



Information Systems Research

Publication details, including instructions for authors and subscription information:
<http://pubsonline.informs.org>

Measuring Information Technology Spillovers

Prasanna Tambe, Lorin M. Hitt

To cite this article:

Prasanna Tambe, Lorin M. Hitt (2014) Measuring Information Technology Spillovers. Information Systems Research 25(1):53-71.
<http://dx.doi.org/10.1287/isre.2013.0498>

Full terms and conditions of use: <http://pubsonline.informs.org/page/terms-and-conditions>

This article may be used only for the purposes of research, teaching, and/or private study. Commercial use or systematic downloading (by robots or other automatic processes) is prohibited without explicit Publisher approval, unless otherwise noted. For more information, contact permissions@informs.org.

The Publisher does not warrant or guarantee the article's accuracy, completeness, merchantability, fitness for a particular purpose, or non-infringement. Descriptions of, or references to, products or publications, or inclusion of an advertisement in this article, neither constitutes nor implies a guarantee, endorsement, or support of claims made of that product, publication, or service.

Copyright © 2014, INFORMS

Please scroll down for article—it is on subsequent pages



INFORMS is the largest professional society in the world for professionals in the fields of operations research, management science, and analytics.

For more information on INFORMS, its publications, membership, or meetings visit <http://www.informs.org>

Measuring Information Technology Spillovers

Prasanna Tambe

Leonard N. Stern School of Business, New York University, New York, New York 10012, ptambe@stern.nyu.edu

Lorin M. Hitt

The Wharton School, University of Pennsylvania, Philadelphia, Pennsylvania 19104, lhitt@wharton.upenn.edu

The measurement of the impact of IT spillovers on productivity is an important emerging area of research. Studies of IT spillovers often adopt a “production function” approach commonly used for measuring R&D spillovers, in which an external pool of IT investment is modeled using weighted measures of the IT investments of other firms, industries, or countries. We show that when using this approach, measurement error in a firm’s own IT inputs can exert a significant upward bias on estimates of social returns to IT investment. This problem is particularly severe for IT spillovers because of the high levels of measurement error in most available IT data. The presence of the bias term can be demonstrated by using instrumental variable techniques to remove the effects of measurement error in a firm’s own IT inputs. Using panel data on IT investment, we show that measurement error corrected estimates of IT spillovers are 40% to 90% lower than uncorrected estimates. This bias term is increasing in the correlation between the IT pool and firms’ own IT investment. Therefore, estimates from models of spillover pools are less sensitive to the issues identified in this paper when the spillover paths minimize the correlation between a firm’s own IT investment and the constructed external IT pool. Implications for researchers, policy makers, and managers are discussed.

Keywords: IT spillovers; IT productivity; measurement error; business value of IT

History: Michael Smith, Senior Editor; Chris Forman, Associate Editor. This paper was received on June 16, 2011, and was with the authors 14 months for 2 revisions. Published online in *Articles in Advance* October 31, 2013.

1. Introduction

Understanding the relationship between IT and productivity has been an important area of economic research for several decades. Although most prior literature has focused on the productivity of a firm’s own IT investments, many researchers are shifting attention toward estimating social returns (i.e., “spillovers”) from IT investment, both in the United States and abroad (Dedrick et al. 2003, Van Reenen et al. 2010). To assess the contribution of IT spillovers to productivity and growth, a common empirical strategy is to embed the aggregate pool of external IT investment from which a firm captures productivity spillovers into a production function, along with conventional inputs and a firm’s own IT investment levels. This approach is derived from a large and influential literature focused on estimating the impact of R&D spillovers on productivity (e.g., see Griliches 1992). A survey of the IT spillovers literature revealed more than 30 recent papers—of those that measure the contribution of IT spillovers to productivity, all but four use a variant of this approach.¹

This paper demonstrates that the application of this empirical framework to the IT context can produce estimates that significantly overstate the size of IT spillovers. It focuses on a form of the errors-in-variables problem in which measurement error is present in a firm’s own IT capital investment, and the external pool is modeled by aggregating the IT investments of other firms. It is well known that when there is measurement error in the IT input, estimates of the productivity of a firm’s own IT investments will be biased toward zero. Because of (a) the problems with measuring and valuing IT and (b) the importance of accurately measuring the contribution of IT to productivity, this measurement problem has occupied a central role in the empirical literature on the productivity of IT investments (Brynjolfsson 1993, Barua et al. 1995, Hitt and Brynjolfsson 1996, Greenan and Mairesse 1996, Barua and Lee 1997, Brynjolfsson and Hitt 2003). However, because the

Guo 2004; Daveri and Silva 2004; Dutta and Otsuka 2004; Becchetti and Adriani 2005; O’Mahoney and Vecchi 2005; Tanuwidjaja 2006; Cheng and Nault 2007; Park et al. 2007a, b; Barker et al. 2008; Gholami et al. 2009; Chou et al. 2009; Lopez-Pueyo et al. 2009; Venturini 2009; Zhang and Lee 2009; Acharya and Basu 2010; Chang and Gurbaxani 2010; Han et al. 2011; Severgnini 2010; Van Reenen et al. 2010; Cheng and Nault 2011; Kooshki and Ismail 2011; Tambe and Hitt 2014; and Huang et al. 2012.

¹ Recent papers discussing IT-related spillovers include Quah 2001; Mun and Nadiri 2002; Stiroh 2002; Dedrick et al. 2003; Van der Wiel and Van Leeuwen 2003; Gelb et al. 2003; Ganley et al. 2003; Nguyen and Atrostic 2004; Rincon-Aznar and Vecchi 2004; Lee and

external IT pool measure is often highly correlated with firms' IT investments, measurement error can result in a significant upward bias being transmitted to the IT spillover coefficient. The transmitted bias is especially large when the IT pool is highly correlated with the firms' own IT inputs. For example, an IT pool constructed from the investments of other firms within the same industry tends to be highly correlated with a firm's own IT investments because firms within an industry share common opportunity and factor costs and often make similar investments.

This problem is particularly serious because of the measurement error in most available IT data that even in the most commonly used data sets has been estimated to be as high as 30% to 40% of the total measure variance (Brynjolfsson and Hitt 2003, Tambe and Hitt 2012). Relative to data on R&D expenditures, which captures a larger fraction of R&D spending, most IT data capture a small fraction of total IT expenditures, which in addition to hardware expenses can include organizational restructuring, software, IT wages, and training. This paper demonstrates the effects of this measurement error on IT spillover estimates by removing the measurement error from IT capital or by constructing spillover measures less subject to measurement error biases.

The classic approach to removing measurement error involves the use of instrumental variables (IV), in which measurement error in the instrument is uncorrelated with measurement error in the primary measure. In this study, we use alternative measures of IT investment to remove measurement error in the IT measure and demonstrate the effects of this source of bias on our IT spillover estimates. After correcting measurement error in the IT input, the magnitudes of our IT spillover estimates fall by 40% to 90% in both cross-sectional and panel estimates.

We also consider how different modeling choices affect the size of the bias term produced by IT measurement error. The bias term is increasing in the size of the measurement error in the IT input and in the covariance between a firm's own IT investment and the constructed IT pool, and it is decreasing in the variance of the IT pool. The use of different types of transmission path data to construct the spillover pool can either decrease or increase the size of the bias term, depending on whether the use of these data reduces or increases the covariance between the IT pool and measures of own IT investment. This is noteworthy because unlike R&D spillovers, which tend to flow between technologically similar firms, IT-related innovations are likely to spill across industry boundaries through a variety of mechanisms because of the "general-purpose" nature of information technology. We show that the size of the bias term is sensitive to whether the particular transmission

path under investigation reduces or increases the correlation between the spillover measure and a firm's own IT investment. When researchers do not benefit from observable linkages that describe the direction of spillovers, they often use IT spillover models with undefined pathways that we show are also highly correlated with own IT expenditures and therefore vulnerable to the transmission of bias.

These issues are important for a number of reasons. The measurement of IT spillovers has been identified as a promising area for future research because accurate estimates are critical for developing a thorough understanding of how IT affects growth (Dedrick et al. 2003, Van Reenen et al. 2010). As an economic externality, IT spillovers are an issue of special interest to economists and policy makers. For example, in sectors where IT spillovers are economically important, policy makers may consider applying public funds to stimulate growth. Accurate estimates of the sizes of spillovers produced by IT investments are a prerequisite for effectively allocating these subsidies, and subsidies based on inflated estimates of IT spillovers will not have the desired effects. Establishing the size of IT spillovers also has implications for understanding variation in IT returns and the allocation of IT value among firms, both of which are of interest to IT researchers and managers.

The paper contributes to a large and established literature on the measurement of IT productivity. Over the past two decades, the IT productivity literature has benefited greatly from studies that examine how modeling and measurement decisions affect estimates of IT value (some notable examples are Barua et al. 1995, Mooney et al. 1996, Barua and Lee 1997, Hitt and Brynjolfsson 1996, Chan 2000, Santhanam and Hartono 2003, Zhu and Kraemer 2003, Devaraj and Kohli 2003, Brynjolfsson and Hitt 2003, Burton-Jones and Straub 2006). This paper is in the same spirit but is among the first to be focused on issues related to the measurement of social rather than private returns to IT investments. Given the broad difficulties that researchers have faced measuring R&D spillovers, this is likely to be a fertile topic—in fact, recent and influential papers in the management and economics literature have focused on the challenges associated with accurately estimating R&D spillovers (Breschi and Lissoni 2001, Knott 2008, Knott et al. 2009, Bloom et al. 2011). However, to the best of our knowledge, this paper is the first to focus on measurement problems for IT spillovers and is the first to focus on this particular source of measurement-error related bias in either the IT or R&D context.

2. Background Literature

One of the primary challenges faced by IT value researchers has been the lack of robust, consistently

available measures of IT investment. Unlike data on R&D investments, regularly collected samples of firm-level IT data have been difficult to obtain. Instead, researchers have based IT productivity studies on ad hoc data sets that are incomplete or available only over short time periods. For example, researchers have used IT capital stock data collected by marketing firms (Dewan and Min 1997, Brynjolfsson and Hitt 2003), IT employment data collected through surveys (Lichtenberg 1995, Brynjolfsson and Hitt 1996), IT asset allocations collected from managerial surveys (Bharadwaj 2000, Aral and Weill 2007, Saunders 2010, Mithas et al. 2011), and IT labor data from archival sources (Tambe and Hitt 2012). Most of these data sources—for example, estimates of IT capital stock from managerial surveys—contain significant errors because of the difficulties associated with estimating and valuing IT capital stock (Brynjolfsson 1993, Dedrick et al. 2003).

IT assets are difficult to measure for several reasons. First, because of the widespread use of information technologies within the firm, a large fraction of IT hardware purchases may be transacted without the knowledge of IT personnel, making it difficult for an IS manager to assess the total value of computer capital stock in the firm. Second, rapid changes in the quality-adjusted pricing of computer hardware have made it difficult to accurately estimate the value of information technologies within the firm. This problem is further aggravated because IT assets can differ along many intangible dimensions, and assigning values to all of these can be difficult (Brynjolfsson and Hitt 2000). Finally, a considerable portion of the value created by IT, such as the creation of software, IT-enabled business processes or databases, is not recorded by conventional measurements of IT capital stock, but may represent a large fraction of a firm's IT investment. To the extent that the ratio of hardware to these types of IT assets varies across firms, further measurement error is introduced. Brynjolfsson and Hitt estimate that in the CITDB IT capital stock data set, probably the largest and most widely used data set for IT productivity research, the error variance may be as large as 30% to 40% of the total measure variance (Brynjolfsson and Hitt 2003).

