Economies before Scale: Survival and Performance of Young Plants in the Age of Cloud Computing*

Wang Jin[1]
MIT Sloan School of Management

Kristina McElheran[2]
University of Toronto and MIT Sloan

This Draft: July 12, 2017

Abstract

Young firms are central to productivity and job growth in the United States, yet they fail at high rates. We examine how a recent rise in firms' ability to access information technology as a service affected the survival and performance of young establishments in the U.S. manufacturing sector. Using detailed Census Bureau data, we track a large representative sample of plants from 2006 to 2014. We find that the ability to "rent" IT as needed – in particular, via cloud computing – was associated with significantly higher survival and growth among young plants. This contrasts with investments in traditional IT capital, which increased their likelihood of failure. Conditional on survival, young plants also exhibited much higher productivity than older plants from IT services expenditure. The effect was more important in IT-intensive and high-variance industries, consistent with a greater option value from reductions in the cost of experimenting with new IT. Also consistent with a learning-based mechanism, the effects are related more to age than to size, although small young plants benefitted the most from the cloud. Overall, our study provides empirical evidence suggesting that cloud computing changed how growth-oriented young plants learn about their IT requirements and benefit from *shared* economies of scale before they achieve significant experience and scale of their own.

**This paper benefited from early feedback from Kevin Bryan, Ricard Gil, Avi Goldfarb, Jorge Guzman, Bill Kerr, Ramana Nanda, Mu-Jeung Yang, and participants in the Workshop on Information Systems and Economics, the NBER Productivity Lunch, and the Queens Economics of Entrepreneurship and Innovation Conference. Outstanding support was provided by Jim Davis at the U.S. Census Bureaus' Boston Research Data Center. This research was conducted while the authors were Special Sworn Status researchers at the U.S. Census Bureau Research Data Centers in Boston and Madison. Any opinions and conclusions expressed herein are those of the authors and do not necessarily represent the views of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed. All errors are our own.*

[1] MIT Sloan School of Management, Initiative on the Digital Economy, 245 1st Street, Room E94, Cambridge, MA 02142. jwangjin@mit.edu
[2] University of Toronto, 105 St. George Street, Toronto, ON M5S 3E6, k.mcelheran@utoronto.ca

I. INTRODUCTION

Young firms have been a long-standing focus of economic research and policy for good reason. They drive a significant portion of U.S. employment and output growth (Haltiwanger et al. 2013, 2016) and innovation (Hansen 1992, Kortum and Lerner 2000). Yet young firms are more likely to die (Decker et al. 2014), and there is widespread concern about declining "business dynamism" in the U.S. due to lower startup entry and survival in many industries (Hathawy and Litan 2014a, b & c; Haltiwanger et al. 2016). Comfortingly, not all small businesses contribute equally to economic performance, and ventures with particularly high growth potential have increased in recent years (Guzman and Stern 2015, 2016).

A popular narrative accompanying this trend is that younger firms are taking better advantage of new technologies that have emerged over the past couple of decades. In particular, near-ubiquitous high-speed internet infrastructure, new software tools for storage and server consolidation ("virtualization") as well as application deployment ("containers"), and robust entry of service providers into data centers and other "cloud computing" technologies have made it possible and affordable for firms to access frontier information technology (IT) rapidly – and on an as-needed basis. Anecdotal evidence abounds that startups, in particular, are taking advantage of cloud computing to experiment and scale rapidly in ways that may be transformative for the economy as a whole (e.g., Machi 2010, Manyika et al. 2011, Ewens et al. forthcoming).[3] Yet cloud investment is typically confounded with other type IT or operating expenditures, and young firms are difficult to observe and track in large numbers. Thus, systematic analysis of this phenomenon remains scarce.

We provide large-scale empirical evidence that dramatic increases in firms' ability to access IT as a service in recent years is associated with significant changes in the survival and performance of young manufacturing plants in the United States. Older plants have continued to invest robustly in technology of

[3] The *MIT Entrepreneurship Review* reported, "Even for non-computationally-intensive start-ups, cloud computing and cloud services have their niche. The economic goals of start-ups are often more geared toward short-term survival rather than long-term financial efficiency. With cloud computing, electricity costs, real estate expenses to house hardware, and IT administrator fees are largely eliminated. Moreover, the economic benefit extends beyond the direct cost of capital equipment. That is, clouds allow companies to become more agile, with respect to changes in IT infrastructure." (Machi 2010).

all kinds. However, young plants exhibit disproportionately greater performance gains associated with their investments in new IT services. At first glance, this is a surprising result, given the longstanding association of IT productivity with both firm size and experience in prior work (e.g. Tambe and Hitt 2012, McElheran 2015) We argue, however, that the capabilities made possible by the rise of the cloud are particularly beneficial to firms facing high uncertainty, thereby making young firms particularly likely beneficiaries of this particular technological innovation.

To understand the phenomenon, we first note that cloud computing diffused very rapidly in the United States, dramatically and rather suddenly changing the possibilities for how firms access IT. However, it was not a straightforward shock to either performance or price: the changes it ushered in involved important tradeoffs. On one side, higher and irreversible fixed costs yet better control and customization with traditional "owned" hardware and software; on the other, lower-fixed cost yet higher variable costs and undifferentiated inputs from IT "rented" via the cloud. The time horizons involved also differ sharply between the two ways of accessing IT, with owned-IT capital involving much longer planning and implementation cycles.

To understand how this might matter differently for young versus old firms, we sketch arguments from mainstream models of firm lifecycle dynamics to predict how these tradeoffs will shift as firms age. Our primary hypothesis is that they matter differently over the firm lifecycle due to differences in option value. Early in a firm's life, external uncertainty about demand and supply conditions and internal uncertainty about how best to organize production are much higher than later on, after time has passed and firms have gained experience. Often, however, experience at sufficient scale or speed is impossible to acquire, and experiments can help young firms learn about which projects are worth pursuing before making risky investments (Kerr et al. 2014, Ewens et al. forthcoming). Lowering the costs of experimentation tends to increase the likelihood that projects with the highest real option values will receive investment. We argue that this logic governs not only core product market decisions, but also decisions over complementary inputs like IT. Thus, the significantly lower up-front commitment and the ability to rapidly scale IT expenditure (either up or down) via cloud technologies will tend to be more valuable early in a firm's life, when option value is high. Later, when uncertainty is resolved, firms can better maximize expected value by investing in more-specific – and hence, more-productive – owned IT capital.

Note that the ability to benefit from firm-specific IT capital may depend in part on scale, in order to cover high fixed costs, and firms also tend to grow as they age. Thus, we work to disentangle the role of firm size in shifting the tradeoffs between owned IT capital and IT services in our analysis.

We test our hypotheses using a representative panel of roughly 26,000 establishments[4] over nine years (2006-2014) from the U.S. Annual Survey and Census of Manufactures. Unexplored dimensions of these rich micro data offer information on IT investment, as well as the details necessary to model survival, employment growth, and multi-factor productivity at the plant level. The data offer many other advantages, as well. The reasonably long and representative panel allows us to present new facts about IT investment in recent years that relate to firms of all ages. We can map our findings – in both time and space – to the important technological changes represented by "cloud computing." In addition, we are able to employ a variety of econometric approaches to address endogeneity and measurement error. Finally, the ability to link our core sample to other Census data sets makes it possible to more deeply explore the specific mechanisms we propose.

Our results show that expenditures on IT services are associated with better survival, growth, and productivity in the critical first five years of a plant's life. Conversely, traditional IT capital, while correlated with productivity conditional on survival, actually is associated with a greater likelihood of failure among the young. Older plants, with greater experience and greater need for firm-specific technology, do not appear to benefit greatly from IT services.

Of course, IT investment is a choice variable and productivity analysis is fraught with identification challenges. To address these issues, we explore a number of well-known techniques for

---

[4] We hypothesize that much of the relevant uncertainty may be quite localized (i.e., varying by industry, geography, production technology, and even plant-level "know how"); thus, we conduct our main analysis at the establishment level and then explore whether this varies for multi-establishment firms.

dealing with endogeneity in productivity estimation (e.g., Blundell and Bond 2000; Ackerberg, Caves, & Frazer 2006). We also instrument for the use of IT services at the plant using lagged local intensity of data service providers (NAICS 518210) and non-local industry demand for IT services. These approaches tend to strengthen our results, consistent with the presence of measurement error in our data and heterogeneous treatment effects in our population. However, some of the results are implausibly large. We also rule out reverse causality by showing that productivity due to IT services only shows up after expenditure. Ultimately, we lean on a "collage of evidence" approach, including within-plant differences over time, robust time-varying controls, and rich mechanism tests to suggest a causal link between cloud use and young plant performance.

As we expect, the conditional correlations are much greater in certain production settings. For firms at the top quartile of expenditure, the returns are concentrated in industry contexts characterized by a greater reliance on IT capital in production prior to the cloud (2005). These are the industries where learning about the firm's specific IT needs was both important and costly prior to the new technology, and therefore where cloud-based option values will be greater. They are also somewhat higher in industries characterized by higher competition as measured by a Lerner index, consistent with greater benefits when overall continuation risks are higher. Importantly, returns from IT services expenditures are primarily concentrated in industries where learning is more difficult: taking variation in quarterly plant capacity utilization as a measure of uncertainty (about demand, internal production, supply, or any combination thereof), we find that even older plants show some productivity gains from IT services (though the benefits are still less than from traditional IT capital and significantly less than those found for young plants).

Controlling for age, we find scale – or, rather, the lack thereof – to be a relevant factor: smaller young plants benefit more from IT services than larger young ones; among older plants, only the smaller ones show benefits from the cloud. However, among the older plants of any size, owned IT capital is the primary driver of IT-related productivity. It is also worth pointing out that, to the extent that the smaller young plants are also the ones with the fastest growth rates, IT services appear particularly important for mitigating uncertainty in *high-growth* manufacturing, a sector of particular interest in the literature (Guzman and Stern 2015, 106; Haltiwanger et al. 2016).

We find almost no support for alternative explanations. Barring differential adjustment costs, if the cloud reduced IT costs with no tradeoffs, we would expect older firms – with richer complementary organizational capital, among other advantages – to benefit at least as much as the young. This is not the case. To the extent allowable with our data, we explore explanations related to adjustment costs such as legacy IT or management practices that may be systematically absent among young plants; findings here are a bit mixed but cannot account for the magnitude of the young-old divide. Financial frictions that might make it more attractive for young firms to substitute higher variable costs for up-front fixed costs are unlikely to explain our findings, insofar as young plants belonging to larger, multi-establishment firms (which have better access to both internal and external financing) show the same pattern of results as young single-unit firms. In contrast, we do find additional support for learning-based benefits from cloud investments. It takes a while, and only applies to relatively early-adopters in our sample who survive, but lagged IT services expenditure is positively and significantly associated with even higher returns to IT capital and software investment later in life (*pending disclosure review*).

These findings are significant in a research and policy context that has expressed concern over the short lifespan of new ventures (e.g. Haltiwanger et al. 2013, Decker et al. 2014), yet has had to navigate with almost no visibility into how significant technological change has affected the youngest and smallest firms in the economy. The magnitude and timing of our results are suggestive that cloud computing and similar IT services are rapidly providing a means for the young to achieve better performance before they achieve experience and scale of their own. This phenomenon is recent, understudied, and has far-reaching implications for productivity and economic growth in the years ahead.

Like most early studies of novel phenomena, our study has limitations, and a couple key details of our research design are worth emphasizing up front. To begin, while this paper speaks to a rich and burgeoning literature on the relationship between entrepreneurship and economic growth,[5] this paper is not a study of firm founding. Our data include a robust sub-sample of very young plants, yet we do not observe founding or incorporation, early financing events, nor failure prior to hiring the first employee. Moreover, we do not restrict our attention exclusively to young *firms*, per se. Many of the young plants in our analysis sample belong to larger multi-unit entities. The fact that the patterns we observe are not sensitive to this distinction is informative about how localized these mechanisms appear to be.

Also, our study takes place during and after the Great Recession. This is helpful to the extent that it provides some variation related to mechanisms of potential interest (e.g., financial constraints). However, intensified selection dynamics might have pushed more marginally productive young firms out of our sample, and we are limited in our ability to control for this selection process. We discuss this in more detail, later in the paper.

By providing a lifecycle perspective on how firms benefit from new technologies, our study makes a few contributions to different literatures. First, it contributes to the economics of entrepreneurship literature focused on experimentation and the importance of real options in young firms (Kerr et al., 2014 Ewens et al. forthcoming). One key difference is that that our focus is on how firms learn about complementary investments – rather than about their core product market – under uncertainty. Also, we are able to show that improving firms' abilities to learn about complementary inputs is associated not only with survival and growth, but also productivity at the micro level. Thus, standard concerns that market- or sector-level performance might diminish due to a reduction selection pressure (e.g., Lee and Mukoyama 2015) are not borne out. Consistent with prior work (Ewens et al. forthcoming) we find that the impact of the cloud is restricted to certain environments (i.e., this is not a broad-based phenomenon); however, we also find that the benefits extend beyond the initial startup costs of the firm –

[5] See, for example, Haltiwanger et al. (2013, 2016) and Guzman and Stern (2015, 2016, 2017) and works cited therein.

at least in manufacturing – and are associated with improved performance of investments made a number of years later.

Despite the fact– or, perhaps, because— our findings hold equally for young single-unit firms and young plants of existing firms, they contribute to the economics and strategy literature on Schumpeterian dynamics. Our evidence on growth and productivity are consistent with popular accounts that cloud computing can replace certain types of IT expertise and functionality that historically resided primarily within large firms. Moreover, it can do so quickly in a way that may lower the cost of entry (Ewens et al. forthcoming) but also may promote survival and faster, more productive growth. This can favor entrants in ways that traditional IT capital does not, and may represent a new engine for "creative destruction" (Schumpeter 1934, 1942) in the economy as a whole. Also, taken at face value, our results on plants belonging to multi-unit firms also provide insight into how incumbent firms may best seek to take advantage of new technologies – a topic of active debate in this literature and in practice (Henderson 1993; Christensen 1997; Bresnahan, Greenstein, and Henderson 2011; Gans 2016).

Due to data limitations, economic evidence disentangling age from size is relatively scarce (see Keung, et al. 2016). Yet the distinction matters because the key mechanisms – and frictions –are very different for age-based versus size-based models of firm dynamics. In our context, if the main barrier to IT-related productivity had been the scale or liquidity to cover fixed costs, then salient policy interventions would specifically target startup growth and financial constraints (e.g., subsidizing small business IT loans). However, our results point to core frictions rooted in uncertainty and learning – and the unavoidable riskiness of making irreversible investments when market and production conditions are new or highly variable. The policy solution to this challenge is less straightforward. One perspective on our findings is that the cloud provides a technological – rather than a policy – solution. Thus, at the highest level, our findings contribute a new perspective to the rich debate about the value of – and interventions required by – young firms in the U.S. economy (e.g. Hurst and Pugsley 2011). In a more academic vein, our results offer new facts to a growing stream of research emphasizing the distinction between size and age in both theoretical and empirical work (e.g., Kueng, et al 2016; Hsieh and Klenow 2014; Haltiwanger, et al. 2013 & 2016).

While size plays a secondary role in this phenomenon, this study nevertheless speaks to a gap in the IT productivity literature concerning the behavior of small firms. Prior work has tended to focus primarily or even exclusively on large – often public – incumbent firms. Research on small firms is flagged as an important area for new work (Dedrick, et al. 2003), and recent contributions (e.g., Tambe and Hitt 2012) have relied on data with a somewhat larger proportion of small and medium-sized firms. Yet, an essential swath of the firm age and size distribution remains largely absent from our understanding of how organizations of different ages and sizes take advantage of IT.

Finally, this study updates our insights on firm use of IT in the wake of rapid and fundamental technological change. The central IT adoption and productivity studies (e.g., Bresnahan and Greenstein 1996; Bresnahan, et al. 1996; Hubbard 2000; Brynjolfsson and Hitt 2003; Forman, et al. 2005, 2008, &2012; Aral, et al. 2006; Tambe and Hitt 2012; and studies cited therein) pre-date, abstract away from, or conclude with the diffusion of the commercial internet.[6] Yet "turnkey" solutions now available via the cloud have advanced the speed, scalability, and modularity of IT services dramatically over the past handful of years (Bryne and Corrado 2016). This raises important questions about how this new type of IT may or may not be productive in today's firms, in which specific organizational environments, and why. Our results suggest that established intuitions may not apply without some tweaking in the age of cloud computing. In particular, the importance of scale and intangible organizational investments may be changing and only still apply at certain times and in certain settings.

[6] See the useful review of this large literature in Forman and Goldfarb (2016).

II. PHENOMENON: THE RISE OF CLOUD COMPUTING

"IT outsourcing" of some form has been available for decades (Dibbern et.al. 2004), but, until relatively recently, it still typically required a certain level of up-front investment, both in terms of money and time. Customized engagements, often with long-standing IT outsourcing partners, helped move certain costs such as hardware and data processing infrastructure off firms' "books" – but usually with an increase in internal coordination costs (Poppo and Zenger 1998).

The mid-2000s marked a dramatic departure.[7] The "on-demand delivery of computing power, database storage, applications, and other IT resources through a cloud services platform via the internet with pay-as-you-go pricing"[8] became dramatically more available with the launch of Amazon Web Services (AWS). In 2006, Amazon started offering its beta version of Elastic Compute Cloud (EC2) services and set the price for an instance (1.7GHz Xeon processor/1.74 GB of RAM) at $0.10 per hour –a much more appealing price for startups and small- and medium-sized businesses compared to previously available products. In 2007 and 2008, Amazon introduced several larger-scale cloud computing services with much higher CPU power and more storage and RAM. From this point on, firms' ability to access quite sophisticated IT services without incurring high fixed costs was transformed (Bryne and Corrado 2016). A wider array of offerings made it possible for cloud customers to "mix and match" their requirements for infrastructure, processing capabilities, storage, and software – and to do so very quickly (McKendrick 2011).

Price Declines 2010 and onward

The price of cloud computing services experienced a sharp decline from roughly 2010 onward due to growth of AWS and entry of new providers. In late 2009, Amazon introduced various new

[7] Some would argue for an earlier inflection point: the introduction of the commercial internet introduced a lower-cost – and, in particular, a more variable-cost – IT model beginning in the mid 1990's. Even so, there was an important distinction between "basic internet" and "advanced" applications– the former being relatively cheap but also less important for firm productivity, and the latter requiring greater co-invention and complementary organizational inputs (Forman, Goldfarb, and Greenstein, 2005, 2008, 2012).

[8] This definition can be viewed from Amazon website at: https://aws.amazon.com/what-is-cloud-computing/

products and payment methods with deep discounts on cloud services. For instance, AWS introduced EC2 Reserved Instances in 2009 and lowered its price later in the same year. Meanwhile, there was also a reduction for existing on-demand EC2 services in the same year in November (Barr 2009a, 2009b, and 2009c). Around the same time, Microsoft fully deployed its own "Azure" cloud computing platform. NASA and Rackspace Hosting launched their joint open-source cloud software project in 2010. IBM launched its cloud computing series in 2011, and Oracle subsequently deployed its Oracle cloud in 2012 (Hauger 2010). The heated market competition among these large vendors stimulated dramatic price declines; the cost per "computing unit" per hour for Amazon compute services alone fell from to as low as a few cents by 2010 (see Table 1).

Cost-Benefit Tradeoffs for Different Types of IT

The key benefits of this new way of accessing IT services were how it allowed businesses to leverage large and powerful computing services very quickly, and according to changes in demand and production volume. The ability to scale this activity up *and down* as needed – and as uncertainty about market demand and internal production was resolved – is frequently cited as a key benefit for young firms that are still struggling to define their product offerings, business models, and target customers (e.g., Machi 2010).

The ability to learn from experiments with products and services on the cloud platform was probably greatest for ventures offering IT-intensive services (Ewens et al. forthcoming). Yet anecdotal evidence points to benefits to manufacturing firms from cloud-based solutions for a range of activities. Some examples include sales and marketing operations, enterprise resource planning and supply chain management, and payments. Also, manufactured products increasingly rely on bundled IT services that require data collection, storage, analysis, and communication (Columbus 2013).

