Digital Capital and Superstar Firms*

Prasanna Tambe
U. Pennsylvania, Wharton

Lorin M. Hitt
U. Pennsylvania, Wharton

Daniel Rock
U. Pennsylvania, Wharton

Erik Brynjolfsson
Stanford University and NBER

June 10, 2021

Abstract

General purpose technologies typically require complementary firm-specific investments. In the case of information technology, these complementary investments produce a form of intangible capital which we call digital capital. We apply Hall's Quantity Revelation Theorem to panel data collected from LinkedIn to compute historical prices and quantities of digital capital. We find that 1) digital capital quantities grew rapidly and accounted for at least 25% of firms' assets by the end of our panel, 2) digital capital prices varied significantly over time, 3) digital capital has disproportionately accumulated in a small set of "superstar" firms, and 4) digital capital accumulation predicts future firm-level productivity. We conclude that the growing importance of superstar firms in the US economy reflects in part the growing importance of digital capital.

Keywords: Digitization, intangible capital, superstar firms, productivity
JEL codes: D24, D25, O32, O33, M21

*Correspondence: erikb@stanford.edu. Digital Economy Laboratory, 366 Galvez St, Rm 328. Stanford, CA 94305, USA. We are grateful for valuable feedback from Guy Berger, Di Mo, Diego Comin, Frank Nagle, Chad Syverson, Jonathan Haskel, and Lynn Selhat and from seminar participants at MIT, the NBER Conference on Research on Income and Wealth, the Workshop on Information Systems and Economics, the International Conference on Information Systems, LinkedIn, the Brookings Institute Initiative on Productivity Measurement, the NBER Conference on Artificial Intelligence, and the INFORMS Conference on Information Systems and Technology. We thank the LinkedIn Economic Graph Research team for providing access to firm-level skills and employment data.

1 Introduction

Superstar firms, unique in their capabilities to scale up innovations, have become increasingly important to the US economy (Autor et al., 2020; Hall, 2018; Van Reenen, 2018; De Loecker et al., 2020). Investments related to digital technologies are likely to play a particularly important role, reflecting, among other things, economies of scale and network effects (Bessen, 2020). For instance, as of May 2021, the five most valuable firms in the S&P 500 were all firms from the tech industry – Microsoft, Google, Facebook, Amazon, and Apple. Furthermore,the share of stock market capitalization represented by the five most valuable firms has been increasing since 2010, and especially rapidly in the last few years (Figure 1).

Much of the rise in the concentration of value in these firms has been attributed to intangible investments (Crouzet and Eberly, 2018; Ayyagari et al., 2019; Covarrubias et al., 2019; Farboodi et al., 2019; Rock, 2019; Bessen et al., 2020; Hubmer and Restrepo, 2021). For digitally-focused firms, investments in the intangible assets needed to realize value from new technologies – like cumulative investment in skills training, new decision-making structures within the firm, management practices, and software customization – can account for significantly greater total costs than the technologies themselves. These assets comprise digital intangible capital (hereafter, referred to in this paper as "digital capital") (Hall, 2001; Brynjolfsson et al., 2002).

In this paper, our definition of "digital capital" includes factors of production that 1) are complementary to recorded investments in IT assets (such as hardware and software), 2) generate returns that accrue to owners of the firm, and 3) are not otherwise recorded on a firm's balance sheets. Examples include the implementation of business processes and other forms of organizational transformation required to support or use new information technologies as well as firm-specific human capital (employee training) related to these technological systems (Milgrom and Roberts, 1990; Black and Lynch, 2001; Bresnahan et al., 2002).[1] There are many forms of

[1]Prior work has proposed different categories of intangible capital, which can include R&D, organizational capital, brand equity, customer relationships, human resources, and other assets that are valuable to the firm but traditionally go unmeasured (Lev, 2000). It has also distinguished intangibles by their unique economic characteristics (Lev, 2000; Haskel and Westlake, 2018). Corrado et al. (2009a) categorize intangible capital into three major categories: computerized information, innovative property (including R&D), and economic competencies (including brand equity and firm-specific resources). Our conception of digital capital overlaps with computerized information, software, and firm-specific resources (firm-specific human capital and organizational structure) correlated with digital investment.

intangible capital that do not fall under the umbrella of digital capital, such as brands and intellectual property. If these types of capital are not (conditionally) correlated with the firm's IT investment, they will not be measured by the empirical approach we use below and therefore are not measured as digital capital. Furthermore, some types of valuable digital complements, such as open source software and some human capital, are not appropriable by the firm and our firm-based measurement does that count these assets as part of total digital capital value (Greenstein and Nagle, 2014; Kogan and Papanikolaou, 2019).

