# IT Knowledge Spillovers and Productivity: Evidence from Enterprise Software*

Peng Huang
R.H. Smith School of Business
University of Maryland
College Park, MD 20742
huang@umd.edu

Marco Ceccagnoli, Chris Forman and D. J. Wu
Scheller College of Business
Georgia Institute of Technology
800 West Peachtree Street, NW
Atlanta, GA 30308
{marco.ceccagnoli;chris.forman;dj.wu}@scheller.gatech.edu

**Abstract**

We evaluate the value to firms of participating in an Internet-enabled network of practice in which IT professionals exchange ideas and help each other resolve problems related to the implementation and use of enterprise software. We show that IT spillovers that take place in these online community networks contribute positively and significantly to firm productivity: by our estimates, a one percent increase in inward IT knowledge spillovers translates to a $0.48 million increase in added value for an average firm in our sample. Further, we highlight the limits of the extent to which firms can benefit from participating in these Internet-enabled networks of practice, by showing that the value of spillovers arises primarily from flows of technical rather than business knowledge related to the enterprise software's modules.

*Key words:* information systems; IT knowledge spillovers; productivity; enterprise software; online community networks

* The authors thank Bin Li for outstanding research assistance. We thank Sanjeev Devan, Bin Gu, Vijay Gurbaxani, Paul Hofmann, Karim Lakhani, Nishtha Langer, Mike Miller, Ivan Png, Vandana Ramachandran, Prasanna Tambe, Marshall Van Alstyne, Feng Zhu, and seminar participants at Boston University, Harvard Business School, the University of California Irvine, University of Utah, and the 2012 Workshop on Information Systems and Economics (WISE) for helpful comments, suggestions and advice. We gratefully acknowledge funding from SAP Labs, LLC. We also thank Harte Hanks Market Intelligence for supplying data. All errors are our own.

## 1. Introduction

The existence of knowledge spillovers are fundamental to endogenous growth theory (Aghion and Howitt 1997, Romer 1986) and have significant implications for the productivity of firms (Griliches 1992). Their importance has given rise to a body of literature on the measurement and economic implications of research and development (R&D) spillovers (Alcácer and Chung 2007, Ceccagnoli 2005, Jaffe 1986, Jaffe et al. 1993, Nadiri 1993). More recently, researchers have sought to measure the existence and implications of knowledge spillovers embodied in IT investments, labeled IT spillovers. IT spillovers arise because the implementation of IT investments within firms involves complementary business process innovation that involves changes to work practices and the structure of organizations (e.g., Bresnahan and Greenstein 1996, Bresnahan et al. 2002). Knowledge of how to implement new IT and the accompanying business process innovation can also spillover to other firms in the economy, mediated by, for example, outsourcing transactions (Chang and Gurbaxani 2012a, b) or worker mobility and employment relationships (Tambe and Hitt 2012a).

There is increasing anecdotal evidence that knowledge flows related to the innovation required for IT investments can be mediated through the use of technology-enabled community networks such as online networks of practice (NoPs). Such networks of practice refer to informal social networks that facilitate exchange of information between persons with practice-related goals (Brown and Duguid 2001). By reducing communication and coordination costs among industry participants, and facilitating knowledge discovery and codification, networks of practice hold the potential to increase knowledge flows between industry participants. However, there may exist significant challenges to their effective use. For one, participants may not have the incentives to make adequate investments to make quality contributions (Boudreau et al. 2011, Huang et al. 2011). Moreover, the requisite knowledge is often tacit and may require adaptation for effective reuse, making it more difficult to transfer via IT-enabled means (Zander and Kogut 1995). Despite the uncertainty about the efficacy of these institutions, at present we have little systematic empirical evidence whether firms and their employees benefit from participating in NoPs, and under what conditions they mediate knowledge flows most successfully.[1] Given that technology firms have invested heavily in building such online networks of practices and firm employees have invested significant time and effort participating in them, this is a surprising gap in understanding.

We seek to measure the value of IT spillovers from one such NoP. In particular, we study an Internet-enabled NoP where IT practitioners exchange knowledge and ideas on the implementation and use of enterprise software. Adoption of enterprise software can improve firm financial and operational performance; however its implementation is difficult, with benefits coming with a lag and sometimes not at all because of the technical and organizational adaptations required (Hitt et al. 2002, McAfee 2002). Thus, flows of knowledge on how to implement these systems are potentially valuable. The NoP that we study, the SAP Community Network, has experienced rapid growth: the total number of registered users grew to around 199,000 in its first five years and there have been over 1.1 million discussion threads (Q&A conversations) posted during the same period. On average, about a quarter of all the questions raised by knowledge seekers are solved by the collective efforts of the community users, and the average time it takes to get a correct solution is less than 5 days.

We measure the effects of knowledge inflows through the NoP on the productivity of 275 Fortune 1000 firms that are SAP adopters and find that IT knowledge flows mediated by the NoP reflect IT spillovers. User firms that experience an increase in knowledge inflows through their employees' participation in the community network also experience an increase in productivity, without bearing the cost required to generate the underlying valuable IT knowledge. By our estimates, a one percent increase in the inward knowledge spillovers translates to a $0.48 million increase in added value for an average firm in our sample.

[1] For research that measures the benefits of IT-enabled knowledge sharing within firms, see Aral et al. (2012) and Di Maggio and Van Alstyne (2011).

We find that some types of IT knowledge spillovers are more valuable than others. We differentiate between two types of IT knowledge: technical knowledge and business functional knowledge. The first type of knowledge consists primarily of expertise in technical domains such as operating systems, programming languages, database management systems, networks, and telecommunications, while the second type emphasizes changes to organization and business processes that are enabled by IT. As noted above, these latter changes to business organization and practice are costly and difficult, but are often a major factor influencing whether a firm benefits from its IT investments (Bloom et al. 2012, Bresnahan et al. 2002). However, because they cannot be written out in terms of general principles and must frequently be customized to local environments, they are also more difficult to codify and transfer between firms (Arora and Gambardella 1994). We find that the productivity gains derived from IT knowledge spillovers are primarily driven by technical knowledge flows rather than knowledge flows about business and operational processes. Thus, our results demonstrate that such Internet-enabled networks of practice have significant limitations in aiding firms with the most costly and valuable aspects of new IT implementation.

Our baseline analyses rely on the identification assumption that changes in spillover inflows over time are uncorrelated with changes to other unobserved factors that influence productivity. We probe this assumption through a series of falsification exercises and an instrumental variables analysis. In particular, we show that the timing of productivity gains coincides with the timing of knowledge inflows from the NoP. Further, we show that IT spillovers complement a firm's investment in IT capital, but do not similarly increase the value of other factors of production such as non-IT capital or non-IT labor. Finally, we instrument for IT spillovers using variables that capture the potential benefits of the NoP based on the firm's installed base of application software and longitudinal variation in the introduction of related forums on the SAP Community Network.

The remainder of the paper is organized as follows: Section 2 presents a brief overview of the literature on knowledge spillovers in IT and hypotheses on their relationship to productivity. Section 3 introduces the research context from which our measure of IT spillovers is derived. The data and empirical methods are introduced in Section 4. We present the results of data analyses in Section 5. In Section 6 we summarize the findings and discuss their implications.

## 2. Knowledge Spillovers, Information Technology, and Productivity

The effective implementation of IT within organizations has emphasized the view of IT as an enabler of business process innovation. Business process innovation requires a range of investments in computing hardware and software and in a range of organizational practices. For example, it may involve changes in the discretion given to employees and retraining, as well as changes to the information flows within an organization (Bresnahan et al. 2002). Prior research has emphasized that business process innovation involves co-invention, the post-adoption invention of complementary business processes and adaptations that make investments in IT useful (Bresnahan and Greenstein 1996).

External knowledge flows of how to implement business process innovation are valuable, and can occur through a variety of means. Third party service providers like IT outsourcing firms frequently assist firms with business process innovation, and knowledge of how to implement new systems is frequently mediated through communication between outsourcing providers and their clients (Chang and Gurbaxani 2012b). Knowledge of how to conduct business process innovation can also spillover between firms that face similar technology opportunities. For example, through the acquisition of new IT workers, firms can obtain access to knowledge gained by these workers through their training at their previous employer (Tambe and Hitt 2012a).

The R&D spillovers literature has frequently described how spillovers can by mediated through conversation and collaboration among researchers; however these “in the air” spillovers (Marshall 1890) are difficult to measure directly and so are frequently measured indirectly using geographic or technological proximity (e.g., Furman et al. 2006, Moretti 2004).[2] An emerging new conduit of knowledge flows through which these types of information exchanges can take place is virtual, Internet-enabled networks of practice. Such networks of practice typically consist of a large, loosely knit, group of individuals involved in a shared practice area. Typically they coordinate through third parties such as professional associations (Wasko and Faraj 2005) and their interactions are facilitated through advances in IT-enabled communications. Participation in such networks is typically open and voluntary, and participants are usually strangers (Wasko and Faraj 2005). These platforms have the potential to facilitate knowledge sharing among individuals and, compared to a traditional closed model, can result in faster and higher quality solutions to technical questions raised by users (Füller et al. 2007). Notably, interactions within these communities provide a “paper trail” that researchers can use to directly measure knowledge flows through online collaboration (Fershtman and Gandal 2011).

