each educational level. To sum up, as the educational level goes up, the spillover effects are stronger. Therefore, *Hypothesis 2 is supported*.

Table 2 - 6. Marginal Effect of Productivity Spillover Effect by Educational Levels

	Marginal Effect (Dollar amount per additional worker)
Bachelor_LHR	930,903*
Master_LHR	4,046,608***
Phd_LHR	19,549,614**
Bachelor_nonLHR	-42,247**
Master_nonLHR	31,278
Phd_nonLHR	223,186

Note: The calculation of marginal effects is based on results in Model (4) of Table 2-5.

In sharp contrast, almost all of the coefficients of the non-leading firms are not significant. This finding also serves as a robustness check i.e., our main finding for H1 does not purely result from the fact that there are more Masters or PhD degree holders working for leading software firms. For employees from non-leading firms, even Masters and PhD degree holders do not contribute significant spillover effects.

Some readers may still have concerns that leading software firms may only hire from elite universities (such as Ivy leagues) which could be the reason for the spillover effects observed. However, at this time, we did not conduct empirical analysis to extract out the university names, majors, grades of individuals from the database. This is because the text mining results are not robust due to the variation in the forms that each of these information are present. Also, even if we could extract the information robustly, there is no foolproof way to compare the educational quality of programs across majors and universities. *Fortunately, our analysis in the next section suggests that our baseline finding does not result from the unobservable quality of the employees hired from leading software firms.*

3.6.2. Work Experience

The results of testing H3 are reported in Models (1) - (4) in Table 2-7. The coefficients for experienced workers hired from leading software firms are significant and positive across all models. In sharp contrast, all other spillover coefficients are not significant except the OLS case. This finding suggests that for employees who do not have long enough work experience, even if they worked at the leading software firms, they do not produce statistically significant spillover effects.

This case also serves as a robustness check for H1. Conceptually, ex-employees of leading firms could be more productive because of three reasons: rigorous hiring, better training, and on-the-job learning at leading firms. Specifically, readers may suspect that H1 results from the fact that leading software firms simply have recruited better talent. As a result, it is the high-quality human capital, not the training or on-the-job learning, that contributes to the spillover empirical results. Our results on work experience eliminate this alternative explanation. Specifically, our results show that without more than 5 years of work experience, the former employees of leading firms (presumably equally talented with excellent educational backgrounds too) do not carry spillover effects. Therefore, *H3 (and again H1) is supported*.

Table 2 - 7. Estimation Results of Productivity Spillover Effect by Work Experience

Models	(1) OLS	(2) RE	(3) FE	(4) Wooldridge
Dependent Variable	Value Added	Value Added	Value Added	Value Added
Capital (K)	0.287***	0.139***	0.113***	0.211***
	(0.012)	(0.020)	(0.022)	(0.047)
Labor (L)	0.489***	0.690***	0.701***	0.695***
	(0.025)	(0.037)	(0.046)	(0.087)
High Experience - LHR	0.034***	0.011***	0.009***	0.020***
	(0.004)	(0.003)	(0.003)	(0.007)
Low Experience - LHR	0.016***	-0.001	-0.003	0.004
	(0.006)	(0.003)	(0.003)	(0.007)
High Experience - nonLHR	0.060***	0.020	0.018	0.050
	(0.018)	(0.012)	(0.013)	(0.036)
Low Experience - nonLHR	-0.014	-0.000	0.000	-0.000
	(0.009)	(0.005)	(0.005)	(0.012)
Observations	1,915	1,915	1,915	1,512
R^2	0.882	0.864	0.859	
No. of Firms	242	242	242	226

Note: *** p<0.01, ** p<0.05, * p<0.1; Robust standard errors in parentheses; Year dummies are included but omitted for brevity.

Our analysis with respect to H1, H2, and H3 together suggest that the aggregated spillover effect found in the literature (Parrotta and Pozzoli 2012; Tambe and Hitt 2014a) may result from very unique groups of employees in the IT industry. The spillover effect mostly results from well-educated and experienced IT workers moving from a few leading software firms.

3.6.3. Robustness Checks

An immediate sensitivity analysis of H3 is to examine other thresholds of work experience. The first four columns of Table 2-8 report the results from using two years of work experience as the threshold and the results are consistent with the results in Table 2-7. Since there is no well-accepted definition of long work experience, “two years” is chosen because in practice, two years could be long enough for an worker to learn significant new skills or knowledge at a new employer (e.g., Grant 2011).

Similar to Table 2-7, only spillover effect from workers with at least two years’ work experience is significant, whereas all other spillover effects are insignificant. This serves as a robustness check that the training and other human capital development activities at the leading firms did create spillover for the recipient IT firms.

At the same time, we also conduct the regressions analysis of Models (1) - (4) (OLS/RE/FE/Wooldridge) in Table 2-7 using only a subsample of workers who have a postgraduate degree. The purpose is to use a sub-sample of superior human capital to examine whether work experience still matters among workers with superior human capital. The rationale is that the spillover effects may be due to workers with superior human capital, but not their work experience at the leading firms. Therefore, to rule out this possibility, we conducted this robustness check. The results are consistent with our baseline cases and are reported in the right half of Table 2-8. Even among workers who have a postgraduate degree, the spillover effects are only positively significant for those who have more than 5 years of work experience. In contrast, workers who have a postgraduate degree and have less than 5 years of work experience do not lead to significant spillover effects in the recipient firms.

Table 2 - 8. Robustness Check of Productivity Spillover Effect by Work Experience

	Full Sample - All Degrees				Superior Human Capital - Postgraduate Degrees			
Models	(1) OLS	(2) RE	(3) FE	(4) Wooldridge	(1) OLS	(2) RE	(3) FE	(4) Wooldridge
Dependent Variable	Value Added	Value Added	Value Added	Value Added	Value Added	Value Added	Value Added	Value Added
Capital (K)	0.289***	0.138***	0.112***	0.219***	0.286***	0.137***	0.112***	0.188***
	(0.012)	(0.020)	(0.022)	(0.044)	(0.012)	(0.020)	(0.022)	(0.040)
Labor (L)	0.510***	0.719***	0.732***	0.743***	0.492***	0.697***	0.708***	0.718***
	(0.032)	(0.041)	(0.050)	(0.092)	(0.021)	(0.036)	(0.044)	(0.067)
High Experience - LHR	0.034***	0.008***	0.005**	0.014**				
(2 yrs)	(0.004)	(0.002)	(0.002)	(0.006)				
Low Experience – LHR	0.027***	0.001	-0.000	0.019				
(2 yrs)	(0.009)	(0.004)	(0.004)	(0.013)				
High Experience - nonLHR	0.025	-0.008	-0.011	0.030				
(2 yrs)	(0.032)	(0.025)	(0.025)	(0.049)				
Low Experience - nonLHR	-0.004	0.001	0.001	0.006				
(2 yrs)	(0.005)	(0.003)	(0.003)	(0.010)				
High Experience - LHR					0.037***	0.014***	0.012***	0.011*
(5 yrs)					(0.005)	(0.003)	(0.003)	(0.006)
Low Experience - LHR					0.031***	0.000	-0.001	0.003
(5 yrs)					(0.009)	(0.005)	(0.005)	(0.010)
High Experience - nonLHR					0.028***	0.012***	0.011**	0.013
(5 yrs)					(0.008)	(0.004)	(0.004)	(0.008)
Low Experience - nonLHR					0.004	0.000	0.000	-0.004
(5 yrs)					(0.004)	(0.002)	(0.002)	(0.005)
Observations	1,915	1,915	1,915	1,512	1,909	1,909	1,909	1503
R^2	0.880	0.863	0.858		0.881	0.865	0.860	
No. of Firms		242	242	226		242	242	225

Note: *** p<0.01, ** p<0.05, * p<0.1; Robust standard errors in parentheses; Year dummies are included but omitted for brevity.