The general economic importance of these types of investments has been well documented (e.g., see Lev, 2000; Bresnahan et al., 2002; Black and Lynch, 2001; Corrado et al., 2009b; McGrattan and Prescott, 2010; Corrado et al., 2012; Eisfeldt and Papanikolaou, 2013; Sculley et al., 2014; Haskel and Westlake, 2018; Crouzet and Eberly, 2018; McGrattan, 2020; Wu et al., 2020; Abis and Veldkamp, 2020). As the economy becomes increasingly digitized, these assets can be expected to grow even further in importance. For instance, there has been a wave of interest in the potential of data science and artificial intelligence (AI) to become the next important general purpose technology that drives economic growth and business value (Brynjolfsson et al., 2018).

The economics of these assets have both similarities to and differences from those of physical capital. Firms make costly investments in both physical capital and in digital capital to increase their capacity to produce output in future years. For instance, a manufacturer seeking to double its output might build a second factory made of bricks, mortar and machinery, adding to its physical assets, or it might invest in new IT-enabled business processes and human capital that enable it to double the throughput of its existing factory, adding to its digital capital. Either of these approaches take time to build and the market value of firms should reflect the expected net present value of the cash flows they can generate in the future. Like physical capital, digital capital can depreciate over time and must be replenished through additional investment. However, unlike physical capital, intangible capital is not directly observable and may be closely tied to a particular firm and sensitive to external economic conditions making the value more volatile and quantities more difficult to measure.

Generally accepted accounting principles typically do not require separate identification of or permit capitalization of the costs of accumulating intangible capital such as the investments in developing and deploying new business processes. Consequently, characteristics of a firm's intangible capital – such as gross investment or depreciation rates – are unknown, limiting the use of conventional capital accounting methods (e.g. see Hall, 1993). Intangible capital is difficult, if not impossible, to trade in secondary markets. This limits the use of most market-based measures of value. Even when the total value for bundles of these intangible assets can be estimated, fundamental distinctions, such as their price and quantity, remain elusive. This price-quantity distinction is especially important because it is the stock (*quantity*) of capital, not its value (*price x quantity*), that contributes to a firms' productivity and output capacity. When the price of digital capital rises, higher market values may reflect little or no actual improvements in aggregate productive capacity, with most of the increase in market value reflecting rents; if the price of digital capital is stable, increasing values reflect genuine improvements to productive capacity. As digital capital becomes a greater component of overall capital stock, differences in the quantities of digital capital across firms may increasingly explain differences in performance between new digitally-focused firms and older firms.

This paper uses an alternative approach to create measures of changes in the prices and quantities of digital capital in US firms over the last three decades. Earlier work has shown that under certain assumptions, the quantities of a firm's capital stock–and specifically, its intangible capital stock–can be inferred from the value of its securities (Solow et al., 1960; Baily, 1981). This approach has been used to describe the aggregate accumulation of intangible assets in US firm (Hall, 2001). One advantage of using this approach is that it separates fluctuations in price from changes in quantity, thereby addressing some of the measurement issues that arise when investors face difficulties valuing intangibles (Cummins, 2005). To the best of our knowledge, however, no prior work has applied these methods to firm-level data. This precludes the application of these intangible capital stock measures to explaining the distribution of productivity among firms.

While Hall (2001) computed his estimates on aggregate (economy-level) data, it has previously not been possible to perform these estimates at the firm level because the recursive nature of the approach requires a relatively long time series. Existing data on firm-level digital investments such as the CI Technology Database database (or CITDB, also known as the Harte-Hanks database) is too short and lacks the necessary year-to-year consistency needed for Hall's approach.[2] We resolve this measurement issue by developing a new firm-level data series on IT labor, a key input to many forms of digital capital.[3]

To compute these measures, we use LinkedIn profiles to track the employment of information technology workers by firms, and we make adjustments to account for sampling differences across occupations, regions, and industries.[4] The data series that we create extends and improves upon prior approaches that computed IT labor inputs from employment data (e.g., Tambe and Hitt (2012b)). With this longer and more precise series on IT labor, combined with other, more conventional, firm-level financial data, we can recover estimates of digital capital quantities and prices from a time series of changes in digital capital values.