This type of IT measurement error exerts a downward bias on estimates of the productivity of IT investments, which is one reason that it proved difficult for many years to provide evidence of positive returns to IT investment (Brynjolfsson 1993). Because of the important role played by IT measurement error in the IT “productivity paradox,” contributions related to modeling and measurement error have played a key role in advancing the IT productivity literature (Barua et al. 1995, Greenan and Mairesse 1996, Barua and Lee 1997, Hitt and Brynjolfsson 1996,

Brynjolfsson and Hitt 2003). To address the effects of measurement error on IT estimates, researchers have used several different approaches, including using long differences to reduce the effects of random measurement error (Brynjolfsson and Hitt 2003), using instrumental variables to correct measurement error (Tambe and Hitt 2012), restricting the sample to a homogeneous set of firms (Barua and Lee 1997), and using statistical models of measurement error to correct biased estimates (Greenan and Mairesse 1996). In general, these studies find that correcting or reducing measurement error significantly raises the estimated returns to IT investment, as predicted by the classic errors-in-variables framework.

Because the IT productivity paradox has largely been resolved through the use of better data and improved methods, some of the attention in the IT value literature has turned toward estimating social returns to IT investment, which has been identified as an important area for future research and is necessary for developing a full understanding of how IT investments affect growth (Dedrick et al. 2003). Firms may derive benefits from the IT investments of other firms through mechanisms that facilitate the transfer of know-how related to new technologies, standards, and practices. Given the uncertainty and costs associated with matching new work practices to new technologies, it may be easier for firms to imitate other firms when adopting new IT-enabled production methods rather than having to discover the right combinations of practices on their own through trial and error (Brynjolfsson and Hitt 2000). Other potential sources of IT externalities include benefits derived from networked applications that enable superior coordination with other firms including suppliers and customers (Cheng and Nault 2007, 2011). The productivity benefits derived from rising levels of external IT investment, therefore, may be substantial and may have important implications for explaining a firm's own productivity as well as other economic phenomena such as regional variation in IT adoption and IT returns (Forman et al. 2005).

The literature providing econometric evidence of IT spillovers, however, is in its relative infancy. It has primarily used methods derived from the literature on R&D spillovers in which researchers embed a measure of the public R&D pool into a production function along with other factors of production to estimate how much the public pool of R&D knowledge contributes to productivity. However, even for estimating R&D spillovers, where investment data are more systematically and comprehensively collected, the circumstantial nature of the approach presents a number of econometric challenges. Surveys of the empirical R&D spillovers literature demonstrate significant variation in R&D spillover estimates (Griliches 1992,

Nadiri 1993), and these surveys emphasize the importance of better data and techniques for creating more accurate estimates of R&D spillovers. Concerns about the interpretation of R&D spillover estimates have also motivated a recent literature focused on the limitations of the production function approach for assessing the contributions of R&D spillovers to productivity and growth (Breschi and Lissoni 2001, Knott 2008, Knott et al. 2009, Bloom et al. 2011). These studies have focused attention on how differences in the way R&D spillover pools are measured and modeled can lead to large differences in estimates as well as on the importance of using data on the microfoundations of knowledge diffusion to improve the accuracy with which spillovers can be studied.

Like R&D spillovers, IT spillovers also have important implications for growth policy and strategic decision making and therefore merit the same attention to factors that affect the accuracy of estimates. Although the measurement of IT spillovers shares many econometric challenges in common with the measurement of R&D spillovers, the issues with measuring and valuing IT have been a longstanding problem in the IT literature and introduce unique difficulties for the measurement of IT spillovers. In the next section, we analytically demonstrate why IT measurement error—which led to understated IT estimates in the earlier IT productivity literature—may lead to inflated estimates of IT spillovers.

3. Analysis

3.1. Framework

We begin by describing the production function approach commonly used for estimating the relationship between productivity and external IT investments, motivated by the literature on the productivity of R&D spillovers. The basic framework, pioneered by Griliches (1979), relates a firm’s output to conventional inputs (X_{it}), its own investments in the knowledge-generating asset (K_{it}), aggregate investment in the industry (K_{at}), and total factor productivity (B):

$$Y_{it} = BX_{it}^{1-\gamma} K_{it}^{\gamma} K_{at}^{\mu} \quad (1)$$

The productivity of a firm depends not only on its own investment in the knowledge-generating asset but also on investments in the knowledge-generating asset that are external to the firm. Implicit in this model are “transmission paths” through which firms capture productivity benefits from the investments of other firms. For example, know-how from other countries may spill in through trade or foreign direct investment, and firm-level productivity spillovers may occur through the transfer of knowledge from patent disclosures, supplier innovations, competitor

analyses, interfirm employee movements, customer collaborations, product market interaction, or other types of network linkages. Proxies for transmission path strength, such as geographic proximity, may also be informative about the direction and strength of knowledge flow between institutions.

Variants of this model have been widely used with a variety of different dependent variables. For example, a firm’s own IT investments and the aggregated investment of other firms in the industry have been tested against observable measures of innovation, such as patent counts, market value, firm sales, R&D investment, Tobin’s q , or combinations of these measures as dependent variables.² In this paper, we focus on estimating the contribution of spillovers at the firm level, and we emphasize results based on this production framework. Our results, however, are also applicable to studies at more aggregate levels, especially in contexts in which input measurement is subject to significant error. In this paper, following the most common approach, we analyze a Cobb-Douglas specification in which output (Y) at time t of firm i in industry j is related to conventional inputs (capital C and labor L), knowledge capital (K), and the aggregated knowledge capital of other firms (K_a). A log-levels version of this model can be written as follows:

$$y_{it} = a(i, j, t) + \beta_c c_{it} + \beta_l l_{it} + \beta_k k_{it} + \beta_{k_a} k_{at} + \varepsilon_{it}, \quad (2)$$

where lowercase variables denote logs and the coefficients are output elasticities. This model forms the basis of much of the empirical work that appears in the literature and in the rest of this paper.

3.2. Errors-In-Variables Bias

The biases discussed in this paper arise from measurement error in inputs, and to the best of our knowledge, this is the first treatment of this issue in either an IT or R&D setting.

It is well known that measurement error in an independent variable leads to inconsistent estimates. The estimate of the coefficient on the mismeasured variable is biased toward zero (the “attenuation bias”), and this bias can be transmitted to other variables, with the direction of the transmitted bias term determined by the variance-covariance matrix of the observations (e.g., see Levi 1973). Since spillover models are inherently “network” models, there is likely to be a high degree of correlation in the IT spillover pool and own IT measures due to homophily in networks—the link structure of the network tends to cluster firms according to factor costs, opportunities, and investment behavior. Therefore, measures of a firm’s own investments in IT or R&D will tend

² See, e.g., Griliches (1992) for a survey of this literature.

to co-vary with the aggregate pool of investment in these factors in specifications such as the one shown in (2). For example, because firms in a given industry experience similar factor costs and technological opportunities, firms in industries with high IT intensity tend also to invest heavily in IT. In these cases, some of the bias from the mismeasured IT input will be transmitted to the coefficient on the IT spillover term, creating an upward bias on that coefficient. The size of the bias term is increasing in the covariance between these inputs and, because of the small sizes of spillover estimates relative to other production inputs, can be of the same order as the estimates themselves. Several pathways have been commonly used in the literature—e.g., supply chain weights, industry weights, and trading weights—to construct measures of IT spillover pools. Our results generally suggest an upward bias whenever a positive correlation exists between own IT investment and the spillover pool.

More formally, consider a model in which, for simplicity, we assume only two inputs to production,³ own investment (x_k) and the aggregate investment of other firms that benefits the focal firm through spillovers (x_s), and suppose that the relation

$$y = \beta_k x_k + \beta_s x_s + \mu \quad (3)$$

is known to hold between the true values of y , x_k , and x_s . We also make standard homoskedasticity assumptions as well as assume that $E(\mu) = 0$, that all variables are measured as deviations from means (eliminating the need for a constant), and that μ is independent of both x_k and x_s . Furthermore, suppose that the first independent variable, x_k , is measured with error

$$x_k^* = x_k + \varepsilon_k \quad (4)$$

and that x_s , constructed as the weighted average of the mismeasured input of other firms, is also measured with error

$$x_s^* = \sum_{j \neq i}^{j \in I(i)} w_j x_{jk}^* = \sum_{j \neq i}^{j \in I(i)} w_j (x_{jk} + \varepsilon_{jk}) = x_s + \varepsilon_s. \quad (5)$$

Ordinary least squares (OLS) estimates can be written

$$p \lim \hat{\beta} = (\Sigma + \Omega)^{-1} \Sigma \beta, \quad (6)$$

³ We restrict the analysis to only these two inputs because the relationship of particular interest involves the correlations between the two variables (IT and the industry aggregate IT measure) and measurement error in the IT variable. Other factors (such as labor) may include some error, which may also be transmitted to the spillover estimate, but the impact is likely to be less severe because of weaker correlations and less measurement error. Moreover, the two input model provides the insights required to motivate the empirical analysis while keeping complexity to a minimum.

where Σ is the covariance matrix of the observations x , and Ω is the covariance matrix of the error terms on the independent variables. In our two-variable case, this can be expanded to⁴

$$p \lim \begin{bmatrix} \hat{\beta}_k \\ \hat{\beta}_s \end{bmatrix} = \frac{1}{(\sigma_{x_k^*}^2 \sigma_{x_s^*}^2 - \sigma_{x_k x_s}^2)} \begin{bmatrix} \sigma_{x_k}^2 \sigma_{x_s^*}^2 - \sigma_{x_k x_s}^2 & \sigma_{x_k x_s} \sigma_{\varepsilon_s}^2 \\ \sigma_{x_k x_s} \sigma_{\varepsilon_k}^2 & \sigma_{x_s}^2 \sigma_{x_k^*}^2 - \sigma_{x_k x_s}^2 \end{bmatrix} \cdot \begin{bmatrix} \beta_k \\ \beta_s \end{bmatrix}. \quad (7)$$

The coefficient estimates on each term are a linear combination of their true values, attenuated by measurement error, and an upward bias transmitted from measurement error in the other input.

$$p \lim \hat{\beta}_k = \frac{(\sigma_{x_k}^2 \sigma_{x_s^*}^2 - \sigma_{x_k x_s}^2)}{(\sigma_{x_k^*}^2 \sigma_{x_s^*}^2 - \sigma_{x_k x_s}^2)} \beta_k + \frac{(\sigma_{x_k x_s})}{(\sigma_{x_k^*}^2 \sigma_{x_s^*}^2 - \sigma_{x_k x_s}^2)} \sigma_{\varepsilon_s}^2 \beta_s, \quad (8)$$

$$p \lim \hat{\beta}_s = \frac{(\sigma_{x_s}^2 \sigma_{x_k^*}^2 - \sigma_{x_k x_s}^2)}{(\sigma_{x_k^*}^2 \sigma_{x_s^*}^2 - \sigma_{x_k x_s}^2)} \beta_s + \frac{(\sigma_{x_k x_s})}{(\sigma_{x_k^*}^2 \sigma_{x_s^*}^2 - \sigma_{x_k x_s}^2)} \sigma_{\varepsilon_k}^2 \beta_k. \quad (9)$$

When spillover values are positive, the attenuation bias and the bias transmitted from mismeasurement in own IT capital move in competing directions. An analysis of the conditions under which the net effect results in an upward bias on the spillover estimate is presented in Appendix A. We also present numerical tests indicating the parameter values required for this net effect to turn negative. In general, the effects of downward attenuation are smaller than the upward bias transmitted from own IT capital in all conditions except where the covariance between the two variables is close to zero and the true value of the spillover coefficient is very high relative to own IT capital. In the conditions characterizing most existing studies on IT spillovers, the net effect is likely to be in an upward direction. We explicitly test this in our empirical analysis below by removing each of the types of measurement error from the two measures to estimate how they each impact the spillover estimate.