Another possible measure to identify workers with better human capital is to check whether the worker has a gap in his/her resume. A significant gap in the resume is usually considered as a disadvantage since it suggests that the worker may have been forced to leave his/her original employer, i.e., he/she involuntarily left the employer. Jackofsky (1984) argued that involuntary turnover is likely to be high among under-performing employees because most organizations may choose to lay off employees with worse performance than their peers. In contrast, voluntary turnover was likely to be high among employees with better performance because they may have more job opportunities outside. In other words, voluntary turnovers may be correlated with higher quality of human capital relative to involuntary turnovers. Therefore, we utilize the gap in the resume as another proxy variable for human capital quality as an additional robustness check. We set the time gap to three months in Model (1) – (4) and six months in Model (5) - (8) in Table 2-9. Again, results in both cases support our conjectures and confirm our main results i.e., work experience at leading software firms is more correlated with spillover effect. In contrast, better human capital (assessed by the gap in the resume) at other firms is not significantly correlated with spillover effects in the recipient firms.

Table 2 - 9. Robustness Check – Gaps in Resume

Models Dependent Variable	(1) OLS Value Added	(2) RE Value Added	(3) FE Value Added	(4) Wooldridge Value Added	(5) OLS Value Added	(6) RE Value Added	(7) FE Value Added	(8) Wooldridge Value Added
Capital (K)	0.287***	0.140***	0.115***	0.263***	0.287***	0.140***	0.115***	0.281***
	(0.012)	(0.020)	(0.022)	(0.051)	(0.012)	(0.020)	(0.022)	(0.056)
Labor (L)	0.528***	0.712***	0.724***	0.607***	0.524***	0.708***	0.720***	0.386***
	(0.025)	(0.037)	(0.045)	(0.101)	(0.025)	(0.037)	(0.045)	(0.131)
WithoutGap LHR	0.028***	0.006**	0.005*	0.013**				
	(0.005)	(0.002)	(0.002)	(0.006)				
WithGap LHR	0.023***	0.005	0.003	0.015				
	(0.005)	(0.003)	(0.003)	(0.011)				
WithoutGap nonLHR	0.022	0.007	0.006	0.036				
	(0.013)	(0.007)	(0.007)	(0.023)				
WithGap nonLHR	-0.023	-0.008	-0.008	-0.043**				
	(0.020)	(0.006)	(0.006)	(0.020)				
WithoutGap LHR (6 months)					0.030***	0.008***	0.007***	0.013**
					(0.005)	(0.003)	(0.003)	(0.007)
WithGap LHR (6 months)					0.021***	0.004	0.003	0.011
					(0.005)	(0.003)	(0.003)	(0.013)
WithoutGap nonLHR (6 months)					0.024*	0.008	0.006	0.023
					(0.014)	(0.008)	(0.008)	(0.024)
WithGap nonLHR (6 months)					-0.020	-0.006	-0.006	-0.026
					(0.019)	(0.006)	(0.006)	(0.021)
Observations	1,923	1,923	1,923	1,539	1,923	1,923	1,923	1,539
R^2	0.881	0.865	0.860		0.881	0.865	0.861	
No. of Firms	242	242	242	226	242	242	242	226

Note: *** p<0.01, ** p<0.05, * p<0.1; Robust standard errors in parentheses;
Year dummies are included but omitted for brevity.

Last, we use an alternative measure of productivity spillover by constructing an external human resource pool (Tambe and Hitt 2014a; Tambe and Hitt 2014b). The knowledge available to the recipient firm *i* is modeled as the weighted sum of the knowledge of the four leading software firms (T_j), where the weights between firms *i* and *j* reflect the amount of knowledge leakage between the two firms. Following Tambe and Hitt (2014a), T_j is measured as the employment level of leading firm *j*, and ϕ_{ij} is defined as the HR flow between the recipient firm *i* and each leading firm *j*, divided by the total number of employees of leading firm *j on LinkedIn*. The weights are determined in this way since we assume that the transfer of knowledge and skills occur through the movement of workers. The estimation results are shown in Table 2-10.

$$S_{it} = \sum_{j \neq i} \phi_{ijt} T_{jt}. \tag{12}$$

Table 2 - 10. Estimation Results of External HR Pool

Models	(1) OLS	(2) RE	(3) FE
Dependent Variable	Value Added	Value Added	Value Added
Capital (K)	0.294***	0.145***	0.116***
	(0.011)	(0.012)	(0.012)
All Labor	0.513***	0.688***	0.698***
	(0.016)	(0.020)	(0.023)
External HR Pool	0.121***	0.045***	0.036***
	(0.011)	(0.009)	(0.009)
Observations	1,933	1,933	1,933
R^2	0.880	0.866	0.860
No. of Firms	242	242	242

Note: *** p<0.01, ** p<0.05, * p<0.1; Standard errors in parentheses; Year dummies are included but omitted for brevity.

The significant and positive coefficient of the external HR pool shows that our previous results with respect to productivity spillover are robust and the coefficient is similar to that of Tambe and Hitt (2014a). In this way, we add to the literature by providing a deeper understanding of the different sources within the aggregated HR spillover effect documented in the literature.

3.6.4. Limitations and Future Works

This study also has a few limitations. First, data on the exact amount of HR investments of the IT firms is not available, to the best of our knowledge. Therefore, our study followed prior studies (Parrotta and Pozzoli 2012; Tambe and Hitt 2014a) in using the headcounts of employees to examine productivity spillover effects. If the data on HR investments were available, researchers can further divide the source of spillover effect into two parts: tangible HR investments and intangible HR assets such as the on-the-job learning because of developing cutting-edge software products. Second, our study examines the spillover effects to publicly listed IT firms for which Compustat data is available. It would be useful to analyze the magnitude of spillover effect to private IT firms, particularly new startups, if their accounting data were available. Third, we chose 4 leading software platform owners based on their market capitalization, platform they provide, and the NAICS codes closest to software publishers (code 511210). Thus Google and Facebook were chosen with code 519130, while Microsoft and Oracle were chosen with code 511210. Other companies that could be examined in future include Apple which is currently excluded as its code 334111 (Electronic Computers Manufacturing) focuses more on its hardware products.

A related topic for future research is to analyze how much the leading IT firm's ex-employees can contribute to the success of new startups by fueling the innovation engine of startups. Last, it is also valuable to explore the effect of HR flow network structure on an IT firm's productivity and the associated firm-level financial performance. If we consider the movements of IT workers as a network of HR flows, it is probable that an IT firm's position in this HR network will be salient to its productivity. The proposed dataset from LinkedIn provides an opportunity for such kind of research in the future.

3.7. Contributions and Conclusion

This study employs a novel productivity analysis method (Wooldridge 2009) to empirically examine the productivity spillover effects from four large software platform providers (Google, Microsoft, Oracle and Facebook) on other publicly listed IT firms. We hypothesize that work experience at these four software firms may contribute more spillover effect than at other firms because these four software firms develop and design the major software platforms that are used by many other IT firms. We find empirical evidence to support our hypothesis. Our findings suggest that leading software firms' HR investment and development activities would also benefit other recipient IT firms that recruit employees from these leaders.