The central contribution of this work then, is in using new IT investment data along with Hall's approach to provide the first firm-level measures of prices and quantities of digital intangible capital. In contrast to prior firm-level work which has investigated the value of IT-related intangibles (Bharadwaj et al., 1999; Brynjolfsson et al., 2002; Dewan et al., 2007), the decomposition of digital capital into price and quantity enables the study of the distribution of asset quantities, their contribution to differences in firm productivity, and the implications digital capital accumulation for important macroeconomic trends. This work also builds on prior research connecting IT investments to differences in the productivity of US firms (Dewan and Min, 1997; Stiroh, 2002; Brynjolfsson and Hitt, 2003), and the important role in this relationship played by complementary organizational investments (Bresnahan et al., 2002). The approach in this paper has the advantage of creating an aggregate measure of the overall stock of digital capital and its economic contribution without having to identify and measure all of its constituent investments.

Our analysis reveals four important facts about the role of digital capital in explaining features of the modern economy. First, the market value of digital capital rose sharply during the late 1990's but then fell in the early 2000's, reflecting changes in the price of digital capital during the dot-com boom and subsequent bust.

[2]This reflects, among other things, significant changes in data collection methods and sampling.

[3]Our measure is based on employment data from from LinkedIn, a popular online professional networking platform (http://www.linkedin.com). Such a measure potentially misses some forms of digital capital, such as open source software repositories. Later in the paper, we are explicit about the mapping between a firm's IT labor measure and its digital capital.

[4]Details of this measure construction and benchmark tests of the data are provided later in the paper.

Second, the value of digital capital began to rise again from 2010 onward, with the timing coinciding with a wave of innovations based on mobile technologies, cloud computing, big data, data science, and most recently, artificial intelligence (AI). With the important exception of the period corresponding with the dot-com boom and subsequent bust, the long run increases in value can largely be attributed to changes in digital capital quantities, rather than prices. Quantities of digital capital rise substantially over the course of our nearly thirty-year panel. By the end of the panel, digital capital quantities account for 20-25% of the levels of physical capital for firms in our sample.

Third, there is substantial heterogeneity among firms in terms of quantities of digital capital they own, with the majority of the increase in quantities concentrated in a subset of "superstar" firms, which we define as those within the top decile of our sample in terms of market value. This concentration contrasts with patterns of accumulation of physical assets (e.g., property, plant, and equipment) and these differences are increasing over time.

Fourth, by creating firm-level measures of digital capital quantities, we can estimate how the accumulation of this form of capital contributes to productivity and growth. In productivity regressions, we estimate the contributions of our digital capital quantity measures alongside IT capital stock measures. We find that the contribution of digital capital to output growth during this period was approximately double that of conventional IT capital stock. Moreover, this contribution occurs with a two- to three-year lag, indicating that that while its value is reflected in firm's stock prices immediately (reflecting the forward-looking decisions of investors), it takes several years to fully install into the firm and become productive.

In sum, because we have access to technology and skill data at the firm level, we can estimate not only how prices and quantities of digital capital have been changing, but also how these quantities vary across firms, and how the growing importance of digital capital is connected to the growing importance of superstar firms in the US economy.

2 Theoretical framework: Asset values and adjustment costs

Our goal is to reveal differences in firms' quantities of digital capital. The notion that the value of a firm's securities reveals its capital stock has appeared in several forms in the literature (Solow et al., 1960; Baily, 1981).

In this paper, we rely principally on Hall's formulation, which is applied to intangible assets. Hall argues that under assumptions of 1) competitive markets, 2) constant returns to scale production, and 3) full factor adjustment, prices and quantities of tangible and intangible capital can be recovered from the value of a firm's securities (Hall, 2001). The main difference in our approach is that we estimate key parameters with firm-level data, using a time-series of investment into a correlate of the firm's intangible investment. In addition, rather than using the total value of a firm's securities as our dependent variable, we use the component of its market value correlated with IT investment. After separating digital capital values into prices and quantities, we use the quantities to estimate the contribution of digital capital to productivity.

From Hall (2001), we have two equations relating investment and market value in the presence of adjustment costs:

$$V_t/K_t = q_t \tag{1}$$

$$q_t = 1 + c'(i_t/K_{t-1}) \tag{2}$$

The first equation indicates that the ratio of market value (V) to dollar of capital value (K) is Tobin's q (1). Then, the adjustment cost condition is that the marginal q value is 1 plus the marginal cost of adjustment at that amount of investment growth ($\frac{i_t}{K_{t-1}}$) (2). The 1 in the equation represents the numeraire price of capital purchase costs and the adjustment cost function is assumed to be convex in time. We follow the convention in Hall (2001) (following Hayashi (1982)) that demonstrates an equivalence between the average q in (1) and marginal q in (2) given our assumptions.