A couple of concrete examples are useful. For instance, certain types of manufacturing rely heavily on computer-aided design (CAD) and computer-aided manufacturing (CAM) technologies that allow product designers and industrial engineers to model product features and design interdependencies, experiment with different design choices, and link closely to physical production equipment. Other authors (e.g., Kerr et al. 2014) have described how manufacturers that are still developing their core products may benefit from computing resources and software applications that allow them to experiment with product details before investing in actual production resources. Additionally, it is important to consider that they need to *experiment with the types of complementary IT applications and related processes that might work for them.* Young plants may be innovating based on design approaches that are not well supported by all software vendors, but the extent of the functionality gap may be difficult to assess without using the technology. Also, a key application of CAD/CAM software is collaboration between supply chain partners. Yet young operations may also be unsure about which suppliers and customers they will want to cultivate in the long run, many of whom may have different – and incompatible – systems in place. Owning the IT required for all of those experiments up-front may be prohibitively expensive; "renting" generic solutions via the cloud until specific uses and collaborations are more certain may provide better productivity and also inform later investments – i.e., creating real options that support short-term survival *and* long-term performance.

Another example comes from the increasingly prevalent practice of bundling IT functionality with physical products. Consumer electronics and cars increasingly have embedded features such as voice-activated controls, on-board diagnostics, and push notifications from manufacturers that require robust and rapidly scalable IT infrastructure. Quickly and flexibly scaling up generic capabilities until specific details are "fine-tuned" may be particularly important for young plants learning about consumer demand for these new add-ons. Again, learning from generic inputs without committing significant resources.

Why wouldn't all firms of all ages take advantage of these new capabilities? Despite the many benefits, cloud-based IT services come with a number of important limitations. The most important limitation for our hypothesis development is that the offerings are generally quite standardized and not necessarily well-tailored to important core business functions within an individual firm.[9] Relatedly, outsourced cloud solutions frequently do not allow the adopting firm to have tight control over its data or software upgrade schedule. Data security is an oft-cited concern; for certain firms or certain activities, concerns about data privacy dominate questions of cost or efficiency (Rahid 2016). In addition, cloud performance depends highly on internet connectivity and speed, which also lie outside the adopting firm's control. Unforeseen and uncontrollable downtimes do occur and have proven costly (e.g. AWS has had outages in recent years that caused thousands of businesses' websites to be unavailable for a time – see Weise 2017). Finally, at least for large-scale operations, the unit costs of the cloud were also not competitive with owned data center investments. A prominent example from outside the manufacturing sector is Dropbox, which stored all of its files on Amazon's servers until 2015, when it moved to its own servers to improve their unit economics. In a prominent interview about the switch, their vice president noted, "Nobody is running a cloud business as a charity. There is some margin somewhere."(Metz, 2016). Thus, there are many reasons for firms to prefer their own IT capital infrastructure, conditional on knowing what they need and being able to afford it.

III. BEHAVIORAL MECHANISMS AND RELATED LITERATURE

How might the tradeoffs presented by cloud computing matter differently for firms of different ages? While the price declines in IT services over this period, and the shift from fixed to variable costs are notable, we argue that it is the flexibility of the cloud services model that mattered most for young manufacturing operations, and the lack of specificity that mattered most for older ones.

[9] Bruce Schneier at the Berkman Center for Internet & Society at Harvard Law School offers a useful culinary analogy: "The downside is that you will have limited customization options. Cloud computing is cheaper because of economies of scale and — like any outsourced task — you tend to get what you get. A restaurant with a limited menu is cheaper than a personal chef who can cook anything you want."
(https://www.schneier.com/blog/archives/2015/06/should_companie.html)

Irreversible Investments under Uncertainty

Much prior work on investment dynamics emphasizes the costs to firms of any age from making irreversible investments under uncertainty (e.g., Marschak 1949; Bernanke 1983; Pindyck 1990; Dixit & Pindyck 1994). While this has been established in the context of traditional (or simply undifferentiated) capital investment, the logic extends naturally to IT capital. Prior to the rise of the cloud, firms that required significant IT in their production process had to make bets – potentially quite large ones – on hardware and software in advance of uncertain growth and with limited ability to adjust as market and operational requirements changed. The lead-time for these bets is long: building a bespoke data center can take up to 18 months (*citation pending approval by sources*). Exacerbating the tension in our context is the high rate of depreciation and obsolescence of IT assets due to rapid technological advances. Those bets could easily fail to pay off – potentially leading to the death of the firm.

Anticipating these risks, firms have incentives to underinvest in order to maintain flexibility, or "real options" (Dixit and Pindyck 1994). This will tend to promote an inefficiently low level of IT investment, on average. Moreover, this distortion will be greater when firms perceive a high option value from waiting. Uncertainty exacerbates investment delays (Guiso and Parigi 1999; Bloom, Bond, and Van Reenen 2007), and uncertainty tends to be highest in the early stages of a firm's life (e.g., Knight 1957, Schumpeter 1934).

The canonical model in this vein, due to Jovanovic (1982), assumes that entrepreneurs lack even private information about their own future profit opportunities and require time to learn about their productivity. Foster et al. (2016) explore how firms learn about external demand, showing that this typically can take a long time in manufacturing. Classic learning-by-doing models (see Bahk and Gort 1993 and Levitt, List, and Syverson 2013 and works cited therein) capture internal learning dynamics. Firms learn how to become more efficient as they produce their products, accumulating critical knowledge about their processes, equipment, employees, effective managerial practices, supply chain partners, and so on.

One way to learn more quickly without making costly initial investments is to conduct experiments. The cloud reduces the costs of experimenting with a range of things that businesses need to learn about (Ewens et al. forthcoming). In particular, they can learn about what types of IT work with the processes, customers, and partners they have – or might want to to have – in place. As long as experiments are not too costly, they generate real options for firms to reduce uncertainty (Nanda et al. 2014). Leveraging the more-affordable learning acquired in the cloud, firms can then make better IT investment decisions designed to maximize expected value of the firm.

Thus, all else equal, we hypothesize that young firms would tend to benefit disproportionately – in terms of both productivity and survival – from new opportunities to access “good-enough” functionality, delay firm-specific investments, and affordably conduct experiments to learn about their IT needs. Moreover, we would expect the benefits of better IT experimentation to be higher in industries where: IT knowledge is particularly valuable (i.e., IT-intensive industries), the baseline risk to survival higher (i.e., there is less scope for failed or costly experiments, for instance when competition keeps profit margins very thin), or when learning is very difficult (i.e., when variance is very high, creating a greater noise-to-signal ratio in the experiments that are conducted and, hence, a need for more of them or more time to conduct them). Finally, if experimentation is useful, we should see higher productivity of later investment choices compared to firms that did not invest in this type of learning.

While firms are young, the value of the learning might justify the costs of undifferentiated IT services. Over time, however, older firms face less uncertainty and can optimize the expected value of the firm through their investment choices. Thus, they will benefit more from IT investments that either start out as firm-specific or become so through investments in adapting the technology to the firm’s needs (Bresnahan and Greenstein 1996). We hypothesize that owned IT capital will therefore be more productive for older firms compared to younger ones and also compared to cloud-based IT inputs.

Financial Frictions

Not all young firms will choose to wait, experiment, or underinvest in IT. The entrepreneurship literature has extensively explored the role of financial frictions in constraining firms' early investment choices (see Kerr and Nanda 2009). In the face of financial constraints, therefore, young firms might find it disproportionately difficult to make essential investments in IT. To the extent that the cloud allows young firms to substitute higher variable costs for fixed costs, this could alleviate financial frictions in the short term. This should improve survival in industries where technology is a key input. However, the predictions on relative productivity are ambiguous, as the improved IT capability comes at a less-attractive price point and delivers less-specific IT services to the firm.

Economies of Scale

We have taken care thus far to separate mechanisms related to age from mechanisms related to size. However, much prior work has emphasized that firms operating at greater scale have an advantage in being able to spread the fixed costs of IT capital across greater output (Forman and Goldfarb 2005, Tambe and Hitt 2012, McElheran 2015). Financing may be related to profitability and revenues, which will also favor larger firms.

A key benefit of the cloud is its ability to pool resources across a wide range of firms to achieve *shared* economies of scale in IT services. This dramatically drives down the per-unit cost of providing IT services. However, the distinction between *cost* and *price* matters in our context. Large firms that own their own IT capital can internalize all of the cost savings achieved through scale. Smaller firms that can only access these economies through the cloud must pay a price that includes a margin for the IT service provider. While this difference may disappear over time (as cloud providers grow and surpass the scale achievable by all but the very largest firms), the price differential is likely still relevant for the early years in which our study takes place. This, combined with the benefits of customization and differentiation for large incumbents firms, will drive a wedge between the benefits of the cloud for young (typically smaller) versus older (typically larger firms), but for reasons unrelated to age, per se.

The distinction between age and size, while empirically challenging to pin down, is theoretically important (see Keung, et al. 2016 for a nice discussion) both for understanding the phenomenon and for making longer-run predictions. For instance, if the benefits of the cloud arise primarily from experimentation and option value, the relationships we study here are likely to be quite stable: uncertainty for startup firms is likely a permanent fixture of the economic landscape. However, if the benefits derive primarily from providing greater economies of scale than a stand-alone firm can achieve, we might expect larger firms to move to cloud-based services in the future as the scale of cloud suppliers grows over time. We test this distinction to the extent possible in data that is largely confined to the early years of this phenomenon. If costs are going down primarily due to economies of scale, we should see a positive productivity shock regardless of age, and varying strongly by firm size.

IV. DATA AND EMPIRICAL MODEL

We take advantage of a novel data set to observe this technological shift as it unfolds in the U.S. manufacturing sector and to gain purchase on the economic implications for plants of different ages and sizes. Our core data set comes from linking rich establishment-level information collected by the U.S. Census Bureau with the Annual Survey of Manufactures (ASM), the quinquennial Census of manufactures (CMF), and the Longitudinal Business Database (LBD). The first two allow us to observe critical inputs to the production function (labor, materials, energy) in order to estimate revenue-based total factor productivity. Important changes to the manufacturing surveys in 2006 make it possible to observe disaggregated IT expenditure for the first time, and the recurrence of the questions allows us to construct a large panel from through 2014 at the establishment level.

The LBD tracks the plants from the time at which they have at least one employee to the time of their failure, regardless of ownership change[10], providing critical data on the age and survival of the

[10] Note that mergers and acquisitions in our data are treated as continuing operation of the plant. "Exit," here, refers to cessation of operations.

plant.[11] Figure 1 shows the birth rate and participation rate of young plants in the entire U.S. manufacturing sector for the years covered by our sample. Notably, the birth rate of young firms dropped precipitously during the Great Recession and recovered, with a strong up-tick in 2014. The decline of young firm participation in the sector stops by 2012 and turns around by 2014, consistent with reports of very recent increases in high-growth entrepreneurship (e.g. Guzman and Stern 2015, 2016).[12]

Measuring Different types of Information Technology Investment

The Census Bureau actually started to collect data on IT investment in the manufacturing sector as early as 1982 in the CMF. The population-wide CMF required establishments at that time to report their expenditure on computers and peripheral data processing equipment as a separate line-item in the section on machinery and equipment expenditures. This question was restricted to the CMF until 2000, when it was permanently added to the ASM, which goes every year to a representative 10% sample of manufacturing plants. From 2001 on (in both the ASM and CMF), this question was moved to the section of the survey concerned with capital expenditures and depreciable assets, reflecting the realization that IT was entering the production function as another form of capital expenditure – more like production equipment – which could be accumulated and depreciated over time.

Software and data processing services became observable for the first time in 1992. This question on expensed IT was modified in 2002 to include "expensed computer hardware and supplies and purchased computer services (software, data transmission, processing services, web design, and expensed computer purchases)." However, everything was combined into one question until 2006, when the form provided separate lines for the following types of IT: 1) capital expenditures on "computers and peripheral data processing equipment," 2) operating expenses on "expensed equipment – expensed computer hardware and other equipment," 3) operating expenses on "expensed purchases of software –

[11] Note that "exit" in our data truly means closing down, and not acquisition by another firm.
[12] All of our findings are robust to truncating the panel in 2013.

purchases of prepackaged, custom coded or vendor customized software," and 4) operating expenses on "data processing and other purchased computer services."

We calculate traditional IT capital stock using the first measure, accumulating it over all the years we observe in the data (2002, onward) and depreciating at the BEA rate for computer hardware. Our measure for IT services comes from the last measure and begins in 2006. This is where cloud investment will be recorded, potentially in addition to other IT services, but excluding telecom and internet access. The presence of other costs will tend to increase measurement error in our core measure, which will tend to be magnified in our fixed-effects models and alleviated when instrumental variables are employed.

We focus on the sharp contrast between IT services and owned IT capital to test our hypotheses, but include controls based on the other two items in all of our specifications: "Equipment" is based on the second measure (to distinguish it from the hardware in the first) and "Software" is based on the third. However, we refrain from inferring too much from the coefficients on these investments. There are too many different things potentially included in "Equipment" (which ranges from lab equipment, to CPUs and monitors, to fax and copy machines), making it difficult to classify, *ex ante*. Software, despite being classified with the variable-cost expenditures on the ASM survey form, may sometimes have very high up-front costs in terms of technology acquisition and, perhaps more important, customization to specific firm processes.

Sample Frame

For technical reasons, we restrict our sample to observations in the ASM and CMF from 2006 to 2014[13] with complete information on inputs (including cost of materials, energy, and employment), output (total value of shipments), value added, and the IT variables. In addition, we restrict our attention to establishments that have positive value-added, positive employment, and positive imputed non-IT capital stock. This procedure leaves us with about 460,000 establishment-years over the 9 years. We further

[13] 2014 data became available after we submitted our results for disclosure review and approval. Extending the panel by one year does not materially change our findings (*pending disclosure review*).

restrict our analysis to observations with reported (as opposed to imputed) IT data, though this turns out not to matter too much in our robustness checks. Our final analysis contains more than 26,000 establishments per year over 9 years. Table A1 in the Appendix shows the descriptive statistics for the ASM-based sample before dropping imputed observations.

Industry Information

We take advantage of the large sample size and fine-grained industry classifications (down to the 6-digit NAICS level) to control for industry heterogeneity in factor shares. In most specifications, in fact, we control for interacted year and industry effects; this effectively controls for industry-specific deflators in our productivity estimation.

Young Plant Coverage

While we have visibility to a large number of establishments that are within the first year of their life, we focus our analysis on plants that are five or fewer years old. This subsample represents around 18.5% of the data that meet our aforementioned restrictions, totaling approximately 5,000 plant observations per year. Prior work (e.g., Haltiwanger 2013, 2016) finds the five-year cutoff to be meaningful; our empirical results also support this cutoff in our setting. When exploring growth models, the sample size reduces mechanically, since this requires establishments to exist for at least two consecutive years. Fixed-effects estimates also are identified only off of the plants that persist for two consecutive years and change their investment year-over year.

For comparison with prior work, it is useful to note that the young plants are also small and more likely to die. The average size of a young plant is fewer than 50 employees –much smaller than is typically observed in related studies (e.g., Tambe and Hitt 2012). The average annual failure rate over our panel for young plants is 3.33% - compared to 1.35% for older plants (see Table 2 for descriptive statistics).

Productivity Estimation

We take a conventional approach to modeling the plant production function (e.g., Brynjolfsson and Hitt 2003, Bartelsman and Doms 2000, Bloom et al. 2012, Tambe and Hitt 2012). Consider a production function that is Cobb-Douglas as given in equation (1):

$$Y_{it} = A_{it} K_{it}^{\alpha} IT_{it}^{\beta} L_{it}^{\lambda} M_{it}^{\gamma} X_{it}^{\mu} \tag{1}$$

Where Y_{it} is total revenue, A_{it} is technical productivity, K_{it} denotes the establishment's non-IT capital stock at the beginning of the period, IT_{it} is the establishment's IT capital stock at the beginning of the period L_{it} is labor input, M_{it} is the establishment's consumption of material and energy inputs, and X_{it} is a vector of additional factors such as industry and age of the plant.[14] We take revenues as our main dependent variable, but also explore value-added and other measures of performance in light of concerns about capital mismeasurement in young establishments (Bartelsman and Doms, 2000) and interest in other measures of firm performance (e.g., Haltiwanger, et al. 2013). Equation 1 can be considered a first-order approximation of a more general production function (such as translog), which we examine later in our robustness checks.

Taking logs provides a tractable form to take to the data:

$$Log(Y_{it}) = \alpha \log(K_{it}) + \beta \log(IT_{it}) + \lambda \log(L_{it}) + \gamma \log(M_{it}) + \mu X_{it} + p_i + \varepsilon_{it} \tag{2}$$

where the productivity term can be decomposed into a set of plant fixed effects p_i and an added stochastic term, ε_{it}.

We innovate by separating out the different types of IT investment into traditional IT capital (ITk_{it}) and IT services ($ITpurch_{it}$), while controlling for other observed IT-related investments ($ITother_{it}$). Also, because we are primarily interested in how the coefficients on productivity vary over the lifecycle, our core

[14] We include the X_{it} controls in exponential form for the convenience of including them in levels rather than logs.

specifications interact the input variables with an indicator of being *YOUNG* or split the sample according to this distinction. Equation 3 (or equivalent sample splits) represents our core estimating equation:

$$\begin{aligned}Y_{it} = \alpha_0 + \beta_1 ITk_{ijt} + \beta_2 ITpurch_{ijt} + \beta_3 ITother + \beta_4 YOUNG_{ijt} + \\ \beta_5 YOUNG_{ijt} \times ITk_{ijt} + \beta_6 YOUNG_{ijt} \times ITpurch_{ijt} + \beta_7 YOUNG_{ijt} \times ITother_{ijt} \\ + \alpha_1 X_{ijt} + \alpha_2 YOUNG_{ijt} \times X_{ijt} + year_t \times ind_j + p_i + \varepsilon_{ijt}\end{aligned} \tag{3}$$

Where *i* denotes the plant, *j* denotes the industry, and *t* denotes the year. X_{ijt} is a vector of plant-level time-varying controls and inputs to the production function, including capital, labor, and materials. Lower-case letters denote the log transform of the variable.

A primary concern in this literature had centered on the opportunity for unobserved demand or productivity shocks to simultaneous boost output and IT investment, creating a spurious relationship and upwardly biased estimates of the productivity impact of IT on economic output. We leverage a number of econometric techniques (Blundell and Bond 2000; Levinsohn and Petrin 2003; and Ackerberg, Caves, and Frazer 2006) that, while demanding a great deal from the data, leverage dynamic panel data estimators to identify the coefficients of interest. We also explore the timing of the effects, in order to at least rule out reverse causality

Instrumental Variables

To try to disentangle the causal impact of IT services in our data, we explore two instruments. Based on rich prior work showing that technology diffusion has a local geographic component (e.g., Griliches 1956; Rogers 1995; Baptista 2000; No 2008; Forman et al. 2005; Conley and Udry 2010; Tambe 2014), we take advantage of the exhaustiveness of the LBD data on U.S. non-farming business to calculate the percentage of establishments in a focal plant's county that are classified as being in the data hosting and processing industry (NAICS 518210). This industry contains the providers of a variety of IT services including application service providers (ASPs), automated data processing, computer data storage, computer time leasing, and computer time sharing services, to name a few. All aforementioned services constitute the main services provided by Amazon Web Services and other larger cloud-computing vendors.[15]

We take the location of these service providers as conditionally exogenous with respect to local demand based on an influential report published at UC Berkeley's Reliable Adaptive Distributed Systems Laboratory in 2009 (Armbrust et al. 2009) arguing that the location of large-scale data centers and cloud services providers is more cost-efficiency-driven than demand-driven. In particular, they provide evidence that key drivers of location choices for large cloud-computing providers centered on the cost efficiency of electricity, cooling, labor, and other operating costs. Other descriptions also cite real-estate costs as a driver of location choice (*citation pending source approval*). Among those, electricity and cooling play a dominant role as these two factors alone account for approximately one-third of the total costs. We are able to control, separately for energy expenditure at the plant level (which will incorporate local energy prices) to control for this. To quote the report, "Physics tells us it's easier to ship photons than electrons; that is, it's cheaper to ship data over fiber optic cables than to ship electricity over high-voltage transmission lines." (Armbrust et al. 2009). We furthermore lag the measure by two years to reduce the likelihood of simultaneity.

Figure 2 shows a map of the county-level intensity of data-services providers as a percentage of all local non-farm establishments from 2010 (*2006 and 2010 coming soon*). Notably, the areas of greatest concentrated are sprinkled all over the U.S., many of them in unexpected locations, such as northern Virginia and Quincy, Washington.