In keeping with prior research on the impact of knowledge spillovers (e.g., Griliches 1992), we test whether inflows of knowledge influence the total factor productivity of firms. In particular, we investigate whether firms receiving knowledge inflows related to technological or business process innovation practices experience higher total factor productivity (TFP) than firms who do not receive such spillovers.

*H1: Other things equal, firms that receive greater inward IT knowledge spillovers are more productive than those who do not.*

We further investigate the nature of IT spillovers that will generate the greatest productivity benefits for firms. We consider two types of knowledge spillovers. Technical knowledge spillovers are primarily related to a variety of domains such as operating systems, programming languages, and database

[2] Spillovers that are mediated by technological or geographic distance are of two types: those that are embedded in improvements to quality-adjusted prices and those that occur through knowledge flows between market participants. While many studies are unable to identify between these two mechanisms, we believe our setting allows us to focus on the impact of knowledge flows. For further details, see Cheng and Nault (2007), Nadiri (1993), and Griliches (1992).

management systems. In contrast, business functional knowledge is related to the adaptation of business processes and functions in the implementation of information systems and the use of IT to achieve business goals (Ko et al. 2005, Lee et al. 1995).

As noted above, adoption of enterprise IT requires significant changes to business processes and organizational practices. When implementing enterprise IT systems, such changes are widely believed to be more costly than investments in hardware and software and any related technical investments (Brynjolfsson et al. 2006). Further, variation in the extent and manner in which these changes are implemented are widely believed to be responsible for a significant fraction of the variance in returns to IT investment that have been observed in the data (e.g., Bloom et al. 2012, Bresnahan et al. 2002). Thus, acquiring knowledge that facilitates these adaptations has the potential to engender significant productivity improvements for firms. This suggests that the number of problems requiring business functional knowledge, and the value of this type of knowledge to firms adopting enterprise IT, are both high.

However, while the value of business functional knowledge is high, such knowledge is highly context dependent, and so may be less codifiable and portable than technical knowledge. Such "knowledge stickiness" (Li and Hsieh 2009, Szulanski 1996) may be caused by the interactions between the characteristics of the knowledge, the source, the recipient, and the context in which the transmission takes place. In other words, solutions to business functional problems are specific to the idiosyncratic organizational environment in which they are embedded. In contrast, solutions to technical problems are broadly applicable across different organizational environments, and therefore such questions are more easily posed by seekers and addressed by users of the community network. This suggests that the number of questions and answers in online networks of practice will be greater for technical knowledge than for business functional knowledge, and that the benefits of inflows of technical functional knowledge will also be greater than for business knowledge. Thus, our second hypothesis is

*H2: Increases in inflows of technical knowledge will have a greater effect on productivity than will increases in inflows of business functional knowledge.*

## 3. Research Context

Our research setting is the online community network run by SAP AG, the largest enterprise software vendor by revenue. As part of its platform strategy, SAP established in 2003 an Internet-based network of practice, the SAP Developer Network (SDN). The SDN was later expanded to include a community for business process experts, and was expanded still further over time to incorporate other communities that interface with SAP's products. Given this increase in breadth, the SDN is now known as the SAP Community Network (SCN).

The SCN serves as a resource repository and a platform for SAP users, developers, architects, consultants and integrators to collaborate and exchange knowledge on the adoption, implementation and customization of SAP solutions. It hosts forums, expert blogs, a technical library, article downloads, a code sharing gallery, e-learning catalogs, wikis and other facilities through which its users contribute their knowledge. As of 2008 there were over 199,000 active users from 224 different countries in the community.

We choose enterprise software as the background for measuring IT knowledge spillovers for several reasons. First, investment in enterprise software and its implementation accounts for a significant portion of total business-related IT spending (Brynjolfsson et al. 2002) — in some cases accounting for as much as 75% of corporate IT investment (McAfee and Brynjolfsson 2008), and adoption of enterprise software has been shown to be associated with significant improvement in firm financial and operational performance (Hitt et al. 2002). Second, as noted above, the adoption of complex IT platforms such as enterprise software often requires complementary business process innovation.

A unique feature of the SAP Community Network is that its users' knowledge contribution to the community can be quantified. The community has a contributor recognition program (CRP) that awards points to community users for each technical article, code sample, video, wiki contribution, forum post, and weblog authored. For example, when users participate in forum discussion, they can receive points for posting solutions to existing discussion threads marked as questions, if the answer is deemed helpful by the person who asks the question. SAP publicly recognizes its most active contributors. For example, on the "Top Contributors" page, the top 50 contributors measured by total reward points are listed. On each discussion forum page, the top three contributors to that forum are listed with their total points received. In addition, SAP identifies and provides special status to exceptional and high-value contributors by granting them the title of "SAP Mentor". SAP Mentors are offered access to SAP senior management, early access to information on products and programs, and greater visibility in the on-line communities as well as at SAP events such as the SAP Tech Ed conference.

Participation in the community network is completely voluntary and anyone can register as a user by providing basic personal information. In particular, users can optionally list the firm that employs them. Using this piece of information, it is possible to aggregate total knowledge inflows to firms whose employees actively participate in SCN (details on the measurement of knowledge flows will be introduced in the next section). The user's profile also lists the country that the user comes from, her relationship to SAP, email address, phone number, expertise, and LinkedIn profile page. Figure 1 presents a sample user profile.

[Insert Figure 1 Here]

To track knowledge flows between the users of SCN, we focus on user interactions through the most frequently used communication format: the discussion forums. Although users of SCN may access knowledge through other formats such as wikis, blogs, and articles, the number of active participants in these other formats is smaller than in the discussion forums, and knowledge flows arising from their use are unfortunately not measurable. The primary purpose of the discussion forums is to provide an avenue for conversations between the community users to help one another solve problems encountered during the implementation, deployment and use of SAP software (Fahey et al. 2007). The forums are organized according to domains of knowledge or expertise, each of which usually corresponds to a technical domain (such as database or operating system), a particular SAP software module or the application of SAP to a particular industry. Examples of forums include SAP on SQL server, data transfers, ERP manufacturing, product life cycle management, CRM-interaction center, and SAP for automotive solutions.

Conversations in each forum are organized by discussion thread. Each thread is initiated by a knowledge seeker, who posts a specific question in a topic forum of her choice. Knowledge contributors, on the other hand, post responses that are attempts at answering the question. A discussion thread is comprised of a list of *messages*, and each message (either a question or an answer attempt) contains the information about the member who posts the message, the body of the message, and a time stamp. Once a correct answer (determined by the knowledge seeker) is received, the discussion thread is closed. Figure 2 presents a sample discussion thread in the forum of SAP ERP Manufacturing - Production Planning, with a question, a correct answer, and a helpful answer.

[Insert Figure 2 Here]

We developed a web scripting tool and obtained the complete history of SAP forum discussions from 2004 to 2008. The dataset includes about 1.1 million discussion threads with 5.0 million messages posted in 209 topic-specific forums. Table 1 presents some summary statistics of the evolution of the SCN over our sample period, including numbers of registered users, topic forums and the discussion threads posted in these forums. Overall, we find that the online community has experienced rapid growth since its establishment. Further, our data are consistent with prior anecdotal evidence that online networks of practice are an effective means of using the "wisdom of the crowd" for peer support (Doan et al. 2011):

by the end of our sample roughly one quarter of all questions raised are solved by the collective effort of the community users, and the average time to obtain a correct solution is less than five days.

[Insert Table 1 Here]

## 4. Data and Methods

### 4.1. Estimation models

We adopt a production function approach and extend it by introducing our measure of IT-related spillovers. A typical production function relates firm output to factors of input (Hulten 2001). For example, a simple form of a three-factor Cobb-Douglas production function has been widely used in prior studies on IT productivity (Brynjolfsson and Hitt 1996, Dewan and Min 1997, Mittal and Nault 2009):

$$Y = AK^{\alpha}L^{\beta}C^{\gamma} \tag{1}$$

Where *Y* is the quantity of production output, *K* is the stock of non-IT capital, *L* is the stock of labor, *C* is the stock of IT-capital, and *A* denotes the total factor productivity (TFP). TFP is defined as the output contribution that is not explained by the factor inputs and is often interpreted as technological progress. In this case, the output elasticity of IT-capital, $\gamma = \partial lnY/\partial lnC$, represents the percentage increase in output due to a one percent increase in IT capital. To incorporate the role of knowledge spillovers, we follow the literature on R&D spillovers by adding to (1) a term that captures the knowledge capital stock obtained from elsewhere (e.g., Chang and Gurbaxani 2012a, Griffith et al. 2006):

$$Y_{it} = Ae^{(\theta lnS_{it})}{K_{it}}^{\alpha}{L_{it}}^{\beta}{C_{it}}^{\gamma} \tag{2}$$

where *S* denotes the measure of inward IT-related knowledge spillovers. In this equation, *i* and *t* index firm and time period, respectively. The Cobb-Douglas production function can be employed to estimate the factor productivities by implementing the following stochastic model:

$$lnY_{it} = a + \alpha lnK_{it} + \beta lnL_{it} + \gamma lnC_{it} + \theta lnS_{it} + \epsilon_{it} \tag{3}$$

OLS estimates of IT spillover effects are likely to suffer from unobserved firm heterogeneity that is correlated with inward spillovers. To address this issue, we introduce a set of firm- and year- fixed effects. This amounts to the modification of (3)

$$lnY_{it} = a + \alpha lnK_{it} + \beta lnL_{it} + \gamma lnC_{it} + \theta lnS_{it} + \mu_i + \delta_t + \epsilon_{it}, \tag{4}$$

Our choice of the estimators follow closely the mainstream methods used in the IS literature (e.g., Tambe and Hitt 2012b). In particular, our relatively short panel does not permit us to employ recent methods that use lagged inputs or outputs to achieve identification (e.g., Ackerberg et al. 2006, Arellano and Bond 1991, Levinsohn and Petrin 2003).