We also expect the bias transmitted from the spillover coefficient to the own IT coefficient estimate to be smaller than that transmitted to the spillover pool because the error variance of the spillover term is likely to be smaller than the error variance of own IT capital because of the convexity of the weighting terms in (5) and because most existing studies suggest that the output elasticity of own IT capital is greater than that of IT spillovers. Therefore, we expect the

⁴ The full derivation is in Appendix A. The derivation follows notational convention for measurement error proofs in a number of econometric texts (see, e.g., Johnston and Dinardo 1997).

second term on the right-hand side of (9) to be the most pernicious of the biases in (8) and (9) and to have the largest impact on the own IT and spillover estimates. This bias term is increasing in both the covariance between the two inputs and the error in the mismeasured input. When there is no measurement error, we recover the unity matrix, so OLS estimates converge to their true values. If the two inputs are uncorrelated, the estimates are attenuated versions of their true values, but there is no transmission of bias from the other input. In practice, however, if the spillover term is constructed as in (5), we expect the two inputs to co-vary because firms collocated in industries with high IT intensity, such as finance, are likely to invest heavily in IT. The size of the bias on the spillover coefficient is also decreasing in the variance of the spillover pool. All other things being equal, therefore, the largest bias terms will result from spillover pools that co-vary closely with the measure of own investment and that exhibit little variation.

The above analysis applies to cross-sectional models, but many studies use panel estimators to account for individual effects. When using fixed-effects estimators, estimates of the contribution of IT capital stock are produced by variation in own IT investment within a single firm at different points in time, which tends to be smaller than the variation in IT investment occurring across firms. Although the ultimate effect on the estimates depends upon on the structure of the measurement error, measurement error problems are generally thought to be exacerbated by the use of fixed-effects estimators (Solon 1985, Wooldridge 2001) because in the presence of measurement error, these estimators often decrease the “signal-to-noise” ratio between variation in the measure and variation in the error term (Griliches and Hausman 1986, Mairesse 1992). Random effects estimators use both cross-firm and within-firm variation to produce estimates, so the effects of the bias term on random-effects estimates should lie somewhere in between that of OLS and fixed-effects estimates. In our analysis below, we demonstrate that the bias term from the OLS case persists when using fixed-effects or random-effects estimators.

3.3. Using Instrumental Variables to Demonstrate Biases Caused by IT Measurement Error

Our strategy is to demonstrate the effect of the bias term discussed in this paper by removing measurement error from the own IT capital measure. The most common method for addressing this type of errors-in-variables problem is to use a second measure of IT investment as an instrumental variable, where errors in this second measure are uncorrelated with errors in the original IT investment variable (e.g., Hausman 2001, Chen et al. 2007 provide a recent

survey of measurement error models). In our empirical analysis, we discuss the alternative measures of IT investment used in this analysis as instrumental variables for our primary measure of IT capital as well as why they meet the conditions necessary to fix the errors-in-variables problem. Removing these errors eliminates the second term in (9).⁵

Our approach is subject to several caveats. Difficulties in producing bias-corrected estimates of the impact of own IT investment or R&D spillovers on productivity have been extensively documented in the literature, and as a practical matter, finding effective instruments to deal with all of the potential omitted variable problems for IT investments in productivity equations has proven to be difficult (Aral et al. 2006 discuss this issue at length). For example, understanding the effects of unobserved organizational complements on estimates of IT productivity is a topic of ongoing interest in the academic literature (e.g., Saunders 2010 is a recent example), and these unobserved organizational assets are also likely to produce an upward bias on IT spillover estimates if some of these complements can generate productivity spillovers.⁶ Estimating spillover effects also raises additional econometric challenges—for example, the well-known “reflection problem” makes it difficult to identify the parameters of interest in R&D and IT spillover frameworks (Manski 1993).

The measurement error corrected estimates that we report, therefore, are still subject to the other common biases that impact IT productivity and IT and R&D spillover estimates. These biases, however, are likely to be present in most, if not all, existing IT spillover estimates, so our findings can be understood as improving on conventional methods being used to estimate IT spillover effects, conditional on these estimates being subject to other sources of bias that are likely to be identified and treated in future research. This approach is similar to that used by papers in an emerging literature on econometric difficulties with estimating R&D spillovers (Knott 2008, Knott et al. 2009, Bloom et al. 2011). These papers generally advance the literature on R&D spillovers by focusing on one or more difficulties with the interpretation of existing R&D spillover estimates. However, the estimates in these papers are still subject to other longstanding econometric issues with estimating R&D spillovers.

⁵ The spillover estimate is still likely to be attenuated, but the size of the attenuation bias is expected to be small because of the smaller error variance of the spillover term.

⁶ Directly addressing this bias in the IT spillover context is a complicated and interesting issue because some subset of the IT, organization, and human capital nexus that is generally proxied for by IT spending can generate economic spillovers, and some cannot. This is likely to be an interesting area for future research.

Nevertheless, recent studies suggest that the endogeneity bias in IT investment appears to be relatively low (e.g., Tambe and Hitt 2012), so measurement error related biases are likely to be among the larger sources of bias. Moreover, the bias on the spillover term caused by measurement error in own IT capital is upward so it is particularly important to address because it can lead to spurious estimates of the impact of IT spillovers on productivity.

4. Data

4.1. Primary Data

Our primary IT data set is constructed by combining computer stock data from Computer Intelligence Info-Corp (CII) with financial information from Compustat. Capital rental prices are provided by the Bureau of Labor Statistics (BLS), and price deflators are obtained from government sources. CII collects data for the 1,000 largest firms in the United States (Fortune 1000). Our panel consists of 3,118 observations over the eight-year period 1987–1994 after omitting firms with incomplete data and those that had missing data other than at the beginning or end of the measurement period.

Sample statistics for these firms are shown in Table 1 and the progression of how the means vary from year to year for these firms is shown in Appendix B. The firms in the sample are large, averaging \$1 billion in value-added. Within the sample, 57% of the firms are from the manufacturing industry; 41% from service; and 2% from mining, construction and agriculture. Some service industries—banking and insurance—are largely excluded because many of the firms in these industries do not report ordinary capital stock on Compustat. Because these industries are particularly computer intensive, the firms in our sample are somewhat less computer intensive than the economy as a whole. Otherwise, our sample appears to be broadly representative of large firms in the U.S. economy, and firms in the sample account for about 15% of total U.S. economic output over our sample period. Correlations between key regression variables are shown in Table 2.

Table 1 Descriptive Statistics for Key Variables

	Mean	Std. deviation	<i>N</i>
<i>Value-Added</i>	7.06	1.11	3,118
<i>Non-IT Capital</i>	7.75	1.13	3,118
<i>Labor</i>	6.48	1.11	3,118
<i>IT Capital (CII)</i>	2.87	1.46	3,118
<i>IT Pool</i>	3.58	0.72	3,118
<i>IT Capital (IDG)</i>	1.33	2.03	1,013
<i>IT Labor</i>	5.66	1.32	1,626

Note. All variables are in logs.

4.2. Data for Instrumental Variables

To test the hypothesis that measurement error in own IT capital produces an upward bias in IT spillover estimates, we remove the measurement error from own IT capital by using alternative measures of IT capital stock as instrumental variables. The first alternative IT measure was collected from the marketing research firm International Data Group (IDG). To remove the effects of measurement error in our primary IT capital measure, the measurement error in the IDG-provided instrumental variable must be uncorrelated with the measurement error in our IT capital stock measure as provided by CII. The independence of the error terms of these two different measures derives from the different methods through which these two market research firms value IT capital.

CII conducts surveys to track specific pieces of computer equipment at the site level and interviews information systems managers, at intervals ranging from monthly to annually, to obtain detailed information on a site’s IT hardware assets. The interview process includes checking on hardware reported in previous interviews to make accurate comparisons. CII assesses the market value of each piece of hardware and aggregates the numbers to form a measure of total hardware use at the firm. As mentioned above, these valuation data omit software, stored data, information system staff, and telecommunications equipment. Market valuation is performed by a proprietary algorithm developed by CII that takes into account market-based rental prices and machine configurations in determining an estimate.

IDG, by comparison, uses very different methods to collect its data. IDG surveys a single officer in the firm. The officer is asked to report the “market value of central processors” and the PCs and terminals in that firm. The number of PCs and terminals is multiplied by an estimated value, determined by the average nominal PC price over 1989–1991 in Berndt and Griliches’ (1990) study of hedonic prices for computers. Thus, both of these approaches include potential sources of error. However, because the CII data is collected at the site level from a number of managers, and the IDG data is collected from a single officer through survey questions, there is little reason to believe that errors between the two measures are correlated. Finally, because of higher accuracy from CII’s more rigorous methods for counting, tracking, and valuing the assets, we use its data as our primary variable.

Unfortunately, of the 3,118 firm-year observations available from CII, only 1,013 have corresponding values for the IDG-provided instrumental variable. We begin our empirical analysis by reporting results

using only the subsample of observations for which the CII and IDG data are both available. However, the size of this sample limits the precision at which many of the key coefficients, and especially the panel coefficients on IT capital and the IT spillover term, are estimated.

To improve our estimates, we take two approaches. First, we report results when using the full CII data set, where we set the missing values of the IDG variable to zero and include a dummy variable in the first stage regression to indicate when the instrument is not available. This approach uses the variation in the available instrument values to resolve the errors-in-variables problem. This results in a weaker instrument but has the advantage of retaining the full sample to produce the remaining parameter estimates. In particular, the larger sample is helpful to increase the precision of both the IT capital and the IT spillover estimates both in cross-sectional and panel estimates. In Appendix C, we report the first-stage results from our instrumental variable regressions to demonstrate the trade-off between the strength of the first-stage regression and the sample size. One caveat to this approach is that firms that appear in IDG may be systematically more productive than firms that do not. However, the consistency of the results across the different instrumental variable sets makes this somewhat less of a concern.

We also supplement the missing IDG data with a second instrumental variable based on IT labor. IT labor measures were obtained from the data and measure construction described and benchmarked in extensive detail in other published work (Tambe and Hitt 2012). This data set is based on firm-level IT employee counts created from the employment histories of a large sample of information technology workers collected through a partnership with a leading online job-search website and include information for each worker on employer name, job title, and dates of employment for every position ever held by that worker. Employment histories at this website are posted by close to 10 million unique individuals who are passively or actively seeking jobs and were aggregated to the employer level to create measures based on IT personnel counts. As with the IDG data, because of the independent methods used to create these measures, there should be little correlation in measurement error between the primary CII capital stock data and the IT investment measures based on the IT labor data set. These IT labor data are a weaker instrument than the IDG capital data but are used to show that our results are not particularly sensitive to the use of the IDG-based instrument.

5. Measurement of IT Spillovers

5.1. Overall Approach

We first use conventional methods to estimate the magnitudes of IT spillovers. We estimate a Cobb-Douglas production function in log-levels using data on value-added, capital, labor, IT, and measures of aggregate IT investment as described in Equation (2) above. Measures of IT capital stock are computed from the data sources described above using the same methods as other papers that use the same data sources (e.g., see Brynjolfsson and Hitt 2003). IT spillover pools at the industry level are constructed using the aggregate IT capital stock of other firms in the same four-digit Standard Industry Classification (SIC) industry.

We demonstrate that using instrumental variables to remove the effects of measurement error in a firm’s own IT inputs significantly reduces the magnitude of the spillover estimate, which is consistent with the presence of a bias term that is transmitted to the spillover estimate from measurement error in own IT capital. We repeat this exercise using a number of different samples and estimators. The estimated spillover coefficient changes across all of these tests, but removing measurement error from own IT capital in each model consistently lowers the overall value of the spillover estimate. Finally, we use different spillover models to show that the covariance between a firm’s own IT investments and the measure of the external IT pool affects the size of the bias transmitted to the spillover pool from the measurement error in own IT capital.