For our second instrument, following recent studies (e.g., Forman, Goldfarb, and Greenstein 2008, 2012; Augereau, Greenstein, and Rysman 2006; Forman. and McElheran., 2016), we introduce a measure of the IT services expenditure of plants in the same industry outside the focal plant's county. This will shift demand for IT services as a function of common industry-level input shares and supplier-driven spillovers (Cheng and Nault 2007), but not due to local geographic spillovers.

---

[15] For more details about NAICS 518210, see
https://www.census.gov/econ/isp/sampler.php?naicscode=518210&naicslevel=6#

V. Results

Investment Levels across IT Types and Age Groups

Before looking at survival and performance, we first examine how investment in these different types of IT changed over the period covered by our sample, for both young and old plants. The descriptive statistics alone are informative. Looking at mean investment levels of IT capital (computers and data processing equipment, winsorized at the 1st and 99th percentiles of the distribution)[16] by age group over time in Figure 3, we observe that older firms, on average, cut back on investment in new IT capital at the time of the financial crisis in 2008 more readily than young firms. Young firms delayed their cutbacks and also cut less in absolute magnitudes. Over time, the combined effects of lower level of annual spend, having less time to accumulate IT capital (due to their age), and deferred investment show up in dramatically lower and falling accumulated IT capital stock for younger plants, as shown in Figure 4.[17] In contrast, expenditure on IT services increases (though moderately) for young plants, while initially falling for older plants as the recession hits, and then recovering and climbing over time (Figure 5). It is worth emphasizing that these are nominal expenditure levels. To the extent that prices for IT are declining – and precipitously so for cloud computing in later years – this represents a significant increase in the *quantity* of IT purchased by manufacturing firms.

Figure 6 shows IT capital levels and IT services expenditure on a per-employee basis for young and old plants. A central take-away from this figure is that, adjusted for size, the expenditure levels between young and old are not very different, and all are trending up over time – or, in the case of IT capital, in the years following the Great Recession. There is a noticeable increase in 2014 in IT

[16] Winsorization is used to address outliers in the empirical distribution for investment and calculated capital stock variables from the Census survey. See Giroud (2015) for an example in the prior literature.

[17] This figure reports the winsorized IT capital stocks for young and old firms created using a conventional perpetual inventory method and using deflators provided by the US Bureau of Economic Analysis (e.g., Bloom, et al. 2013).

expenditure, particularly for young plants. (We verified that none of our findings in any of tables to follow are sensitive to this late trend in our data.)

To examine whether these differences between young and old are significant – and to control for potential differences between young and old in terms of industry composition – Table 3 presents results from multiple regression (OLS) with an indicator "Young" for being five or fewer years old regressed on different categories of IT expenditure. These specifications include unreported industry controls (at the 6-digit NAICS level),[18] as well as either unreported year controls (columns 1-2) or an indicator for being in the "Late" years in our sample (2008-2014 – columns 3 through 4d). The latter highlights trends during and after the financial contraction of 2008-2009.[19] As in Figure 6, this table describes expenditure on a per-employee basis to better compare firms of different sizes. Consistent with Figure 6, the differences, while statistically significant at the 1% level, are economically quite small. Column 1 indicates that the average (winsorzied) accumulated IT capital stock per employee for plants that are five or fewer years old ("Young") was 5% lower[20] on average than for their older counterparts. Annual *expenditure* (on IT of all kinds) in column 2 is also lower by about 2%.[21]

Columns 3 and 4 explore how much this changes over our sample timeframe. Consistent with the descriptive statistics, we observe that even within industries, IT capital accumulates faster than it depreciates, ending up 12% higher on average in the later years (2008-2014), though it remains lower for the young plants on average by 6% percent. In column 4, IT flows (all kinds combined) per employee go down, on average, about 8% in the later years. The differences between the young and old when we look at average flows are not statistically different from zero.

---

18 An example of the narrowness of the North American Industry Classification (NAICS) at the 6-digit level would be the distinction between motor and generator manufacturing (NAICS 335312) and switchgear and switchboard manufacturing (NAICS 335315) within the category of Electrical Equipment Manufacturing (NAICS 3353).

19 We report the coefficient for all of the later years combined for ease of exposition, noting that the results are similar for both the recessionary years (2008-2009) and for those immediately following (2010-2014). This is consistent with research showing the persistent effect of the Great Recession on investment (Fort et al. 2013).

20 This percentage change is calculated by taking the coefficient of 0.063 and dividing by the mean of IT capital stock per employee for the combined sample, which is 1.27 (not reported in Table 2).

21 The sample mean for IT expenditure of all kind per employee is 1.06 (not reported in Table 2).

However, differences appear when we break the IT flows down by type in columns 4a-4d. In columns 4a and 4c we observe that, controlling for industry, young plants spend slightly more on a per-employee basis than the old on IT capital and equipment (though only the IT capital coefficient is statistically significant, at the 10% level). The young actually spend significantly less – about 9% less – on a per-employee basis on IT services than the old do; similarly for software expenditure, at 16% less (se also Table 2). The overall trend, regardless of age or industry, is towards less IT capital and equipment in later years, with more IT services and software, suggesting a substitution away from hardware towards software and services across the board. We explore this in more detail, below.

The interaction effect of being young and later in the sample is consistently negative in these models, but never statistically significant. This may be due, in part, to large reductions in young-firm employment levels during the Great Recession (Zarutskie and Young 2016) that may have exceeded the rate of IT expenditure contraction in the later years. Another interpretation of this pattern is that, while all plants cut back their level of investment in the recessionary years, startups may face basic initial investments that are difficult to defer. This has implications for young firm survival, which we explore later.

Expenditure Shares

Given the importance of lifecycle dynamics for young firms, it is an open question whether meaningful comparisons may be found primarily in expenditure *levels*, even adjusting for headcount at the firm. In our sample, the young are, on average 2.5 years old, while the older group has a mean age of 25 years (see table 2). Thus, we also explore the shares of observed IT spend across the different types of IT for both young and old in Table 4.

Controlling for industry and year fixed-effects, column 1 of table 4 shows that both young and old plants allocated about 5.4 percentage points less of their IT spend to accumulating computer and data processing equipment in the recessionary and post-recession years.[22] Conversely, both groups also allocated higher percentages of their budgets to IT services (2.1 percentage points) and software (3.6 percentage points) in later years. These results are significant at the one-percent level and consistent to what we observed in Table 3. That said, young plants exhibit a small but statistically significant greater shift into IT services as a percentage of expenditure, and it shows up in the years associated with the diffusion and price decline of the cloud.

A key missing piece of the puzzle is how many "units of IT services" firms are able to acquire through different types of IT expenditures, and how much they may be able change them via these modest adjustments. Considering the observed reduction in prices for cloud-based IT during our sample period, and combined with anecdotal evidence concerning the rapid increases in speed, quality, and reliability of cloud computing over this time, it is likely that the expenditure levels and allocation underrepresent the actual change in use of IT services. The actual shift in IT input mix might exceed the reallocation of dollars. Unfortunately, quantities of IT inputs are beyond the reach of our data.

For completeness, it is worth recognizing that the cost of IT capital may be falling over this period, as well, as Moore's Law continues; however, the cost savings from this channel should show up for both in-house IT and outsourced IT services. To best understand how changes in price of the new technology impacted performance, we turn to measures of survival, growth, and productivity.

Survival

To test hypotheses about how different types of IT might relate to survival, we model the survival process directly using a Cox proportional hazard model, with failure as the dependent variable (other related survival models produce similar results). Table 5 presents these findings. We split the sample between young and old for ease of interpretation, so comparisons run across pairs of columns. Recall that numbers above 1 indicate an increased chance of failure; numbers below 1 indicate improved survival.

[22] Again, since the results are similar for both Great Recession (2008-2009) and post-Recession (2010-2014) periods, we simplified the exposition by combing the two periods into one "Late" category. *Results available upon request.*

We start with traditional IT capital stock to establish a baseline for comparison. Strikingly, columns 1a and 1b indicate that traditional IT capital investments are associated with an *increase* in the likelihood that a young plant will fail (decreased chance of survival); while they appear to be negatively correlated with failure in older firms. The magnitude of the coefficient is 1.046 for young, which suggests that a one-unit increase in logged IT capital stock for young (about a $23,000 increase) is associated with a roughly 4.6% *higher* risk of failure. In contrast, a one-unit increase in logged IT capital stock for old (about $77,000) is associated with a 3.2% lower risk of failure.[23] This is compared to a baseline yearly failure rate of 3.3% for the young sample and 1.3% for the older sample (Table 2).

In contrast, IT services are negatively correlated with the failure of both young and old establishments. A one-unit increase at the mean of the logged IT services distribution is associated with a 4.9% (or 1.6 percentage points) lower risk of exiting for young plants and a 10% (0.13 percentage points) lower risk of exiting for older plants.

Columns 2a and 2b explore the results by using "high IT" indicators, rather than continuous expenditure measures. The cutoff for the indicator is that the plant be in the top quartile for that type of expenditure within the same NAICS4 industry (this also addresses potential concerns surrounding mismeasurement of IT that could downwardly bias the other results). The results are consistent with the first two columns, though the higher IT investment thresholds increase the magnitudes of the coefficients. For instance, young establishments with high IT capital stock have a roughly 26.3% higher risk of failure, while old establishments with high IT capital stock have a roughly 15.1% *lower* risk. In addition, all else equal, high investment in IT services lowers the exiting risk by 17.6% and 27.3%, respectively for young and old.

[23] The dollar value of this effect for young is calculated based on the mean of the logged IT capital stock distribution for young at 2.6 (reported in table 2), which is around $13,000. A one-unit increase from the mean of the logged IT capital stock is 3.6, indicating a roughly $23,000 increase. Similarly, the dollar value of the effect for old is calculated based on the mean of the logged IT capital stock distribution for old of 3.8. One-unit change in the logged value is 4.6 indicating a $77,000 increase in IT capital stock.

Growth

In addition to being concerned about survival, the business dynamics literature highlights the importance of young firms for aggregate employment growth (Haltiwanger et al. 2013, 2016). In columns 3a and 3b we observe that increases in IT investment of all kinds – with the notable exception of software – are associated with increases in employment. The differences between young and old, however, are primarily seen in the IT services expenditure. Table A3 shows the results when the combined sample is analyzed using a model that interacts all of the explanatory variables with an indicator for being young; the coefficient on Young interacted with IT services is 0.008 and significant at the 5% level. Thus, any output gains associated with these technologies are not coming at the expense of jobs in our sample. In short, IT services has a disproportionately greater association with both survival and growth for plants that are five are fewer years old, compared to older establishments.

IT Productivity in Young vs. Older Plants

Table 6a presents the results of an OLS estimation of the standard Cobb-Douglas production function described in equation (3) (columns 1, 2, 3, 4, & 7), as well as analogous models including plant fixed effects (columns 5 & 6). All columns focus on aggregate output (sales) as the dependent variable, controlling for a rich set of inputs including the cost of materials, the cost of energy (electricity and fuel combined), labor (both expenditure on temporary employees and the count of regular employees at the plant), accumulated non-IT capital stock, and year-industry fixed effects. We do not separately observe quantities and prices, so this represents a revenue-based measure of total factor productivity ("TFPR" as discussed in Foster et al. 2008). Note that the number of employees is included as an input, thus all results in this table are conditional on size measured in terms of employment.

We draw attention to the average coefficients for the different types of IT in column 1 to make a couple of observations. The first is that, despite concerns about slowing IT productivity in the wake of the internet boom of the late 1990s (Stiroh 2008), we find no evidence of slowing. All types of IT investments captured in the US Census data have positive coefficients, suggesting that they contribute to revenues in excess of their marginal input costs (subject to the standard concerns about causality and omitted input factors, which we address in more detail, below). This holds across all sizes and ages of plants.[24]

In column 1, the coefficient for IT capital stock is about 0.015 and significant at the one-percent level. The log-transformed production function allows us to interpret elasticities directly: a one-percent increase in IT capital stock is associated with a 0.015% increase in the total value of shipments, all else equal. This result is largely consistent with previous estimates of the productivity of traditional IT capital (e.g., Tambe and Hitt 2012, and Brynjolfsson and McElheran 2016).[25]

IT services, in sharp contrast, has a smaller association with productivity on average. The coefficient is not even statistically distinguishable from zero in the overall sample.

Columns 2 and 3 of table 6a are identical to column 1, but split our sample by an indicator of being five or fewer years old. *This is the core finding of the paper:* for the subsample of firms that are five or fewer years old, IT services are significantly correlated with productivity. Specifically, a one-percent increase in young firms' spending on IT services is associated with a roughly 0.01% increase in sales.

Column 4 shows results for the complete sample, but fully interacted with the Young indicator. This provides quick evidence that the difference between the IT Services coefficients for young versus older plants is statistically significant, at the one-percent level. The equivalent result for IT capital is quite the opposite. The effect of being young on the returns to traditional IT capital is negative, but not statistically significant. We cannot reject that the young receive, on average, the same productivity benefit

[24] The IT coefficients in column 1 are robust to including or excluding plant and firm age controls and indicators of multi-unit status (*not reported*).

[25] We also tested a specification with value added as our dependent variable and estimated the production function similar to Brynjolfsson and Hitt (1995). The magnitude of our estimate on IT capital stock in the value-added specification is positive and significant but slightly smaller than the results they report. Our study differs in important ways that may explain this. First, our sample includes many more smaller and younger establishments, which tend to have smaller returns to IT capital. In addition, we are controlling for more fine-grain industry codes at the level of 6-digit NAICS. Finally, we are measuring IT productivity for recent years while their sample is constructed from 1988 to 1992, so the technology in question is also different across studies.

as the old from their investments in traditional IT capital. Note that this is conditional on survival, however, which is likely upwardly biased by the increased likelihood of failure associated with IT capital expenditure observed in Table 5.

Columns 5 & 6 include plant fixed effects, which identifies the effects of interest off of *changes* within the plant. This will differ in important ways from estimates based on levels. In particular, any benefits to plants with high initial investments who stay at that level will not contribute to the estimation. It also is based on a smaller underlying number of observations, as it only uses plants that persist for at least two consecutive years (roughly 2/3 of the young sample). Perhaps more importantly, it removes the effects of unobserved time-invariant organizational capital that may serve as an important complement to IT in the production function. In Column 5, it is interesting to find that the coefficient on IT services changes hardly at all for young plants. Among older plants (column 6), the returns to IT services are now positive and statistically significant from zero, suggesting that any long-lived organizational characteristics are actually unhelpful for IT services productivity (i.e., might generate adjustment costs). However, they are still smaller than the returns that young firms enjoy from IT services, and the difference remains statistically significant.

Selection

To the extent that young plants may be making large bets on IT, unsuccessful plants will exit the sample without contributing to our productivity estimates. As discussed, we expect that this will exert a systematic upwards bias, contributing to high “excess returns” for owned IT capital. For IT services, the direction of the bias is harder to sign. To the extent that the cloud helps the marginal firm survive when it might otherwise fail, this will exert downward pressure on the observed average productivity of young investors in the cloud. On the other hand, if other selection pressures are causing the cohort of surviving young firms to be unusually productive during and after the Great Recession (Lee and Mukoyama 2015), the relationship between IT services and productivity in young plants which weather these difficult financial times -- while also taking advantage of new cloud technology -- might be upwardly biased and our estimates need to be interpreted with this selection process in mind.

Organizational Complements

Comparing the coefficients on IT capital and IT services in the models with and without fixed effects is informative about the importance of organizational capital in IT productivity from these different types of IT. To the extent that organizational capital may be time-invariant (e.g., Tambe et al. 2012), plant fixed effects will strip these effects from the productivity estimations. Consistent with prior studies, unobserved organizational capital appears to matter a great deal for the returns to traditional IT capital. The point estimate on IT capital drops by over 75% when plant-level fixed effects are included. The coefficients remaining very precisely measured, consistent with high levels of unmeasured organizational inputs interacting positively with the accumulated IT stocks.

In sharp contrast, however, the coefficient for IT services *changes almost not at all,* on average, when plant fixed effects are included (and is not statistically lower than the estimate without fixed effects). The point estimate becomes noisier due to the demands that this specification places on the data, but the pattern suggests that accumulated plant "know-how" or other organizational capital is less of a contribution to the productivity of IT services, both on average and among younger plants. This finding is consistent with anecdotal evidence that cloud-based IT is relatively standardized and therefore not reliant on the "co-invention" of processes and technology at the plant (Bresnahan and Greenstein, 1996).

How IT Productivity has Changed in the Age of Cloud Computing

An interesting question we can answer with these results is whether young or old plants appear to be more productive with their IT, *overall,* in the age of cloud Computing. Given the indistinguishable IT capital productivity between the two age groups and the distinct advantage for young plants in IT services, the combined returns show that the young are able to use generic IT in the cloud to close the productivity gap – and even overtake the older ones (conditional on survival). Table 6b reports on joint tests of linear combinations of the coefficients from Table 6a. Two observations are important to note.

The first is that failing to account separately for IT services would misrepresent the magnitude and source of IT productivity in the U.S. economy in recent years. In table A.6, we find that omitting IT services does not inflate the IT capital coefficient to the point of accounting for this omitted variable. In short, *studies that do not measure the cloud are missing a critical part of the modern IT productivity story.*

The second observation of note is that the magnitude and age distribution of the effect we estimate for the cloud is fundamentally shifts our understanding of which types of firms are productive in the age of the cloud. The differences between the young and the old – with or without plant-fixed effects – is statistically significant and redounds disproportionately to young plants (Table 6b). Thus, the combined effect of traditional IT capital and new IT services is sufficient to make *young plants the leaders when it comes to returns on IT investments*. To the extent that this is a recent trend, young establishments may be contributing more to aggregate productivity growth in ways that have been systematically missing from prior studies.

Causal Identification

Table 7 shows how these results change when we use a range of techniques developed to account for endogenous investment decisions in productivity estimation. Column 1 shows our OLS estimates from table 6a for comparison. Column 2 uses dynamic panel structural estimation relying on 2-period lagged differences for all variable investments to instrument for current-period investment levels (Blundell and Blond 200). We find that the coefficient on IT services is higher and significant at the one-percent level in this specification. Column 3 reports estimates developed by Levinsohn and Petrin (2003), using expenditure on intermediate inputs – in this case, cost of temporary employees, though results are consistent when we use cost of materials – to instrument for the unobserved productivity shock. This yields a coefficient much closer to our OLS estimate. Ackerberg-Caves-Frazer (2006) discuss some limitations of this approach; using their estimator, the effect in column 4 is again larger than our OLS estimate. The results of this table suggest that the results in table 6a may actually be biased downward.

Table 8 shows the results of our instrumental variables estimation, which again yields higher estimates, though some implausibly so. Column 1 uses only the 2-year lagged data center intensity (percentage of number of data centers to total establishments) in the local county as the instrument. It has a very high first-stage coefficient which is significant at the one-percent level, and passes all tests including the weak identification, under identification, and endogeneity tests. Column 2 instruments for IT Services expenditure using the average industry IT Services adoption outside the firm and county of the focal establishment. The results in column 2 pass both the weak and under identification tests but fail to reject the null hypothesis that IT services is exogenous. Column 3 combines the two. The coefficients in all columns are larger and are precisely measured in columns 1 and 3. Taken at face value, they are consistent with non-trivial measurement error in the IT services variable (which is plausible, given that the survey includes a number of different expense categories in its definition). They are also consistent with strong local average treatment effects (Angrist and Pischke 2009), whereby the instrument is picking up a stronger productivity response among plants that are also sensitive to the instrument – i.e., whose IT services expenditure is strongly affected by the presence of local suppliers and/or by industry-level factors. That said, the dramatic jump in the magnitude of the coefficients is reason to take these estimates with a healthy dose of skepticism

Table 9 explores the timing of these effects to further explore whether these estimates may be interpreted as causal. Following the approach for a standard Granger Causality Test, (Granger 1969) it shows that the timing of the effects runs from IT services investment to productivity, not vice versa. In columns 1-3, we show models of lagged IT investments regressed on current sales (including the usual controls for other inputs). Column 1 shows a productivity correlation with lagged IT variables. The results here indicate that expenditure on lagged IT services is positively and significantly correlated with current sales, all else equal. In column 2, we further include lagged sales in the model as a control – and find that the effect of lagged IT services disappears. This is consistent with IT services having a contemporaneous (at least within the same year) effect on sales that disappears once we control for the productivity effect from that year. In addition, much less variation is left to explain once we control for lagged sales, as output is highly serially correlated (Foster et al. 2016). Column 3 goes on to include forward IT services in the model. For reverse causality to be a problem, we would expect a relationship between past sales and future IT investments, but we do not observe this to be the case for IT services. Interestingly, we see the potential for reverse causality with traditional IT capital, which would make sense if past performance is required to access the financial liquidity needed to make future investments or if learning is important for productivity from this type of investment.