Our study contributes to the literature in the following ways. Adding to the prior studies conducted at the aggregate level (Parrotta and Pozzoli 2012; Tambe and Hitt 2014a), our results show that in the IT industry, the labor productivity spillovers to recipient firms result from very specific sources: Masters and PhD degree holders as well as employees with more than 2 years' work experience from 4 leading software firms. Our study also adds to the productivity analysis literature in showing the importance of productivity spillovers from human resource investments, given that most existing IS studies focused solely on the spillover effects from IT capital investments. Further, to the best of our knowledge, this paper is a first attempt in the IS literature to use the generalized OP method and joint GMM method (Wooldridge 2009) to analyze productivity. This novel method has been shown to be more general than AB, OP, and LP methods and is also easy for researchers to implement in Stata (one of the most popular econometrics software). Last, this study also demonstrates the utility of a recent data collection method by using computer programs to scrape millions of profiles from online professional network for econometrics analysis.

The results provide several important implications for HR managers of IT firms. First, hiring former employees from leading software firms would benefit recipient IT firm's productivity. Understanding the magnitude of such spillover effects is critical for developing the recipient firm's hiring strategy. For example, the large marginal effect of hiring one additional worker from leading software firms also imply that it is worth paying more to workers from leading software firms than those from other firms. Second, workers from leading software firms could have different productivity spillovers to the recipient firm depending on their human capital. Specifically, our findings show that the educational level, work experience, and even lack of gap in resumes are correlated with larger spillover effects of employees from leading software firms.

Our study also provides implications for policy makers. The significant productivity spillover effects from leading software firms could drive the growth of the IT industry by increasing the productivity of other IT firms. With this in mind, for example, non-poaching agreements among technology companies (Rosenblatt and Gullo 2013) may not be suitable for the industry. Policy makers should aim to protect free mobility of labor within the IT industry because this can improve recipient firms' productivity and thereby the IT industry's total productivity. Governments can also consider subsidizing training efforts of leading firms such that the entire IT industry can benefit from these investments. For IT workers, our analysis reveals the high value of work experience at major software platform providers, which they could consider when they search for and choose their jobs.

4. GENERAL CONCLUSION

This dissertation aims to investigate the impacts of two important phenomena pertaining to firm productivity in the IT industry. We conduct two empirical studies to assess their impacts on firm productivity. Study 1 examines the economies of scale of the SaaS firms. Study 2 pioneers the use of LinkedIn data to examine the productivity spillover effects through worker mobility from major software providers. Both studies use advanced econometric methods for productivity analysis. Especially, Study 2 is the initial effort in the IS literature to employ the method proposed by Wooldridge (2009) as the identification strategy.

The findings from Study 1 show that SaaS firms exhibit smaller economies of scale than traditional software firms. This property may result from the business model of SaaS firms. In contrast to traditional software firms, SaaS vendors may incur significant variable costs: they need to invest in IT infrastructure and hire more staff to serve more corporate buyers. Therefore, the well-known economies of scale of traditional software businesses arising from zero variable cost disappear in the SaaS model. The implication is that larger SaaS vendors may not have cost advantages over smaller competitors, as is the case in the traditional software industry. Leveraging a larger customer base may not be an effective strategy in the SaaS industry anymore. Instead, SaaS firms should focus more on differentiating and customizing their products as the main vehicle for competition. From the buyers' point of view, corporate clients of SaaS services could focus on choosing SaaS vendors that provide the most effective service, without worrying that smaller SaaS vendors might ultimately be forced out of the market due to the economies of scale of larger vendors.

The findings from Study 2 show that leading software firms' HR investment and development activities would also benefit other recipient IT firms through worker mobility. Study 2 also provides the magnitudes of such productivity spillover effects. In Study 2, the large marginal effect of hiring one additional worker (around 1.1 million US dollars) from leading software firms implies that it is worth paying more to workers from leading software firms than those from other firms. The findings from Study 2 also show that the productivity spillover effects through worker mobility are mainly contributed by experienced and well-educated workers from leading software firms. Therefore, Study 2 provides the source and the magnitude of productivity spillover in the IT industry, which are critical for developing an efficient hiring strategy in the context of IT talent war. According to the findings, IT firms should actively recruit well-educated workers with more than five years' experience from leading software firms. At the same time, government policy makers should also protect the free mobility of workers in the IT industry to maximize social welfare and facilitate productivity growth of the industry.

In conclusion, the findings from this dissertation contribute to the IT productivity literature by showing that two phenomena beyond IT investment could also significantly and substantially affect the productivity of IT firms. Specifically, the two studies examine two important phenomena of the IT industry, i.e., SaaS model and worker mobility. This research thus provides several implications. First, researchers in the field of IT productivity may need to continuously assess the productivity impacts of the emerging phenomena in the IT industry as this industry is fast evolving. Second, even some widely cited characteristics of the IT industry may not have been verified by rigorous empirical studies, and thus may not be scientifically true. Therefore, special attentions need to be paid to such issues. Third, this dissertation provides an example of using new types of data which are becoming available as the digital economy grows. Before the introduction of online professional social network, using public data to accurately trace and analyze the mobility of millions of workers is rather difficult. Researchers could utilize these opportunities of exploring and analyzing new public data on the Internet. Fourth, Study 2 also provides a rigorous analysis of the “value of the employees”. It could be easily modified to guide corporate HR practices such as deciding the appropriate salary and compensation packages for different types of workers.

REFERENCES

Ackerberg, D., Caves, K., and Frazer, G. 2006. "Structural Identification of Production Functions." Working Paper.

Ackerberg, D., Lanier Benkard, C., Berry, S., and Pakes, A. 2007. "Econometric Tools for Analyzing Market Outcomes," *Handbook of econometrics* (6), pp. 4171-4276.

Agrawal, A., and Tambe, P. 2013. "Private Equity, Technological Investment, and Labor Outcomes," *International Conference on Information Systems (ICIS)*, Milan, Italy.

Aigner, D., Lovell, C., and Schmidt, P. 1977. "Formulation and Estimation of Stochastic Frontier Production Function Models," *Journal of Econometrics* (6:1), pp. 21-37.

Almeida, P., and Kogut, B. 1999. "Localization of Knowledge and the Mobility of Engineers in Regional Networks," *Management Science* (45:7), pp. 905-917.

Altinkemer, K., Ozcelik, Y., and Ozdemir, Z.D. 2011. "Productivity and Performance Effects of Business Process Reengineering: A Firm-Level Analysis," *Journal of Management Information Systems* (27:4), pp. 129-162.

Ang, S., Slaughter, S., and Ng, K.Y. 2002. "Human Capital and Institutional Determinants of Information Technology Compensation: Modeling Multilevel and Cross-Level Interactions," *Management Science* (48:11), pp. 1427-1445.

Aral, S., Brynjolfsson, E., and Van Alstyne, M. 2012a. "Information, Technology, and Information Worker Productivity," *Information Systems Research* (23:3), pp. 849-867.

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

Arellano, M., and Bond, S. 1991. "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations," *The Review of Economic Studies* (58:2), pp. 277-297.

Argote, L., and Ingram, P. 2000. "Knowledge Transfer: A Basis for Competitive Advantage in Firms," *Organizational behavior and human decision processes* (82:1), pp. 150-169.