5.2. Measurement Error and IT Spillovers

First, we produce estimates of spillovers using the specification described in Equation (2) and then demonstrate that correcting the measurement error in own IT capital reduces the sizes of these estimates by removing this source of bias. Our spillover pools are computed as the average IT investment of other firms in the same four-digit SIC code, although in a later section we examine how using data on spillover transmission paths affects the size of the bias term. Table 3 shows the results of estimating (2) with firm-level data on the 1,013 observations for

Table 2 Correlations Between Key Regression Variables

	1	2	3	4	5
1. <i>Non-IT Capital</i>	1.00				
2. <i>Labor</i>	0.703	1.00			
3. <i>Materials</i>	0.712	0.762	1.00		
4. <i>IT Capital (CII)</i>	0.664	0.732	0.634	1.00	
5. <i>IT Pool</i>	0.138	0.231	0.016	0.353	1.00

Note. $N = 3,118$.

Table 3 Impact of Measurement Error Correction on OLS Estimates of IT Spillovers

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
DV: Log(Value-Added)	OLS	IV	OLS	IV	OLS	IV	OLS	IV	IV
Sample	All firms	All firms	All firms	All firms	Manuf only	Manuf only	All firms	All firms	All firms
Instrumental variable		IDG capital		IDG capital		IDG capital		IDG capital	IDG capital, IDG IT pool
Log(Labor)	0.777*** (0.0203)	0.751*** (0.0251)	0.774*** (0.0203)	0.752*** (0.0251)	0.820*** (0.0366)	0.760*** (0.0411)	0.764*** (0.0400)	0.733*** (0.0519)	0.733*** (0.0484)
Log(Non-IT Capital)	0.189*** (0.0172)	0.178*** (0.0195)	0.192*** (0.0169)	0.180*** (0.0201)	0.168*** (0.0217)	0.141*** (0.0296)	0.186*** (0.0263)	0.149*** (0.0385)	0.149*** (0.0390)
Log(IT Capital)	0.0269** (0.0119)	0.0632** (0.0253)	0.0241* (0.0124)	0.0589** (0.0275)	0.0225 (0.0216)	0.115** (0.0560)	0.0293 (0.0208)	0.0988 (0.0605)	0.0984* (0.0567)
Log(IT Pool)			0.0199 (0.0181)	0.0110 (0.0196)	0.0762** (0.0373)	0.0432 (0.0434)	0.0430 (0.0395)	0.0306 (0.0511)	0.0340 (0.111)
Controls	Industry year	Industry year	Industry year	Industry year	Industry year	Industry year	Industry year	Industry year	Industry year
Hausman <i>t</i> -statistic (IT Capital)		1.86		1.63		2.73		1.23	1.33
Hausman <i>t</i> -statistic (IT Pool)				1.59		2.44		0.80	0.10
First-stage <i>R</i> ²		0.712		0.717		0.659		0.715	App C ^a
First-stage <i>F</i> -statistic		67.07		55.06		26.35		18.20	App C ^a
Prob > <i>F</i>		0.000		0.000		0.000		0.000	App C ^a
Observations	1,013	1,013	1,013	1,013	444	444	262	262	262
<i>R</i> -squared	0.968	0.967	0.968	0.967	0.952	0.946	0.963	0.960	0.960

Notes. Robust standard errors are shown in parentheses and are clustered on firm; columns (2), (4), (6), and (8) use the unmodified IDG IT capital measures as an instrument for the main CII based IT capital measure. Column (9) also uses the IT pool constructed using IDG data as an instrument for the main IT pool measure constructed using the CII data. Hausman tests are for a significant change in the IT capital or IT spillover coefficient after applying the instrumental variable. Columns (1) through (4) include the sample for which IDG and CII IT capital measures are both available. Columns (5) and (6) are the observations within that sample in manufacturing industries. Columns (7)–(9) are the observations in the sample used in columns (1)–(4) for which an IT pool instrument could also be created.

^aSee Appendix C for full first-stage regression results.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

which CII and IDG IT capital data are both available. Columns 1 and 2 show estimates of the contribution of IT capital with and without measurement error correction. Application of the IDG based instrument, in column 2, increases the size of the coefficient on IT capital and reduces the contribution of other inputs. The increase in the own IT capital coefficient is consistent with removal of the attenuation bias caused by measurement error in the independent variable. For the instrumental variable regressions in column 2, we also report results from the first-stage regressions that suggest that the IDG instrument is highly correlated with the CII IT capital measure.⁷

Columns 3 and 4 show OLS estimates using the same sample with the IT pool included. The coefficient estimate on the IT spillover pool's contribution to value-added is positive but not significant,⁸ but like the other inputs, the magnitude is reduced

by application of the instrument (see column 4). In columns (5) and (6), we show the results from similar regressions on a subset of firms in manufacturing industries—a sector in which firms are significantly larger and IT measures may therefore contain more error. Column 5 shows that the coefficient estimate on the spillover term is positive and significant and is generally estimated with more precision than the estimate on the contribution of a firm's internal IT capital stock. The point estimate on the own IT capital coefficient is not significantly different from zero, but it is consistent in magnitude with the coefficient on own IT capital from the full sample. After applying the instrumental variable in (6), the magnitude of the spillover estimate falls substantially, and Hausman tests indicate that the coefficient changes are significant for both own IT capital and the IT pool.

The changes in these estimates after application of the instrument are consistent with the argument

⁷Coefficient estimates from first-stage regressions are reported in Appendix C.

⁸Our goal in this study is to provide evidence of an econometric issue with estimating IT spillovers rather than provide evidence for these IT spillovers. Therefore, we are more interested in changes to the size of these coefficients, rather than the coefficient sizes

themselves. The emerging stream of literature cited in the introduction is focused on providing such evidence of these spillovers, and indeed, the magnitude of these spillovers is likely to be sensitive to the sample, depending on industry, time period, and other factors.

Table 4 Robustness Checks with Alternative Sets of Instrumental Variables

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
DV: Log(<i>Value-Added</i>)	OLS	IV	OLS	IV	OLS	IV	IV
Sample	All firms	All firms	Manuf	Manuf	All firms	All firms	All firms
Instrumental variable	IDG capital + Missing		IDG capital + Missing		IT labor	IT labor + IDG capital	
Log(<i>Labor</i>)	0.738*** (0.0175)	0.654*** (0.0468)	0.759*** (0.0300)	0.560*** (0.106)	0.781*** (0.0198)	0.597*** (0.155)	0.716*** (0.0394)
Log(<i>Non-IT Capital</i>)	0.203*** (0.0121)	0.168*** (0.0209)	0.192*** (0.0155)	0.124*** (0.0410)	0.177*** (0.0144)	0.0927 (0.0740)	0.147*** (0.0210)
Log(<i>IT Capital</i>)	0.0314*** (0.00845)	0.149*** (0.0565)	0.0299** (0.0150)	0.287** (0.129)	0.0360*** (0.00969)	0.299 (0.216)	0.130*** (0.0461)
Log(<i>IT Pool</i>)	0.0237 (0.0152)	-0.0104 (0.0249)	0.0512* (0.0287)	-0.0573 (0.0707)	-0.0389* (0.0213)	-0.0901* (0.0533)	-0.0572** (0.0240)
Controls	Industry year	Industry year	Industry year	Industry year	Industry year	Industry year	Industry year
Hausman <i>t</i> -statistic (<i>IT Capital</i>)		3.08		3.32		2.89	3.24
Hausman <i>t</i> -statistic (<i>IT Pool</i>)		2.94		3.16		2.32	2.57
First-stage <i>R</i> ²		0.671		0.699		0.639	0.648
First-stage <i>F</i> -statistic		11.50		5.33		5.13	9.07
Prob > <i>F</i>		0.000		0.006		0.020	0.000
Observations	3,118	3,118	1,387	1,387	1,626	1,626	1,626
<i>R</i> -squared	0.967	0.959	0.959	0.924	0.974	0.931	0.969

Notes. Robust standard errors are shown in parentheses and are clustered on firm; columns (1) and (2) are the expanded sample using the modified IDG based instrumental variable. Columns (3) and (4) are the observations from the expanded sample in manufacturing industries. Columns (5) and (6) are the observations for which the IT labor measure is available. Column (7) uses both IDG IT capital and the IT labor measures as instrumental variables using the same sample used in (5) and (6). Hausman tests are for a significant change in the IT capital or IT spillover coefficient after applying the instrumental variable.
 p* < 0.1, *p* < 0.05, ****p* < 0.01.

that mismeasurement in own IT capital transmits an upward bias to the spillover pool, but this form of measurement error in the spillover pool can also attenuate the spillover estimate. The analysis in Appendix A suggests that a downward attenuation bias for positive spillover estimates is likely to be dominated by an upward bias transmitted from own capital, but we can explicitly test this by removing the measurement error from the spillover term. In columns (7) through (9), we report results from regressions in which we apply instrumental variables to both the IT capital and the IT pool to estimate the size of the attenuation bias relative to the size of the upward bias transmitted from the mismeasured IT input. The instrumental variable for the IT spillover pool is constructed using the same methods as in Equation (5), except that we use IDG IT capital data instead of the CII IT capital data.

The number of observations for which we can construct an industry level IT pool is fairly restricted because of limitations with the IDG data, so our estimates are subject to that caveat. However, in (7), we report regression estimates from the small sample of firms for which the data are available to construct IDG based IT pool measures. In (8), we apply IDG IT capital as an instrument for CII IT capital, which lowers the estimate on the IT pool, as in the prior regressions. In (9), we apply instruments for both the IT capital and the IT pool term. The changes

in the coefficients are generally consistent with the argument in Appendix A. Applying both instruments together in (9) increases the value of the spillover pool estimate by removing the attenuation bias, but this effect is smaller than the downward effect of removing measurement error in IT capital. The bias transmitted in the reverse direction from the spillover pool to the IT capital measure is also minimal, which is consistent with the analysis presented above. These estimates are generally in line with the argument that the attenuation effect of measurement error in the spillover pool is small relative to the upward bias from measurement error in own IT capital.

In the regressions in Table 3, the sample size somewhat limits the precision of some key estimates, and Hausman test results of the changes to the coefficients on IT capital and the IT spillover pools are close to significant but inconsistent. In Table 4, we use larger samples to improve the quality of our estimates and to serve as a baseline for the panel estimates, which require a larger sample for more within-firm variation. We report estimates from regressions using an alternative set of instrumental variables that expands the sample size in exchange for a weaker first stage in the IV regression results.⁹ First, we expand the sample by setting the missing values of the IDG variable

⁹ First-stage regression results for key instrumental variable regressions are shown in Appendix C.

to zero and include a dummy variable in the first stage regression to indicate when the instrument is not available. Although this weakens the instrument, the first-stage results, shown in Table 4, are reasonably high, and expanding the sample produces more precisely estimated coefficients.

Columns (1) and (2) replicate the results from columns (3) and (4) in Table 3, using a larger sample of firms with the expanded instrument set. The changes to the coefficient estimates on own IT capital and the IT pool after applying the instrument in column (2) are similar to those in Table 3. Columns (3) and (4) of Table 4 replicate the results from (5) and (6) of Table 3 using only the sample of manufacturing firms. Although the results are similar to those in Table 3 when using only manufacturing firms, use of the larger sample in Table 4 produces a significant estimate on own IT capital, and the Hausman tests are significant for both own IT capital as well as the IT pool, although this is to some extent due to the larger increase in the IT capital coefficient caused by the weaker first stage. In columns (5) and (6) of Table 4, we use IT labor as an instrument instead of IDG capital measures. Using IT labor as an instrument has a similar effect on the estimates on the IT capital and IT spillover terms. It increases the own IT capital coefficient and produces a corresponding fall in the spillover estimate, both of which are comparable in size to the effects when using the IDG measures as the instrumental variable for the CII measures. However, it has a very weak first stage, so some of the change in coefficients is due to inflating the estimates on own IT capital. Nevertheless, the similar results generated when using IT labor as the instrument support the argument that our findings are not particularly sensitive to using the IDG capital variable as an instrument. In column (7), we strengthen the first stage from column (6) by augmenting the IT labor measures with the IDG capital measures where available, and the results are largely the same as in the prior regression. The change in coefficient estimate is smaller, but the Hausman tests are significant. Overall, the estimates in Table 4 suggest that our results are not particularly sensitive to our use of the IDG data to remove the effects of measurement error in our primary IT capital measure.