Once capital is installed in a firm, its value to shareholders may be greater than its purchase price. This difference reflects the adjustment costs that competitors would need to incur to duplicate the same productive capacity. For instance, simply having a computer delivered to a company's loading dock will not increase the firms output, but installing productivity software, integrating it securely into the firm's computer network, and training a worker to use it effectively may. These additional steps are costly, often far more costly than the hardware itself, and thus lead to values of q that exceed 1, perhaps even by several multiples. Furthermore, sunk adjustment costs can lead to an ex-post quasi-rent of installation equal to the difference between the competitive value of installed capital less the firm's specific adjustment cost. The sum of these values accrues to the shareholders to produce the identity:

$$V_t = \sum_{j=1}^{J} q_{j,t} K_{j,t} \tag{3}$$

The market value of the firm at time t is the sum of the values of its constituent assets (indexed by j), priced at their respective q values. Solving the system of equations above leads to a convenient way to recover q and $\mathbf{K}$ at some time t if we know the market value and previous period's capital stock of the firm.

Not knowing the adjustment cost function precisely, we have that:

$$c'(i_t/K_{t-1}) = c'(\frac{K_t - (1-\delta)K_{t-1}}{K_{t-1}}) = q_t - 1 \tag{4}$$

We can observe market values, and we have to pick a starting capital value to solve the recurrence relation above. With last period's capital stock and the current market value, we can infer q and $\mathbf{K}$ in the current period. This also requires knowing the adjustment cost function. If we assume a quadratic function for adjustment costs (Holt et al., 1960; Hamermesh and Pfann, 1996; Belo et al., 2014), we can solve for the adjustment cost function with a similar recurrence relation.

$$c(\frac{i_t}{K_{t-1}}) = \frac{\alpha_t}{2}(\frac{K_t - K_{t-1}}{K_{t-1}})^2 \tag{5}$$

$$c'(\frac{i_t}{K_{t-1}}) = \alpha_t(\frac{i_t}{K_{t-1}}) = q_t - 1 \tag{6}$$

These two equations are taken directly from Hall (2001).

Importantly, α can vary across firms, capital varieties, industries, regions, or other relevant groupings of capital and may differ based on whether stock is increasing or decreasing. For our analyses, we assume a value for α and later test robustness of our conclusions to alternative assumptions.[5]

The marginal adjustment cost at a given level of net investment per unit of capital stock in the previous period is equal to the q value of investment in the current period minus the numeraire. From equation 1, we have that q is equal to the market value of a given asset divided by the asset replacement value. The point where the marginal adjustment cost function intersects the hyperbola $V_t = q_t K_t$ defined by the market value sets the equilibrium price and quantity of capital (Hall (2001) illustrates this as shown in Figure 2). With measures of a capital panel K_t and net investment i_t (or gross investment and asset-specific depreciation rates), these equations can be solved to yield q_t and K_t.

Note that the q curve can be interpreted as the full market value of the firm or the q values of specific types of assets. Cash, for example, has low adjustment costs, so the supply curve of cash is close to flat with a price of one dollar. Property, plant, and equipment has a higher installed price reflecting these adjustment costs or rents associated with successful deployment.

Rather than solving directly for q, we can estimate the following hedonic regression from observed capital stocks and observed market values to recover the q value as a vector of coefficients:

$$V_{i,t} = \mathbf{K}'_{i,t}\beta + \mathbf{X}'_{i,t}\xi + \gamma_i + \nu_t + \epsilon_{i,t} \tag{8}$$

The value of a firm is decomposed into its constituent assets $\mathbf{K}$ with controls $\mathbf{X}$ and fixed effects for firm i and time t. If all assets are measured perfectly, then the coefficient vectors return the asset q values. Note that the controls and fixed effects are not necessarily required in this case.

[5] With the additional assumption that q is stable, the second order condition is useful for recovering an approximation of α_t as follows:

$$c''(\frac{i_t}{K_{t-1}}) = \alpha_t \tag{7}$$

It is possible, however, to treat the slope of the marginal adjustment cost function as an unknown instead with by assuming the investment rates instead and solving equation 6. Since we are interested in the extent to which digital capital assets have been accumulated in recent history and not the convexity of the adjustment cost function, this paper does not implement such an approach. For more easily measured asset varieties where investment rates are easier to observe, an empirically estimated adjustment cost convexity parameter $\hat{\alpha}_{it}$ for some observational unit i might be a useful measure of rent durability.