Armbrust, M., Fox, A., Griffith, R., Joseph, A.D., Katz, R., Konwinski, A., Lee, G., Patterson, D., Rabkin, A., and Stoica, I. 2010. "A View of Cloud Computing," *Communications of the ACM* (53:4), pp. 50-58.

Arora, A., and Nandkumar, A. 2012. "Insecure Advantage? Markets for Technology and the Value of Resources for Entrepreneurial Ventures," *Strategic Management Journal* (33:3), pp. 231-251.

Banker, R., Wattal, S., Liu, F.C., and Ou, C.S. 2008. "Value Contributed by Education in IT Firms," *International Conference on Information Systems*, Paris, France.

Banker, R.D., Chang, H., and Kemerer, C.F. 1994. "Evidence on Economies of Scale in Software Development," *Information and Software Technology* (36:5), pp. 275-282.

Banker, R.D., Charnes, A., and Cooper, W.W. 1984. "Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis," *Management Science* (30:9), pp. 1078-1092.

Banker, R.D., and Slaughter, S.A. 1997. "A Field Study of Scale Economies in Software Maintenance," *Management Science* (43:12), pp. 1709-1725.

Bardhan, I., Krishnan, V., and Lin, S. 2013. "Business Value of Information Technology: Testing the Interaction Effect of IT and R&D on Tobin's Q," *Information Systems Research* (24:4), pp. 1147-1161.

Bardhan, I.R., Demirkan, H., Kannan, P., Kauffman, R.J., and Sougstad, R. 2010. "An Interdisciplinary Perspective on IT Services Management and Service Science," *Journal of Management Information Systems* (26:4), pp. 13-64.

Battese, G.E., and Coelli, T.J. 1995. "A Model for Technical Inefficiency Effects in a Stochastic Frontier Production Function for Panel Data," *Empirical economics* (20:2), pp. 325-332.

Becker, G.S. 1962. "Investment in Human Capital: A Theoretical Analysis," *The journal of political economy* (70:5), pp. 9-49.

Berry, S., Levinsohn, J., and Pakes, A. 1995. "Automobile Prices in Market Equilibrium," *Econometrica* (63:4), pp. 841-890.

Bhargava, H.K., and Sundaresan, S. 2004. "Computing as Utility: Managing Availability, Commitment, and Pricing through Contingent Bid Auctions," *Journal of Management Information Systems* (21:2), pp. 201-227.

Blundell, R., and Bond, S. 1998. "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models," *Journal of econometrics* (87:1), pp. 115-143.

Boh, W.F., Slaughter, S.A., and Espinosa, J.A. 2007. "Learning from Experience in Software Development: A Multilevel Analysis," *Management Science* (53:8), pp. 1315-1331.

Bonvanie, R. 2007. "SaaS: The New Distress Call," in: *Businessweek*. Bloomberg.

Bresnahan, T.F., Brynjolfsson, E., and Hitt, L.M. 2002. "Information Technology, Workplace Organization, and the Demand for Skilled Labor: Firm-Level Evidence," *The Quarterly Journal of Economics* (117:1), pp. 339-376.

Brynjolfsson, E., and Hitt, L.M. 1996. "Paradox Lost? Firm-Level Evidence on the Returns to Information Systems Spending," *Management Science* (42:4), pp. 541-558.

BusinessInsider. 2011. "Infographic: The Incredibly Difficult Process of Getting Hired at Google." Retrieved 2 March 2013, from http://www.businessinsider.com/infographics-process-of-getting-hired-at-google -2011-10

Campbell-Kelly, M. 2009. "Historical Reflections the Rise, Fall, and Resurrection of Software as a Service," *Communications of the ACM* (52:5), pp. 28-30.

Carr, N.G. 2005. "The End of Corporate Computing," *MIT Sloan Management Review* (46:3), pp. 67-73.

Casper, S., and Murray, F. 2005. "Careers and Clusters: Analyzing the Career Network Dynamic of Biotechnology Clusters," *Journal of Engineering and Technology Management* (22:1), pp. 51-74.

Chang, Y.B., and Gurbaxani, V. 2012a. "The Impact of IT-Related Spillovers on Long-Run Productivity: An Empirical Analysis," *Information Systems Research* (23:3-Part-2), pp. 868-886.

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

Chen, P.Y., and Wu, S.Y. 2013. "The Impact and Implications of On-Demand Services on Market Structure," *Information Systems Research* (24:3), pp. 750-767.

Cheng, H.K., and Koehler, G.J. 2003. "Optimal Pricing Policies of Web-Enabled Application Services," *Decision Support Systems* (35:3), pp. 259-272.

Cheng, Z., and Nault, B.R. 2007. "Industry Level Supplier-Driven IT

Spillovers," *Management Science* (53:8), pp. 1199-1216.

Cheng, Z., and Nault, B.R. 2012. "Relative Industry Concentration and Customer-Driven IT Spillovers," *Information Systems Research* (23:2), pp. 340-355.

Choudhary, V. 2007. "Comparison of Software Quality under Perpetual Licensing and Software as a Service," *Journal of Management Information Systems* (24:2), pp. 141-165.

CIA. 2013. "The World Factbook."

Coelli, T., Rao, D., O'Donnell, C., and Battese, G. 2005. *An Introduction to Efficiency and Productivity Analysis*. Springer.

ComputerWeekly. 2009. "Service-Oriented Architecture in and beyond the Credit Crunch." from http://www.computerweekly.com/feature/ Service-oriented-architecture-in-and- beyond-the-credit-crunch

Cope, K. 2012. "Employees: The Best Companies Know They're the Foundation of Success." Retrieved 2 March 2013, from http://www.tlnt.com/2012/02/20/employees-the-best-companys-know-t heyre-the-foundation -of-success/

Currie, W.L., and Parikh, M.A. 2006. "Value Creation in Web Services: An Integrative Model," *The Journal of Strategic Information Systems* (15:2), pp. 153-174.

Del Gatto, M., Di Liberto, A., and Petraglia, C. 2011. "Measuring Productivity," *Journal of Economic Surveys* (25:5), pp. 952-1008.

Demirkan, H., Cheng, H.K., and Bandyopadhyay, S. 2010. "Coordination Strategies in an SaaS Supply Chain," *Journal of Management Information Systems* (26:4), pp. 119-143.

Desisto, R.P. 2010. "Saas Industry Report Card," Gartner.

Dewan, S., and Kraemer, K.L. 2000. "Information Technology and Productivity: Evidence from Country-Level Data," *Management Science* (46:4), pp. 548-562.

Dewan, S., and Min, C. 1997. "The Substitution of Information Technology for Other Factors of Production: A Firm Level Analysis," *Management Science* (43:12), pp. 1660-1675.

Dewan, S., and Ren, F. 2007. "Risk and Return of Information Technology Initiatives: Evidence from Electronic Commerce Announcements," *Information Systems Research* (18:4), pp. 370-394.

Dewan, S., and Ren, F. 2011. "Information Technology and Firm Boundaries: Impact on Firm Risk and Return Performance," *Information Systems Research* (22:2), pp. 369-388.

Fan, M., Kumar, S., and Whinston, A.B. 2009. "Short-Term and Long-Term Competition between Providers of Shrink-Wrap Software and Software as a Service," *European Journal of Operational Research* (196:2), pp. 661-671.

Garg, R., and Telang, R. 2012. "To Be or Not to Be Linked on Linkedin: Job Search Using Online Social Networks," *NBER Summer Institute*

Working Paper.