Our initial identification relies on the assumption that changes in spillover inflows are uncorrelated with other unobserved factors that may be changing over time and influencing productivity. A particular concern with our measure of knowledge spillovers is that there may exist other forms of knowledge inflows that are unmeasured and correlated with our main variable, such as knowledge inflows obtained by the employees of the firm through active search of existing discussion threads, reading articles on wikis and blogs, or downloading sample code, which could lead to inflated estimate of the spillover effect measured by forum activities.

Cognizant of the challenges of obtaining identification in our setting, we provide a variety of evidence that is suggestive of a causal relationship between knowledge inflows and productivity. First, we control for a variety of alternative measures of activity in the SAP Community Network, such as the number of registered users employed by a firm in SCN, and the number of questions posted by them. Second, we present instrumental variable regressions that use information on the software modules installed by a firm, and the timing of introduction of new forums related to SAP modules as instruments for the stock of IT spillovers. As we discuss in further detail below, the introduction of new forums related to SAP modules that the firm is using should increase spillover inflows, but should be uncorrelated with other unobserved factors—including other knowledge inflows—influencing productivity. Last, we present the results of a falsification exercise in which we examine whether the productivity benefits of knowledge inflows exhibit the "right" timing, by showing that leads of our spillover term have no productivity benefits, and these benefits also display the "right" type of complementarity with other types of production inputs: they enhance the value of IT investments, but not non-IT capital or labor.

**4.2. Data**

We construct a dataset of publicly traded firms that are SAP adopters. Our data come from a variety of sources. To identify SAP adopters, we obtained a detailed list of all installations of SAP product modules in the United States prior to the end of year 2004 from SAP AG. We use the Harte Hanks Computer Intelligence (CI) Technology database to collect firm-level IT investment data. The CI database records detailed information about IT infrastructure for most of the Fortune 1000 firms, including data on the quantity of mainframes, peripherals, minicomputers, servers and PC systems, as well as other IT hardware stocks. The CI database has been widely used by prior studies to investigate issues related to IT productivity (e.g., Brynjolfsson and Hitt 2003, Chwelos et al. 2010, Dewan et al. 2007). The CI data were matched with Standard and Poor's Compustat database to obtain financial data that we use to construct measures of production output, non-IT capital stock and labor expenses.

**4.2.1. Sample**

We use several steps to define the sample. Our sample begins in 2004, when the SCN started, and ends in 2008, which is the last year for which we have IT data.[3] To obtain the data for our sample, we first retrieve the set of firms that were on the Fortune 1000 list at least once during 2004-2008 and match them to Compustat data. We then match these firms with the CI database. Because we are interested in

[3] While SCN originated in 2003, it was re-launched in 2004 with a new reward system so we use 2004 as the starting point for our analysis.

knowledge spillovers related to the implementation and the related business process innovation in SAP products, we further restrict our sample to those firms that had installed at least one SAP module prior to the end of 2004. The final sample is an unbalanced panel of 275 firms with 1240 observations over a 5-year period.

### 4.2.2 Variables

*IT Knowledge Spillovers*

We measure inward IT-related knowledge flows, $S_{it}$, from forum conversations that took place on the SAP Community Network. For each question that is posted, the rules of the SAP reward program specify that the knowledge seeker can use her discretion to judge the quality of answers posted by knowledge contributors and distribute reward points as follows: 10 reward points for correct answers (at most 1 answer can be evaluated as correct), 6 points for very helpful answers (at most 2 answers), and 2 points for helpful answers (no limit on number). We define a knowledge inflow as an incident where a knowledge seeker gives reward points to knowledge contributors in recognition of their quality responses. As noted above, we use a crawler program to identify user information such as location and firm. Next, we select all the users that reside in the United States, and match them to firms in our sample by examining their employer affiliations and domains of their email addresses.

For each user *a* who is an employee of firm *i*, we retrieve all the discussion threads that were initiated by *a* in year *t*, and examine the history of the answers posted by other forum users. If *a* received any correct, very helpful, or helpful answers in year *t*, the total number of reward points she gave to the knowledge contributors are used as a proxy for inward IT spillovers to *a*. The reward points are then aggregated across all the threads posted by *a* in year *t* to derive an individual-level knowledge inflow, $S_{at}$. The firm level spillover variable is defined as the sum of knowledge inflows of all the individuals who are employees of the firm:

$$S_{it} = \sum_{a \in F_i} S_{at}$$

where $F_i$ is the set of users who are employees of firm *i*.

While we note that our measure of knowledge inflow is likely to suffer from measurement error due to missing data on the knowledge seekers who did not report their employers,[4] such measurement error would most likely result in attenuation biases. In addition, we observe no systematic differences in knowledge inflow between questions asked by knowledge seekers who reported their employers and those asked by seekers who didn't reveal their employers: the average inflow per question per year is 2.91 for non-reporting seekers and 3.08 for reporting seekers, and the difference is not statistically significant (p=0.35). To further address the possible estimation biases introduced by this measurement error, we also use instrument variable method in which we instrument knowledge inflow using variables that are unlikely to be correlated with the measurement errors.

*IT Capital*

Our measure of IT capital is derived from the CI database. The information in the database covers major categories of IT hardware investments made by firms, such as personal computing, systems and servers, networking, software, storage and managed services. Historically the CI database has provided direct measures of IT capital stock, but this measure is not available over the years of our sample. As a proxy, we adopt the method used by Brynjofsson and Hitt (1995), Hitt and Brynjofsson (1996), Dewan and Min (1997), and Gu et al. (2008) and measure the IT capital stock using an estimate of the market value of the IT hardware systems plus three times the current year IT labor expenses. Inclusion of IT labor expense in the calculation of IT capital is justified by the fact that a large fraction of IT labor expenses is dedicated to the development of computer software, which is a capital good. The assumption that underlies this

[4] For example, among all the discussion threads that are initiated by US knowledge seekers during our sample period, only 48% (23,973 out of 49,977) of them have a seeker who reported her employer.

method is that the current IT labor spending is a good proxy for the IT labor expenses in the recent past, and IS staff "stock" depreciates fully in three years (Brynjolfsson and Hitt 1995).

The first component of this variable is equal to the market value of total PCs and servers currently owned by the firm, converted to constant 2005 US dollars. To be specific, we collect market prices of PCs and servers in the United States from two report series produced by the *Gartner Dataquest Market Statistics* database—Gartner Worldwide Server Forecast and Gartner Worldwide PC Forecast—from 2004 to 2008. These two report series present detailed statistics on the number of shipments, prices, vendor revenues and other related information about PC and Servers, broken down to the level of each geographic region and market segment. [5] Our market prices for PCs and servers are calculated as the average user price across their respective market segments within the United States. These prices are then multiplied by the quantities of PCs and servers owned by the firm to derive the market value of the IT computer assets for each firm. Finally, we deflate the market value by the Bureau of Economic Analysis (BEA) price index for computers and peripheral equipment.

The second component of IT capital stock is IT-related labor expenses. The CI database provides the number of IT employees of the sample firms at the site level, where a site represents a particular firm-location, much like the concept of establishment in Census data. We aggregate the site-level employee numbers to the firm level to derive the total number of IT-related employees hired by the firm.[6] IT labor prices are obtained from the Occupational Employment and Wage Estimates series from the Bureau of Labor Statistics (BLS) Occupational Employment Statistics (OES), and we use the mean annual wage of

[5] Gartner Dataquest defines PC market segments as: desk-based, mobile, professional, and home. Server market segments are defined by CPU types, which include x86, IA64, RISC, and other. The database covers the global regions of Asia/Pacific, Eastern Europe, Latin America, Middle East & Africa, and Western Europe. Several country level statistics are also available, including the United States, Canada, and Japan.

[6] The CI database actually records a range of IT employees at each site. The ranges are defined as: 1-4, 5-9, 10-24, 25-49, 50-99, 100-249, 250-499, and 500 or more. For each range, we take the middle value of the range as the number of IT employees.

computer and mathematical occupations as the average labor price for IT employees. As the wage reported by the OES series does not reflect benefits, we multiply the wage number by the ratio of total compensation to salary, which is obtained from BLS Employer Costs for Employee Compensation (ECEC) series. The IT labor expense is then deflated by the BLS Employment Cost Index (ECI) for private industry workers.