To complete the analysis, we reverse the dependent and explanatory variables in columns 4 and 5, putting IT services on the left-hand side. Here, we see that there is serial correlation in IT investment, but that the relationship between past sales and current investment in IT services is negative and of arguable statistical significance; in fact it loses statistical significance in column 5 when we separately control for forward sales (which is positive and significant and in line with the causality in the behavioral model). In addition to the estimation procedures discussed above, we take this as reasonable evidence for a causal relationship between IT services investment and productivity in our data.

Is this Really the Cloud?

Given our concerns about measurement error, we further probe the extent to which we can attribute these effects to cloud computing, per se, as opposed to other outsourced IT services (such as IT consulting). Table 10 shows that the effects do not show up until the price declines and greater diffusion of cloud computing from roughly 2010 onward. To our knowledge, other types of outsourced IT services did not experience similar price shocks (if anything, the Great Recession should have pushed prices for consulting and related services down, earlier, but this did not have an effect we can observe in our data). Column 1 reports on our continuous measure for IT services, column 2 uses a “High IT” indicator to capture expenditure that is in the top quartile for the plant’s industry. Both columns show that our effects are largely confined to the later years, when cloud computing had diffused further and fallen in price. For the higher levels of expenditure, the effects show up a bit earlier, during the recession years. There is no variation in the IT capital stock coefficients across these periods.

We also correlated our IT services measure with an external data set that specifically provides information on cloud computing use. Our measures are very highly correlated at the county-industry level (*results pending disclosure review*).

Mechanism Tests and Nuances: Industry Context Variation

Table 11 further explores the nuances and underlying drivers of these patterns. Again, restricting attention to the subsample of young plants, columns1 and 2 interact the IT variables with an indicator of whether the plant is in an industry that intensively relies on IT as an input. This is constructed by identifying the 3-digit NAICS[26] industries with above-sample mean IT capital in 2005 (pre-cloud). Examples of typically IT-intensive industries (Jorgenson et al. 2007) include printing, semiconductors, instruments manufacturing, aerospace, and other transportation equipment. We consider these to be industry settings where learning about IT would tend to be particularly important. This does not account for shifts over time in the IT-intensity of industries, but has the virtue of being uncontaminated by industry variation in cloud diffusion (which is likely to be more endogenous).

For the continuous measure of investment in column 1, the interaction is noisy. For indicators of top-quartile (by NAICS 4) expenditure (which applies for all of the IT variables in the table, not just IT services) in column 2, the interaction is large at 0.062 and significant at the five-percent level. All of the correlation of IT services with plant productivity in this specification shows up where our hypotheses predict it would matter the most.

Column 3 of table 11 interacts the IT variables with an indicator of being in a "high-competition" industry.[27] This indicator includes roughly 200 6-digit NAICS industries and includes a wide range of activities from fluid milk manufacturing to boxes to fabricated metal and motor vehicle metal stamping. We would expect continuation risks to be higher in more-competitive settings were margins are thinner,

[26] 3-digit NAICS codes are required to keep plants in well-defined categories throughout our sample, when there were some NAICS industry classifications at finer levels of aggregation.

[27] We calculate the Lerner Index for the plant's 4-digit NAICS industry, taking the bottom quartile as indicative of being in a more competitive industry context.

and thus where the consequences of failed or costly experiences would be greater. The interaction term is again positive, but significant at only the ten-percent level. The main effect persists at a similar level of statistical significance. We interpret this as mixed evidence for our expectation that we would see that the effect of cloud-based show up disproportionately in settings where firms face lower profit margins.

Next, we explore the impact of market and production variability on these estimates. When variance is higher, learning is more difficult. Columns 4 and 5 show that the effects of IT services on productivity are dramatically increased in industries where the yearly variance in plant capacity utilization for the industry (6-digit NAICS) is higher. For the young, the coefficient on the interaction term is 0.015 and significant at the five-percent level. Interestingly, in these specific industry contexts, even older plants demonstrate productivity gains from the cloud. The coefficient is much smaller at 0.005, but also statistically significant. We interpret this as confirmatory evidence for our hypothesized importance of uncertainty in driving this overall patter in the data.

Age vs. Size

Next, we explore the extent to which these effects are due to age versus size, a distinction that has become increasingly important in the firm lifecycle literature (Keung et al. 2016). We test this distinction by interacting Young, size indicators, and the same set of IT variables in Table 12.[28] While the previous productivity results all control for size measured as total employment, here, we explore a sharper cutoff to further differentiate the effects. We construct an indicator for "Large" that is equal to one when the plant has a total number of employees greater than or equal to the sample median for their 4-digit NAICS industry. Unsurprisingly, large plants are much more productive in our sample, on average.

The interaction effects between the different IT measures and size are surprising in light of prior work showing that IT tends to be more productive in large firms (e.g., Tambe and Hitt 2012, McElheran 2015). These coefficients are not statistically different from zero, except for IT services, which are

[28] Inputs to the production function are all included, but not interacted in order to preserve degrees of freedom.

significantly (in both the economic and statistical meaning of the word) *less* productive in large plants compared to small ones.

To get a clearer sense of how age and size interact, we estimate linear combinations of the coefficients and test their joint statistical significance in table 12. Some meaningful patterns emerge. When it comes to traditional IT capital, the effects are primarily about age: old plants of any size enjoy higher IT capital productivity benefits than young plants of any size. These results conform to a model of firm lifecycle dynamics based only on age effects, where learning and co-invention take time.

The results on IT services require some care to interpret. If we organize plants into size-age categories: young and small, young and large, old and small, old and large, we find that the plants that benefit the most from IT services are *both young and small (YS).* The joint coefficient, which is significant at the one percent level, shows an output elasticity of 2.2%. Plants that are *young and large (YL)* benefit, too, though not as much. Their output elasticity is 0.9% and statistically undifferentiated from the comparison group, which is *old and small (OS)* and has an output elasticity of 1.2%. Plants that are both *old and large (OL)* derive some positive benefit from their investments in IT services, but the magnitude of the effect is considerably smaller at 0.2%.

The first and last results are the easiest to interpret. *YS* plants are the ones we would expect to have the most frictions according to all of the lifecycle models. While these results provide the strongest evidence for the impact of cloud computing on young plants, they are in some respect the least informative, in that they do little to disentangle the underlying mechanisms. The *OL* plants make the most sense from an economic equilibrium perspective, because the diffusion of the technology over time and the lack of differentiation should make this sort of technology less useful for this population of firms. Adjustment costs beyond the reach of our data may also play a role.

The results in the middle provide some progress towards disentangling mechanisms. Their interpretation benefits from a lens emphasizing variation in “entrepreneurial quality” (e.g., Guzman and Stern 2015). Old small plants are plants that have had time to grow and signal their quality to the market – and yet they remain small, suggesting either that they have no aspirations to grow or limited ability to do so. For this group, some cheap (and getting cheaper) generic IT may be better than no IT, hence the association with higher productivity. To the extent that entrepreneurial quality is difficult to observe, this omitted factor would go far to explain the weaker performance of IT in smaller firms in prior studies.

The *YL* plants clearly have growth aspirations and capabilities – in fact, they had to have either entered at scale or grown very quickly to become large in five or fewer years. If size were the primary consideration, therefore, they should be more productive with all types of IT. We interpret their *weaker* performance with traditional IT capital compared to older plants and the weaker performance with IT services compared to smaller young plants as consistent with having made some big bets already (lower option value) but still being behind the older plants when it comes to co-invention (Bresnahan and Greenstein 1996) . They may also be subject to unobservable adjustment costs associated with organizational size.

Multi-Unit Status

Our analysis thus far has concerned plants, but not firms. However, roughly 70% of both young and old samples are made up of plants that belong to multi-establishment firms. All of our results are robust to the inclusion of a multi-unit indicator and to models fully interacted multi-unit status with all other inputs. Table 13 presents this evidence. This mimics table 12, but substituting an indicator for belonging to a multi-unit firm (MU) with the age indicators and all of the IT variables. The joint tests of the combined coefficients show that there is no statistical difference between young single-unit plants and young multi-unit plants. This was surprising to us, as one would expect learning from the parent firm to diffuse to new extensions of the organization. To test this intuition, we further interacted the MU indicator with an indicator for being in a different industry from the parent firm.[29] Within the young firm sample (to avoid a four-way interaction), we learned that the higher IT services productivity is concentrated *entirely*

[29] Firm industry is determined by taking the 6-digit NAICS classification of the oldest plant in the firm; ties are broken based on revenues and employment.

in the subsample of plants that are doing something different from the founding establishment of the firm (*results pending disclosure review).*

This finding is consistent with our hypothesis about learning and provides a strong contradiction to mechanisms rooted in financial frictions. To the extent that multi-unit plants should be able to rely on internal capital markets to finance IT investment (Kuppuswamy and Belen 2010), there should be a sharp distinction between these types of plants where none exists.. Conversely, to the extent that learning about the operating system, its IT requirements, and how to align the two of them to each other requires very localized learning-by-doing or may be consistent with very localized uncertainty in supply of demand conditions, this lack of a distinction is precisely what our hypothesizing would predict.

IT Interactions

Table 14 explores whether these different types of IT are complements or substitutes in the production functions for both young and old plants and provides evidence (*pending disclosure review*) consistent with learning. Controlling for industry at the 6-digit NAICS level, we observe in columns 1 and 2 that there is some substitution between traditional IT capital and IT services for both young and old plants. Although the coefficients are relatively modest; they are statistically significant at the 5% level. This is useful for interpreting the investment patterns observed at the beginning, where we some reallocation of expenditure from traditional IT capital to cloud-based IT in later years. If the types of IT were good substitutes for each other and cloud was simply cheaper, we might expect this effect to be bigger. The substitution is weaker, consistent with differing willingness to tradeoff between these two types of IT for plants of different ages.

More fine-grained data would be needed to understand the details of how firms are combining these new technologies. An important concern, however, is that firms *cannot* combine them well, and that older firms with robust legacy IT systems are unable to adjust to the cloud-based platform. Column 3 lags

IT capital stock by one period, and interacts it with all of the other IT variables.[30] The coefficient is negative, but noisy, providing no strong evidence that prior IT capital stock interacts negatively with cloud-based IT.

Columns 4 and 5 (*pending disclosure review*) lag IT services by several periods to see if our hypotheses about the benefits of cloud-based experimentation lead to better IT investments, later in a plant's life. The lag required varies by the type of IT, and can only be observed for the subset of plants that a) use the cloud relatively early, b) survive a number of years. Column 4 reports that 5 or 6 years later, the early use of IT service is associated with high productivity for IT capital. The coefficient is positive and significant at the 5% level. Within 3 to 4 years, prior IT services expenditure has a positive interaction with software investments; again, this is significant at the 5% level. While there is surely selection underlying these results, the pattern is highly suggestive.

Robustness Checks

Table 15 provides evidence that our findings are robust to a range of other econometric choices. Columns 1 shows the results for a translog specification. The statistical significant on IT services disappears, potentially in part due to the tremendous demands this specification places on the data. Other results are very consistent: including interactions and squares for all of the input variables (column 2); including regional controls in addition to industry-year fixed-effects (column 3), and using a quality-adjusted measure of labor input[31] (column 4). Columns 5 and 6 test robustness to our data cleaning choices: we get similar results if we actually include observations with imputed IT values (and more than double the sample size). We conclude that our productivity findings are also robust to outliers (probably due to the log-transformed production function), as we we see similar results using the non-winsorized values (column 6). The core of the paper centers on winsorized results because the descriptive statics

[30] We accidentally disclosed only the young subsample. However, the results for the old are qualitatively similar.

[31] Calculated by multiplying the total number of production-worker hours times the ratio of total production worker wages to total salaries and wages. See Foster, Grim and Haltiwanger (2016).

actually are sensitive to this decision and this keeps the underlying data consistent throughout the core analysis.

VI. Conclusion

This paper provides the first large-scale evidence related to the survival and performance prospects for U.S. firms in the age of cloud computing. Based on arguments from the entrepreneurial finance literature, we hypothesize that the key benefit of the cloud comes from a new ability to experiment with rapidly available, flexible, yet relatively generic IT early in a firm's life. Over time, as uncertainty resolves, firms should benefit more from firm-specific IT capital investments. The conditional correlations are strongly consistent with our hypotheses. Young firms show a much higher association between IT services expenditure and survival, growth, and performance. Older firms show little or no benefit, except in certain specifications and industry settings, and even so the estimates are much lower.

It is important to contextualize these results, as they stand in contrast to prior evidence that IT tends to be most productive in large firms. How can we reconcile these findings? To begin, we have unusually good visibility to a representative sample of young firms, along with accurate age data. Thus, we can more precisely estimate coefficients for this tail of the distribution and can disentangle age and size. Also, we explore both productivity and survival. A key insight is that firm-specific IT investments are risky for young firms – they can be productivity-enhancing conditional on survival, but they tend to promote exit as well. Firms that survive to become large will be those that have won their early IT bets.

Moreover, if we had restricted our attention only to traditional IT capital, we would have come to similar conclusions as prior studies. While our main take-away is that learning and co-invention remain central to traditional IT productivity, these mechanisms dramatically favor older, more established firms – which also tend to be large. What has fundamentally shifted is the technological landscape – and our ability to measure it.

Our results also strongly contradict common assumptions that cloud benefits are largely to the easing of financial frictions faced by young firms. We are not claiming that there is no benefit for capitalconstrained early stage ventures from substituting higher variable costs for fixed costs via this new model. However, many of the patterns in our data (in particular, the results on plants belonging to multi-unit firms) strongly contradict this as the main channel at work. Detailed mechanisms tests also support our experimentation and learning-based explanations.

Like many studies of new phenomena, our findings are not free of limitations. In particular, we worry about measurement error. While the Census data is remarkably detailed, we cannot precisely disentangle cloud-specific expenditure from other IT services. The timing and profound reversal in core patterns of IT productivity advantage just as the cloud diffuses and becomes cheaper is striking. We also corroborate our measure with and external data set on cloud use, but only at the county-industry level, Ideally, we would see IT services separated by cloud and non-cloud expenditure at the plant level. This would be a useful direction for future survey development.

Also, as we discuss throughout, standard techniques for addressing endogeneity of many types are of limited effectiveness (though the evidence against reverse causality is relatively strong). Ultimately, we lean on the rapid diffusion of the technology, the sharpness of the timing, controlling for unobserved time-invariant plant characteristics, and detailed mechanism tests to suggest that there was a causal link between this important technological change and meaningful firm outcomes.

Finally, another limitation of our study is that it tells us about IT performance over the lifecycle only in the short term. It becomes difficult to track the long-term effects of these investments in young firms due to their high baseline exit rates. Moreover, the changes are still quite recent. It would be useful to replicate this study as time passes and we can observe longer-term effects.

We can only sketch out the implications for what may follow in the long run. Uncertainty is a permanent feature of the economic landscape, and the association with firm age and certain industry settings is unlikely to change. Thus, we expect that these associations might be long-lived. This has important implications for market-level dynamics including entry, the basis of firm competition, and equilibrium market structure. Unpacking how this may play out for industrial organization over time would be an interesting area for future research.

REFERENCES

Acemoglu, D., Aghion, P., Lelarge, C., Van Reenen, J. and Zilibotti, F., (2007). "Technology, Information, and the Decentralization of the Firm." The Quarterly Journal of Economics, **122**(4), pp.1759-1799.

Ackerberg, D., Caves, K. and Frazer, G., (2006). "Structural identification of production functions."

Angrist, J.D. and Pischke, J.S., (2008). Mostly harmless econometrics: An empiricist's companion. Princeton university press.

Aral, S., Brynjolfsson, E., and Wu, D. J. (2006). "Which Came First, IT or Productivity? The Virtuous Cycle of Investments and Use in Enterprise Systems." Paper presented at the 27th International Conference on Information Systems, Milwaukee.

Armbrust, Michael; Armando Fox; Rean Griffith; Anthony D Joseph; Randy H Katz; Andrew Konwinski; Gunho Lee; David A Patterson; Ariel Rabkin and Ion Stoica. 2009. "Above the Clouds: A Berkeley View of Cloud Computing." Technical Report UCB/EECS-2009-28, EECS Department, University of California, Berkeley.

Attewell, Paul., (1992). "Technological Diffusion and Organizational Learning: The Case of Business Computing." Organization Science **3**(1):1-19.

Augereau, A., Greenstein, S. and Rysman, M., (2006). "Coordination versus differentiation in a standards war: 56k modems." The RAND Journal of Economics, **37**(4), pp.887-909.

Bahk, B.-H., and Gort, M., (1993). "Decomposing Learning by Doing in New Plants." Journal of Political Economy, **101**(4): 561-583.

Baldwin, W., and Scott, J., (2013). Market structure and technological change. Taylor & Francis.

Barr, J., (2009a). "Amazon EC2 – Now an Even Better Value." Amazon Corp. Blog, Amazon Web Services.

Barr, J., (2009b)., "Lower Pricing for Amazon EC2 Reserved Instances." Amazon Corp. Blog, Amazon Web Services.

Barr, J., (2009c)., "Amazon EC2 Spot Instances – And Now How Much Would You Pay?." Amazon Corp. Blog, Amazon Web Services.

Bartelsman, E.J. and Doms, M., (2000). "Understanding productivity: Lessons from longitudinal microdata." Journal of Economic literature, **38**(3), pp.569-594.

Baptista, R., (2000). "Do innovations diffuse faster within geographical clusters?" International Journal of industrial organization, **18**(3), pp.515-535.

Bernanke, B. S., (1983). "Irreversibility, Uncertainty, and Cyclical Investment." The Quarterly Journal of Economics, **98**(1): 85-106.

Bloom, N., Brynjolfsson, E., Foster, L., Jarmin, R. S., Saporta-Eksten, I., and Van Reenen, J., (2013). Management in America, Center for Economic Studies Discussion Papers #CES 13-01: U.S. Census Bureau.

Bloom, N., Bond, S., and Van Reenen, J., (2007). Uncertainty and investment dynamics. The Review of Economic Studies, **74**(2): 391-415.

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

Blundell, R. and Bond, S., (2000). "GMM estimation with persistent panel data: an application to production functions." Econometric Reviews, **19**(3), pp.321-340.

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

Bresnahan, T. and Greenstein, S., (1996). "Technological Progress and Co-Invention in Computing and in the Uses of Computers." Brookings Papers on Economic Activity. 1-78.

Bresnahan, Timothy F; Shane Greenstein and Rebecca M Henderson. 2011. "Schumpeterian Competition and Diseconomies of Scope: Illustrations from the Histories of Microsoft and Ibm," The Rate and Direction of Inventive Activity Revisited. University of Chicago Press, 203-71.

Bryne, David and Carol Corrado (2016). "ICT Prices and ICT Services: What do they tell us about Productivity and Technology?" Economics Working Program Working Paper Series EPWP#16-05, The Conference Board.

Brynjolfsson, E., and Hitt, L. M., (2003). Computing Productivity: Firm-Level Evidence. Review of Economics & Statistics, **85**(4): 793-808.

Brynjolfsson, E., Hitt, L. M., and Yang, S. (2002). Intangible assets: Computers and Organizational Capital. Brookings Papers on Economic Activity(1): 137-198.

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

Brynjolfsson, E. and McElheran, K., (2016). "Data in Action: Data-Driven Decision Making in U.S. Manufacturing", Center for Economic Studies Discussion Papers #16-06,

Brynjolfsson, E. and McElheran, K., (2016). "The Rapid Rise of Data-Driven Decision Making." American Economic Review Papers and Proceedings **106**(5): 133-139.

Brynjolfsson, E. and Saunders, A., (2010). Wired for innovation.

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

Christensen, C. M., (1997). The Innovator's Dilemma. New York: Harper Collins.

Conley, T.G. and Udry, C.R., (2010). “Learning about a new technology: Pineapple in Ghana.” The American Economic Review, **100**(1), pp.35-69.

Cool, K., Dierickx, I. and Costa, L.A., (2012). Diseconomies of time compression.

Cooley, T.F. and Quadrini, V., (2001). “Financial markets and firm dynamics.” American Economic Review, pp.1286-1310.

Davenport, T. H., (1993). Process Innovation: Reengineering Work Through Information Technology. Cambridge, MA: Harvard Business School Press.