Many spillover studies use fixed-effects or random-effects estimators to account for unobservable firm effects. In columns 1 through 8 of Table 5, we report panel estimates using the sample for which the IDG capital and IT labor measures are both available because of the higher sample size and because the IT labor data exhibit greater panel variation. Column 1 shows fixed-effects estimates without the spillover term. The coefficient estimate on IT capital is consistent with fixed-effects estimates produced by prior

research using these data. Column 2 shows estimates from the same model after applying an instrumental variable. Column 3 shows elasticities for a fixed-effects specification after introducing the spillover term. The coefficient estimate on the spillover pool is positive and significant and is larger in magnitude than the estimate on own IT capital. Column 4 shows the measurement error corrected results after using the IDG and IT labor data as instruments for private IT capital stocks. The decline in the IT pool coefficient is consistent with the coefficient changes when using OLS in Tables 3 and 4.

Columns 5 through 8 perform a similar analysis when using random-effects specifications, with the decline in the IT spillover term being of the same order as in the fixed-effects model after correcting measurement error using IDG IT capital measures and IT labor as instrumental variables. Generally, the estimates in Table 5 suggest that correcting the measurement error in own IT capital increases the coefficient estimates on private IT capital stocks and lowers the estimate on the IT pool when using panel estimators, but Hausman tests do not indicate a significant difference in the spillover coefficient when using our panel specifications. However, this is in part due to the limited power of these data and instruments for panel regressions—neither the change in IT capital nor the spillover coefficient is significantly different when using our instruments in panel regressions.

Overall, the observations on the other coefficients in these models are largely consistent with theoretical values (in the noninstrumented regressions) and exhibit comparable behavior to that reported in previous work. One result of note is that the direct coefficient on information technology capital rises significantly in the instrumental variables regressions, especially in the panel estimates. At least part of this increase is due to the correction for downward bias in IT capital contribution due to measurement error. However, the higher coefficient estimate may also be because the application of a second measure as an instrument accentuates the endogeneity of computer investment—that is, at least some of the correlation between our two measures of computer capital is due to short-run shocks driving up firm-level investment that may also be correlated with output. A closely related possibility is that the higher coefficient estimates may be reflecting a local average treatment effect—for example, cases where managers know their IT capital spending better may be better managed and more productive.¹⁰ In both cases, the effects of omitted variable biases may be accentuated when they affect both the original variable and the

¹⁰ We thank an anonymous editor for making this point.

Table 5 Measurement Error Correction on Panel Estimates of IT Spillovers

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Variables	Fixed effects	Fixed effects, IV	Fixed effects	Fixed effects, IV	Random effects	Random effects, IV	Random effects	Random effects, IV
Instrumental variable		IDG capital + IT labor		IDG capital + IT labor		IDG capital + IT labor		IDG capital + IT labor
Log(<i>Labor</i>)	0.758*** (0.0211)	0.677*** (0.159)	0.755*** (0.0211)	0.662*** (0.170)	0.762*** (0.0150)	0.634*** (0.130)	0.761*** (0.0151)	0.607*** (0.140)
Log(<i>Non-IT Capital</i>)	0.169*** (0.0184)	0.160*** (0.0268)	0.169*** (0.0184)	0.158*** (0.0281)	0.186*** (0.0127)	0.146*** (0.0418)	0.187*** (0.0127)	0.137*** (0.0468)
Log(<i>IT Capital</i>)	0.0279*** (0.00753)	0.153 (0.245)	0.0270*** (0.00753)	0.176 (0.268)	0.0305*** (0.00651)	0.210 (0.182)	0.0302*** (0.00653)	0.249 (0.200)
Log(<i>IT Pool</i>)			0.0555** (0.0257)	0.0252 (0.0619)			0.0118 (0.0180)	−0.0208 (0.0443)
Controls	Industry year	Industry year	Industry year	Industry year	Industry year	Industry year	Industry year	Industry year
Hausman <i>t</i> -statistic (<i>IT Capital</i>)		0.510		0.556		0.989		1.10
Hausman <i>t</i> -statistic (<i>IT Pool</i>)				0.536				0.805
Observations	1,626	1,626	1,626	1,626	1,626	1,626	1,626	1,626
Number of firms	222	222	222	222	222	222	222	222

Notes. Robust standard errors are shown in parentheses and clustered on firm; all instrumental variable regressions use both the IDG IT capital measure and IT labor measures as instruments on the sample for which IT labor is available. Hausman tests are for a significant change in the IT capital or IT spillover coefficient after applying the instrumental variable.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

instrument. Although this highlights the need for better instruments for IT, it is notable that the decline in the spillover term occurs despite these biases. Furthermore, if we set IT capital to its theoretical value, we see a similar reduction in the spillover coefficient, which suggests that changes in the direct IT coefficient by itself in the IV estimates are not responsible for the reduction in the spillover term estimates.

The changes in the coefficient estimates with capital and labor set to factor share imply an error variance for IT capital measurement that is slightly less than 35% of the total variance in IT capital. The implied error variance for the spillover measurement is about one-third of this number. Thus, the relative sizes of the measurement errors required to produce the biases that we observe are in line with earlier studies that have found potential for significant error in IT capital measurement (Brynjolfsson and Hitt 2003, Tambe and Hitt 2012). Moreover, these calculations coincide with the intuition that when the spillover term is highly correlated with own IT investment, it is essentially representing a lower variance estimate of direct IT returns and thus picking up some of the direct IT effect. Overall, these results in both pooled OLS and panel specifications demonstrate that (1) measurement error in firm-level inputs can create spurious increases in the estimated effects of externalities, (2) these problems also occur for panel methods, and (3) instrumental variables can be used to remove

the measurement error and the effects of this source of bias from the spillover estimate.

This bias can potentially account for a substantial amount of variation in IT spillover estimates. In Table 6, we show that under reasonable assumptions, using estimated values from the empirical analysis above, the ratio of the elasticities of the spillover pool to the elasticities of a firm's own IT investment can exhibit wide variation when we allow measurement error in the firm's own investments to bias the spillover term. To create the values in Table 6, we use the relationships derived in (7) to compute how measurement error in the IT input, along with changes to the covariance between the spillover pool and own IT investment, affects the ratio of their estimated output elasticities. For reasonable values of the true ratio, the swing in values is large enough to accommodate a large amount of variation and demonstrates that the magnitude of spillover estimates relative to the estimates of own IT capital are very sensitive to measurement error in the input data.

5.3. How Modeling Affects the Impact of Measurement Error on Spillover Estimates

In the preceding section, we used instrumental variables to demonstrate that removing measurement error from own IT capital can significantly reduce the size of IT spillover estimates. In this section, we focus on how construction of the spillover pool measure affects the severity of the bias term. As mentioned

Table 6 Computed Output Elasticity Ratios, IT Spillovers to Own IT

σ_{ks}	True output elasticity ratio, IT pool to own IT investment			
	0.5	0.2	0.1	0.05
0.10	0.53	0.25	0.16	0.11
0.20	0.59	0.33	0.23	0.19
0.30	0.67	0.41	0.32	0.27
0.40	0.76	0.51	0.43	0.39
0.50	0.89	0.67	0.59	0.55
0.60	1.09	0.91	0.85	0.81
0.70	1.47	1.42	1.40	1.38
0.80	2.60	3.19	3.47	3.63

Notes. Values illustrate how the ratio of the estimated output elasticities of the IT spillover pool and firm's own IT investments exhibits variation when measurement error in the firm's own IT stock can bias the spillover term. Computations assume values for IT pool variances, IT stock variances, and error variances that are in accordance with estimates reported above. Left-hand column represents covariance between own IT investment and IT spillover pool. The variance of the measurement error is fixed at 25% of the total variance of the IT measure.

above, larger biases result from spillover measures that (1) co-vary with a firm's own investment levels and (2) have smaller variances. Therefore, even in the absence of instruments, more accurate spillover estimates will be obtained from models using data on spillover pathways that decreases the covariance between the measures of own IT investment and IT spillover pools and result in a larger variance for the spillover pool. On the other hand, the bias will be relatively more severe when the relevant mechanisms increase the covariance between a firm's own IT investment and the IT spillover pool or reduce the variance of the spillover pool.

We use a variety of data sources on transmission paths to show that the effects of using different data sources to create spillover models can either enlarge or reduce the bias term, depending on whether the use of these data increases or decreases the covariance between private investments and the spillover pool. The first data set we use comes from the same sources described above, CII data describing information technology investments at the intra-firm establishment level and other financial measures at the firm level from Compustat databases. We exploit elements of the data at the establishment level to model proximity. Specifically, we take advantage of the establishment-level data that assign SIC codes to different establishments within each firm. Firms in the sample have an average of almost 69 establishments per firm. In addition, and most important for this study, each firm, through its establishments, occupies a variety of industry positions. On average, firms in this sample occupy more than four different two-digit SIC categories. Thus, firms that look similar at the aggregate firm level, in which they are generally assigned to a single SIC category, are quite

different if compared at the establishment level. If technological proximity plays an important role in knowledge spillovers, then each firm may be best described as a collection of establishments that in turn operate in respective different technological areas and with respective access to different spillover pools. The firm's access to IT spillovers, then, may be best modeled by accounting for diversity in its constituent establishments.

Whereas firm-level analyses often model the external knowledge capital as the industry average weighted sum of the investments of other firms, we use these additional measures of proximity in order to examine the role of covariance in bias. Our first comparison measure uses establishment-level data to model spillover pools available to a firm through its constituent establishments. The spillover pool available to a firm is computed as the weighted sum of the spillover pools available to each of its establishments, and lowercase and uppercase indices represent establishments and firms, respectively,

$$s_i = \sum_{i \in I} \left(\frac{c_i}{c_I} \right) s_i.$$

The weights are determined by the ratio of IT capital at the establishment level to total IT capital at the firm level. Therefore, conditional on the size of the spillover pool, corporate sites that are larger and more invested in information technologies will transfer more know-how from these spillover pools into the larger organization. The spillover pool of each establishment i is computed as the average of the information technology investments of all other establishments that occupy the same SIC industry $I(i)$, not including other establishments that are in the same firm $F(i)$

$$s_i = \frac{1}{N} \sum_{j \in I(i)} c_j.$$

Therefore, the spillover pool of each establishment within a firm is driven by the SIC code in which the establishment operates, independent of the technological position of the parent firm. This spillover measure, therefore, differs from traditional firm-level measures because use of establishment-level data provides a different measure of technological position.

The second modeling structure we test in this section uses data on IT labor flows, which in prior work has been argued to be a specific transmission path for IT spillovers (Dedrick et al. 2003). The IT labor flow data we use are based on a partnership with a leading online jobs board and are described in detail in other published work (Tambe and Hitt 2014). Our approach is to compute the IT pool as the IT intensity of all other firms from which a firm hires at least 5% of its

Table 7 Computed Multiplier on Spillover Estimate Using Different Transmission Paths

	Covariance with IT capital stock	Variance	Upward bias ^a
<i>IT Capital</i> stock		1.62	
<i>IT Pool</i> (Industry)	0.471	0.689	0.032
<i>IT Pool</i> (Establishment)	0.344	0.408	0.053
<i>IT Pool</i> (Labor)	0.084	0.929	0.018

^aEstimate of upward bias on spillover coefficient with given covariance and variance values, computed using the same methods as in Appendix A.

new IT workers, which implies that a firm receives larger productivity spillovers from organizations from which they hire many technical employees. For the purposes of this paper, the most notable fact about these IT labor flow data is that for the sample under investigation, IT workers appear to commonly move across industries; thus, the spillover pool accessible through labor flows is not as highly correlated with a firms' own IT investments, suggesting that the impact of measurement error on the spillover coefficient estimate will be less than when using the spillover pool measures based on firm or establishment industry.