*Production Output*

We follow prior literature (Brynjolfsson and Hitt 2003, Dewan and Min 1997) and use added value as the measure of production output, which equals deflated sales less deflated materials. Compared to sales, added value is said to be less noisy and more comparable across industry sectors (Dewan and Min 1997). Annual sales numbers are retrieved from Compustat, and we deflate them using industry-specific (at 2-digit NAICS sector) price deflators from BEA *Gross Output and Related Series by Industry*. Materials are calculated by subtracting undeflated labor and related expenses (Compustat data item XLR) from undeflated total operating expenses (Compustat data item XOPR), and deflating by the BLS Producer Price Index (PPI) for intermediate materials, supplies, and components.

*Non-IT Capital*

The calculation of total capital stock is similar to that in Brynjofsson and Hitt (2003) for ordinary capital. Specifically, the gross book value of capital stock (property, plant and equipment (Total-Gross), Compustat data item PPEGT) is deflated by an industry-specific capital investment deflator reported in BLS 1987-2010 Detailed Capital Measures.[7] In order to apply the deflators, the average age of capital stock is calculated as the ratio of total accumulated depreciation (Compustat data item DPACT) to current depreciation (DP). We then subtract the deflated computer capital from deflated total capital to get the value of non-IT capital.

[7] Retrieved from http://www.bls.gov/mfp/mprdload.htm.

*Non-IT Labor*

Consistent with prior studies on IT productivity (Bresnahan et al. 2002, Brynjolfsson and Hitt 2003), total labor expense is either obtained directly from Compustat labor and related expenses (data item XLR), or calculated as the product of a firm's reported number of employees (Compustat data item EMP) and industry-average labor cost per employee, and deflated by the BLS Employment Cost Index (ECI) for private industry workers. Average labor cost per employee is obtained from *National Sector NAICS Industry-Specific estimates* series of BLS OES. To account for the fraction of benefits in total compensation, we multiply the wage number by the ratio of total compensation to salary, which is obtained from BLS *Employer Costs for Employee Compensation* (ECEC) series. Non-IT labor is defined as the difference between deflated total labor expense and IT labor expense.

Table 2 reports the summary statistics of the variables. The average firm in the sample has sales of $16.65 billion, added value of $5.49 billion, and 41,167 employees, consistent with our sample being large publicly-traded, Fortune 1,000 SAP adopters. In addition, firms in our sample invest heavily in IT capital, which has a mean level of $97.49 million and maximum of $1.18 billion. Table 3 provides the correlation matrix among the key variables.

[Insert Table 2 and Table 3 Here]

Table 4 presents a breakdown of the sample firms by vertical industries, which is based on 2-digit NACIS sectors. It is notable that firms in manufacturing industry account for the majority (66%) of the sample, followed by utilities firms (8%).

[Insert Table 4 Here]

## 5. Results

In this section we first establish a link between inward IT spillovers and productivity through a series of panel data models. We next probe the assumptions of our baseline model through a series of robustness checks and instrumental variables analyses. Last, we examine the difference in results for technical and business functional knowledge spillovers.

*Baseline Spillover Analyses*

We consider the role of IT knowledge spillovers in driving variations in total factor productivity. In Columns 1 and 2 of Table 5 we report the result from fixed effects and random effects models, respectively. The coefficients for the spillover term are significant in both models ($p<0.05$), indicating that firms with greater inward IT spillovers produce more output, given the same amount of investment in capital, labor and IT. Particularly, results from the fixed effects model imply that a one percent increase in the amount of inward spillovers is associated with 0.009 percent increase in the added value produced by a firm. Considering that the added value of an average firm in our sample is $5.491 billion, this translates into a $0.48 million increase in production output.

In columns 3 and 4 we explore alternative ways of measuring our IT spillovers variable. One potential issue with our baseline spillover measure is that knowledge seekers may not pay enough attention to the answers posted, or simply lack the expertise to judge the quality of the answers, leading to mismeasurement of the spillover variable. Moreover, we may mismeasure the magnitude of spillovers if a knowledge seeker rewards too many knowledge contributors by marking their posts as helpful (since, unlike correct answers and very helpful answers, there is no limit on the number of helpful answers per thread).

Such measurement errors, if they exist, would most likely result in an attenuation bias in the estimates.[8] However, to assess how measurement error might influence our results, we perform two separate

[8] In addition, we note that while providing feedback is not mandatory for knowledge seekers, SAP does give incentives encouraging them to do so: the knowledge seeker also receives 1 point when rewarding knowledge contributors (limit to 1 point per thread). Further, to the extent that seekers lack the ability to judge the quality of the answers, forum moderators (usually an SAP employee) have the authority to assign reward points on behalf of the seeker if an answer is deemed correct by the moderator.

analyses. First, we calculate the spillover variable using reward points from only correct and very helpful answers. We present the result of this analysis in Column 3 of Table 5. Second, we construct the spillover variable by simply counting the number of questions that are resolved (questions that received either a correct answer or at least a very helpful answer). The result is presented in Column 4 of Table 5. We find that the findings are robust to different ways of measuring the spillover variable.[9]

As noted earlier, a central identification assumption for our analysis is that there exist no unobserved factors that are correlated both with inward IT spillovers and firm productivity. In the reminder of this section, we probe the validity of this assumption by adding additional controls, through instrumental variable analysis, and through falsification exercises.

We first probe whether our inward spillover variable reflects the value of other kinds of unobserved firm activity in the SCN. To proxy for other kinds of firm activity, in Column 5 we add a control for the cumulative number of registered users who are the seeking firm's employees in the SCN, and in Column 6 we present model results that explicitly include the total number of questions that are raised by a firm's employees (recall that a question is usually the first message that initiates a discussion thread) in a year as a control. In Column 7 we present a model that incorporates both the number of registered users and number of questions as controls. In all cases, the coefficient of our inward spillover variable remains positive and economically and statistically significant. Thus, even when controlling for other types of activity in the network, there appears to be something about inward IT spillovers that is positively affecting productivity.

[Insert Table 5 Here]

*Instrumental Variable Analysis*

[9] To further alleviate concerns over measurement errors in our spillover variable, we also perform a test in which we remove the outliers in the sample (observations with a spillover value among the top 1%). Our findings are robust to all these additional analyses.

Although the use of firm fixed effects in the production function estimation is helpful in teasing out the effects of time-invariant, unobserved firm heterogeneity, our spillover estimates may still suffer from omitted variable bias arising from changes in unobserved firm characteristics over time. Particularly, firm productivity may be influenced by unobservable factors that are correlated with our measure of inward knowledge spillovers, such as changes in the complexity of a firm's IT architecture that drives up the need for support related to SAP products, variation in the maturity of the installed SAP modules, how well these modules work with other existing IT solutions in the user environment, or the degree to which the existing business processes need to be modified. While it is impossible to control for all these unobserved factors, we use instrumental variables regressions to increase our confidence in establishing a causal relationship between IT spillovers and productivity.

As with any instrument, we seek to find variables that exogenously shift the likelihood of receiving inward knowledge spillovers and that at the same time are uncorrelated with the unobserved error component mentioned above. To find such variables, we make use of the SAP installation data and take advantage of the variations in the different product modules installed by the firms in our data.[10] The logic behind our instrument is that firms with different product modules may derive varying degrees of benefits from the use of particular SCN forums. Moreover, the number and type of forums has increased over time. This variation yields our first instrumental variable: the number of forums that are useful for a firm in a particular year.

To construct this variable, we observe that the online forums are usually organized so that each forum is dedicated to a specific topic, which often corresponds to an SAP product module or an industry solution. In addition, SAP introduced the forums gradually over time, with only 57 forums in 2004 and 209 forums in 2008 (see details in Table 1). To identify which forums are most useful for a firm, we create a mapping

[10] We thank SAP for providing this data. Some of the most frequently installed SAP product modules include financial accounting (FI), materials management (MM), business intelligence and data warehouse (BIW).

table that associates each SAP product module with the most relevant forum.[11] Using the modules installed at each firm in 2004, we compute a firm-to-forum correspondence and measure the number of forums that are relevant to a firm in a particular year. Finally, we create a binary variable *high_forum* that is set equal to 1 if the number of relevant forums is greater than the sample median and to 0 otherwise. Since the forums are introduced at different times, this variable is time-varying. In addition, the introduction of the forums reflects SAP decisions that are exogenous to the characteristics of the firms that may drive changes in the firm's productivity.

Our second instrumental variable is derived from the characteristics of firms that belong to the same industry as the focal firm. Firms in the same industry often face similar technology opportunities and deploy similar IT infrastructures to perform similar business functions and processes (Tambe and Hitt 2012a), and therefore tend to be exposed to similar knowledge spillovers. For each firm-year observation, we calculate the average inward knowledge spillovers of all other firms in the same industry (defined using 3-digit SIC codes) in that year, and use it as our second IV. This variable is expected to be correlated with the focal firm's inward knowledge spillovers, but is unlikely to be correlated with other unobserved firm characteristics that affect productivity.