Gartner. 2009. "The Real Truth About Cloud, SaaS and Saving Money Now." from http://my.gartner.com/it/content/1046400/1046412/ july15_cloudcomputing_darylplummer_webinar.pdf

Gawer, A. 2009. *Platforms, Markets and Innovation*. Cheltenham, UK and Northampton, US: Edward Elgar.

Ge, C., Huang, K.W., and Png, I.P.L. 2014. "Engineer/Scientist Careers: Patents, Online Profiles, and Misclassification Bias," Working Paper.

Grant, A. 2011. "How to Be a Smart Job-Hopper," in: *U.S.News*.

Grant, R.M. 1996. "Toward a Knowledge-Based Theory of the Firm," *Strategic Management Journal* (17), pp. 109-122.

Greene, W.H. 2011. *Econometric Analysis*, (7 ed.). Prentice Hall.

Griliches, Z. 1979. "Issues in Assessing the Contribution of Research and Development to Productivity Growth," *The Bell Journal of Economics* (10:1), pp. 92-116.

Griliches, Z. 1994. "Productivity, R&D, and the Data Constraint," *The American Economic Review* (84:1), pp. 1-23.

Guillory, J., and Hancock, J.T. 2012. "The Effect of Linkedin on Deception in Resumes," *Cyberpsychology, Behavior, and Social Networking* (15:3), pp. 135-140.

Han, K., CHaNg, Y.B., and Hahn, J. 2011a. "Information Technology Spillover and Productivity: The Role of Information Technology

Intensity and Competition," *Journal of Management Information Systems* (28:1), pp. 115-146.

Han, K., Kauffman, R.J., and Nault, B.R. 2011b. "Research Note—Returns to Information Technology Outsourcing," *Information Systems Research* (22:4), pp. 824-840.

Hatch, N.W., and Dyer, J.H. 2004. "Human Capital and Learning as a Source of Sustainable Competitive Advantage," *Strategic Management Journal* (25:12), pp. 1155-1178.

Hayward, M.L., and Hambrick, D.C. 1997. "Explaining the Premiums Paid for Large Acquisitions: Evidence of CEO Hubris," *Administrative Science Quarterly* (42:1), pp. 103-127.

He, L. 2013. "Google's Secrets of Innovation: Empowering Its Employees," in: *Forbes*.

Henderson, J.V. 2007. "Understanding Knowledge Spillovers," *Regional Science and Urban Economics* (37:4), pp. 497-508.

Hitt, L.M., and Brynjolfsson, E. 1996. "Productivity, Business Profitability, and Consumer Surplus: Three Different Measures of Information Technology Value," *MIS quarterly* (20:2), pp. 121-143.

Hitt, L.M., Wu, D., and Zhou, X. 2002. "Investment in Enterprise Resource Planning: Business Impact and Productivity Measures," *Journal of Management Information Systems* (19:1), pp. 71-98.

Hitt, M.A., Bierman, L., Shimizu, K., and Kochhar, R. 2001. "Direct and

Moderating Effects of Human Capital on Strategy and Performance in Professional Service Firms: A Resource-Based Perspective," *Academy of Management journal* (44:1), pp. 13-28.

Holmes, K. 2012. "Grow or Add? SaaS Application Scaling Basics." Retrieved 15th Sep 2012, from http://www.rackspace.com/blog/grow-or-add-saas-application-scaling-basics/

Huang, K.W., and Sundararajan, A. 2011. "Pricing Digital Goods: Discontinuous Costs and Shared Infrastructure," *Information Systems Research* (22:4), pp. 721-738.

Huang, K.W., and Wang, M. 2009. "Firm-Level Productivity Analysis for Software as a Service Companies," *International Conference on Information Systems*, Phoenix, USA.

Jackofsky, E.F. 1984. "Turnover and Job Performance: An Integrated Process Model," *Academy of management review* (9:1), pp. 74-83.

Kantrowitz, A. 2013. "Google Reader Founder: I Never Would Have Founded Reader inside Today's Google," in: *Forbes*.

Katz, M.L., and Shapiro, C. 1994. "Systems Competition and Network Effects," *The Journal of Economic Perspectives* (8:2), pp. 93-115.

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

Knight, F.H. 1944. "Diminishing Returns from Investment," *The Journal of Political Economy* (52:1), pp. 26-47.

Knott, A.M., and Posen, H.E. 2005. "Is Failure Good?," *Strategic Management Journal* (26:7), pp. 617-641.

Kudyba, S., and Diwan, R. 2002. "Research Report: Increasing Returns to Information Technology," *Information Systems Research* (13:1), pp. 104-111.

Levinsohn, J., and Petrin, A. 2003. "Estimating Production Functions Using Inputs to Control for Unobservables," *The Review of Economic Studies* (70:2), pp. 317-341.

Li, S., Shang, J., and Slaughter, S.A. 2010. "Why Do Software Firms Fail? Capabilities, Competitive Actions, and Firm Survival in the Software Industry from 1995 to 2007," *Information Systems Research* (21:3), pp. 631-654.

Lichtenberg, F.R., and Siegel, D. 1991. "The Impact of R&D Investment on Productivity: New Evidence Using Linked R&D-LRD Data," *Economic Inquiry* (29:2), pp. 203-229.

Lieberman, M.B., Lau, L.J., and Williams, M.D. 1990. "Firm-Level Productivity and Management Influence: A Comparison of US and Japanese Automobile Producers," *Management Science* (36:10), pp. 1193-1215.

Lin, W.T., and Shao, B. 2000. "Relative Sizes of Information Technology

Investments and Productive Efficiency: Their Linkage and Empirical Evidence," *Journal of the Association for Information Systems* (1:7), pp. 1-35.

Lord, R.G., and Maher, K.J. 1990. "Alternative Information-Processing Models and Their Implications for Theory, Research, and Practice," *Academy of management review* (15:1), pp. 9-28.

Loveman, G.W. 1994. "An Assessment of the Productivity Impact of Information Technologies," in *Information Technology and the Corporation of the 1990s: Research Studies,* T.J. ALLEN and M.S.S. MORTON (eds.). Cambridge, MA: MIT Press, pp. 84-110.

Ma, D., and Seidmann, A. 2007. "ASP On-Demand Versus MOTS In-House Software Solutions," Working Paper.

Ma, D., and Seidmann, A. 2008. "The Pricing Strategy Analysis for the “Software-as-a-Service” Business Model," *Grid Economics and Business Models* (52:6), pp. 103-112.

Madsen, T.L., Mosakowski, E., and Zaheer, S. 2003. "Knowledge Retention and Personnel Mobility: The Nondisruptive Effects of Inflows of Experience," *Organization Science* (14:2), pp. 173-191.

Malmendier, U., and Tate, G. 2008. "Who Makes Acquisitions? CEO Overconfidence and the Market's Reaction," *Journal of Financial Economics* (89:1), pp. 20-43.

Mas-Collel, A., Whinston, M.D., and Green, J. 1995. *Microeconomic Theory*.

Oxford university press, Oxford.

Meeusen, W., and Van Den Broeck, J. 1977. "Efficiency Estimation from Cobb-Douglas Production Functions with Composed Error," *International Economic Review* (18:2), pp. 435-444.

Milanov, H., and Shepherd, D.A. 2013. "The Importance of the First Relationship: The Ongoing Influence of Initial Network on Future Status," *Strategic Management Journal* (34:6), pp. 727-750.