Table 7 shows how the covariance between IT capital stock measures and the IT spillover pool measure impacts the size of the expected bias term in the different models. The use of establishment level data should increase the size of the upward bias on the spillover coefficient, but the use of IT labor flows that tend to cut across industry lowers the multiplier on the measurement error term, suggesting that the use of establishment data will slightly increase the size of the bias term transmitted to the spillover coefficient, and that the use of IT labor flow data will lower it, potentially reducing the upward bias by an order of magnitude. Although the computed biases are unlikely to match the biases we observe in our empirical analysis below because of the simplicity of our analytic model, they provide the basic insight that differences in transmission paths can either increase or decrease the size of the bias term.

Table 8 shows the results of regressions comparing the performance of the firm and establishment level spillover measures. Because of the limited availability of the establishment level and labor flow data, we use the IDG capital instrumental variables modified with the missing dummy variable because this produces the largest overall sample. All regressions are pooled in levels, with controls for industry and year. Columns (1) and (2) show the firm-level results when the spillover term is computed using firms in the same four-digit SIC industry, and the sample is restricted to observations for which the establishment level data are also available. Columns (3) and (4) show the results from the comparable sample when

Table 8 Impact of Measurement Error Correction on IT Spillover Estimates Using Different Spillover Models

Variables	(1)	(2)	(3)	(4)
DV: Log(<i>Value Added</i>)	Pooled OLS	IV	Pooled OLS	IV
Instrumental Variable		IDG capital + Missing		IDG capital + Missing
Log(<i>Labor</i>)	0.734*** (0.0207)	0.652*** (0.0543)	0.741*** (0.0206)	0.654*** (0.0538)
Log(<i>Non-IT Cap</i>)	0.196*** (0.0139)	0.165*** (0.0220)	0.190*** (0.0141)	0.169*** (0.0179)
Log(<i>IT Capital</i>)	0.0363*** (0.00948)	0.147** (0.0627)	0.0328*** (0.00926)	0.142** (0.0574)
Log(<i>IT Pool</i>)	0.0181 (0.0181)	-0.0128 (0.0279)		
Log(<i>IT Pool—Estab</i>)			0.0294** (0.0137)	-0.0120 (0.0255)
Controls	Industry year	Industry year	Industry year	Industry year
Hausman <i>t</i> -statistic (<i>IT Capital</i>)		2.71		2.90
Hausman <i>t</i> -statistic (<i>IT Pool</i>)		2.58		2.81
First-stage <i>R</i> ²		0.660		0.676
First-stage <i>F</i> -statistic		9.24		11.67
Prob > <i>F</i>		0.000		0.000
Observations	2,355	2,355	2,355	2,355
<i>R</i> -squared	0.967	0.959	0.967	0.960

Notes. Robust standard errors are shown in parentheses and clustered on firm; sample is limited to observations for which instrumental variables and data on both types of transmission paths are available. Columns (1) and (2) show estimates using the IT spillover pool constructed using the IT investments of firms in the same industry. Columns (3) and (4) show estimates using the IT spillover pool constructed using IT investments of firms with establishments in the same industry, weighted by establishment size. The instrumental variable in (2) and (4) is the modified IDG IT capital variable with a dummy variable included for missing values. Hausman tests are for a significant change in the IT capital or IT spillover coefficient after applying the instrumental variable.

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

the spillover term is computed using the establishment method. In both cases, the application of the instrumental variable raises the coefficient estimate on IT capital and reduces the spillover estimate, and Hausman test statistics suggest that the changes in the IT capital coefficient and the IT spillover coefficient are statistically meaningful in both cases.

In columns (1) through (6) of Table 9, we restrict the sample to the observations for which all three spillover measures (firm, establishment, labor) can be computed for direct comparability. The uncorrected elasticities of IT capital are similar across all three models. Some of the coefficient estimates on the spillover terms are not significantly different from zero in our sample, most likely due to the reduced

Table 9 Impact of Measurement Error Correction on IT Spillover Estimates Using Different Spillover Models

Variables	(1)	(2)	(3)	(4)	(5)	(6)
DV: Log(Value Added)	Pooled OLS	IV	Pooled OLS	IV	Pooled OLS	IV
Instrumental Variable		IDG IT capital + Missing		IDG IT capital + Missing		IDG IT capital + Missing
Log(Labor)	0.698*** (0.0363)	0.578*** (0.0964)	0.702*** (0.0350)	0.601*** (0.0929)	0.690*** (0.0363)	0.605*** (0.0872)
Log(Non-IT Cap)	0.230*** (0.0274)	0.215*** (0.0288)	0.236*** (0.0269)	0.228*** (0.0284)	0.237*** (0.0274)	0.230*** (0.0275)
Log(IT Capital)	0.0441** (0.0214)	0.204* (0.122)	0.0277 (0.0218)	0.157 (0.111)	0.0409** (0.0206)	0.150 (0.102)
Log(IT Pool—Firm)	−0.0489* (0.0283)	−0.0784* (0.0425)				
Log(IT Pool—Estab)			0.0692** (0.0319)	0.00724 (0.0573)		
Log(IT Pool—Labor)					−0.00765 (0.00875)	−0.0108 (0.00964)
<i>Controls</i>	Industry year	Industry year	Industry year	Industry year	Industry year	Industry year
Hausman <i>t</i> -statistic (IT Capital)		1.40		1.25		1.18
Hausman <i>t</i> -statistic (IT Pool)		1.21		1.23		0.64
First-stage <i>R</i> ²		0.666		0.692		0.663
First-stage <i>F</i> -statistic		3.40		3.48		3.93
Prob > <i>F</i>		0.036		0.033		0.021
Observations	426	426	426	426	426	426
<i>R</i> -squared	0.974	0.962	0.974	0.967	0.973	0.968

Notes. Robust standard errors are shown in parentheses and clustered on firm; sample is limited to observations for which data are available for instrumental variables and all three types of transmission paths. Columns (1) and (2) show estimates using the IT spillover pool constructed using the IT investments of firms in the same industry. Columns (3) and (4) show estimates using the IT spillover pool constructed using IT investments of firms with establishments in the same industry, weighted by establishment size. Columns (5) and (6) show estimates using the IT spillover pool constructed using firms' IT investments, weighted by the IT labor flow network. Instrumental variable in all IV regressions is the modified IDG IT capital variable with a dummy variable for missing values. Hausman tests are for a significant change in the IT capital or IT spillover coefficient after applying the instrumental variable.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

sample size. As expected, however, correcting measurement error in (2) and (4) substantially decreases the magnitude of the spillover estimate. Most importantly, the corrected estimate in (6) moves by a small amount, which suggests that very little of the bias term from IT measurement error is transmitted to the spillover coefficient because of the lower covariance between the two measures indicated in Table 7. The estimates from Table 9 indicate that the size of the bias term transmitted to the spillover coefficient can vary considerably depending on the transmission path being tested and in some contexts may not significantly alter the spillover coefficient at all.

6. Discussion

This paper demonstrates that IT spillover estimates can be too large if biases caused by measurement error in a firm's own IT investment are transmitted to the estimate of the spillover coefficient—our estimates suggest that correcting this source of bias reduces the magnitude of spillover estimates by 40% to 90%. Because of the well-known difficulties in measuring, valuing, and depreciating IT capital stocks, this is

a significant problem when measuring IT spillovers, making these biases an issue worthy of attention.

We apply instrumental variables to empirically demonstrate this presence of this bias in models of IT spillovers. A literature has focused on the errors-in-variables problem and provides guidance on the use of instruments when attempting to correct the effects of measurement error (e.g., see Griliches and Hausman 1986). We also consider factors affecting the size of the bias term. The transmission of the bias to the spillover term depends on covariation between a firm's own IT capital stock and the IT spillover pool as well as the variance of the spillover pool, so the severity of the bias term will be determined by the particular transmission paths that are used when modeling the spillover pool. Spillover mechanisms producing little covariation between IT investment and the spillover pool reduce the size of the bias term. The increasing availability of some kinds of data that are more easily collected using information technologies has created new opportunities for modeling spillover paths, so this is a useful distinction because researchers increasingly use these data to study individual microfoundations of spillovers.

Our findings have implications for policy and management. Understanding the magnitude of IT spillovers is critical for developing a complete understanding of IT productivity and has implications for how new information technologies should be deployed. However, as with R&D spillovers, estimates of the productivity of IT spillovers are subject to a number of econometric pitfalls and therefore should be interpreted with care. Spillover estimates that are too high may lead to ineffective subsidies or unrealized productivity benefits for managers hoping to capture IT spillovers. Our findings also suggest the importance of data collection on IT investments. The problems discussed in this paper are a result of the absence of accurate data describing firms' digital investments. Better data, therefore, not only facilitates better estimates of IT returns but would also improve the accuracy of IT spillover measurement.^t

Finally, although we have shown that some existing estimates of the contributions of productivity spillovers may be overstated, our points address methods, not economics. IT externalities are potentially significant and may be important in size. However, care must be taken when modeling IT spillover pools in the face of measurement errors in the constituent data, and our results suggest the importance of understanding these sources of bias.

Acknowledgments

The authors are indebted to comments from Marshall Van Alstyne as well as seminar participants at the Workshop for Information Systems and Economics and the International Conference on Information Systems.

Appendix A. Derivation of Equation (7)

Starting with a two variable specification where variables are expressed as deviations from means:

$$y = \beta_k x_k + \beta_s x_s + u.$$

Application of the OLS estimator $\hat{\beta} = (X'X)^{-1}X'Y$ yields

$$p \lim \hat{\beta} = \left[\begin{matrix} x_{k^*} \\ x_{s^*} \end{matrix} \right] \left[\begin{matrix} x_{k^*} & x_{s^*} \end{matrix} \right]^{-1} \left[\begin{matrix} x_{k^*} \\ x_{s^*} \end{matrix} \right] \left[\begin{matrix} \beta_k & \beta_s \end{matrix} \right] \left[\begin{matrix} x_k \\ x_s \end{matrix} \right],$$

$$p \lim \hat{\beta} = \left[\begin{matrix} \sigma_{x_k^*}^2 & \sigma_{x_k x_s^*} \\ \sigma_{x_k x_s^*} & \sigma_{x_s^*}^2 \end{matrix} \right]^{-1} \left[\begin{matrix} x_{k^*} \\ x_{s^*} \end{matrix} \right] \left[\begin{matrix} x_k & x_s \end{matrix} \right] \left[\begin{matrix} \beta_k \\ \beta_s \end{matrix} \right],$$

$$p \lim \hat{\beta} = \frac{1}{(\sigma_{x_k^*}^2 \sigma_{x_s^*}^2 - \sigma_{x_k x_s^*}^2)} \left[\begin{matrix} \sigma_{x_s^*}^2 & -\sigma_{x_k x_s^*} \\ -\sigma_{x_k x_s^*} & \sigma_{x_k^*}^2 \end{matrix} \right] \left[\begin{matrix} \sigma_{x_k} & \sigma_{x_k x_s} \\ \sigma_{x_k x_s} & \sigma_{x_s}^2 \end{matrix} \right] \left[\begin{matrix} \beta_k \\ \beta_s \end{matrix} \right],$$

$$p \lim \hat{\beta} = \frac{1}{(\sigma_{x_k^*}^2 \sigma_{x_s^*}^2 - \sigma_{x_k x_s^*}^2)} \left[\begin{matrix} \sigma_{x_s^*}^2 \sigma_{x_k}^2 - \sigma_{x_k x_s^*}^2 & (\sigma_{x_s^*}^2 - \sigma_{x_s}^2) \sigma_{x_k x_s} \\ (\sigma_{x_k^*}^2 - \sigma_{x_k}^2) \sigma_{x_k x_s} & \sigma_{x_s}^2 \sigma_{x_k^*}^2 - \sigma_{x_k x_s^*}^2 \end{matrix} \right] \left[\begin{matrix} \beta_k \\ \beta_s \end{matrix} \right],$$

$$p \lim \left[\begin{matrix} \hat{\beta}_k \\ \hat{\beta}_s \end{matrix} \right] = \frac{1}{(\sigma_{x_k^*}^2 \sigma_{x_s^*}^2 - \sigma_{x_k x_s^*}^2)} \left[\begin{matrix} \sigma_{x_k}^2 \sigma_{x_s^*}^2 - \sigma_{x_k x_s^*}^2 & \sigma_{x_k x_s} \sigma_{x_s^*}^2 \\ \sigma_{x_k x_s} \sigma_{x_s^*}^2 & \sigma_{x_s}^2 \sigma_{x_k^*}^2 - \sigma_{x_k x_s^*}^2 \end{matrix} \right] \left[\begin{matrix} \beta_k \\ \beta_s \end{matrix} \right].$$