We run an instrumental variable model with fixed effects using these two IVs, and present the results in Column 1 of Table 6. We find that our main result is robust to this test. The first stage Wald F-statistic is 164.05, well beyond 10, the suggested rule of thumb, and the Stock-Yogo weak identification test critical value of 19.93 (at 10% maximal IV size), indicating that our instruments are not weak. In addition, Hansen J statistic does not reject the null that our overidentification restrictions are valid (p=0.6197). A Hausman test comparing our IV regression with baseline fixed effect regression (Column 1 of Table 5) indicates that we cannot reject the null that the differences in the coefficients are not systematic

[11] The mapping table is available upon request.

(p=0.972). In sum, while in themselves these results are not conclusive, they are supportive of a link between inward IT spillovers and productivity.

*Falsification Tests*

To increase the confidence in our interpretation of the regression results, we conduct a falsification test where we examine whether the timing of the effects of the spillovers are appropriate. If the benefits of spillovers appear before they should - in other words, prior to a firm actually receiving spillovers - then this suggests our results are affected by omitted variable bias. In Column 2 of Table 6 we report a model in which we add a variable that is equal to the spillovers in the next period. If our interpretation is correct, we expect that the future value of spillovers should have no effect on the productivity of the current year. This is indeed the result we find. While any conclusion that we can draw from this model is tentative due to the short length of our panel, these results provide an additional piece of evidence in support of our interpretation of the results.

Second, to the extent that our spillover variable measures knowledge about the implementation and use of enterprise software, a major component of a firm's IT investment, we expect that there is a synergy between IT capital and IT spillovers, in the sense that the output elasticity of IT capital would be higher if a firm receives greater inward spillovers, due to more effective use of its existing IT infrastructure. To examine this assertion in the data, we interact the IT capital variable with the spillover term and enter it into the fixed effects model. The result is presented in Column 3 of Table 6. The coefficient estimate of the interaction term is indeed positive, consistent with the presence of complementarity between IT capital and IT knowledge spillovers. For the purpose of comparison, we also interact the spillover variable with non-IT capital and non-IT labor, and enter them into the fixed effects model. The results from Column 4 of Table 6 suggest that no such complementarity exist for non-IT capital or non-IT labor.

[Insert Table 6 Here]

*Technical vs. Business Functional Knowledge Spillovers*

One open question in the results discussed so far is what type of IT spillover is most valuable to firms. In the context of IT knowledge, prior research has emphasized two distinct dimensions of knowledge that are particularly relevant in the process of adopting an information system: technical and business functional knowledge. As noted above, technical knowledge is primarily related to a domain such as operating systems, programming languages, database management systems, networks, and telecommunications. In contrast, business functional knowledge is related to the adaptation of business processes and functions in the implementation of information systems and the use of IT to achieve business goals (Ko et al. 2005, Lee et al. 1995).

We construct separate measures of technical knowledge spillovers and business functional knowledge spillovers by examining the forum in which a knowledge seeker's question is raised. We define a forum as technical-oriented if the forum is dedicated to topics related to low-level, enabling technologies of an enterprise system such as programming languages, database technologies, data transfer issues, and reporting and formatting tools. In contrast, we define a forum as business-oriented if the discussion topics in the forum focus on the configuration of the enterprise system to implement a particular business function or process, such as monitoring employee performance, coordination of supply chains, consolidating procurement processes, or managing projects.[12] The classification process is done by independent coding from two of the authors who have conducted prior research related to SAP enterprise software, and two research assistants who have experiences using SAP software products. The rare inconsistencies in the classification results across the coders (6 out of 209) were resolved by checking the content of the forums and discussions among them. We remove forums for which we do not have

[12] Some examples of technical-oriented forums are: Java Programming, Form Printing, SAP on SQL Server, Service-Oriented Architecture, and Data Transfers. Some examples of business-oriented forums are: Logistic Materials Management, Sales and Distribution General, ERP Operations – Quality Management, Knowledge Management & Collaboration, and Product Lifecycle Management.

sufficient information to make a classification, or those for which the classification scheme does not apply (for example, forums such as career center, suggestions and comments, etc.). In total, 16 forums out of 209 are excluded for these reasons.

Preliminary statistics reveal some interesting contrasts between the two categories of knowledge that the firms receive from SCN. First, we notice that there is an asymmetry in the demand for the two types of knowledge: there are fewer forums that are dedicated to topics related to business functions (77) than those dedicated to technical questions (116). In addition, during our sample period community users raise more technical questions (658,574) than business functional questions (453,271). Second, on the supply side of the knowledge exchange, we observe that business-related questions in general receive fewer responses from knowledge contributors (4.55 vs. 4.81, $p<0.01$), are more difficult to solve (with the probability of obtaining a correct answer being 0.226 vs. 0.262, $p<0.01$), and therefore the knowledge seekers of business-related questions in general receive fewer inward spillovers (reward points per question being 5.33 vs. 5.72, $p<0.01$). As a result of the interaction of these demand-side and supply-side factors, firms in our sample on average receive more technical knowledge spillovers (mean=6.75 reward points) than business functional knowledge spillovers (mean=3.75 reward points).

To explore the extent to which these two types of knowledge spillovers contribute to productivity gains, we enter technical and business functional knowledge spillovers as separate measures into the productivity regression, and the results are presented in Column 1 and Column 2 of Table 7, respectively. We also run a regression in which both technical spillovers and business functional spillovers are entered together, and present the results in Column 3 of Table 7. In Columns 4 and 5 of Table 7, we show the results when the effects of number of registered users and total number of questions associated with a firm-year are controlled for, respectively. The results consistently show that the productivity benefit that a firm derives from the SAP Community Network mainly comes from technical knowledge spillovers, rather than business functional knowledge spillovers. For example, the results of Column 3 indicate that while the estimated output elasticity of technical knowledge spillover is 0.011 ($p<0.1$), the output elasticity of business knowledge spillover is much lower (0.001) and not statistically different from zero.

To address concerns about potential omitted variable bias, we instrument technical and business functional knowledge spillovers using *high_forum*, industry average technical spillovers, and industry average business functional spillovers as instrumental variables. The first stage Wald F-statistic, 74.38, is well beyond 10, the suggested rule of thumb, and is greater than the Stock-Yogo test critical value of 13.43 (at 10% maximal IV size), rejecting the null hypothesis of weak instruments. In addition, the Hansen J test (with p=0.6020) does not reject the null that our overidentification restrictions are valid. The second stage estimates for business and technical spillovers are similar to those for the regression models without instruments in Column 5. Combined with the observation that business functional questions attract fewer responses and are more difficult to obtain a solution for, our results suggest that business functional knowledge is less codifiable and portable and is more difficult to be transferred to a different business configuration to generate economic benefits.

[Insert Table 7 Here]

## 6. Conclusions and Discussion

In this study, we find that knowledge inflows from other participants of an online community of practice on enterprise software are associated with a significant increase in firm productivity. Further, these knowledge inflows increase the productivity of a firm's own IT investments. However, we find that the benefits of these knowledge inflows occur only from forums related to technical topics such as operating systems or databases. We find little evidence of benefits from spillovers related to business functional knowledge.

Our results demonstrate that IT-enabled platforms can play an important role in mediating knowledge flows that can increase productivity. While prior work has suggested that firms who are located in highly innovative places may have an advantage in obtaining access to productivity-enhancing knowledge flows (e.g., Furman et al. 2006, Moretti 2004), our work suggests that even firms that are located in more economically isolated locations may be able to receive knowledge spillovers through online platforms. In other words, our study suggests that information technology has the capability to bridge the distance created by geography. However, our findings show that the mechanism we study also has significant limitations—they are unable to facilitate knowledge flows in the types of business process innovation that are most costly and most valuable to implement. This suggests that online platforms such as IT-enabled networks of practice may be only a partial substitute for more traditional spillover channels such as geographic proximity or networks of social relationships.

Most of the existing studies that assess the contribution of IT spillovers to firm productivity have measured spillovers by constructing an aggregate pool of external IT investments in related industries (e.g., Cheng and Nault 2007, 2012).[13] While this method is valuable, it has the potential to produce an upward bias on the estimated effects of spillovers (Tambe and Hitt 2012a). In contrast, in this study we introduce a new method of measuring IT-related spillovers by directly observing knowledge exchange. This allows us to create a "paper trail" for IT spillovers in much the same way that patent citations are used to measure spillovers in the R&D literature, and is less likely to be affected by the potential estimation bias that arises from the pooled approach. Related, while prior research on spillovers has highlighted the distinctions between rent spillovers and knowledge spillovers (Griliches 1979, 1992), studies of IT spillovers that use the pooled method face challenges in identifying between these mechanisms due to the way that the spillover variable is constructed. Our method allows us to isolate the effects of knowledge spillovers on firm productivity.

Our results suggest a potential benefit to firms who participate in the types of IT-enabled networks that we study. However, we note two limitations that may cause us to underestimate the benefits of participation. For one, participants who receive larger knowledge inflows from question-and-answer

[13] For exceptions, see Chang and Gurbaxani (2012a, b) and Tambe and Hitt (2012a).

forums will also be those who ask more questions; thus, there may be a negative selection effect because firms who ask more questions may also be less capable and less productive.[14] Second, our results reflect only the short-run response to the IT-enabled network over a five year window. Participants in the network are likely still learning how to use the network most efficiently to receive knowledge inflows, and in what ways they will contribute knowledge to the network.[15] Thus, future spillover benefits may be higher than what we measure here. However, it is very likely that the differences between technical and business functional knowledge are likely to persist. The challenges of representing business process knowledge in terms of abstract and general principles that can be easily transferred across firm boundaries are well known (e.g., Arora and Gambardella 1994), and researchers and practitioners have made slow progress in overcoming them.