Mithas, S., and Krishnan, M.S. 2008. "Human Capital and Institutional Effects in the Compensation of Information Technology Professionals in the United States," *Management Science* (54:3), pp. 415-428.

Mittal, N., and Nault, B.R. 2009. "Research Note—Investments in Information Technology: Indirect Effects and Information Technology Intensity," *Information Systems Research* (20:1), pp. 140-154.

Moretti, E. 2004a. "Estimating the Social Return to Higher Education: Evidence from Longitudinal and Repeated Cross-Sectional Data," *Journal of econometrics* (121:1), pp. 175-212.

Moretti, E. 2004b. "Workers' Education, Spillovers, and Productivity: Evidence from Plant-Level Production Functions," *American Economic Review* (94:3), pp. 656-690.

Morley, B. 2006. "Causality between Economic Growth and Immigration: An Ardl Bounds Testing Approach," *Economics Letters* (90:1), pp. 72-76.

Morrison, C.J., and Berndt, E.R. 1991. "Assessing the Productivity of

Information Technology Equipment in US Manufacturing Industries," *National Bureau of Economic Research Working Paper 3582*.

Nonaka, I. 1994. "A Dynamic Theory of Organizational Knowledge Creation," *Organization Science* (5:1), pp. 14-37.

Olley, S., and Pakes, A. 1996. "The Dynamics of Productivity in the Telecomunications Equipment Industry," *Econometrica* (64:6), pp. 1263-1297.

Oreskovic, A. 2013. "Yahoo's Rigorous Hiring Practices, Imposed by Marissa Mayer, Get Flak within Company," in: *Reuters*.

Palomeras, N., and Melero, E. 2010. "Markets for Inventors: Learning-by-Hiring as a Driver of Mobility," *Management Science* (56:5), pp. 881-895.

Park, J., Shin, S.K., and Sanders, G.L. 2007a. "Impact of International Information Technology Transfer on National Productivity," *Information Systems Research* (18:1), pp. 86-102.

Park, J., Shin, S.K., and Shin, H.H. 2007b. "The Intensity and Externality Effects of Information Technology Investments on National Productivity Growth," *IEEE Transactions on Engineering Management* (54:4), pp. 716-728.

Parrotta, P., and Pozzoli, D. 2012. "The Effect of Learning by Hiring on Productivity," *The RAND Journal of Economics* (43:1), pp. 167-185.

Pepitone, J. 2013. "Apple Throws Recruitment Party for Blackberry

Employees," CNN.

Petrin, A., and Levinsohn, J. 2012. "Measuring Aggregate Productivity Growth Using Plant-Level Data," *The RAND Journal of Economics* (43:4), pp. 705-725.

Pisano, G.P. 1994. "Knowledge, Integration, and the Locus of Learning: An Empirical Analysis of Process Development," *Strategic Management Journal* (15:S1), pp. 85-100.

Pring, B. 2011. "Positioning and Pitching Saas for Optimum Effect," Gartner

Rai, A., and Sambamurthy, V. 2006. "Editorial Notes—the Growth of Interest in Services Management: Opportunities for Information Systems Scholars," *Information Systems Research* (17:4), pp. 327-331.

Rosen, S. 1972. "Learning and Experience in the Labor Market," *Journal of Human Resources* (7:3), pp. 326-342.

Rosenblatt, J., and Gullo, K. 2013. "Apple, Google Must Face Group Antitrust Hiring Lawsuit." from http://www.bloomberg.com/news/2013-10-25/ apple-google-must-face-group-antitrust-hiring-lawsuit.html

Rosenkopf, L., and Almeida, P. 2003. "Overcoming Local Search through Alliances and Mobility," *Management Science* (49:6), pp. 751-766.

Salesforce.com. 2011. "Annual Report for the Fiscal Year Ended January 31, 2011".

Schmidt, K.M., and Schnitzer, M. 2003. "Public Subsidies for Open Source? Some Economic Policy Issues of the Software Market," *Harvard*

Journal of Law & Technology (16:2), pp. 474-502.

Shao, B., and Lin, W.T. 2002. "Technical Efficiency Analysis of Information Technology Investments: A Two-Stage Empirical Investigation," *Information & management* (39:5), pp. 391-401.

SIIA. 2001. "Software as a Service: Strategic Backgrounder."

Smith, K.G., Collins, C.J., and Clark, K.D. 2005. "Existing Knowledge, Knowledge Creation Capability, and the Rate of New Product Introduction in High-Technology Firms," *Academy of Management journal* (48:2), pp. 346-357.

Smith, M.A., and Kumar, R.L. 2004. "A Theory of Application Service Provider (ASP) Use from a Client Perspective," *Information & management* (41:8), pp. 977-1002.

Solow, R.M. 1957. "Technical Change and the Aggregate Production Function," *The Review of Economics and Statistics* (39:3), pp. 312-320.

Somaya, D., Williamson, I.O., and Lorinkova, N. 2008. "Gone but Not Lost: The Different Performance Impacts of Employee Mobility between Cooperators Versus Competitors," *The Academy of Management Journal* (51:5), pp. 936-953.

Song, J., Almeida, P., and Wu, G. 2003. "Learning–by–Hiring: When Is Mobility More Likely to Facilitate Interfirm Knowledge Transfer?," *Management Science* (49:4), pp. 351-365.

Sullivan, J. 2013. "How Google Became the #3 Most Valuable Firm by Using People Analytics to Reinvent HR." Retrieved Feb 11, 2014, from http://www.ere.net/2013/02/25/how-google-became-the-3-most-valuable-firm-by-using-people-analytics-to-reinvent-hr/

Susarla, A., and Barua, A. 2011. "Contracting Efficiency and New Firm Survival in Markets Enabled by Information Technology," *Information Systems Research* (22:2), pp. 306-324.

Susarla, A., Barua, A., and Whinston, A.B. 2003. "Understanding the Service Component of Application Service Provision: Empirical Analysis of Satisfaction with ASP Services," *MIS quarterly* (27:1), pp. 91-123.

Tambe, P. 2012. "How the IT Workforce Affects Returns to IT Innovation: Evidence from Big Data Analytics," *Workshop on Information Systems and Economics (WISE)*.

Tambe, P. 2014. "Big Data Investment, Skills, and Firm Value," *Management Science* (60:6), pp. 1452-1469.

Tambe, P., and Hitt, L.M. 2010. "How Offshoring Affects IT Workers," *Communications of the ACM* (53:10), pp. 62-70.

Tambe, P., and Hitt, L.M. 2012a. "Now IT's Personal: Offshoring and the Shifting Skill Composition of the US Information Technology Workforce," *Management Science* (58:4), pp. 678-697.

Tambe, P., and Hitt, L.M. 2012b. "The Productivity of Information Technology Investments: New Evidence from IT Labor Data," *Information Systems*

Research (23:3), pp. 599-617.

Tambe, P., and Hitt, L.M. 2014a. "Job Hopping, Information Technology Spillovers, and Productivity Growth," *Management Science* (60:2), pp. 338-355.

Tambe, P., and Hitt, L.M. 2014b. "Measuring Information Technology Spillovers," *Information Systems Research* (25:1), pp. 53-71.

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

Teece, D.J. 2010. "Business Models, Business Strategy and Innovation," *Long Range Planning* (43:2), pp. 172-194.

Temple, J.R.W. 2001. "Generalizations That Aren't? Evidence on Education and Growth," *European Economic Review* (45:4), pp. 905-918.