This yields the following:

$$p \lim \hat{\beta}_k = \frac{(\sigma_{x_k}^2 \sigma_{x_s^*}^2 - \sigma_{x_k x_s^*}^2)}{(\sigma_{x_k^*}^2 \sigma_{x_s^*}^2 - \sigma_{x_k x_s^*}^2)} \beta_k + \frac{(\sigma_{x_k x_s})}{(\sigma_{x_k^*}^2 \sigma_{x_s^*}^2 - \sigma_{x_k x_s^*}^2)} \sigma_{x_s}^2 \beta_s,$$

$$p \lim \hat{\beta}_s = \frac{(\sigma_{x_s}^2 \sigma_{x_k^*}^2 - \sigma_{x_k x_s^*}^2)}{(\sigma_{x_k^*}^2 \sigma_{x_s^*}^2 - \sigma_{x_k x_s^*}^2)} \beta_s + \frac{(\sigma_{x_k x_s})}{(\sigma_{x_k^*}^2 \sigma_{x_s^*}^2 - \sigma_{x_k x_s^*}^2)} \sigma_{x_k}^2 \beta_k.$$

The spillover estimate is impacted by an attenuation bias and a bias that is transmitted from mismeasurement in private IT capital stocks, which move in competing directions for positive spillover values. To evaluate the conditions under which the net effect results in an upward bias for positive spillover values, we can write the spillover estimate as the true spillover coefficient adjusted by a bias term:

$$\beta_s + \left[\frac{(\sigma_{x_k x_s})}{(\sigma_{x_k^*}^2 \sigma_{x_s^*}^2 - \sigma_{x_k x_s^*}^2)} \sigma_{x_k}^2 \beta_k - \left(1 - \frac{(\sigma_{x_s}^2 \sigma_{x_k^*}^2 - \sigma_{x_k x_s^*}^2)}{(\sigma_{x_k^*}^2 \sigma_{x_s^*}^2 - \sigma_{x_k x_s^*}^2)} \right) \beta_s \right].$$

Rearranging terms produces

$$\beta_s + \left[\frac{(\sigma_{x_k x_s}) \sigma_{x_k}^2}{(\sigma_{x_k^*}^2 \sigma_{x_s^*}^2 - \sigma_{x_k x_s^*}^2)} \beta_k - \frac{\sigma_{x_k}^2 (\sigma_{x_s}^2 - \sigma_{x_k x_s^*}^2)}{(\sigma_{x_k^*}^2 \sigma_{x_s^*}^2 - \sigma_{x_k x_s^*}^2)} \beta_s \right].$$

This bias term inside the brackets is in the upward direction (positive) if

$$\frac{\sigma_{x_k x_s} \sigma_{x_k}^2}{\sigma_{x_k^*}^2 \sigma_{x_s^*}^2} > \frac{\beta_s}{\beta_k}.$$

We expect this to hold true in most samples because the measurement error variance for the spillover term is significantly less than that of the variance on own IT capital because of the convexity of the weighting terms in the

Table A.1 Bias in Spillover Term (Large Spillover Pod)

$\sigma_{x_k x_s}$	β_s/β_k					
	0.10	0.20	0.30	0.40	0.50	1.00
0.10	0.002	0.002	0.002	0.002	0.002	0.001
0.20	0.004	0.004	0.004	0.004	0.004	0.003
0.30	0.007	0.007	0.007	0.007	0.007	0.006
0.40	0.010	0.010	0.010	0.010	0.010	0.009
0.50	0.015	0.014	0.014	0.014	0.014	0.013
0.60	0.021	0.021	0.020	0.020	0.020	0.019
0.70	0.031	0.031	0.031	0.030	0.030	0.029
0.80	0.052	0.052	0.052	0.051	0.051	0.049
0.90	0.127	0.126	0.125	0.125	0.124	0.120
β_s	0.005	0.010	0.015	0.020	0.025	0.050

Notes. Columns indicate changes in the true ratio of the output elasticities between own IT capital and the spillover pool, and rows indicate changes in the covariance between the two measures. Values greater than 0 indicate an upward bias on the spillover term. The variance of the measurement error is conservatively fixed at 25% of the total variance of the IT measure. The bottom row indicates the value of β_s implied by the parameter values.

Table A.2 Bias in Spillover Term (Small Spillover Pool)

$\sigma_{x_k x_s}$	β_s/β_k					
	0.10	0.20	0.30	0.40	0.50	1.00
0.10	0.001	0.000	-0.001	-0.003	-0.004	-0.010
0.20	0.003	0.002	0.001	-0.001	-0.002	-0.008
0.30	0.006	0.005	0.003	0.002	0.001	-0.006
0.40	0.009	0.008	0.006	0.005	0.003	-0.004
0.50	0.013	0.011	0.010	0.008	0.007	-0.002
0.60	0.019	0.017	0.015	0.013	0.011	0.002
0.70	0.029	0.026	0.024	0.021	0.019	0.006
0.80	0.049	0.045	0.042	0.038	0.034	0.016
0.90	0.120	0.112	0.104	0.096	0.088	0.049
β_s	0.005	0.010	0.015	0.020	0.025	0.050

Notes. Columns indicate changes in the true ratio of the output elasticities between own IT capital and the spillover pool, and rows indicate changes in the covariance between the two measures. Values greater than 0 indicate an upward bias on the spillover term (shaded areas indicate negative values). The variance of the measurement error is conservatively fixed at 25% of the total variance of the IT measure. The bottom row indicates the value of β_s implied by parameter values.

spillover pool construction and because the output elasticity of the spillover term is expected to be less than that of the output elasticity of the spillover pool. More formally, the direction of the inequality depends upon (a) the ratio of the error variance terms, (b) the ratio of the output elasticities, (c) covariance between own IT capital and the spillover pool, and (d) the variance of own IT capital. Using the parameters from the sample in this study for (a) and (d), Table A.1 illustrates how the computed net effect ($\hat{\beta}_s - \beta_s$) changes in response to changes to (b) and (c). For reasonable values, the net effect of the bias falls close to zero for extreme values, but it is always upward. When using the parameter estimates from the OLS full-sample regressions with industry spillover pools reported in the tables,

the computed effects from Table A.1 are consistent with biases of between 40% and 55% even when using the lower value of the two covariance entries in between which the true covariance falls. Therefore, the true bias is likely to be somewhat higher.

Because the ratio of the error variance terms (factor (a)) is driven by the convexity of the weighting terms in the spillover pool construction, it will fall when spillover pools are constructed by weighting smaller numbers of external inputs. In Table A.2, we show the values recomputed by assuming that spillover pools are computed using the inputs of only 3 other firms (rather than 29, which was the mean pool size in our original sample), which brings the error variance of the spillover pool much closer to the error variance in the own IT capital measure. However, even in this case, we only observe a net downward effect (shaded areas indicate a downward bias) where the output elasticity of the spillover pool is very large compared to the output elasticity of own IT capital, and the covariance between the two measures is very low. For most reasonable values, therefore, we expect the net effect on the spillover pool estimate to be in the upward direction.

Appendix B. Summary Statistics by Firm-Year

	1987	1988	1989	1990	1991	1992	1993	1994
<i>Value-Added</i>	7.07	7.10	7.05	7.06	7.03	7.03	7.06	7.09
<i>Non-IT Capital</i>	7.68	7.66	7.65	7.74	7.78	7.81	7.84	7.87
<i>Labor</i>	6.53	6.52	6.49	6.50	6.50	6.49	6.52	6.50
<i>IT Capital (CII)</i>	2.62	2.76	2.66	2.74	2.96	2.77	3.08	3.42
<i>IT Pool</i>	3.24	3.40	3.34	3.41	3.65	3.54	3.86	4.16
Number of observations	344	364	402	402	402	402	401	401

Notes. All figures are mean values of logged variables. Total number of observations is 3,118. Includes full sample for which CII IT capital measures are available.

Appendix C. First-stage Estimates for Instrumental Variable Regressions

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:	Log(CII IT Capital)	Log(CII IT Capital)	Log(CII IT Pool)	Log(CII IT Capital)	Log(CII IT Capital)	Log(CII IT Capital)
Instrumental variables:	IDG Capital	IDG Capital IDG IT Pool	IDG Capital IDG IT Pool	IDG Capital + Missing	IDG Capital + IT Labor	IT Labor
Log(IDG IT Capital)	0.360** (0.044)	0.367** (0.088)	0.031 (0.038)	0.201** (0.037)	0.232** (0.034)	
IT Capital Missing ^a				0.748** (0.164)	0.888** (0.152)	
Log(IT Labor)					0.100** (0.034)	0.091** (0.034)
Log(CII IT Pool)		-0.170 (0.133)	0.310** (0.096)			
Log(Labor)	0.442** (0.075)	0.238** (0.089)	-0.059 (0.038)	0.710** (0.064)	0.562** (0.052)	0.634** (0.052)
Log(Non-IT Capital)	0.263** (0.057)	0.448** (0.063)	0.074 (0.047)	0.273** (0.054)	0.296** (0.032)	0.306** (0.033)

Appendix C (Continued)

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:	Log(CII IT Capital)	Log(CII IT Capital)	Log(CII IT Pool)	Log(CII IT Capital)	Log(CII IT Capital)	Log(CII IT Capital)
Instrumental variables:	IDG Capital	IDG Capital IDG IT Pool	IDG Capital IDG IT Pool	IDG Capital + Missing	IDG Capital + IT Labor	IT labor
Observations	1,013	262	262	3,118	1,626	1,626
R ²	0.71	0.72	0.82	0.66	0.64	0.63
First-stage F-statistic	67.1	38.99	203.5	29.5	9.94	1.33
Prob > F	0.00	0.000	0.000	0.00	0.000	0.25

Notes. First-stage regression results from baseline IV regressions reported in Tables 3 and 4. Standard errors (in parentheses) are clustered on firm.

(1) uses the IDG IT capital measure as an instrument for CII IT capital (from Table 3, Column 2).

(2) uses the IDG IT capital and IDG IT pool measure as instruments for CII IT capital (from Table 3, Column 8).

(3) uses the IDG IT capital and IDG IT pool measure as instruments for the CII IT pool (from Table 3, Column 9).

(4) uses the IDG IT capital and missing data dummy variable as instruments for CII IT capital (from Table 4, Column 2).