The knowledge flows we study are spillovers—the social returns to the investment and implementation of enterprise IT exceed the private returns because of knowledge flows that take place over SCN. A common feature of settings with spillovers is underprovision of investment. In our setting this can take place if firms do not internalize the externalities of their investment in IT. Knowledge flows will also be smaller if participants have insufficient incentives to contribute their knowledge to the community. Future work should study the effects of participation in NoPs on firms' IT investment decisions, and research participant incentives to contribute knowledge to the network. Huang et al. (2012), who study cross-country differences in incentives to contribute to a NoP, is an important step in this direction. We hope that our research will stimulate this and related work in this important area.

[14] For further details on this, see Di Maggio and Van Alstyne's (2011).

[15] For example, in some networks of practice, the number of knowledge providers has been declining over time as potential contributors learn about their ability to contribute to the community (Huang et al. 2011).

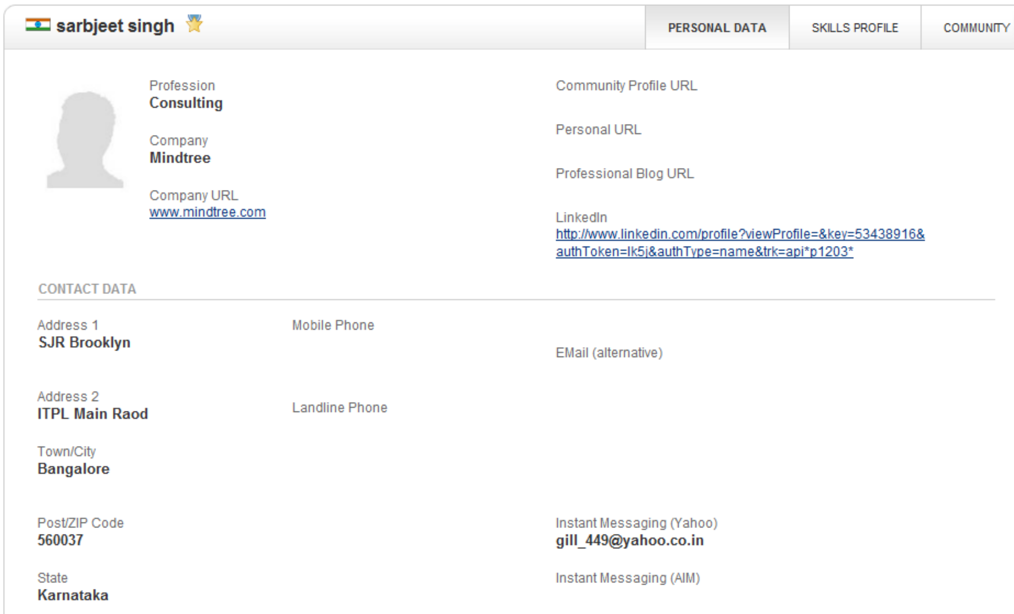


**Figure 1: Sample User Profile**

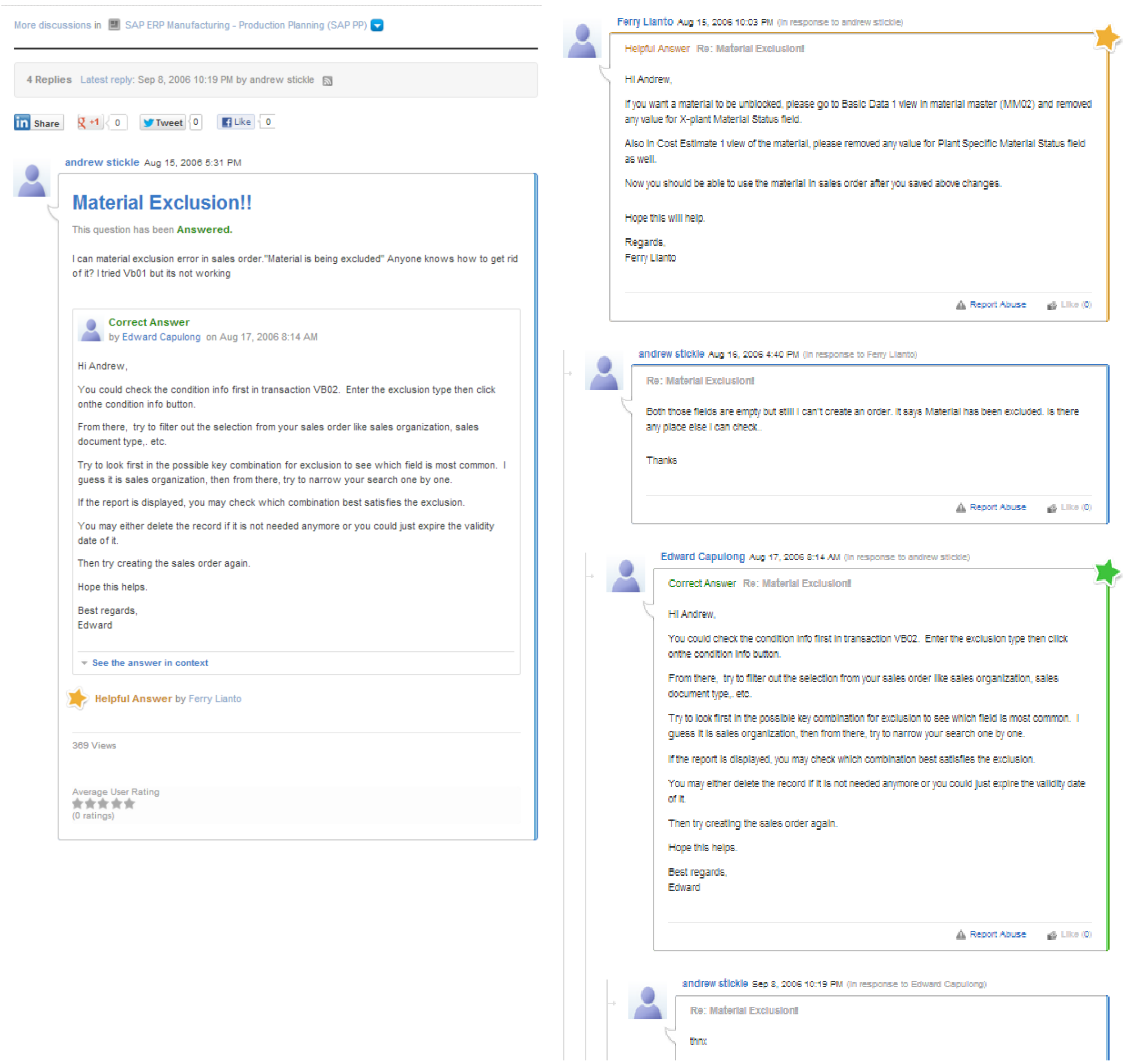


**Figure 2: Sample Discussion Thread**

**Table 1 Evolution of SAP Community Network**

| Year | Cumulative number of registered users | Number of active forums | Number of new threads initiated in the year | Average number of replies for threads initiated in the year | Percentage of questions solved | Percentage received helpful answers | Percentage received very helpful answers | Number of days until correct answer |
|---|---|---|---|---|---|---|---|---|
| 2004 | 19,289 | 57 | 16,296 | 4.679 | 0.107 | 0.073 | 0.098 | 13.378 |
| 2005 | 43,226 | 83 | 67,225 | 5.394 | 0.244 | 0.271 | 0.295 | 4.735 |
| 2006 | 80,981 | 141 | 176,422 | 5.160 | 0.242 | 0.293 | 0.314 | 3.359 |
| 2007 | 137,552 | 179 | 394,183 | 4.731 | 0.227 | 0.260 | 0.287 | 4.219 |
| 2008 | 198,975 | 209 | 463,740 | 4.625 | 0.252 | 0.255 | 0.256 | 4.512 |

**Table 2 Summary Statistics**

| Variable | Obs | Mean | Std. dev | Min | Max |
|---|---|---|---|---|---|
| Annual sales (million $) | 1240 | 16649.51 | 33024.88 | 298.9129 | 364392.4 |
| Added value (million $) | 1240 | 5491.165 | 8811.336 | 118.1137 | 73242.29 |
| Non-IT capital (million $) | 1240 | 12526.01 | 29661.25 | 48.44337 | 321772.7 |
| IT capital (million $) | 1240 | 97.48961 | 138.2879 | 0.000984 | 1181.667 |
| Non-IT labor expense (million $) | 1240 | 2781.929 | 4405.866 | 28.75259 | 40586.13 |
| No. of employees (thousands) | 1238 | 41.16707 | 59.58815 | 0.658 | 428 |
| IT spillover (reward points) | 1240 | 10.71452 | 93.82237 | 0 | 2190 |

**Table 3 Pearson Correlation Matrix of Selected Variables**

| | Variable | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|---|
| 1 | Annual sales | 1.000<br>(-) | | | | | | |
| 2 | Added value | 0.8612<br>(0.000) | 1.000<br>(-) | | | | | |
| 3 | Non-IT capital | 0.8313<br>(0.000) | 0.769<br>(0.000) | 1.000<br>(-) | | | | |
| 4 | IT capital | 0.3398<br>(0.000) | 0.4847<br>(0.000) | 0.3327<br>(0.000) | 1.000<br>(-) | | | |
| 5 | Non-IT labor expense | 0.5479<br>(0.000) | 0.8101<br>(0.000) | 0.3659<br>(0.000) | 0.5767<br>(0.000) | 1.000<br>(-) | | |
| 6 | No. of employees | 0.5236<br>(0.000) | 0.757<br>(0.000) | 0.3716<br>(0.000) | 0.5557<br>(0.000) | 0.9335<br>(0.000) | 1.000<br>(-) | |
| 7 | IT spillover | 0.0166<br>(0.5594) | 0.0443<br>(0.1189) | -0.0014<br>(0.9605) | -0.0023<br>(0.9362) | 0.0593<br>(0.0369) | 0.0349<br>(0.2194) | 1.000<br>(-) |

Significance levels are shown in the parentheses.