Decker, R., Haltiwanger, J., Jarmin, R. and Miranda, J., (2014). The role of entrepreneurship in US job creation and economic dynamism. The Journal of Economic Perspectives, **28**(3), pp.3-24.

Dedrick, J., et al., (2003). "Information technology and economic performance: A critical review of the empirical evidence." ACM Computing Surveys (CSUR) **35**(1): 1-28.

Diamond, Douglas W., (1993). "Seniority and Maturity of Debt Contracts." Journal of Financial Economics **33**(3): 341-68.

Dibbern, J., Goles, T., Hirschheim, R. and Jayatilaka, B., (2004). “Information systems outsourcing: a survey and analysis of the literature.” ACM Sigmis Database, **35**(4), pp.6-102.

Dierickx, I. and Cool, K., (1989). “Asset stock accumulation and sustainability of competitive advantage.” Management Science, **35**(12), pp.1504-1511.

Dixit, A. K., & Pindyck, R. S., (1994). Investment Under Uncertainty: Princeton University Press.

Ericsson, K.A. and Lehmann, A.C., (1996). Expert and exceptional performance: Evidence of maximal adaptation to task constraints. Annual Review of Psychology, **47**(1), pp.273-305.

Evans, D.S. and Jovanovic, B., (1989). “An estimated model of entrepreneurial choice under liquidity constraints.” Journal of Political Economy, **97**(4), pp.808-827.

Ewens, M., Nanda, R. and Rhodes-Kropf, M., (2016). “Cost of experimentation and the evolution of venture capital.”

Forman, C., and Goldfarb, A., (2006). “ICT Diffusion to Businesses.” In T. Henderschott (Ed.), Handbook of Economics and Information Systems, Vol. 1: 1-52: Elsevier.

Forman, C., Goldfarb, A. and Greenstein, S., (2005). “How did location affect adoption of the commercial Internet? Global village vs. urban leadership.” Journal of Urban Economics, **58**(3), pp.389-420.

Forman, C., Goldfarb, A., and Greenstein, S., (2008). “Understanding the Inputs into Innovation: Do Cities Substitute for Internal Firm Resources?” Journal of Economics and Management Strategy, **17**(2): 295-316.

Forman, C., Goldfarb, A., and Greenstein, S., (2012). “The Internet and Local Wages: A Puzzle.” American Economic Review, **102**(1): 556-575.

Fort, T., Haltiwanger, J., Jarmin, R., and Miranda, J., (2013). “How Firms Respond to Business Cycles: The Role of Firm Age and Firm Size.” IMF Economic Review, **61**(3): 521-559.

Foster, L., Haltiwanger, J. and Syverson, C., (2008). “Reallocation, firm turnover, and efficiency: Selection on productivity or profitability?” The American economic review, **98**(1), pp.394-425.

Foster, L., Haltiwanger, J., and Syverson, C., (2016). “The slow growth of new plants: Learning about demand?” Economica, **83**(329): 91-129.

Foster, R. N., (2012). Innosight Executive Briefing: Creative Destruction Whips Through Corporate America: Innosight.

Friedman, T., (2011). One Country, Two Revolutions, New York Times Sunday Review.

Gans, J., (2016). The Disruption Dilemma: MIT Press.

Garicano, L. and Rossi-Hansberg, E., (2012). “Organizing growth.” Journal of Economic Theory, **147**(2), pp.623-656.

Garcia-Macia, D., Hsieh, C.T. and Klenow, P.J., (2016). “How Destructive is Innovation?” (No. w22953). National Bureau of Economic Research.

Gentry, W.M. and Hubbard, R.G., (2004). “Entrepreneurship and household saving.” Advances in Economic Analysis & Policy, **4**(1).

Giroud, X. and Mueller, H.M., (2015). “Firm Leverage, Consumer Demand, and Unemployment during the Great Recession.” mimeo, New York University.

Guiso, L. and Parigi, G., (1999). “Investment and demand uncertainty.” The Quarterly Journal of Economics, **114**(1), pp.185-227.

Guzman, Jorge, and Scott Stern., (2015). "Nowcasting and Placecasting: Entrepreneurial Quality and Performance." In *Measuring Entrepreneurial Businesses: Current Knowledge and Challenges*. University of Chicago Press.

Guzman, J. and Stern, S., (2016). “The State of American Entrepreneurship: New Estimates of the Quantity and Quality of Entrepreneurship for 15 US States, 1988-2014.” (No. w22095). National Bureau of Economic Research.

Haltiwanger, J., et al., (2013). "Who creates jobs? Small versus large versus young." Review of Economics and Statistics **95**(2): 347-361.

Haltiwanger, J., Jarmin, R.S., Kulick, R. and Miranda, J., (2016). High Growth Young Firms: Contribution to Job, Output, and Productivity Growth. In Measuring Entrepreneurial Businesses: Current Knowledge and Challenges. University of Chicago Press.

Hansen, J. A., (1992). "Innovation, Firm Size, and Firm Age." Small Business Economics **4**(1)": 37-44.

Hathaway, I. and Litan, R.E., (2014a). “Declining business dynamism in the United States: A look at states and metros.” Brookings Institution.

Hathaway, I. and Litan, R.E., (2014b). “The other aging of America: the increasing dominance of older firms.” Brookings Institution.

Hathaway, I. and Litan, R.E., (2014c). “What’s Driving the Decline in the Firm Formation Rate? A Partial Explanation.” The Brookings Institution.

Hauger, D., (2010), “Windows Azure General Availability”, Microsoft Corp. Blog, Microsoft Corp.

Henderson, R., (1993). “Underinvestment and incompetence as responses to radical innovation: evidence from the photolithographic alignment equipment industry.” RAND Journal of Economics, **24**(2).

Hsieh, Chang-Tai and Peter J. Klenow., (2014). "The Life Cycle of Plants in India and Mexico ." The Quarterly Journal of Economics **129**(3): 1035-84.

Hubbard, T.N., (2000) “The Demand for Monitoring Technologies: The Case of Trucking,” *Quarterly Journal of Economics* 115: 533-60.

Hurst, E. and Pugsley, B.W., (2011). “What do small businesses do?” (No. w17041). National Bureau of Economic Research.

Jorgenson, D.W., Ho, M.S., Samuels, J.D. and Stiroh, K.J., (2007). “Industry origins of the American productivity resurgence.” Economic Systems Research, **19**(3), pp.229-252.

Jovanovic, Boyan., (1982). "Selection and the Evolution of Industry." Econometrica **50**(3): 649-70.

Kerr, W.R. and Nanda, R., (2009). “Democratizing entry: Banking deregulations, financing constraints, and entrepreneurship.” Journal of Financial Economics, **94**(1), pp.124-149.

Kerr, W.R., Nanda, R. and Rhodes-Kropf, M., (2014). “Entrepreneurship as experimentation.” The Journal of Economic Perspectives, **28**(3), pp.25-48.

Knight, F. H., (1957). Risk, Uncertainty and Profit.(Eighth Impression.): London School of Economics and Political Science.

Kortum, S. and Lerner, J., (2001). “Does venture capital spur innovation?.” In Entrepreneurial inputs and outcomes: New studies of entrepreneurship in the United States (pp.1-44). Emerald Group Publishing Limited.

Kueng, Lorenz, Mu-Jeung Yang, and Bryan Hong., (2016). "Sources of Firm Life-Cycle Dynamics: Size vs. Age Effects."

Lee, Y., & Mukoyama, T. (2015)., “Entry and exit of manufacturing plants over the business cycle.” European Economic Review, **77**: 20-27.

Levitt, S. D., List, J. A., & Syverson, C., (2013). “Toward an understanding of learning by doing: Evidence from an automobile assembly plant.” Journal of Political Economy, **121**(4): 643-681.

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

Machi, Ethel., (2010). "Living in the Cloud: Opportunities for Entrepreneurs." MIT Entrepreneurship Review. http://miter.mit.edu/articleliving-cloud-opportunities-entrepreneurs/.

Macklem, Tiff and Robert Helsley., (2017). "Entrepreneurial Business Judgment Is Canada's Scarcest Resource." *The Globe and Mail,* February 6th.

Manyika, James, Michael Chui, Brad Brown, Jacques Bughin, Richard Dobbs, Charles Roxburgh, and Angela Huang Byers., (2011)." Big Data: The next frontier for innovation, competition, and productivity." McKinsey Global Institute.

Marschak, J., (1949). "Role of liquidity under complete and incomplete information". The American Economic Review, **39**(3): 182-195.

McElheran, K. (2015)., "Do market leaders lead in business process innovation? The case (s) of e-business adoption." Management Science, **61**(6): 1197-1216.

McKendrick, J., (2011). "How Cloud Computing is Fueling the Next Startup Boom." Forbes, November 11, 2011..

Melville, N., Kraemer, K. and Gurbaxani, V., (2004). "Review: Information technology and organizational performance: An integrative model of IT business value." MIS quarterly, **28**(2), pp.283-322.

Metz, Cl, (2016), "The epic story of dropboxs exodus from the amazon cloud empire.", Wired Magazine Business Section, March 14, 2016.

No, Angela., 2008. "Cities and Growth: Knowledge Spillovers in the Adoption of Advanced Manufacturing Technologies." The Canadian Economy in Transition.

Parker, B., (2011). Business Strategy: Cloud Computing in Manufacturing: IDC insights.

Pindyck, R. S., (1990). Irreversibility, uncertainty, and investment: National Bureau of Economic Research.

Poppo, L. and Zenger, T., (1998). Testing altermative theories of the firm: Transaction cost, knowledge-based, and measurement explanations for make-or-buy decisions in information services. Strategic Management Journal, pp.853-877.

Rashid, F.Y., (2016). "The dirty dozen: 12 cloud security threats." InfoWorld, March 11, 2016

Rajan, R.G. and Zingales, L., (2001). "Financial systems, industrial structure, and growth." Oxford Review of Economic Policy, **17**(4), pp.467-482.

Rogers, E.M., 2010. Diffusion of innovations. Simon and Schuster.

Schulz, R., Musa, D., Staszewski, J. and Siegler, R.S., (1994). The relationship between age and major league baseball performance: implications for development. Psychology and Aging, **9**(2), p.274.

Schumpeter, Joseph Alois., (1934). The Theory of Economic Development. Cambridge: Harvard University Press. Original edition, First published in German, 1912. Reprint, New York: Oxford University Press, 1961.

Schumpeter, Joseph. Alois., (1942). Capitalism, Socialism, and Democracy. New York: Harper.

Stiroh, K., (2008). “Information technology and productivity: Old answers and new questions.” CESifo Economic Studies, **54**(3), pp.358-385.

Tambe, P., (2014). “Big data investment, skills, and firm value.” Management Science, **60**(6): 1452-1469.

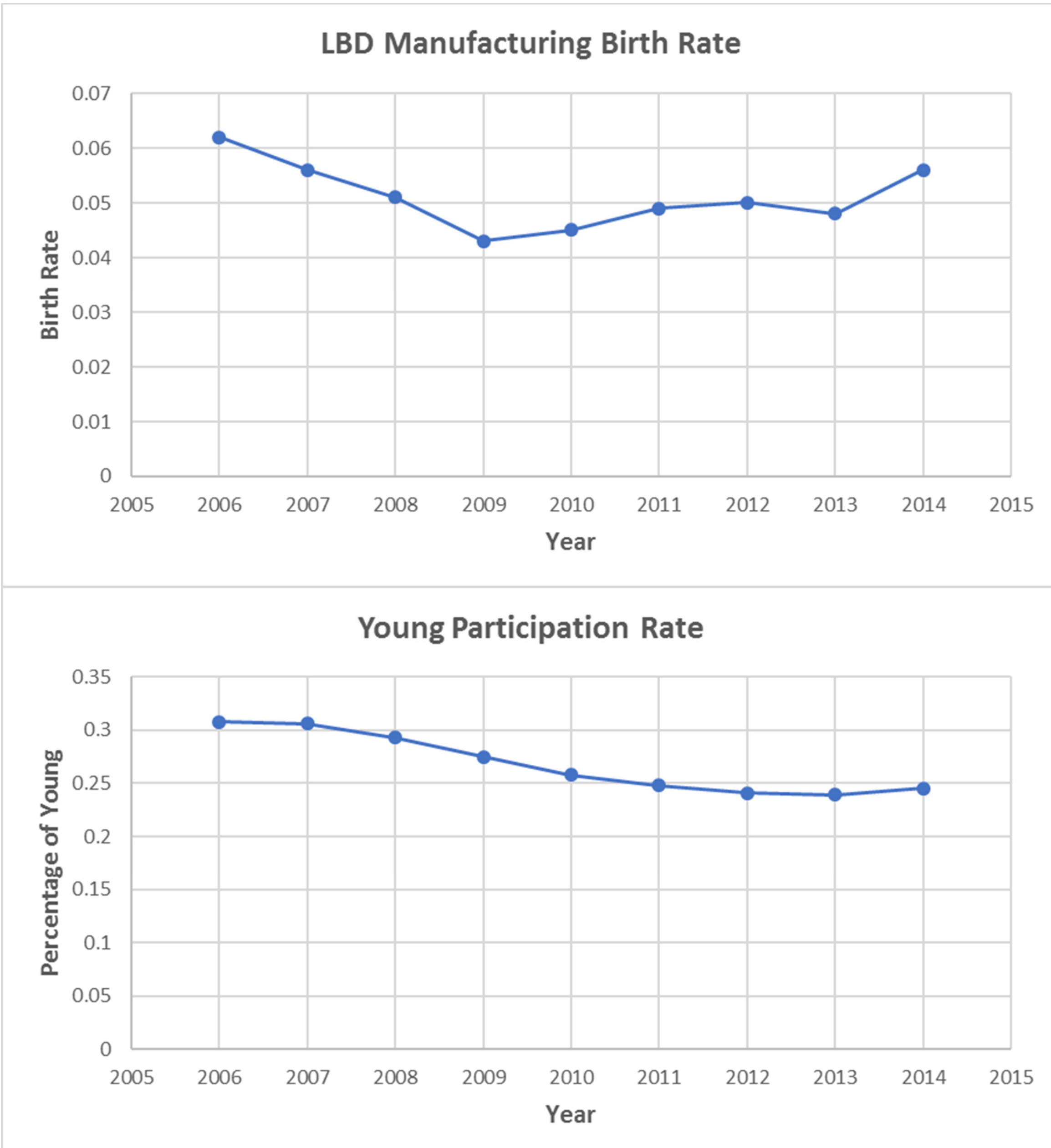
Tambe, P. and Hitt, L. M., (2012). "The productivity of information technology investments: New evidence from IT labor data." Information Systems Research **23**(3-part-1): 599-617.

Weise, E., (2017). “Massive Amazon cloud service outage disrupts sites.” USATODAY, March 1, 2017

White, T. K., Petrin, A., & Reiter, J. P., (2007). Multiple Imputations in the Annual Survey of Manufactures, Working Paper. Duke University: Economics.

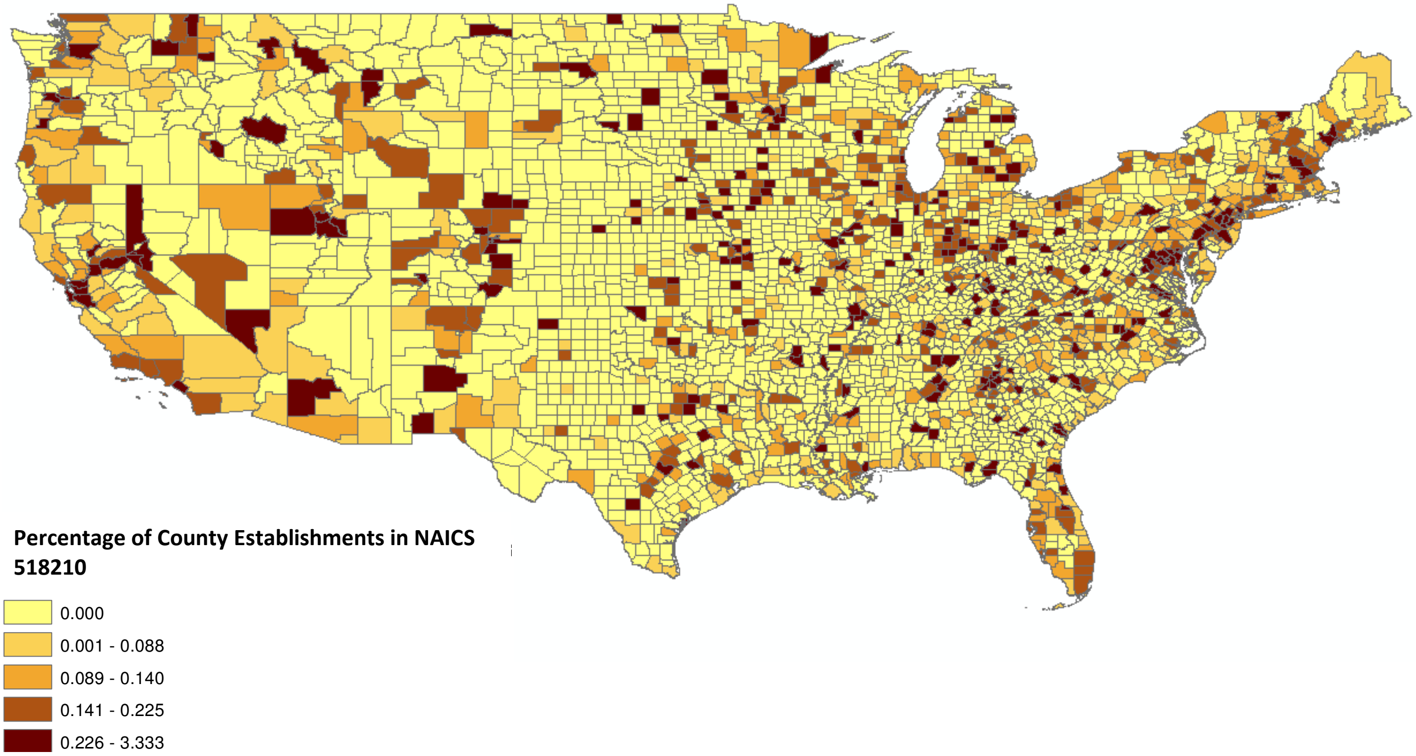
Zarutskie, R. and Yang, T., (2016). How did young firms fare during the great recession? evidence from the kauffman firm survey. In Measuring Entrepreneurial Businesses: Current Knowledge and Challenges. University of Chicago Press.

Figure 1. Population Statistics on Birth Rate and Prevalence of Young Plants in U.S. Manufacturing, 2006-2014.



Note: Based on the U.S. Census Bureau's Longitudinal Business Database for industries in the manufacturing sector, 2006 -2014.

Figure 2. Data Processing, Hosting, and Related Services Intensity by County in the U.S. 2010



Note: Values represent the ratio of establishments in Data Processing, Hosting, and Related Services (NAICS 518210) to the total number of establishments within the county. Based on public County Business Pattern 2010 data from the U.S. Census Bureau. Values reported are in percentage points – i.e., the most concentrated counties have a ratio of 0.23% to 3.33% of establishments belonging to the Data Services industry.