TheNewYorkTimes. 2012. "Power, Pollution and the Internet." Retrieved 1st Oct 2012, from http://www.nytimes.com/2012/09/23/technology/data-centers-waste-vast-amounts-of-energy-belying-industry-image.html

Tirole, J. 1988. *The Theory of Industrial Organization*. MIT press.

Varian, H.R., and Repcheck, J. 2010. *Intermediate Microeconomics: A Modern Approach*. WW Norton & Company New York, NY.

Viega, J. 2009. "Cloud Computing and the Common Man," *Computer* (42:8), pp. 106-108.

Walker, J. 2012. "School's in Session at Google," in: *The Wall Street Journal*.

Walsh, K.R. 2003. "Analyzing the Application Asp Concept: Technologies, Economies, and Strategies," *Communications of the ACM* (46:8), pp. 103-107.

Whitney, L. 2013. "Apple, Google, and the E-Mail Trail in the No-Poaching Case." CNET.com.

Wong, R. 2012. "Report: Apple Hiring Former Google Engineers to Fix Maps App," in: *Fox Business*.

Wooldridge, J.M. 2009. "On Estimating Firm-Level Production Functions Using Proxy Variables to Control for Unobservables," *Economics Letters* (104:3), pp. 112-114.

Xin, M., and Levina, N. 2008. "Software-as-a-Service Model: Elaborating Client-Side Adoption Factors," *Proceedings of the 29th International Conference on Information Systems*, Paris, France.

Yang, S.B., Lim, J.H., Oh, W., Animesh, A., and Pinsonneault, A. 2012. "Research Note-Using Real Options to Investigate the Market Value of Virtual World Businesses," *Information Systems Research* (23:3-part-2), pp. 1011-1029.

APPENDIX

A1-1. List of SaaS Firms

Athenahealth Inc	Rightnow Technologies Inc
Concur Technologies Inc	Salary.Com Inc
Constant Contact Inc	Salesforce.Com Inc
Convio Inc	Soundbite Communications Inc
Dealertrack Holdings Inc	SPS Commerce Inc
Demandtec Inc	Successfactors Inc
Kenexa Corp	Taleo Corp
Kintera Inc	Ultimate Software Group Inc
Liveperson Inc	Visual Sciences Inc
Medidata Solutions Inc	Vocus Inc
Netsuite Inc	Webex Communications Inc
Omniture Inc	

A1-2. Correlation of Variables in Study 1

	Value-Added	Capital	Labor	R&D	Advertising	Firm Size	TE
SaaS Firms							
	Value-Added	Capital	Labor	R&D	Advertising	Firm Size	TE
Value-Added	1.000						
Capital	0.907	1.000					
Labor	0.920	0.871	1.000				
R&D	-0.332	-0.271	-0.292	1.000			
Advertising	-0.006	0.004	-0.028	0.075	1.000		
Firm Size	0.939	0.887	0.872	-0.268	-0.049	1.000	
TE	0.560	0.447	0.286	-0.258	0.055	0.473	1.000
Traditional Software Firms							
	Value-Added	Capital	Labor	R&D	Advertising	Firm Size	TE
Value-Added	1.000						
Capital	0.881	1.000					
Labor	0.940	0.892	1.000				
R&D	-0.234	-0.111	-0.174	1.000			
Advertising	-0.152	-0.093	-0.121	0.824	1.000		
Firm Size	0.938	0.920	0.942	-0.143	-0.102	1.000	
TE	0.574	0.368	0.352	-0.113	-0.068	0.446	1.000

A2-1. Mathematical Detail

According to Wooldridge (2009), the two equations that identify the coefficients of labor and capital (β, γ) are as follows:

$$y_{it} = \alpha + \beta L_{it} + \phi LHR_{it} + \theta nonLHR_{it} + \gamma K_{it} + g(K_{it}, m_{it}, L_{it}) + e_{it}, \qquad (1)$$

and

$$y_{it} = \alpha + \beta L_{it} + \phi LHR_{it} + \theta nonLHR_{it} + \gamma K_{it} + f[g(K_{i,t-1}, m_{i,t-1}, L_{i,t-1})] + u_{it}, \qquad (2)$$

$$u_{it} = a_{it} + e_{it}. \qquad (3)$$

where m_{it} is operationalized as firm capital expenditure (the OP approach) in our study. Capital expenditure of software firms could be used to capitalize software development cost, purchase computers, and other activities that support the business of software firms. We do not use the LP approach in which the cost of goods sold is the proxy variable due to the following two reasons. First, unlike manufacturing firms, capital expenditure is not very lumpy among our sample IT firms. Second, the cost-of-goods sold has larger variation in our sample firms because software and e-commerce services firms have negligibly small variable costs conceptually and have very small proportion of cost of revenue in accounting statements.

Assume $g(K_{it}, m_{it}, L_{it})$ is a third-order polynomial (e.g., Petrin and Levinsohn 2012), and let

$$\begin{aligned} g(K_{it}, m_{it}, L_{it}) = {} & \eta + \pi K_{it} + \psi m_{it} + \delta L_{it} + \pi_2 K_{it}^2 + \psi_2 m_{it}^2 + \delta_2 L_{it}^2 \\ & + \lambda_1 K_{it} m_{it} + \lambda_2 K_{it} L_{it} + \lambda_3 m_{it} L_{it} + \lambda_4 K_{it}^2 m_{it} + \lambda_5 K_{it} m_{it}^2 \\ & + \lambda_6 K_{it}^2 L_{it} + \lambda_7 K_{it} L_{it}^2 + \lambda_8 L_{it}^2 m_{it} + \lambda_9 L_{it} m_{it}^2 \\ & + \pi_3 K_{it}^3 + \psi_3 m_{it}^3 + \delta_3 L_{it}^3, \end{aligned} \qquad (4)$$

v_{it} is the unobserved productivity shock and we assume

$f(v_{i,t-1}) = \tau + \beta_v v_{i,t-1}$. (5) (Wooldridge 2009).

So equations (1) and (2) are transformed to equations (6) and (7):

$$
\begin{aligned}
y_{it} = {} & (\alpha+\eta) + \beta L_{it} + (\gamma+\pi)K_{it} + \psi m_{it} + \delta L_{it} \\
& + \pi_2 K_{it}^2 + \psi_2 m_{it}^2 + \delta_2 L_{it}^2 + \lambda_1 K_{it} m_{it} \\
& + \lambda_2 K_{it} L_{it} + \lambda_3 m_{it} L_{it} + \lambda_4 K_{it}^2 m_{it} \\
& + \lambda_5 K_{it} m_{it}^2 + \lambda_6 K_{it}^2 L_{it} + \lambda_7 K_{it} L_{it}^2 + \lambda_8 L_{it}^2 m_{it} \\
& + \lambda_9 L_{it} m_{it}^2 + \pi_3 K_{it}^3 + \psi_3 m_{it}^3 + \delta_3 L_{it}^3 + e_{it},
\end{aligned} \tag{6}
$$

$$
\begin{aligned}
y_{it} = {} & (\alpha+\eta+\tau) + \beta L_{it} + \gamma K_{it} + \pi K_{i,t-1} + \psi m_{i,t-1} \\
& + \delta L_{i,t-1} + \pi_2 K_{i,t-1}^2 + \psi_2 m_{i,t-1}^2 + \delta_2 L_{i,t-1}^2 \\
& + \lambda_1 K_{i,t-1} m_{i,t-1} + \lambda_2 K_{i,t-1} L_{i,t-1} + \lambda_3 m_{i,t-1} L_{i,t-1} \\
& + \lambda_4 K_{i,t-1}^2 m_{i,t-1} + \lambda_5 K_{i,t-1} m_{i,t-1}^2 + \lambda_6 K_{i,t-1}^2 L_{i,t-1} \\
& + \lambda_7 K_{i,t-1} L_{i,t-1}^2 + \lambda_8 L_{i,t-1}^2 m_{i,t-1} + \lambda_9 L_{i,t-1} m_{i,t-1}^2 \\
& + \pi_3 K_{i,t-1}^3 + \psi_3 m_{i,t-1}^3 + \delta_3 L_{i,t-1}^3 + u_{it}.
\end{aligned} \tag{7}
$$