**Table 4 Industry Segments of the Sample**

| 2-digit NAICS | Description | Freq. | % |
|---|---|---|---|
| 11 | Agriculture, Forestry, Fishing and Hunting | 5 | 0.4 |
| 21 | Mining, Quarrying, and Oil and Gas Extraction | 25 | 2.02 |
| 22 | Utilities | 104 | 8.39 |
| 23 | Construction | 8 | 0.65 |
| 31-33 | Manufacturing | 824 | 66.45 |
| 42 | Wholesale Trade | 51 | 4.11 |
| 44-45 | Retail Trade | 48 | 3.87 |
| 48-49 | Transportation and Warehousing | 17 | 1.37 |
| 51 | Information | 71 | 5.73 |
| 52 | Finance and Insurance | 12 | 0.97 |
| 53 | Real Estate and Rental and Leasing | 10 | 0.81 |
| 54 | Professional, Scientific, and Technical Services | 31 | 2.5 |
| 56 | Administrative and Support and Waste Management and Remediation Services | 10 | 0.81 |
| 62 | Health Care and Social Assistance | 9 | 0.73 |
| 72 | Accommodation and Food Services | 15 | 1.21 |
| Total | | 1,240 | 100 |

## Table 5 Baseline Spillover Models

| Variables | (1) Baseline | (2) Random effects | (3) Alternative measure of spillovers 1 | (4) Alternative measure of spillovers 2 | (5) Includes # of users | (6) Includes # of questions | (7) Includes # of users and # of questions |
|---|---|---|---|---|---|---|---|
| Non-IT capital | 0.109** | 0.241*** | 0.109** | 0.109** | 0.110** | 0.110** | 0.110** |
| | (0.043) | (0.017) | (0.043) | (0.043) | (0.043) | (0.043) | (0.043) |
| IT capital | 0.010* | 0.009* | 0.010* | 0.010* | 0.010* | 0.011* | 0.010* |
| | (0.006) | (0.015) | (0.006) | (0.006) | (0.006) | (0.006) | (0.006) |
| Non-IT labor | 0.726*** | 0.683*** | 0.726*** | 0.725*** | 0.727*** | 0.729*** | 0.728*** |
| | (0.059) | (0.018) | (0.059) | (0.059) | (0.059) | (0.059) | (0.059) |
| IT spillover | 0.009** | 0.011** | 0.009** | 0.016** | 0.011*** | 0.015*** | 0.014** |
| | (0.004) | (0.005) | (0.004) | (0.006) | (0.004) | (0.005) | (0.006) |
| Log(registered users) | | | | | -0.020 | | -0.011 |
| | | | | | (0.013) | | (0.017) |
| Log(questions) | | | | | | -0.015* | -0.009 |
| | | | | | | (0.009) | (0.010) |
| Constant | 1.726*** | | 1.726*** | 1.727*** | 1.683*** | 1.659*** | 1.670*** |
| | (0.452) | | (0.452) | (0.452) | (0.456) | (0.457) | (0.456) |
| Observations | 1,240 | 1,240 | 1,240 | 1,240 | 1,240 | 1,240 | 1,240 |
| Number of firms | 275 | 275 | 275 | 275 | 275 | 275 | 275 |
| R-squared | 0.575 | 0.569 | 0.575 | 0.575 | 0.577 | 0.577 | 0.577 |

All models except Column (2) use firm-level fixed effects and year dummies. Column (2) estimates include 2-digit NAICS dummies. Robust standard errors (clustered by firm) in parentheses except for Column 2. All R-square values are "within" estimates that do not include the explanatory power of the fixed effects. Columns (1), (2), (5) (6) and (7) compute IT spillovers using reward points from correct, very helpful, and helpful answers, Column (3) uses only reward points from only correct and very helpful answers, and Column (4) computes IT spillovers by counting the number of correct answers to questions.
*** p<0.01, ** p<0.05, * p<0.1

## Table 6 IV and Falsification Tests

| Variables | (1) IV estimates | | (2) Spillover timing | (3) Complementarity between spillover and IT capital | (4) Exploring other complementarity |
|---|---|---|---|---|---|
| | First stage – IT spillover | Second stage | | | |
| Non-IT capital | -0.052 | 0.109*** | 0.083* | 0.110** | 0.109** |
| | (0.168) | (0.036) | (0.048) | (0.043) | (0.044) |
| IT capital | -0.049 | 0.010** | 0.009† | 0.009† | 0.008 |
| | (0.042) | (0.005) | (0.006) | (0.006) | (0.006) |
| Non-IT labor | 0.056 | 0.725*** | 0.745*** | 0.725*** | 0.727*** |
| | (0.147) | (0.050) | (0.061) | (0.059) | (0.059) |
| IT spillover | | 0.016* | 0.008** | -0.001 | 0.039* |
| | | (0.009) | (0.003) | (0.007) | (0.021) |
| IT spillover (t+1) | | | 0.004 | | |
| | | | (0.003) | | |
| IT capital X spillover | | | | 0.003* | 0.006** |
| | | | | (0.002) | (0.002) |
| Non-IT capital X spillover | | | | | 0.002 |
| | | | | | (0.003) |
| Non-IT labor X spillover | | | | | -0.010** |
| | | | | | (0.005) |
| High forum | 0.133† | | | | |
| | (0.084) | | | | |
| Industry peer spillover | 0.550*** | | | | |
| | (0.054) | | | | |
| Constant | | | 1.762*** | 1.727*** | 1.732*** |
| | | | (0.477) | (0.452) | (0.455) |
| Observations | 1,227 | 1,227 | 1,206 | 1,240 | 1,240 |
| Number of firms | 262 | 262 | 268 | 275 | 275 |
| R-squared | 0.291 | 0.574 | 0.580 | 0.576 | 0.578 |

All models use firm-level fixed effects and year dummies. Robust standard errors (clustered by firm) in parentheses. Within R-squares are reported for panel data models.
*** p<0.01, ** p<0.05, * p<0.1, † p<0.15

**Table 7 Technical vs. Business-Related Spillovers**

| Variables | (1) Technical spillovers | (2) Business spillovers | (3) Technical + business spillovers | (4) Controlling for registered users | (5) Controlling for registered users + questions | (6) Instrumental variables | | |
|---|---|---|---|---|---|---|---|---|
| | | | | | | First stage – technical spillover | First stage – business spillover | Second stage |
| Non-IT capital | 0.111** | 0.109** | 0.110** | 0.111** | 0.111** | -0.022 | -0.108 | 0.112*** |
| | (0.043) | (0.043) | (0.043) | (0.043) | (0.043) | (0.139) | (0.115) | (0.036) |
| IT capital | 0.010* | 0.010* | 0.010* | 0.010* | 0.010* | -0.004 | -0.038 | 0.010** |
| | (0.006) | (0.006) | (0.006) | (0.006) | (0.006) | (0.032) | (0.032) | (0.005) |
| Non-IT labor | 0.726*** | 0.725*** | 0.725*** | 0.726*** | 0.727*** | -0.053 | 0.171 | 0.725*** |
| | (0.059) | (0.059) | (0.059) | (0.059) | (0.059) | (0.126) | (0.108) | (0.050) |
| Technical spillover | 0.011** | | 0.011* | 0.012** | 0.014** | | | 0.022* |
| | (0.005) | | (0.006) | (0.006) | (0.006) | | | (0.013) |
| Business spillover | | 0.006 | 0.001 | 0.003 | 0.005 | | | 0.004 |
| | | (0.004) | (0.005) | (0.005) | (0.005) | | | (0.012) |
| Log(registered users) | | | | -0.019 | -0.011 | | | |
| | | | | (0.013) | (0.017) | | | |
| Log(questions) | | | | | -0.008 | | | |
| | | | | | (0.010) | | | |
| High forum | | | | | | 0.084 | 0.055 | |
| | | | | | | (0.066) | (0.066) | |
| Peer technical spillover | | | | | | 0.558*** | 0.124*** | |
| | | | | | | (0.066) | (0.045) | |
| Peer business spillover | | | | | | 0.058* | 0.479*** | |
| | | | | | | (0.034) | (0.063) | |
| Constant | 1.717*** | 1.730*** | 1.719*** | 1.676*** | 1.665*** | | | |
| | (0.452) | (0.451) | (0.452) | (0.457) | (0.457) | | | |
| Observations | 1,240 | 1,240 | 1,240 | 1,240 | 1,240 | 1,227 | 1,227 | 1,227 |
| Number of firms | 275 | 275 | 275 | 275 | 275 | 262 | 262 | 262 |
| R-squared | 0.575 | 0.574 | 0.575 | 0.577 | 0.577 | 0.306 | 0.298 | 0.573 |

All models use firm-level fixed effects and year dummies. Robust standard errors (clustered by firm) in parentheses. Within R-squares are reported for panel data models.