Figure 3. Annual Capitalized Investments in Computers and Data Processing Equipment in U.S. Manufacturing, 2006-2014 ($Thousands)
Winsorized at the 99th percentile

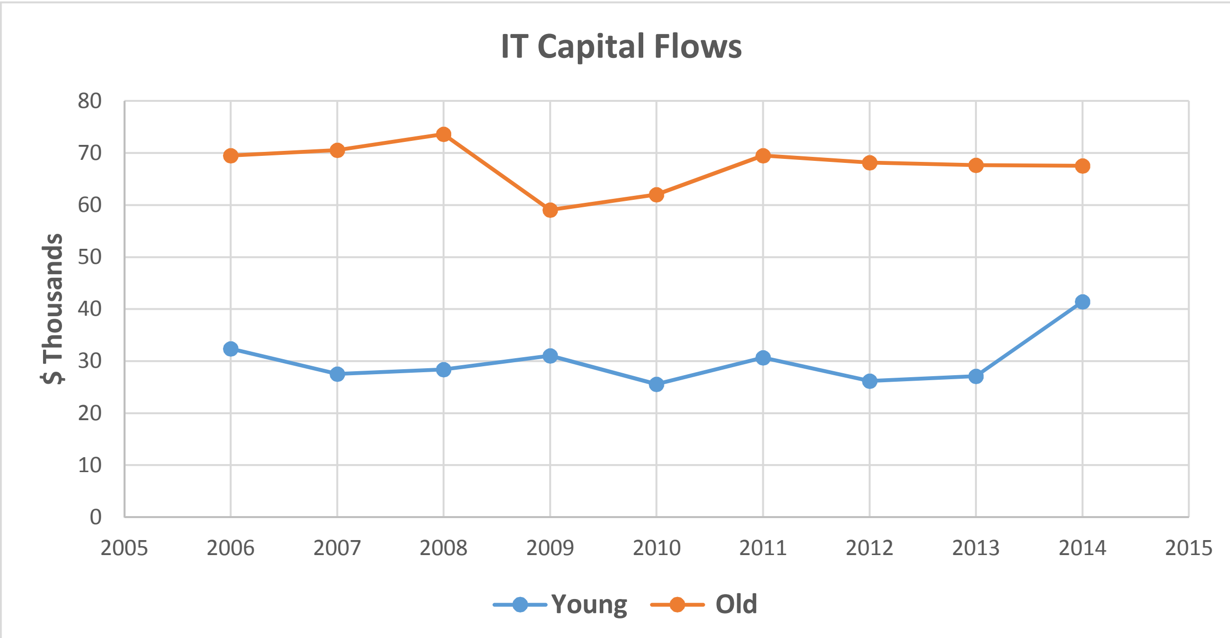


Figure 4. Accumulated and Depreciated IT Capital Stocks (Computers and Data Processing Equipment) in U.S. Manufacturing, 2006-2014 ($Thousands)
Winsorized at the 99th percentile

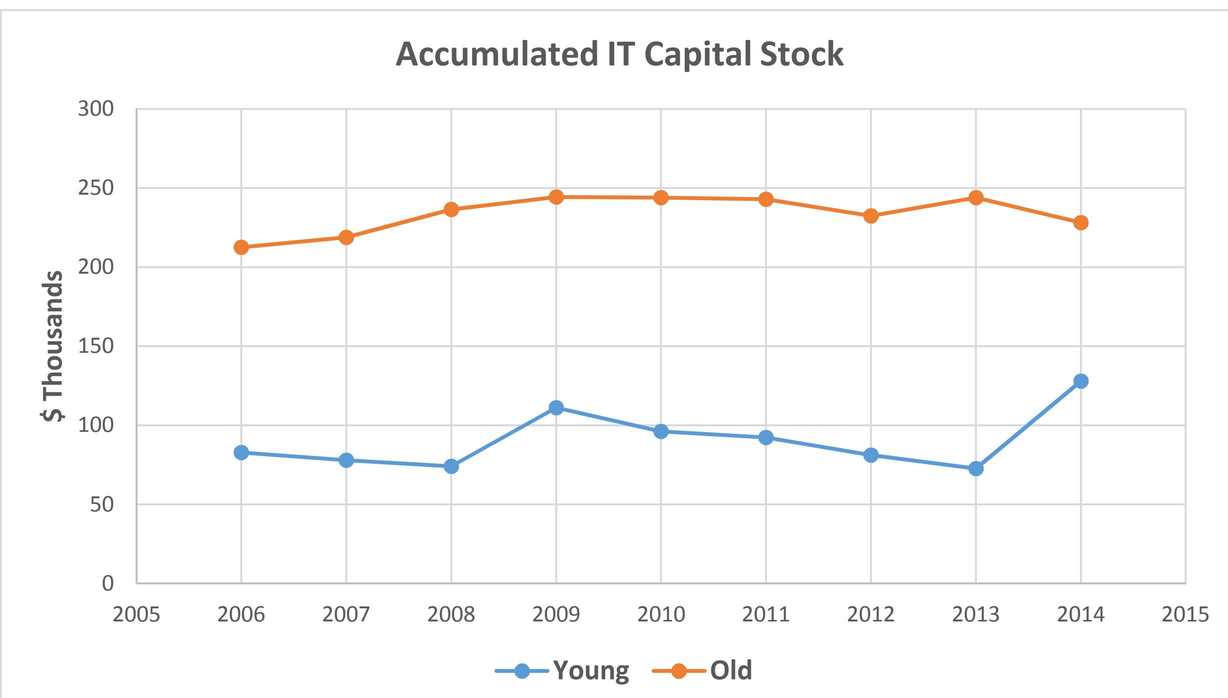


Figure 5. Annual Expenditures on IT Services in U.S. Manufacturing, 2006-2014 ($Thousands)
Winsorized at the 99th percentile

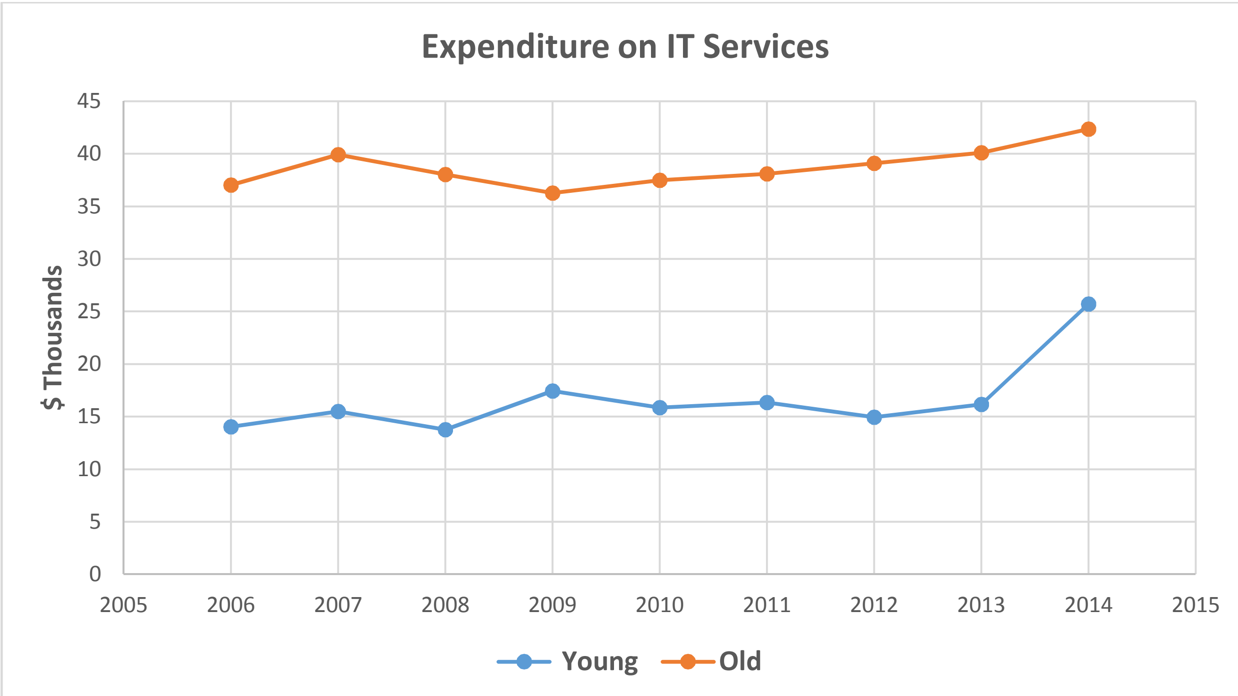


Figure 6. Per-Employee Annual Expenditures on IT Capital flows and IT Services in U.S. Manufacturing, 2006-2014 ($Thousands)
Winsorized at the 99th percentile

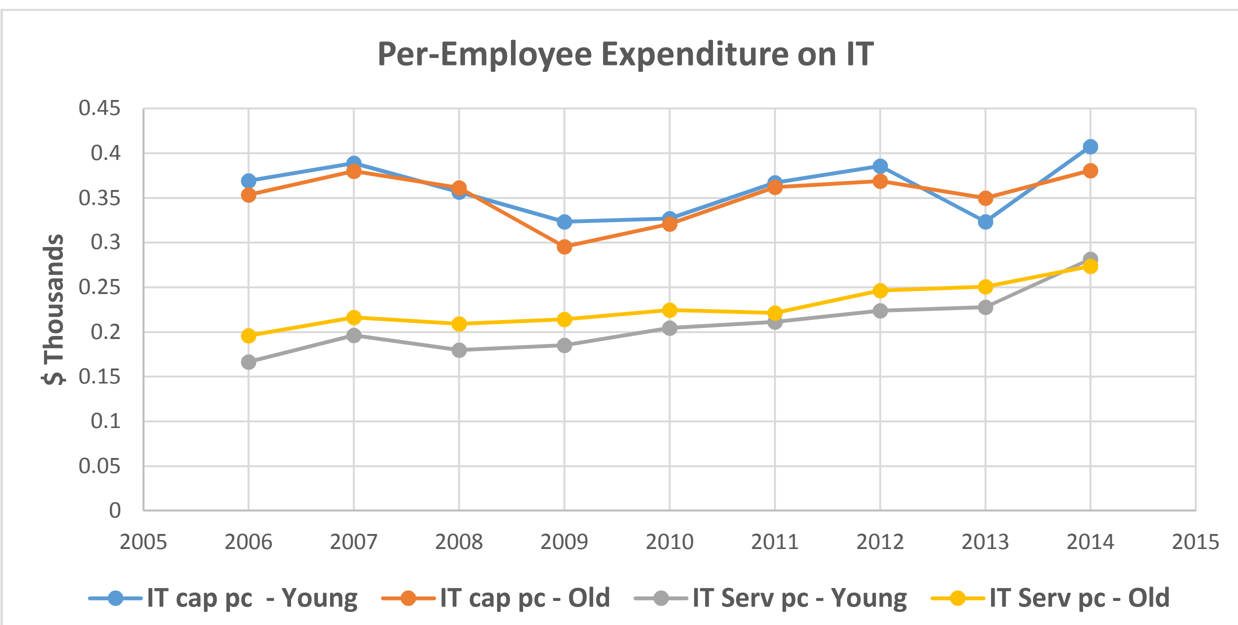


Table 1. Change of Pricing for Amazon EC2 CPU

Service Description	CU	2006 Aug	2007 Oct	2008 May	2009 Oct	2010 Feb	2010 July	2010 Sep	2010 Nov	2011 Nov	Original $ /CU/Hour	Current $ /CU/Hour	Reduction %
Small - "The Original"	1.0	$0.10			$0.09						$0.100	$0.085	15.0%
Large	4.0		$0.40		$0.34						$0.100	$0.085	15.0%
Extra Large	8.0		$0.80		$0.68						$0.100	$0.085	15.0%
High-CPU Medium	5.0			$0.20	$0.17						$0.040	$0.030	15.0%
High-CPU Extra Large	20.0			$0.80	$0.68						$0.040	$0.030	15.0%
High-Memory Double Extra large	13.0				$1.20			$1.00			$0.090	$0.077	17.0%
High-Memory Quad Extra large	26.0				$2.40			$2.00			$0.090	$0.077	17.0%
High-Memory Extra large	6.5					$0.75					$0.012	$0.077	33.0%
Cluster Compute	33.5						$1.60				$0.050	$0.040	19.0%
Micro	0.9							$0.02			$0.020	$0.020	0.0%
Cluster GPU instance	33.5								$2.10		$0.060	$0.060	0.0%
Cluster Compute Eight Extra Large	88.0									$2.40	$0.030	$0.030	0.0%

Price Reduction

Source:http://www.gregarnette.com/blog/2011/11/a-brief-history-cloud-cpu-costs-over-the-past-5-years/i

Table 2. Descriptive Statistics

Variable	Definition	**Young** Less than or equal to 5 years	**Older** Greater than 5 years old
Age	Plant age	2.50 (1.45)	25.3 (9.98)
Exit (failure)	Average annual rate (percent) of plant exit due to failure	3.33 (17.9)	1.35 (11.5)
Number of Employees	Total number of employees	68.5 (166)	189 (421)
Sales	Total value of shipments ($Millions)	30.0 (149.7)	116 (611)
Value Added	Value added ($Millions)	12.4 (66.7)	44.4 (243)
IT Capital Stock†	Stock of IT capital such as computers and data-processing equipment ($Millions)	0.09 (0.33)	0.23 (0.58)
IT Capital annual flow†	Capital expenditure for the year on computers and data processing equipment ($Millions)	0.03 (0.13)	0.07 (0.20)
IT Services†	Operating expenses on data processing and other purchased computer services ($Millions)	0.02 (0.08)	0.04 (0.14)
Software†	Operating expenses on software ($Millions)	0.01 (0.06)	0.03 (0.10)
Equipment†	Operating expenses on equipment ($Millions)	0.02 (0.075)	0.04 (0.13)
Non-IT Capital Stock†	Traditional capital stock on non-IT equipment and structure ($Millions)	10.2 (37.6)	28.0 (65.2)
Multi-Unit Status	Indicator for whether plants belong to a multi-unit firms	0.70 (0.46)	0.69 (0.46)
Log(Sales)	Plant total sales in log terms ($Thousands)	8.88 (1.58)	10.2 (1.56)
Log (IT Capital Stock†**)**	IT capital stock in log terms ($Thousands)	2.60 (1.90)	3.80 (1.98)
Log (IT Capital flows†**)**	IT capital expenditure in log terms ($Thousands)	1.29 (1.73)	1.95 (0.21)
Log (IT Services†**)**	Operating expenses on Data processing and other purchased computer services in log terms ($Thousands)	0.87 (1.49)	1.30 (1.90)
Log (Equipment Expenditure†**)**	Operating expenses on equipment in log terms ($Thousands)	1.13 (1.51)	1.79 (1.89)
Log (Software Expenditure†**)**	Operating expenses on software in log terms ($Thousands)	0.83 (1.37)	1.39 (1.82)
Number of Observations		~41,300	~198,400

Note: Std. Deviations in parentheses; † indicates the variable is winsorized at the 1% and 99% levels.

Table 3. IT Capital and Expenditures per Employee for Young vs. Older Plants in U.S. Manufacturing 2006-2014

	(1)	(2)	(3)	(4)	(4a)	(4b)	(4c)	(4d)
Dependent Variable	**IT Capital Stock per Employee**	**Annual IT Flows (all kinds) per Employee**	**IT Capital Stock per Employee**	**Annual IT Flows (all kinds) per Employee**	**IT Capital Flows per Employee**	**IT Services Spend per Employee**	**Software Spend per Employee**	**Equipment Spend per Employee**
Young	-.063*** (0.012)	-.021*** (0.010)	-0.037* (0.022)	-0.008 (0.019)	0.017* (0.009)	-0.021*** (0.009)	-0.025*** (0.004)	0.006 (0.006)
Late (2008 – 2014)			0.150*** (0.011)	-0.081*** (0.010)	-0.029*** (0.004)	0.018*** (0.003)	0.026*** (0.002)	-0.078*** (0.003)
Young x Late			-0.041 (0.026)	-0.035 (0.022)	-0.011 (0.010)	-0.002 (0.007)	-0.002 (0.005)	-0.020 (0.007)
Industry Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y
Year Fixed Effects	Y	Y	N	N	N	N	N	N
N	~239,700	~239,700	~239,700	~239,700	~239,700	~239,700	~239,700	~239,700
R-Squared	0.094	0.099	0.093	0.097	0.048	0.047	0.085	0.060

Note: Results from columns 1 and 2 are from OLS regressions controlling for year- and industry- (6-digit NAICS) fixed effects. Results from columns 3 to 4d are from OLS regressions controlling for industry-fixed effects (6-digit NAICS). All columns use the entire analysis sample containing both young and old plants from 2006 to 2014. The dependent variables are the IT capital stock and annual IT expenditure variables winsorized at the 99th percentile. In addition, the rate of outliers is not correlated with age as we observe similar percentages in both young and old. **Young** is an indicator for an plant being less than or equal to 5 years old. **Late** is the indicator for the sample years 2008 through 2014. Statistical significance is denoted as follows: * 10%, ** 5%, *** 1%.

Table 4. IT Expenditure Breakdown by Type for Young and Older Plants, 2006 – 2014

	(1)	**(2)**	**(3)**	**(4)**
Dependent Variables	**% Expenditure on IT Capital Flows**	**% Expenditure on IT Services**	**% Expenditure on Software**	**% Expenditure on Equipment**
Young	0.009** (0.004)	-0.004 (0.003)	-0.015*** (0.003)	0.010*** (0.003)
Late	-0.054*** (0.002)	0.021*** (0.002)	0.036*** (0.001)	-0.003* (0.002)
Young x Late	-0.006 (0.004)	0.011** (0.003)	0.002 (0.003)	-0.007* (0.004)
Industry Fixed Effects	Y	Y	Y	Y
N	~239,700	~239,700	~239,700	~239,700
R-Squared	0.152	0.086	0.097	0.131

Note: Results are from OLS regressions controlling for industry (6-digit NAICS) fixed effects. The dependent variables are the percentage of each type of IT spending with respect to the total expenditure reported on IT. All columns include an unreported indicator for whether the plant reported zero IT expenditure (note that these are reported, not imputed zeroes). **Late** is the indicator for the sample years 2008 through 2014. There is no statistical difference between the Great Recession years (2008-2009) and the post-Recession period (2010-2014), so they are combined for ease of exposition. Statistical significance is denoted as follows: * 10%, ** 5%, *** 1%.

Table 5. IT Expenditure, Survival, and Employment Growth in Young vs. Older Plants, 2006-2014

Model Description	**(1a) Survival Young**	**(1b) Survival Older**	**(2a) Survival (High IT dummies) Young**	**(2b) Survival (High IT dummies) Older**	**(3a) Employment Growth Young**	**(3b) Employment Growth Older**
Dependent Variable	Probability of Failure	Probability of Failure	Probability of Failure	Probability of Failure	Change in Employment	Change in Employment
IT Capital stock	1.046** (0.023)	0.968** (0.014)	1.263** (0.147)	0.849*** (0.053)	0.014*** (0.004)	0.008*** (0.002)
IT Services	0.951* (0.026)	0.900*** (0.014)	0.824** (0.073)	0.727*** (0.040)	0.014*** (0.003)	0.008*** (0.001)
Software	0.923** (0.030)	0.846*** (0.016)	1.014 (0.092)	0.777*** (0.046)	-0.141*** (0.017)	-0.092*** (0.012)
Equipment	0.875*** (0.027)	0.823*** (0.014)	0.616*** (0.075)	0.523*** (0.036)	0.026*** (0.003)	0.017*** (0.001)
Non-IT Capital Stock	0.879*** (0.014)	1.007 (0.018)	0.885*** (0.012)	0.972* (0.015)	0.019*** (0.005)	0.035*** (0.009)
Industry Controls	NAICS4	NAICS4	NAICS3	NAICS3	N	N
Industry x Year Fixed Effects	N	N	N	N	Y	Y
# of Plants per Year	~4,900	~22,400	~4,900	~22,400	~2,400	~14,500
# of Years	6	6	6	6	8	8
R-Squared					0.105	0.081

Note: Results in columns 1a, 1b, 2a, and 2b are hazard rates from a Cox proportional hazard model of the likelihood of failure. Columns 3a and 3b report OLS coefficients of a regression on year-over-year change in employment at the plant. The independent variables for the first two columns include levels of IT investment; the next two substitute indicators for being in the 75th percentile of that type of IT investment for that plant's NAICS4 industry. Additional controls include accumulated and depreciated non-IT capital stock in log terms and plant age (not reported but available upon request). For columns 3a and 3b, the explanatory variables are year-over-year changes in IT (calculated using log differences) and changes in non-IT capital stock in log terms. The sample for the survival models from columns 1a to 2b included data from 2006 to 2012 due to limitations in the 2013 LBD preventing identification of plant exit. The sample for the growth models requires that plants persist at least two consecutive years in the sample. Columns 1a and 1b control for 4-digit NAICS code; columns 2a and 2b control for 3-digit NAICS code because IT indicators are constructed based on 4-digit NAICS. Standard errors from specifications in columns 3a and 3b are clustered at the plant level. Statistical significance is denoted as follows: *10%, **5%, ***1%.