The orthogonality conditions for (6) are:

$$E(e_{it} \mid L_{it}, K_{it}, m_{it}, L_{i,t-1}, K_{i,t-1}, m_{i,t-1}, \ldots, L_{i1}, K_{i1}, m_{i1}) = 0, \tag{8}$$

The orthogonality conditions for (7) are:

$$E(u_{it} \mid K_{it}, L_{i,t-1}, K_{i,t-1}, m_{i,t-1}, \ldots, L_{i1}, K_{i1}, m_{i1}) = 0. \tag{9}$$

Equation (6) and (7) can be simultaneously estimated using GMM in Stata.

A2-2. Correlation of Variables in Study 2

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1) ***Y***	1.000						
(2) ***K***	0.878	1.000					
(3) ***L***	0.426	0.396	1.000				
(4) ***LHR***	0.441	0.379	0.583	1.000			
(5) ***Bachelor - LHR***	0.408	0.350	0.521	0.987	1.000		
(6) ***Master - LHR***	0.434	0.373	0.588	0.982	0.955	1.000	
(7) ***Phd – LHR***	0.297	0.252	0.292	0.796	0.811	0.775	1.000
(8) ***Other - LHR***	0.474	0.411	0.658	0.952	0.902	0.920	0.669
(9) ***HighExp - LHR***	0.429	0.368	0.535	0.992	0.981	0.977	0.797
(10) ***LowExp - LHR***	0.431	0.376	0.672	0.924	0.905	0.899	0.711
(11) ***WithoutGap - LHR***	0.403	0.345	0.529	0.990	0.985	0.977	0.815
(12) ***WithGap - LHR***	0.498	0.434	0.671	0.905	0.869	0.876	0.638
(13) ***WithoutGap - LHR (6 months)***	0.407	0.349	0.534	0.992	0.986	0.979	0.816
(14) ***WithGap - LHR (6 months)***	0.496	0.433	0.672	0.876	0.837	0.842	0.598
(15) ***nonLHR***	0.554	0.496	0.858	0.832	0.797	0.823	0.578
(16) ***Bachelor - nonLHR***	0.546	0.491	0.830	0.848	0.819	0.840	0.602
(17) ***Master - nonLHR***	0.522	0.465	0.819	0.870	0.845	0.861	0.645
(18) ***Phd - nonLHR***	0.459	0.406	0.824	0.780	0.743	0.777	0.554
(19) ***Other - nonLHR***	0.559	0.503	0.882	0.757	0.709	0.746	0.483

(20) ***HighExp - nonLHR***	0.559	0.498	0.849	0.827	0.792	0.820	0.574
(21) ***LowExp - nonLHR***	0.518	0.470	0.845	0.808	0.775	0.795	0.563
(22) ***WithoutGap - nonLHR***	0.552	0.491	0.821	0.849	0.820	0.843	0.603
(23) ***WithGap - nonLHR***	0.543	0.493	0.891	0.788	0.746	0.775	0.529
(24) ***WithoutGap - nonLHR (6 months)***	0.552	0.491	0.820	0.850	0.821	0.844	0.605
(25) ***WithGap - nonLHR (6 months)***	0.539	0.491	0.896	0.775	0.732	0.762	0.515
	(8)	**(9)**	**(10)**	**(11)**	**(12)**	**(13)**	**(14)**
(8) ***Other - LHR***	1.000						
(9) ***HighExp - LHR***	0.937	1.000					
(10) ***LowExp - LHR***	0.903	0.869	1.000				
(11) ***WithoutGap - LHR***	0.919	0.991	0.889	1.000			
(12) ***WithGap - LHR***	0.931	0.874	0.913	0.837	1.000		
(13) ***WithoutGap - LHR (6 months)***	0.921	0.991	0.894	0.999	0.845	1.000	
(14) ***WithGap - LHR (6 months)***	0.918	0.843	0.891	0.802	0.991	0.806	1.000
(15) ***nonLHR***	0.840	0.800	0.846	0.798	0.830	0.801	0.818
(16) ***Bachelor - nonLHR***	0.841	0.819	0.852	0.821	0.824	0.823	0.808
(17) ***Master - nonLHR***	0.851	0.843	0.868	0.846	0.834	0.848	0.816
(18) ***Phd - nonLHR***	0.789	0.751	0.794	0.757	0.753	0.758	0.744
(19) ***Other - nonLHR***	0.801	0.720	0.794	0.710	0.803	0.713	0.799
(20) ***HighExp - nonLHR***	0.833	0.802	0.823	0.797	0.816	0.799	0.803
(21) ***LowExp - nonLHR***	0.820	0.763	0.869	0.769	0.828	0.772	0.819

(22) ***WithoutGap - nonLHR***	0.841	0.825	0.842	0.826	0.815	0.828	0.798
(23) ***WithGap - nonLHR***	0.819	0.748	0.834	0.741	0.831	0.745	0.827
(24) ***WithoutGap - nonLHR (6 months)***	0.841	0.826	0.843	0.827	0.815	0.829	0.797
(25) ***WithGap - nonLHR (6 months)***	0.811	0.734	0.826	0.725	0.829	0.729	0.826
	(15)	**(16)**	**(17)**	**(18)**	**(19)**	**(20)**	**(21)**
(15) ***nonLHR***	1.000						
(16) ***Bachelor - nonLHR***	0.993	1.000					
(17) ***Master - nonLHR***	0.986	0.989	1.000				
(18) ***Phd - nonLHR***	0.898	0.885	0.900	1.000			
(19) ***Other - nonLHR***	0.978	0.956	0.941	0.871	1.000		
(20) ***HighExp - nonLHR***	0.995	0.987	0.980	0.895	0.976	1.000	
(21) ***LowExp - nonLHR***	0.970	0.968	0.959	0.867	0.942	0.941	1.000
(22) ***WithoutGap - nonLHR***	0.993	0.993	0.986	0.889	0.960	0.993	0.952
(23) ***WithGap - nonLHR***	0.987	0.972	0.964	0.891	0.981	0.975	0.974
(24) ***WithoutGap - nonLHR (6 months)***	0.993	0.993	0.986	0.889	0.960	0.993	0.952
(25) ***WithGap - nonLHR (6 months)***	0.981	0.964	0.957	0.885	0.978	0.968	0.971
	(22)	**(23)**	**(24)**	**(25)**			
(22) ***WithoutGap - nonLHR***	1.000						
(23) ***WithGap - nonLHR***	0.961	1.000					
(24) ***WithoutGap - nonLHR (6 months)***	1.000	0.961	1.000				
(25) ***WithGap - nonLHR (6 months)***	0.952	0.999	0.951	1.000			