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

## References


Ackerberg, D., K. Caves, G. Frazer. 2006. Structural identification of production functions. UCLA working paper.

Aghion, P., P. Howitt. 1997. *Endogenous growth theory*. MIT press, Cambridge, MA.

Alcácer, J., W. Chung. 2007. Location strategies and knowledge spillovers. *Management Science*. **53**(5) 760-776.

Aral, S., E. Brynjolfsson, M. Van Alstyne. 2012. Information, technology, and information worker productivity. *Information Systems Research*. **23**(3) 849-867.

Arellano, M., S. Bond. 1991. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *The Review of Economic Studies*. **58**(2) 277-297.

Arora, A., A. Gambardella. 1994. The changing technology of technological change: general and abstract knowledge and the division of innovative labour. *Research Policy*. **23**(5) 523-532.

Bloom, N., R. Sadun, J. Van Reenen. 2012. Americans do IT better: US multinationals and the productivity miracle. *The American Economic Review*. **102**(1) 167-201.

Boudreau, K.J., N. Lacetera, K.R. Lakhani. 2011. Incentives and problem uncertainty in innovation contests: An empirical analysis. *Management Science*. **57**(5) 843.

Bresnahan, T., S. Greenstein. 1996. Technical progress and co-invention in computing and in the uses of computers. *Brookings Papers on Economic Activity. Microeconomics*. **1996** 1-83.

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

Brown, J.S., P. Duguid. 2001. Knowledge and organization: A social-practice perspective. *Organization Science*. **12**(2) 198-213.

Brynjolfsson, E., D. Fitoussi, L. Hitt. 2006. The information technology iceberg. MIT working paper.

Brynjolfsson, E., L. Hitt. 1995. Information technology as a factor of production: The role of differences among firms. *Economics of Innovation and New Technology*. **3**(3&4) 183 - 200.

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

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

Brynjolfsson, E., L.M. Hitt, S. Yang. 2002. Intangible assets: Computers and organizational capital. *Brookings Papers on Economic Activity*. **2002**(1) 137-181.

Ceccagnoli, M. 2005. Firm heterogeneity, imitation, and the incentives for cost reducing R&D effort. *The Journal of Industrial Economics*. **53**(1) 83-100.

Chang, Y.B., V. Gurbaxani. 2012a. The impact of IT-related spillovers on long-run productivity: An empirical analysis. *Information Systems Research*. **23**(3) 868-886.

Chang, Y.B., V. Gurbaxani. 2012b. Information technology outsourcing, knowledge transfer, and firm productivity: An empirical analysis. *MIS Quarterly*. **36**(4) 1043-1063.

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

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

Chwelos, P., R. Ramirez, K.L. Kraemer, N.P. Melville. 2010. Does technological progress alter the nature of information technology as a production input? New evidence and new results. *Information Systems Research*. **21**(2) 392-408.

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

Dewan, S., C. Shi, V. Gurbaxani. 2007. Investigating the risk return relationship of information technology investment: Firm-level empirical analysis. *Management Science*. **53**(12) 1829-1842.

Di Maggio, M., M. Van Alstyne. 2011. Information sharing and performance: Evidence from a Japanese bank. MIT working paper, available at SSRN: http://ssrn.com/abstract=1893164

Doan, A., R. Ramakrishnan, A.Y. Halevy. 2011. Crowdsourcing systems on the world-wide web. *Communications of the ACM*. **54**(4) 86-96.

Fahey, R., A.C. Vasconcelos, D. Ellis. 2007. The impact of rewards within communities of practice: a study of the SAP online global community. *Knowledge Management Research & Practice*. **5** 186-198.

Fershtman, C., N. Gandal. 2011. Direct and indirect knowledge spillovers: the “social network” of open-source projects. *The RAND Journal of Economics*. **42**(1) 70-91.

Füller, J., G. Jawecki, H. Mühlbacher. 2007. Innovation creation by online basketball communities. *Journal of Business Research*. **60**(1) 60-71.

Furman, J.L., M.K. Kyle, I.M. Cockburn, R. Henderson. 2006. Public & private spillovers, location and the productivity of pharmaceutical research. *Annales d'Economie et de Statistique*. **2005**(79/80) 167-190.

Griffith, R., R. Harrison, J. Van Reenen. 2006. How special is the special relationship? Using the impact of US R&D spillovers on UK firms as a test of technology sourcing. *The American Economic Review*. **96**(5) 1859-1875.

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

Griliches, Z. 1992. The search for R&D spillovers. *Scandinavian Journal of Economics*. **94**(Supplement) S29-S47.

Hitt, L.M., E. Brynjolfsson. 1996. Productivity, business profitability, and consumer surplus: Three different measures of information technology value. *MIS Quarterly*. **20**(2) 121-142.

Hitt, L.M., D.J. Wu, X. Zhou. 2002. Investment in enterprise resource planning: Business impact and productivity measures. *Journal of Management Information Systems*. **19**(1) 71-98.

Huang, P., A. Tafti and S. Mithas. 2012. Knowledge contribution in online network of practice: The role of IT infrastructure, foreign direct investment and immigration. *Proceedings of the Thirty Third International Conference on Information Systems*, Orlando.

Huang, Y., P.V. Singh, K. Srinivasan. 2011. Crowdsourcing new product ideas under consumer learning. Carnegie Mellon University Working Paper.

Hulten, C.R. 2001. Total factor productivity: a short biography. C.R. Hulten, E.R. Dean, M.J. Harper, eds. *New Developments in Productivity Analysis*. University of Chicago Press, 1-54.

Jaffe, A.B. 1986. Technological opportunity and spillovers of R & D: Evidence from firms' patents, profits, and market value. *The American Economic Review*. **76**(5) 984-1001.

Jaffe, A.B., M. Trajtenberg, R. Henderson. 1993. Geographic localization of knowledge spillovers as evidenced by patent citations. *The Quarterly Journal of Economics*. **108**(3) 577-598.

Ko, D.G., L.J. Kirsch, W.R. King. 2005. Antecedents of knowledge transfer from consultants to clients in enterprise system implementations. *MIS Quarterly*. **29**(1) 59-85.

Lee, D.M.S., E.M. Trauth, D. Farwell. 1995. Critical skills and knowledge requirements of IS professionals: A joint academic/Industry investigation. *MIS Quarterly*. **19**(3) 313-340.

Levinsohn, J., A. Petrin. 2003. Estimating production functions using inputs to control for unobservables. *Review of Economic Studies*. **70**(2) 317-341.

Li, C.-Y., C.-T. Hsieh. 2009. The impact of knowledge stickiness on knowledge transfer implementation, internalization, and satisfaction for multinational corporations. *International Journal of Information Management*. **29**(6) 425-435.

Marshall, A. 1890. *Principles of economics*. Macmillan, London.

McAfee, A. 2002. The impact of enterprise information technology adoption on operational performance: An empirical investigation. *Production & Operations Management*. **11**(1) 33-53.

McAfee, A., E. Brynjolfsson. 2008. Investing in the IT that makes a competitive difference. *Harvard Business Review*. **86**(7/8) 98-107.
Mittal, N., B.R. Nault. 2009. Investments in information technology: Indirect effects and information technology intensity. *Information Systems Research*. **20**(1) 140-154.
Moretti, E. 2004. Workers' education, spillovers, and productivity: evidence from plant-level production functions. *American Economic Review*. **94**(3) 656-690.
Nadiri, M.I. 1993. Innovations and technological spillovers. NBER Working Paper, No. 4423.
Romer, P.M. 1986. Increasing returns and long-run growth. *The Journal of Political Economy*. **94**(5) 1002-1037.
Szulanski, G. 1996. Exploring internal stickiness: Impediments to the transfer of best practice within the firm. *Strategic Management Journal*. **17**(Winter Special Issue) 27-43.
Tambe, P., L.M. Hitt. 2012a. Measuring spillovers from information technology investments. New York University Working Paper.
Tambe, P., L.M. Hitt. 2012b. The productivity of information technology investments: New evidence from IT labor data. *Information Systems Research*. **23**(3) 599-617.
Wasko, M., S. Faraj. 2005. Why should I share? Examining social capital and knowledge contribution in electronic networks of practice. *MIS Quarterly*. **29**(1) 35-57.
Zander, U., B. Kogut. 1995. Knowledge and the speed of the transfer and imitation of organizational capabilities: An empirical test. *Organization Science*. **6**(1) 76-92.