Table 6a. Estimates of IT Productivity for Young vs. Older Plants, 2006-2014

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Model Description	**All**	**Young Only**	**Older Only**	**Sales x Young**	**Sales Young Only + F.E.**	**Sales Older Only + F.E.**	**Young Only (High IT indicators)**
IT Capital Stock	0.015*** (0.001)	0.011** (0.003)	0.014*** (0.001)	0.014*** (0.001)	0.007** (0.004)	0.003** (0.001)	0.068*** (0.013)
IT Services	0.001 (0.001)	0.010*** (0.003)	0.001 (0.001)	0.001 (0.001)	0.009*** (0.003)	0.0014** (0.0007)	0.026** (0.010)
Software	0.002* (0.001)	0.005* (0.003)	0.001 (0.001)	0.001 (0.001)	0.009*** (0.003)	0.003** (0.001)	0.018 (0.011)
Equipment	0.014*** (0.001)	0.018*** (0.003)	0.013*** (0.001)	0.013*** (0.001)	0.012*** (0.003)	0.006*** (0.001)	0.053*** (0.010)
Young	0.013*** (0.005)			0.212*** (0.039)			
IT Capital Stock x Young				-0.003 (0.003)			
IT Services x Young				0.009*** (0.003)			
Software x Young				0.005 (0.003)			
Equipment x Young				0.006** (0.003)			
Non-IT Capital Stock	0.055*** (0.002)	0.027*** (0.003)	0.073*** (0.003)	0.073*** (0.003)	0.011*** (0.004)	0.019*** (0.004)	0.029*** (0.003)
Non-IT Capital Stock x Young				-0.047*** (0.004)			
Inputs: Labor & Materials	Y	Y	Y	Y	Y	Y	Y
Inputs x Young	N	N	N	Y	N	N	N
Industry x Year Fixed Effects	Y	Y	Y	Y	N	N	Y
Plant & Year Fixed Effects	N	N	N	N	Y	Y	N
# of Plants per Year	~26,600	~4,600	~22,000	~26,600	~4,600	~22,000	~4,600
# of Years	9	9	9	9	9	9	9
R-Squared	0.943	0.911	0.947	0.943	0.585	0.600	0.911

Note: Results in columns 1, 2, 3, 4, and 7 are based on weighted OLS regression using ASM sampling weights controlling for year-industry (6-digit NAICS) fixed effects. Columns 5 and 6 are based on plant fixed effect models controlling for year trends. The dependent variable for all columns is total sales in log terms. Production inputs are also controlled for (but not reported) in all models in log terms, including: cost of material, cost of energy, and labor (both expenditure on temporary employees and the count of regular employees). In addition, the coefficients for interaction terms between Young and the production inputs in column 4 are not reported to save space (available upon request). Standard errors for all columns are clustered at the plant level. Results are robust to two-way clustering at county & plant and firm & plant levels, as well (*pending disclosure review*). Statistical significance is denoted as follows: * 10%, ** 5%, *** 1%.

Table 6b. Linear Combinations of IT Productivity Coefficients from Table 6a:

	(1)	(2)	(3)	(4)
Model Description	**Sales Young Only**	**Sales Older Only**	**Sales Young Only + F.E.**	**Sales Older Only + F.E.**
IT Capital Stock	0.011** (0.003)	0.014*** (0.001)	0.007** (0.004)	0.003** (0.001)
IT Capital Stock + Equipment	0.029*** (0.004)	0.027*** (0.002)	0.019*** (0.005)	0.009*** (0.001)
IT Capital Stock + IT Services	0.021*** (0.004)	0.014*** (0.002)	0.016*** (0.005)	0.004*** (0.001)
IT Capital Stock + IT Services + Equipment	0.039*** (0.005)	0.028*** (0.002)	0.027*** (0.006)	0.010*** (0.001)
IT Capital Stock + IT Services + Equipment + Software	0.044*** (0.005)	0.029*** (0.002)	0.036*** (0.007)	0.013*** (0.002)
Industry x Year Fixed Effects	Y	Y	N	N
Plant & Year Fixed Effects	N	N	Y	Y

Note: The coefficients are calculated using the **lincom** command in Stata 13, based on coefficients from Table 6a, columns 4-7 respectively for young and old samples. Statistical significance is denoted as follows: * 10%, ** 5%, *** 1%.

Table 7. Alternative Estimates of IT Performance Benefits for Young Plants, 2006-2014

	(1)	(2)	(3)	(4)
Model Description	**OLS**	**Blundell-Bond**	**Levinsohn-Petrin**	**Ackerberg-Caves-Frazer**
Dependent Variables	**Sales**	**Sales**	**Sales**	**Sales**
IT Capital Stock	0.011** (0.003)	0.005 (0.010)	0.013*** (0.002)	0.029 (0.035)
IT Services	0.010*** (0.003)	0.031*** (0.012)	0.007*** (0.002)	0.022*** (0.007)
Software	0.005* (0.003)	-0.002 (0.013)	0.010*** (0.003)	0.026 (0.022)
Equipment	0.018*** (0.003)	0.032*** (0.012)	0.023*** (0.002)	0.040*** (0.013)
Non-IT Capital Stock	0.073*** (0.003)	0.008 (0.015)	0.031*** (0.011)	0.041*** (0.015)
Inputs: Labor & Materials	Y	Y	Y	Y
Plant & Year Fixed Effects	N	Y	Y	Y
Industry x Year Fixed Effects	Y	N	N	N
# of Plants per Year	~4,600	~4,600	~4,600	~4,600
# of Years	9	7	9	9

Note: Column 1 is identical to Table 6a, column 2. Column 2 employs the system GMM estimator following Blundell and Bond (2000) to address potential endogeneity of IT adoption in the productivity estimation. It uses two-period lagged differences and levels as GMM instruments for IT services expenditure. This specification passes both over-identification and autocorrelation tests. Column 3 follows the approach in Levinsohn and Petrin (2003), using expenditure on intermediate inputs (cost of temporary employees) as proxy for unobservable productivity shocks. Column 4 employs the method developed by Ackerberg, Caves, and Frazer (2006) to further account for collinearity problems when estimating productivity using the Levinsohn-Petrin techniques. Production inputs are also controlled for (but not reported) in all models in log terms, including: cost of materials, cost of energy (both electricity and fuel), and labor (both expenditure on temporary employees and the count of regular employees). Statistical significance is denoted as follows: * 10%, ** 5%, *** 1%. We also tested the semiparametric method developed by Olley and Pakes (1996), which uses capital investment (both structure and equipment) as a proxy for unobservable shocks that could lead to spurious correlation between IT services expenditure and productivity. The results from this method are consistent with the results presented in this table.

Table 8. Instrumental Variable Estimates for Young Plants, 2006-2014

	(1)	(2)	(3)
Model Description	**Data Center Intensity (lagged 2 years)**	**Average Industry IT Services Adoption**	**Both**
Dependent Variables	**Sales**	**Sales**	**Sales**
IT Capital Stock	-0.004 (0.007)	0.010 (0.006)	0.002 (0.005)
IT Services	0.402** (0.146)	0.111 (0.110)	0.281** (0.090)
Software	-0.079* (0.032)	-0.016 (0.024)	-0.053** (0.020)
Equipment	-0.015 (0.014)	0.011 (0.010)	-0.004 (0.009)
Non-IT Capital Stock	0.023*** (0.003)	0.023*** (0.002)	0.023*** (0.002)
Inputs: Labor & Materials	Y	Y	Y
Industry and Year Fixed Effects	Y	Y	Y
Endogeneity Test (Durbin-Wu-Hausman)	16	1	13

First Stage			
Data Center Intensity (lagged 2 period)	29.17*** (8.02)		29.09*** (8.02)
Average IT Services Adoption		0.312*** (0.097)	0.311*** (0.097)
F-test	13.22	10.27	11.58
# of Plants per Year	~4,600	~4,600	~4,600
# of Years	9	9	9

Note: Column 1 used the lagged data center intensity (percentage of number of data centers to total establishments) in the local county as the instrument. Column 2 instrumented IT Services using the average industry (6-digit NAICS) IT Services adoption outside the firm and county of the focal plant. Column 3 reports the results using both instruments. The results in columns 1 and 3 pass all tests including weak identification, under identification, and endogeneity tests. The results in column 2 passed both weak and under identification tests but failed to reject the null hypothesis that IT services is exogenous (the estimated coefficients are similar between the IV estimation and the OLS). Production inputs are also controlled for (but not reported) in all models in log terms, including: cost of materials, cost of energy (both electricity and fuel), and labor (both expenditure on temporary employees and the count of regular employees). Statistical significance is denoted as follows: * 10%, ** 5%, *** 1%.

Table 9. Reverse Causality Tests in Young Plants

Model Description	(1) No lagged TVS	(2) Lagged TVS	(3) Forward IT	(4) Lagged TVS	(5) Forward TVS
Dependent Variables	**Sales**	**Sales**	**Sales**	**IT Services**	**IT Services**
Lagged IT Capital Stock	0.006 (0.004)	-0.006* (0.003)	-0.019*** (0.005)	0.016** (0.007)	0.021** (0.010)
Lagged IT Services	0.013*** (0.004)	0.0004 (0.003)	0.004 (0.005)	0.698*** (0.011)	0.708*** (0.016)
Lagged Software	0.001 (0.005)	-0.005 (0.003)	-0.003 (0.005)	0.043*** (0.011)	0.062*** (0.016)
Lagged Equipment	0.018*** (0.005)	0.007 (0.004)	-0.012** (0.006)	0.008 (0.009)	0.022* (0.013)
Lagged Sales		0.510*** (0.017)	0.527*** (0.019)	-0.033** (0.017)	-0.047 (0.031)
Forward IT Capital Stock			0.017** (0.006)		
Forward IT Services			-0.004 (0.004)		
Forward Software			-0.004 (0.005)		
Forward Equipment			0.019*** (0.005)		
Forward Sales					0.056** (0.027)
Industry x Year Fixed Effects	Y	Y	Y	Y	Y
# of Plants per Year	~2,100	~2,100	~1,200	~2,100	~1,200
# of Years	8	8	7	8	7
R-Squared	0.923	0.954	0.962	0.625	0.641

Note: Results are from the weighted OLS regression using ASM sampling weights. All columns use young sample only and control for industry-year fixed-effects. The dependent variables for columns 1 to 3 are total value of shipment in log terms while the dependent variables for columns 4 and 5 are the IT services in log terms. The sample size is similar to those in the growth models since lagged and forward variables are needed to test for reverse causality and hence the plants in the analysis sample are required to show up two or three consecutive years. Additional controls include cost of material, cost of energy (both electricity and fuel), imputed non-IT capital stock, and labor (both expenditure on temporary employees and the count of regular employees) in log terms. Standard errors are clustered at the plant level. Statistical significance is denoted as follows: * 10%, ** 5%, *** 1%.

Table 10. Timing of IT Productivity Effects, 2006-2014

Model Dependent Variable = Total Revenues	(1) Continuous IT	(2) High IT Indicators (Top Quartile)
IT Capital Stock	0.008* (0.004)	0.077*** (0.018)
IT Services	0.002 (0.004)	0.006 (0.016)
Equipment	0.006 (0.004)	0.029* (0.016)
Software	0.001 (0.005)	-0.011 (0.022)
IT Capital Stock x Great Recession (2008-2009)	0.005 (0.006)	-0.026 (0.025)
IT Services x GR	0.006 (0.006)	0.050** (0.023)
Software x GR	0.010 (0.007)	0.060** (0.029)
Equipment x GR	0.020*** (0.007)	0.018 (0.023)
IT Capital Stock x Post Recession (2010-2014)	0.002 (0.005)	-0.007 (0.025)
IT Services x Post	0.012** (0.005)	0.040* (0.022)
Software x Post	0.004 (0.007)	0.050* (0.027)
Equipment x Post	0.017*** (0.006)	0.041* (0.021)
# of Plants per Year	~4,600	~4,600
# of Years	9	9
R-Squared	0.911	0.911

Note: Results in all columns are based on the weighted OLS regression using ASM sampling. All columns use young sample only and control for industry-year fixed-effects. The dependent variable for all columns is total value of shipment in log terms. GR is defined as the Great Recession year from 2008 to 2009. Post-Recession indicates the years from 2010 to 2014. The coefficients for both period dummies are omitted since we controlled for year-industry fixed-effects. Additional variables including cost of material, cost of energy (both electricity and fuel), imputed non-IT capital stock, and labor (both expenditure on temporary employees and the count of regular employees) in log terms are controlled but not reported to save space (available up on request). In addition, the coefficients for interaction terms between young and other inputs, the interactions between local GR shock and other IT variables, and between local GR shock and contraction periods are also not reported in the table to save space. Standard errors for all columns are clustered at the plant level. Statistical significance is denoted as follows: * 10%, ** 5%, *** 1%.

Table 11. Industry Variation in IT Productivity, 2006-2014

Model Description	(1) IT-Intensive in 2005	(2) IT-Intensive in 2005 (High IT Dummies)	(3) High Competition (Lerner Index)	(4) High-Uncertainty (Young)	(5) High-Uncertainty (Older)
Dependent Variables	**Sales**	**Sales**	**Sales**	**Sales**	**Sales**
IT Capital Stock	0.012*** (0.004)	0.039* (0.016)	0.016*** (0.004)	0.020*** (0.004)	0.020*** (0.002)
IT Services	0.002 (0.004)	-0.014 (0.013)	0.007* (0.003)	0.003 (0.004)	-0.001 (0.001)
Software	0.005 (0.005)	0.016 (0.013)	0.006 (0.004)	0.015*** (0.004)	0.007*** (0.001)
Equipment	0.019*** (0.004)	0.042*** (0.014)	0.027*** (0.003)	0.024*** (0.004)	0.017*** (0.001)
IT Capital Stock x High IT Intensity in 2005	0.001 (0.005)	0.063* (0.025)			
IT Services x High IT Intensity in 2005	0.010 (0.005)	0.062** (0.020)			
Software x High IT Intensity in 2005	0.005 (0.006)	-0.001 (0.022)			
Equipment x High IT Intensity in 2005	0.012 (0.006)	0.035 (0.021)			
IT Capital Stock x High Competition			-0.006 (0.006)		
IT Services x High Competition			0.011* (0.006)		
Software x High Competition			0.012 (0.008)		
Equipment x High Competition			-0.001 (0.007)		
IT Capital Stock x High Uncertainty				-0.012* (0.005)	-0.004 (0.002)
IT Services x High Uncertainty				0.015** (0.005)	0.005** (0.002)
Software x High Uncertainty				-0.014* (0.006)	-0.005* (0.002)
Equipment x High Uncertainty				0.007 (0.006)	0.005* (0.002)
Industry and Year Fixed Effects	Y	Y	Y	Y	Y
# of Plants per Year	~4,600	~4,600	~4,600	~4,600	~22,000
# of Years	9	9	9	9	9

Note: Results in all columns are based on weighted OLS regressions using ASM sampling weights. All columns control for sector (2-digit NAICS) and year fixed effects using the young sample only (except column 5). The dependent variable for all columns is logged total value of shipments. **High IT Intensity in 2005** is an indicator for plants in industries (3-digit NAICS) with above-mean IT capital stock in 2005. **High Competition** is an indicator equal to 1 if the one-period lagged industry (6-digit NAICS) Lerner index is in the bottom 25th percentile for the entire ASM. **High Uncertainty** is an indicator for being in 6-digit NAICS industries with above-mean industry variance in the quarterly plant capacity utilization rate (based on data from the US Census Bureau's Plant Capacity Utilization Survey matched to our sample). Inputs including costs of material, cost of energy (both electricity and fuel), imputed non-IT capital stock, and labor (both expenditure on temporary employees and the count of regular employees) in log terms are controlled for but not reported (available up on request). Standard errors for all columns are clustered at the plant level. Statistical significance is denoted as follows: * 10%, ** 5%, *** 1%.

Table 12. Size vs. Age in IT Productivity, 2006-2014

Dependent Variable	Sales
IT Capital Stock	0.039*** (0.002)
IT Services	0.012*** (0.002)
Software	0.012*** (0.002)
Equipment	0.026*** (0.002)
Young	0.027* (0.014)
Large	0.305*** (0.011)
Large x Young	0.089*** (0.023)
IT Capital Stock x Large	0.002 (0.003)
IT Services x Large	-0.011*** (0.002)
Software x Large	-0.003 (0.003)
Equipment x Large	-0.002 (0.003)
IT Capital Stock x Young	-0.032*** (0.004)
IT Services x Young	0.010* (0.004)
Software x Young	0.014** (0.005)
Equipment x Young	0.018*** (0.005)
IT Capital Stock x Young x Large	-0.002 (0.006)
IT Services x Young x Large	-0.002 (0.006)
Software x Young x Large	-0.010 (0.007)
Equipment x Young x Large	-0.020*** (0.007)
Industry x Year Fixed Effects	Y
# of Plants per Year	~26,000
# of Years	9
R-Squared	0.932

Comparison Group: IT Capital Stock + older + small	0.039*** (0.002)
Linear Combination: IT Capital Stock + young + small	0.006* (0.004)
Linear Combination: IT Capital Stock + young + large	0.006 (0.005)
Linear Combination: IT Capital Stock + older + large	0.040*** (0.002)
Comparison Group: IT services + older + small	0.012*** (0.002)
Linear Combination: IT services + young + small	0.022*** (0.004)
Linear Combination: IT services + young + large	0.009*** (0.004)
Linear Combination: IT services + older + large	0.002 (0.001)

Note: Results of weighted OLS regression using ASM sampling weights. **Young** indicates 5 or fewer years old. **Large** indicates above-median employment compared to the plant's 4-digit NAICS industry in a given year. Additional controls include logged cost of materials, cost of energy (both electricity and fuel), imputed non-IT capital stock, and labor (both expenditure on temporary employees and the count of regular employees); these are not interacted with **Young** or **Large** in this specification. Joint tests of significant for the linear combinations are conducted using the STATA 13 *lincom* command. Standard errors are clustered at the plant level. Statistical significance is denoted as follows: * 10%, ** 5%, *** 1%.

Table 13. Age and Multi-Unit Status in Multi-Factor IT Productivity 2006-2014

Dependent Variable	Sales
IT Capital Stock	0.013*** (0.002)
IT Services	0.007*** (0.002)
Software	0.002 (0.002)
Equipment	0.011*** (0.002)
Young	0.021 (0.017)
MU	0.042*** (0.011)
MU x Young	0.024 (0.021)
IT Capital Stock x MU	0.007*** (0.003)
IT Services x MU	-0.008*** (0.002)
Software x MU	-0.001 (0.002)
Equipment x MU	0.004* (0.002)
IT Capital Stock x Young	-0.019*** (0.005)
IT Services x Young	0.004 (0.005)
Software x Young	0.010* (0.005)
Equipment x Young	0.011*** (0.005)
IT Capital Stock x Young x MU	-0.001 (0.006)
IT Services x Young x MU	0.009 (0.006)
Software x Young x MU	-0.005 (0.007)
Equipment x Young x MU	-0.007 (0.006)
Industry x Year Fixed Effects	Y
# of Plants per Year	~26,000
# of Years	9
R-Squared	0.943

Comparison Group: IT Capital Stock + older + single-unit	0.013*** (0.002)
Linear Combination: IT Capital Stock + young + single-unit	-0.005 (0.005)
Linear Combination: IT Capital Stock + young + MU	0.0005 (0.003)
Linear Combination: IT Capital Stock + older + MU	0.021*** (0.002)
Comparison Group: IT services + older + single-unit	0.007*** (0.002)
Linear Combination: IT services + young + single-unit	0.010*** (0.004)
Linear Combination: IT services + young + MU	0.011*** (0.003)
Linear Combination: IT services + old + MU	-0.001 (0.001)

Note: Results of weighted OLS regression using ASM sampling weights. **Young** indicates 5 or fewer years old. **MU** indicates whether the plants belong to multi-unit firms. Additional controls include logged cost of materials, cost of energy (both electricity and fuel), imputed non-IT capital stock, and labor (both expenditure on temporary employees and the count of regular employees); these are not interacted with **Young** or **MU** in this specification. Joint tests of significant for the linear combinations are conducted using the STATA 13 *lincom* command. Standard errors are clustered at the plant level. Statistical significance is denoted as follows: * 10%, ** 5%, *** 1%.