(5) uses the IDG IT capital, IT labor data, and missing data dummy variable as instruments for CII IT capital (from Table 4, Column 7).

(6) uses the IDG IT labor data as instruments for CII IT capital (from Table 4, Column 6).

* $p < 0.1$, ** $p < 0.01$, *** $p < 0.05$.

References

- Acharya R, Basu S (2010) ICT and total factor productivity growth: Intangible capital or productive externalities? Working paper, Industry Canada.
- Aral S, Weill P (2007) IT assets, organizational capabilities, and firm performance: How resource allocations and organizational differences explain performance variation. *Organ. Sci.* 18(5): 763–780.
- Aral S, Brynjolfsson E, Wu DJ (2006) Which came first, IT or productivity? A virtuous cycle of investment and use in enterprise systems. Working paper, MIT Sloan School of Management.
- Barker G, Fuss M, Waverman L (2008) The contribution of ICT to productivity in Australia. ANU Centre for Law and Economics 3. Working paper 3-2008, Center for Law and Economics Australian National University.
- Barua A, Lee B (1997) The information technology productivity paradox revisited: A theoretical and empirical investigation in the manufacturing sector. *Internat. J. Flexible Manufacturing Systems* 9:145–166.
- Barua A, Kriebel C, Mukhopadhyay T (1995) Information technologies and business value: An analytic and empirical investigation. *Inform. Systems Res.* 6(1):3–23.
- Becchetti L, Adriani F (2005) Does the digital divide matter? The role of information and communication technology in cross-country level and growth estimates. *Econom. Innovation New Tech.* 14(6):435–453.
- Berndt E, Griliches Z (1990) Price indexes for microcomputers: An exploratory study. NBER Working Paper 3378, National Bureau of Economic Research, Cambridge, MA.
- Bharadwaj A (2000) A resource based perspective on information technology capability and firm performance: An empirical investigation. *MIS Quart.* 24(1):169–196.
- Bloom N, Schankermann M, Van Reenen J (2012) Identifying technological spillovers and product market rivalry. Working paper, Centre for Economic Performance, London School of Economics and Political Science.
- Breschi S, Lissoni F (2001) Knowledge spillovers and local innovation systems: A critical survey. *Indust. Corporate Change* 10(4):975–1005.
- Brynjolfsson E (1993) The productivity paradox of information technology. *Comm. ACM* 36(12):67–77.
- Brynjolfsson E, Hitt L (1996) Paradox lost? Firm-level evidence on returns to information systems spending. *Management Sci.* 42(4):541–558.
- Brynjolfsson E, Hitt L (2000) Beyond computation: Information technology, organizational transformation, and business performance. *J. Econom. Perspect.* 14(4):23–48.
- Brynjolfsson E, Hitt L (2003) Computing productivity: Firm-level evidence. *Rev. Econom. Statist.* 85(4):793–808.
- Burton-Jones A, Straub D (2006) Reconceptualizing system usage: An approach and empirical test. *Inform. Systems Res.* 17(3): 228–246.
- Chan Y (2000) IT value: The great divide between qualitative and quantitative and individual and organizational measures. *J. Management Inform. Systems* 16(4):225–261.
- Chang Y, Gurbaxani V (2010) The impact of IT-related spillovers on long-run productivity: An empirical analysis. *Inform. Systems Res.* 23(3, Part 2):868–886.
- Chen X, Hong H, Nekipelov D (2013) Measurement error models. *J. Econom. Literature.* Forthcoming.
- Chen J, Nault B (2007) Industry level supplier-driven IT spillovers. *Management Sci.* 53(8):1199–1216.
- Chen Z, Nault B (2011) Relative industry concentration and customer-driven IT spillovers. *Inform. Systems Res.* 23(2): 340–355.
- Chou Y, Kauffman R, Shao B (2009) Small is beautiful: An empirical study of complementarities, substitution and spillovers in the IT industry of the OECD countries. Working paper, W.P. Carey School of Business, Arizona State University.
- Daveri F, Silva O (2004) Not only Nokia: What Finland tells us about new economy growth. *Econom. Policy* 19(38):117–163.
- Dedrick J, Gurbaxani V, Kraemer K (2003) Information technology and economic performance: A critical review of the empirical evidence. *ACM Comput. Surveys* 35(1):1–28.
- Devaraj S, Kohli R (2003) Performance impacts of information technology: Is actual usage the missing link? *Management Sci.* 49(3):273–289.
- Dewan S, Min C (1997) The substitution of information technology for other factors of production: A firm-level analysis. *Management Sci.* 43(12):1660–1675.
- Dutta D, Otsuka K (2006) An analysis of knowledge spillover from information and communication technology in Australia, Japan, South Korea, and Taiwan. Working paper, University of Sydney.
- Forman C, Goldfarb A, Greenstein S (2005) How did location affect adoption of the commercial Internet? Global village, urban density, and industry composition. *J. Urban Econom.* 58(3): 389–420.

- Ganley D, Kraemer K, Wong P (2003) Spillover effects of production of IT on use of IT. *Electronic Markets* 13(3):271–281.
- Gelb E, Getz D, Oberman G (2003) ICT spillovers in rural areas. Working paper, EFITA Conf. 2003, Debrecen, Hungary.
- Gholami R, Guo X, Higon M, Lee S (2009) Information and communications technology (ICT) international spillovers. *IEEE Trans. Engrg. Management.* 56(2):329–340.
- Greenan N, Mairesse J (1996) Computing and productivity in France: Some evidence. NBER Working Paper 5836, NBER, Cambridge, MA.
- Griliches Z (1979) Issues in assessing the contribution of research and development to productivity growth. *Bell J. Econom.* 10(1):92–116.
- Griliches Z (1992) The search for R&D spillovers. *Scandinavian J. Econom.* 94(Supplement):29–47.
- Griliches Z, Hausman J (1986) Errors in variables in panel data. *J. Econometrics* 31(1):93–188.
- Han K, Chang Y, Hahn J (2011) Information technology spillover and productivity: The role of information technology intensity and competition. *J. Management Inform. Systems.* 28(1):115–146.
- Hausman J (2001) Mismeasured variables in econometric analysis: Problems from the right and problems from the left. *J. Econom. Perspect.* 15(4):57–67.
- Hitt L, Brynjolfsson E (1996) Productivity, business profitability, and consumer surplus: Three different measures of information technology value. *MIS Quart.* 20(2):121–142.
- Huang P, Ceccagnoli M, Forman C, Wu DJ (2012) Network of practice, IT knowledge spillovers, and productivity: Evidence from enterprise software. Working paper, University of Maryland.
- Johnston J, Dinardo J (1997) *Econometric Methods*, 4th ed. (McGraw-Hill Companies, Singapore).
- Knott A (2008) R&D returns causality: Absorptive capacity or organizational IQ. *Management Sci.* 54(12):2054–2067.
- Knott A, Posen H, Wu B (2009) Spillover asymmetry and why it matters. *Management Sci.* 55(3):373–388.
- Kooshki M, Ismail R (2011) The impact of information and communication technology investment externalities on economic growth in newly industrialized countries. *Proc. 2nd Internat. Conf. Bus. Econom. Res., Langkawi, Malaysia*, 1282–1292.
- Lee S, Guo XJ (2004) Information and communication technology (ICT) and spillover: A panel analysis. Working paper, Econometric Society 2004 Far Eastern Meetings, Seoul, Korea.
- Levi M (1973) Errors in the variables bias in the presence of correctly measured variables. *Econometrica* 41(5):985–986.
- Lichtenberg F (1995) The output contributions of computer equipment and personnel: A firm-level analysis. *Econom. Innovation New Tech.* 3(3–4):201–218.
- Lopez-Pueyo C, Sanau J, Barcenilla S (2009) International technological spillovers from ICT-producing manufacturing industries: A panel data analysis. *Internat. Rev. Appl. Econom.* 23(2): 215–231.
- Mairesse J (1992) Time-series and cross-sectional estimates on panel data: Why are they different and why should they be equal? Hartog J, Ridder G, Theeuwes J, eds. *Panel Data and Labor Market Studies* (North-Holland, Amsterdam), 81–95.
- Manski CF (1933) Identification of endogenous social effects: The reflection problem. *Review Econom. Studies* 60(3): 531–542.
- Mithas S, Tafti A, Bardhan I, Goh J (2012) Information technology and firm profitability: Mechanisms and empirical evidence. *MIS Quart.* 36(1):205–224.
- Mahoney M, Vecchi M (2005) Quantifying the impact of ICT on output growth: A heterogeneous dynamic panel approach. *Economica* 72(288):615–633.
- Mooney J, Gurbaxani V, Kraemer K (1996) A process oriented framework for assessing the business value of information technology. *ACM SIGMIS Database* 27(2):68–81.
- Mun S, Nadiri I (2002) Information technology externalities: Empirical evidence from 42 U.S. industries. NBER Working Paper 9272, NBER, Cambridge, MA.
- Nadiri MI (1993) Innovations and technological spillovers. NBER Working Paper 4423, NBER, Cambridge, MA.
- Nguyen S, Atrostic BK (2006) How Businesses Use Information Technology: Insights for Measuring Technology and Productivity. Working Paper 06-15, Center for Economic Studies, U.S. Census Bureau.
- Park J, Shin JS, Sanders G (2007a) Impact of international information technology transfer on national productivity. *Inform. Systems Res.* 18(1):86–102.
- Park J, Shin JS, Shin H (2007b) The intensity and externality effects of information technology investments on national productivity growth. *IEEE Trans. Engrg. Management* 54(4):716–728.
- Parsons N, Phillips M (2007) An evaluation of the federal tax credit for scientific research and experimental development. Working paper, Department of Finance Canada.
- Quah D (2001) ICT clusters in development: Theory and evidence. Working paper, London School of Economics and Political Science.
- Rincon-Aznar A, Vecchi M (2004) The dynamic impact of ICT spillovers on companies' productivity performance. NIESR Working Paper, National Institute of Economics and Social Research, London, UK.
- Santhanam R, Hartono E (2003) Issues in linking information technology capability to firm performance. *MIS Quart.* 27(1):125–153.
- Saunders A (2010) Valuing IT-related intangible capital. *ICIS 2010 Proc.*, Paper 243.
- Severgnini B (2010) Is ICT a jack-in-the-box? A counterfactual approach for identifying productivity spillovers. Working paper, Copenhagen Business School, Denmark.
- Solon G (1985) Benefits and limitations of panel data: Comment. *Econometric Rev.* 4(1):183–186.
- Tambe P, Hitt L (2012) The productivity of information technology investments: New evidence from IT labor data. *Inform. Systems Res.* 23(3-Part-1):599–617.
- Tambe P, Hitt L (2014) Job hopping, information technology spillovers, and productivity growth. *Management Sci.* 60(2): 338–355.
- Tanuwidjaja E (2006) Technology diffusion: The case of information and communication technologies. Working paper, National University of Singapore.
- Van der Wiel G, Van Leeuwen H (2003) Do ICT spillovers matter? Evidence from Dutch firm-level data. CPB Discussion Paper. CPB Netherlands Bureau for Economic Policy Analysis.
- Van Reenen J, Bloom N, Draca M, Kretschmer T, Sadun R (2010) The economic impact of ICT. Center for Economic Performance.
- Venturini F (2009) The modern drivers of productivity. Working paper, University of Perugia.
- Wooldridge JM (2001) *Econometric Analysis of Cross Section and Panel Data* (MIT Press, Cambridge, MA).
- Zhang JJ, Lee ST (2009) A time series analysis of international ICT spillover. Hunter MG, Tan FB, eds. *Handbook of Research on Information Management and the Global Landscape*, 132–145.
- Zhu K, Kraemer K (2003) E-commerce metrics for net-enhanced organizations: Assessing the value of e-commerce to firm performance in the manufacturing sector. *Inform. Systems Res.* 13(3):275–295.