Table 14. IT Complementarities for both Young and Older Plants

Model Description	(1) Same-Year Interaction	(2) Same-Year Interaction	(3) Lagged IT Capital	(4) Lagged IT Services	(5) Lagged IT Services
Sample	Young	Older	Young	All	All
Dependent Variables	Sales	Sales	Sales	Sales	Sales
IT Capital Stock	0.013*** (0.004)	0.014*** (0.002)	0.006 (0.006)		
IT Services	0.015*** (0.005)	-0.0001 (0.003)	0.011 (0.008)		
Software	-0.016** (0.006)	-0.006* (0.003)	-0.031** (0.010)		
Equipment	0.027*** (0.005)	0.012*** (0.003)	0.030*** (0.009)		
IT Capital Stock x IT Services	-0.003** (0.001)	-0.001** (0.0006)	-0.003 (0.002)		
IT Capital Stock x Lagged† IT Services				+**	
Software x IT Services	0.004*** (0.002)	0.002*** (0.001)	0.005* (0.002)		
Software x Lagged†† IT Services					+**
Equipment x IT Services	-0.0001 (0.001)	0.001 (0.001)	0.001 (0.002)		
IT Capital Stock x Software	0.004*** (0.002)	0.001 (0.001)	0.007** (0.002)		
IT Capital Stock x Equipment	-0.003** (0.001)	0.0002 (0.001)	-0.003 (0.002)		
Non-IT Capital Stock	0.027*** (0.003)	0.073*** (0.003)	0.043*** (0.008)		
Other inputs	Y	Y	Y		
Industry x Year Fixed Effects	Y	Y	Y		
# of Plants per Year	~4,600	~22,000	~2,000		
# of Years	9	9	8		
R-Squared	0.911	0.947	0.923		

†Lagged 5-6 years; other lags are noisy.
††Lagged 3-4 years; other lags are noisy.

Note: Results in all columns are based on the weighted OLS regression using ASM sampling weights. Column 1 uses young sample only and column 2 uses the older sample. Column 3 uses young sample and interacts other variables with one period lagged IT capital stock. All specifications control for industry-year fixed-effects. Additional controls include cost of material, cost of energy (both electricity and fuel), and labor (both expenditure on temporary employees and the count of regular employees) – all in log terms. The coefficients for these variables are not reported to save space (available up on request). Standard errors for all columns are clustered at the plant level. Statistical significance is denoted as follows: * 10%, ** 5%, *** 1%.

Table 15. Robustness Tests

Models	(1) Translog Production Function	(2) Regional Controls	(3) Different Labor calculation	(4) Including Imputed Data	(5) Non-Winsorized
Dependent Variables	**Sales**	**Sales**	**Sales**	**Sales**	**Sales**
IT Capital Stock	0.017 (0.015)	0.013*** (0.003)	0.025*** (0.003)	0.013*** (0.002)	0.011*** (0.003)
IT Services	0.026 (0.025)	0.010*** (0.003)	0.018*** (0.003)	0.019*** (0.002)	0.010*** (0.003)
Software	0.017 (0.018)	0.005 (0.003)	0.021*** (0.003)	0.011*** (0.003)	0.005* (0.003)
Equipment	0.076*** (0.018)	0.016*** (0.003)	0.029*** (0.003)	0.024*** (0.002)	0.018*** (0.003)
Industry x Year Fixed Effects	Y	Y	Y	Y	Y
# of Plants per Year	~4,600	~4,600	~4,600	~11,300	~4,600
# of Years	9	9	9	9	9
R-Squared	0.933	0.914	0.904	0.924	0.911

Note: Column 1 estimates a translog production function including interactions and the squares for all inputs. To account for geographical differences, column 2 further controls for reginal, industry and year fixed effects. Column 3 addresses the concern of omitted variable bias on the quality of the labor input by using the quality adjusted labor measure (calculated by multiplying the total production hours with the ratio of the production worker wage to the total salary in log term following Foster, Grim, and Haltiwanger 2016). In column 4, the specification utilizes the observations with the Census imputed values for IT variables to test the sensitivity of our results to the Census imputation. Finally, column 5 re-estimates the production function using the non-winsorized variables from the ASM and CMF. Results for columns 1 to 5 are based on the weighted OLS regression using ASM sampling weights controlling for industry-year fixed-effects. Additional variables including cost of material, cost of energy, imputed non-IT capital stock, and labor (both expenditure on temporary employees and the count of regular employees) – all in log terms are controlled for all columns. Young sample is used for all columns in the table. Standard errors for all columns are clustered at the plant level. Statistical significance is denoted as follows: * 10%, ** 5%, *** 1%.

Appendix Table A.1 Descriptive Statistics for ASM-based Sample (Young vs. Older)

Variable	Definition	ASM-based Sample Young	ASM-based Sample Older
Missing no IT data	All four IT variables are reported	0.41 (0.49)	0.56 (0.50)
Missing only IT Services data	Only the IT services expenditure is missing	0.01 (0.09)	0.02 (0.13)
Missing all Expensed IT	Missing all IT operating expenses	0.50 (0.50)	0.31 (0.46)
Missing all IT data	Missing all IT variables	0.47 (0.50)	0.28 (0.45)
Age	Plant age (truncated due to the start of the LBD in 1967).	2.13 (1.58)	24.9 (10.1)
Number of Employees	Total number of employees	56.1 (184)	178 (419)
Sales	Total value of shipment (in $millions)	23.9 (164)	102 (532)
Sales per employee	Total value of shipment per employee	0.41 (1.35)	0.50 (1.52)
Value Added	Value added (in $millions)	10.4 (89.8)	41.1 (234)
Value-Added per employee	Value added per employee (in $thousands)	176 (518)	201 (624)
IT Capital Stock†	Traditional IT capital stock (accumulated and depreciated using a perpetual inventory method) (in $thousands)	70.9 (285)	220 (543)
IT Capital Stock per employee†	Accumulated and depreciated IT capital stock per employee (in $thousands)	1.25 (2.25)	1.33 (2.22)
IT Capital flows†	Capital expenditure on computers and peripheral data processing equipment (in $thousands)	24.0 (103)	62.0 (181)
IT Capital flows per year per employee†	Capital expenditure on computers and peripheral data processing equipment per employee (in $thousands)	0.39 (0.85)	0.36 (0.79)
IT Services†	Operating expenditure on data processing and other purchased computer services (in $thousands)	11.5 (57.7)	32.9 (111)
IT Services per employee†	Operating expenditure on data processing and other purchased computer services per employee (in $thousands)	0.19 (0.48)	0.21 (0.53)
Software†	Operating expenditure on purchased software, including prepacked, custom coded or vendor customized software (in $thousands)	8.63 (42.20)	26.8 (80.64)

Software per employee†	Operating expenditure on software per employee (in $thousands)	0.13 (0.31)	0.16 (0.35)
Equipment†	Operating expenditure on equipment (expensed computer hardware and other equipment such as copiers, fax machines, telephones, shop and lab equipment, CPUs, monitors) (in $thousands)	12.4 (56.4)	36.2 (107)
Equipment per employee†	Operating expenditure on equipment per employee (in $thousands)	0.20 (0.47)	0.22 (0.48)
Non-IT Capital Stock†	Traditional (non-IT equipment and structure) capital stock. Accumulated and depreciated using a perpetual inventory method) (in $millions)	9.60 (38.4)	24.6 (55.7)
Non-IT Capital Stock per employee†	Traditional (non-IT equipment and structure) capital stock per employee (in $thousands)	151 (275)	160 (249)
Multi-Unit Status	Indicator for whether plants belong to a multi-unit firms	0.62 (0.49)	0.67 (0.47)
Log (IT Capital Stock†)	Traditional IT capital stock in log terms	2.41 (1.85)	3.76 (1.95)
Log (IT Capital flows†)	IT capital expenditure in log terms	1.32 (1.63)	1.93 (2.08)
Log (IT Services†)	Operating expenditure on data processing and other purchased computer services in log terms	0.92 (1.34)	1.36 (1.83)
Log (Equipment Expenditure†)	Operating expenditure on equipment in log terms	1.08 (1.36)	1.76 (1.83)
Log (Software Expenditure†)	Operating expenditure on purchased software in log terms	0.84 (1.24)	1.39 (1.75)

Note: Std. Deviations in parentheses; † indicates the variable is winsorized at the 1% and 99% levels.

Appendix Table A.2. Testing Age Cutoffs

Model	(1) Age Cutoff
Sample	All
Dependent Variables	Sales
IT Capital Stock	0.011*** (0.003)
IT Services	0.010*** (0.002)
Software	0.006* (0.003)
Equipment	0.018*** (0.003)
Age quintiles (2, 3, 4, and 5)	Y
IT Services x Age 6-15	-0.006* (0.003)
IT Services x Age 16-25	-0.008** (0.003)
IT Services x Age 26-35	-0.010*** (0.003)
IT Services x Age 36+	-0.011*** (0.003)
Industry x Year Fixed Effects	Y
Other IT x Age Indicators	Y
# of Plants per Year	~26,000
# of Years	9
R-Squared	0.943

Note: Results are based on the weighted OLS regression using ASM sampling weights. The dependent variable is total sales log terms. Additional controls include cost of material, cost of energy (both electricity and fuel), imputed non-IT capital stock, and labor (both expenditure on temporary employees and the count of regular employees) in log terms. The coefficients for these variables are not reported to save space (available up on request). Standard errors for all columns are clustered at the plant level. Statistical significance is denoted as follows: * 10%, ** 5%, *** 1%. Age groups one is defined for the plants from new entries to less than 6 years old; age group two contains the plants from 7 to 15 years old; age group three contains the plants from 16 to 25 years old; age group four contains the plants from 25 to 34 years old; the rest of older plants are considered age group five. These cutoffs are selected to make sure there are roughly equal number of plants in each group. Results are generally consistent and robust across various age cutoffs and grouping methods.

Appendix Table A3. Employment and Output Growth (Young vs. Older Plants)

Models	(1) Employment Growth	(2) Output Growth
Dependent Variables	Change in Employment	Change in Sales
IT Capital Stock	0.008*** (0.002)	0.001 (0.002)
IT Services	0.008*** (0.001)	0.003*** (0.001)
Software	-0.092*** (0.012)	0.039*** (0.008)
Equipment	0.017*** (0.001)	0.008*** (0.001)
Young	0.013*** (0.004)	0.008* (0.004)
IT Capital Stock x Young	0.004 (0.004)	0.007* (0.004)
IT Services x Young	0.008** (0.004)	0.001 (0.003)
Software x Young	-0.034 (0.021)	0.062*** (0.022)
Equipment x Young	0.007* (0.004)	0.001 (0.003)
Non-IT Capital Stock	0.034*** (0.008)	-0.015 (0.013)
Industry x Year Fixed Effects	Y	Y
# of Plants per Year	~16,900	~16,900
# of Years	8	8
R-Squared	0.076	0.525

Note: Results in both columns reported the coefficients from the employment and output growth models. The common independent variables for columns 1 and 2 are the changes in IT variables (calculated using log differences) and changes in non-IT capital stock in log term. Note that all the changes in IT except those for IT services are likely to be a bit lumpy. Column 2 contains additional controls for input changes including changes in cost of material, cost of energy, and changes in the cost of temporary employees and total number of employment. The coefficients for these controls are omitted to save space but available upon request. The sample for the growth models requires the plants to show up at least two consecutive years in the analysis sample. Standard errors from specifications in both columns are clustered at the plant level. Statistical significance is denoted as follows: * 10%, ** 5%, *** 1%.

Appendix Table A.4. Pairwise Correlation

	E-AGE	Log EMP	Log TVS	Log VA	Log nonIT-KST	Log ITK	Log IT Serv	Log ITK Flow	Log Exp SW	Log Equip	Log Ctemp	MU	IT-Using	High Comp
Plant Age	1													
Log Total Employment	0.350	1												
Log Total Value of Shipment	0.324	0.826	1											
Log Value Added	0.321	0.822	0.916	1										
Log non-IT K stock	0.364	0.689	0.760	0.706	1									
Log IT K stock	0.239	0.598	0.551	0.558	0.555	1								
Log IT Services	0.109	0.279	0.247	0.255	0.201	0.280	1							
Log IT Capital Flow	0.137	0.436	0.388	0.408	0.318	0.557	0.259	1						
Log Expenditure on Software	0.150	0.392	0.349	0.370	0.292	0.404	0.319	0.375	1					
Log Expenditure Equipment	0.152	0.452	0.420	0.435	0.346	0.398	0.280	0.352	0.505	1				
Log Cost of Temporary Emp	0.109	0.442	0.467	0.460	0.380	0.349	0.218	0.273	0.293	0.331	1			
Multi-Unit Status	0.005	0.168	0.316	0.258	0.278	0.065	-0.024	-0.007	-0.001	0.082	0.182	1		
IT-Using Industries	-0.009	0.057	-0.034	0.016	-0.062	0.119	0.0799	0.111	0.125	0.108	0.055	-0.080	1	
High-competitive Industries	0.006	0.063	0.167	0.043	0.085	-0.036	-0.030	-0.046	-0.052	-0.036	-0.013	0.100	-0.138	1

Note: All correlations are significant at least the 5% level except the correlations between MU status and plant age, and between MU status and log expenditure on software. The correlations are based on non-imputed sample.

Appendix Table A.5. Correlation between IT Services and Other Inputs

Models	(1) Cost of Temp Employee	(2) Cost of Material	(3) Cost of Energy	(4) Total number of Employee	(5) IT software	(6) Equipment	(7) All
Dependent Variables	**IT Services**	**IT Services**	**IT Services**	**IT Services**	**IT Services**	**IT Services**	**IT Services**
Cost of Temporary Employee	0.061*** (0.011)						0.043*** (0.011)
Cost of Material		0.085*** (0.018)					0.033* (0.019)
Cost of Energy			0.036* (0.019)				-0.003 (0.019)
Number of Total Employee				0.189*** (0.037)			0.080** (0.038)
Expenditure on IT software					0.147*** (0.022)		0.117*** (0.022)
Expenditure on Equipment						0.105*** (0.018)	0.069*** (0.017)
Non-IT Capital Stock							-0.006 (0.017)
IT Capital Stock							0.026 (0.017)
Plant & Year Fixed Effects	Y	Y	Y	Y	Y	Y	Y
# of Plants per Year	~4,600	~4,600	~4,600	~4,600	~4,600	~4,600	~4,600
# of Years	9	9	9	9	9	9	9
R-Squared	0.654	0.653	0.651	0.653	0.657	0.655	0.663

Note: Results in columns 1 to 6 reported the correlations between IT services and other key inputs from the basic OLS model controlling for plant and year fixed-effects respectively. Column 7 also controls for plant and year fixed-effects but include all key inputs in the same specification. Standard errors for all specifications are clustered at the plant level. Statistical significance is denoted as follows: * 10%, ** 5%, *** 1%.

Appendix Table A.6. Additional Robustness Tests

Models	(1)	(2)	(3)
Dependent Variables	**Value-Added per Employee**	**Markup**	**Sales (High IT dummies)**
IT Capital Stock	0.071*** (0.008)	0.001 (0.001)	0.049*** (0.005)
IT Services	0.102*** (0.013)	-0.002** (0.0004)	0.004 (0.004)
Software	0.126*** (0.015)	-0.001 (0.001)	0.006 (0.004)
Equipment	0.204*** (0.013)	0.004*** (0.001)	0.050*** (0.004)
Young	0.421*** (0.041)	0.032*** (0.009)	0.201*** (0.040)
IT Capital Stock x Young	-0.009 (0.017)	-0.001 (0.001)	-0.048*** (0.004)
IT Services x Young	0.116*** (0.026)	0.003** (0.001)	0.035*** (0.010)
Software x Young	0.085** (0.034)	0.0001 (0.001)	0.015 (0.010)
Equipment x Young	0.032 (0.028)	0.001 (0.001)	0.011 (0.011)
Non-IT Capital Stock	0.170*** (0.005)	0.005*** (0.001)	0.076*** (0.003)
Non-IT Capital Stock x Young	-0.069*** (0.007)	-0.002 (0.001)	-0.048*** (0.004)
Other Inputs x Young	Y	Y	Y
Industry x Year Fixed Effects	Y	Y	Y
# of Plants per Year	~26,000	~26,000	~26,000
# of Years	9	9	9
R-Squared	0.289	0.091	0.943

Note: Results in all columns are based on weighted OLS regressions using ASM sampling weights. All columns use the entire sample and control for industry-year fixed-effects. All variables in column 1 are in per employee basis except young dummy. The dependent variable for column 2 is the markup (calculated by dividing operating profit by total value of shipment). The operating profit is calculated by subtracting the total operating cost including cost of material, energy, and labor from the total value of shipment. Columns 1 and 2 control for total number of employees in log term. Column 3 employs specification similar to table 5 column 2 but with IT dummies instead of continuous IT measures. Additional inputs including costs of material, cost of energy (both electricity and fuel), and labor (both expenditure on temporary employees and the count of regular employees) in log terms are controlled for but not reported (available up on request). Standard errors for all columns are clustered at the plant level. Statistical significance is denoted as follows: * 10%, ** 5%, *** 1%.

Appendix Table A.7. Estimates of IT Productivity – I.T. CAPITAL ONLY -- for Young vs. Older, 2006-2014

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Model Description	**Sales All**	**Sales Young Only**	**Sales Older Only**	**Sales x Young**	**Sales Young Only + F.E.**	**Sales Older Only + F.E.**	**Sales x Young + F.E.**
IT Capital Stock	0.017*** (0.001)	0.013** (0.003)	0.016*** (0.001)	0.016*** (0.001)	0.007** (0.004)	0.003** (0.001)	0.002** (0.001)
Young	0.013*** (0.005)			0.206*** (0.039)			0.084** (0.039)
IT Capital Stock x Young				-0.002 (0.003)			0.005 (0.003)
Non-IT Capital Stock	0.055*** (0.002)	0.027*** (0.003)	0.073*** (0.003)	0.073*** (0.003)	0.011*** (0.004)	0.019*** (0.004)	0.021*** (0.004)
Non-IT Capital Stock x Young				-0.047*** (0.004)			-0.010*** (0.004)
Inputs: Labor & Materials	Y	Y	Y	Y	Y	Y	Y
Inputs x Young	N	N	N	Y	Y	Y	Y
Plant and Year Fixed-Effects	N	N	N	N	Y	Y	Y
Industry x Year Fixed-Effects	Y	Y	Y	Y	N	N	N
# of Plants per Year	~26,600	~4,600	~22,000	~26,600	~4,600	~22,000	~26,600
# of Years	9	9	9	9	9	9	9

Note: Results in columns 1, 2, 4, and 5 are based on weighted OLS regression using ASM sampling weights controlling for year-industry fixed effects. Columns 3, 6 and 7 are based on plant-fixed effect models controlling for year trends. The dependent variable for all columns is total sales in log terms. Production inputs are also controlled for (but not reported) in all models in log terms, including: cost of material, cost of energy, and labor (both expenditure on temporary employees and the count of regular employees). In addition, the coefficients for interaction terms between Young and the production inputs in columns 2, 3, 6 & 7 are not reported to save space (available upon request). Results are also robust to inclusion of an indicator for and/or restriction of the sample to single-unit status (not shown). Standard errors for all columns are clustered at the plant level. Statistical significance is denoted as follows: * 10%, ** 5%, *** 1%.

Appendix Table A.8. IT Expenditure Breakdown by Type for Young and Older (IT Intensive vs. Not IT Intensive Industries, 2006 – 2014)

	(1)	(2)	(3)	(4)
Dependent Variables	**% Expenditure on IT Capital Flows**	**% Expenditure on IT Services**	**% Expenditure on Software**	**% Expenditure on Equipment**
Young	0.005** (0.005)	-0.011** (0.004)	-0.013*** (0.003)	0.020*** (0.005)
Late	-0.062*** (0.003)	0.024*** (0.002)	0.036*** (0.002)	0.002 (0.002)
Young x Late	-0.006 (0.006)	0.014** (0.005)	0.005 (0.004)	-0.012* (0.006)
Young x high IT Capital Stock in 2005	0.008 (0.008)	0.015* (0.006)	-0.002 (0.005)	-0.021** (0.007)
high ITK in 2005 x Late	0.016*** (0.004)	-0.007* (0.003)	0.001 (0.003)	-0.009** (0.004)
Young x high IT Capital Stock in 2005 x Late	0.003 (0.009)	-0.006 (0.007)	-0.005 (0.006)	0.011 (0.008)
Industry Fixed Effects	Y	Y	Y	Y
N	~239,700	~239,700	~239,700	~239,700
R-Squared	0.150	0.086	0.097	0.131

Note: Results are from OLS regressions controlling for industry (6-digit NAICS) fixed effects. The dependent variables are the percentage of each type of IT spending with respect to the total expenditure reported on IT. All columns include an unreported indicator for whether the plant reported zero IT expenditure (note that these are reported, not imputed zeroes). **Late** is the indicator for the sample years 2008 through 2014. The **High ITK in 2005** is an indicator for the industries with higher than the sample mean IT capital stock in 2005 at the 3 digit NAICS level. Statistical significance is denoted as follows: * 10%, ** 5%, *** 1%.