3 IT investment time-series

3.1 IT labor measures

Applying this approach to intangible capital measurement requires a measure that allows us to estimate the firm's IT capital stock. Because IT employment is highly complementary to IT capital, we use it as a proxy measure for IT capital.[6]

Our IT labor measures are constructed from data on the employment histories of US IT workers, which are used to infer the distribution of IT employment in US firms. The primary data were obtained from LinkedIn, the leading online professional network upon which individuals post employment histories, which include data on each job they have held. These data provide a time series for each employee of each employer and occupation over the span of their career.[7] To measure IT employment, we identify employees who have job titles identifying them as IT workers, link them to firms and then aggregate by year. Our LinkedIn panel covers the years from 2000 onwards. We supplement our LinkedIn IT employment panel with data from prior work that used a similar approach on databases from different Internet job search boards that covered earlier timer periods (Tambe and Hitt, 2012b).

The details of the data series construction and validation of these measures using other data sets with known sampling properties (e.g., the BLS Occupational Employment Survey) are in Data Appendix A. A brief summary of the procedures is as follows. We extract from each LinkedIn profile the firm, occupation (based on LinkedIn's internal taxonomy), and time of each employment spell which we aggregate by firm, after reconciling firm names to common identifiers. This provides firm-level occupation-year counts. We then adjust for imperfect sampling on LinkedIn by rescaling these counts to reflect the difference between a) the ratio of observed employees to total firm employment in a given year from Compustat, and b) the observed

[6]Our analysis is based on a hedonic regression of the firm's market value on its capital assets. We relate IT labor to the firm's IT capital stock by making the assumption that a dollar of IT capital requires a fixed amount of IT labor to maintain it. That is, a firm that doubles its IT capital would require twice the IT labor to maintain it. This assumption allows us to use IT labor expenditure as a proxy measure for the firm's aggregate IT capital stock. Lichtenberg (1995) was perhaps the first to use IT labor as a primary measure of IT investment. Similar approaches have been used in more recent work such as (Tambe and Hitt, 2012b; Rock, 2019) Prior work that has examined correlations between IT capital and IT employment at the firm level finds a correlation of 0.56 for those measures in levels (versus 0.62 for those measures in logs) (Tambe and Hitt, 2012a).

[7]The data were obtained in early 2018, but we only use the years through 2016 because the lag in workers updating their profiles could otherwise add significant noise to our measures.

distribution of occupations in LinkedIn and the expected distribution of corresponding occupations from BLS-OES employment statistics for Compustat firms overall. [8] These two adjustments account for the differing propensity of workers in some firms and in some occupations to participate on LinkedIn.

Among other advantages, this approach allows a longer and much more consistent time series than is available in data on IT capital stock such as the CITDB database (e.g., see Bresnahan et al. (2002)). The length of the time series permits the study of digital capital in a consistent manner through the last decade of technology spending, and makes it possible to use the dynamic recursion methods we describe to estimate digital capital. Finally, the use of labor based measures may be more closely tied to the creation of digital capital as firms can deploy new business processes on existing hardware and the rise of cloud computing has placed more of the IT hardware investment outside the firm.

3.2 Measurement Error of IT Investment

The data we use are collected through a partnership with LinkedIn, not scraped, so we are not subject to the types of measurement problems that arise when using scraped or incomplete data from a website. Nonetheless, there are gaps. Many workers do not participate on LinkedIn, and some of those who do might report inaccurate or incomplete employment history information. This inaccurate or incomplete reporting can generate error in our IT employment measures. The high participation rate of IT workers within the US on the LinkedIn platform suggests that measurement error should be less than what we might otherwise observe in the CITDB IT capital data or other comparable IT data sources. Prior work (Tambe and Hitt, 2012b) has indicated that employment data, as a proxy for IT investment, are subject to less measurement error than alternative sources such as CI capital where error variance may be as high as 30-40% of the total measure variance (Brynjolfsson and Hitt, 2003).

Another source of measurement error when using IT labor as an IT input measure is IT employment outsourcing, which can include investments in cloud computing. If firms outsource a significant fraction of their IT employment, they would appear to derive the full value of IT investment from their in-house and outsourced labor

[8] We require an assumption of identical employment composition for profiles in the U.S. and abroad that may not be correct. If European workers are less likely to be software engineers, for example, we might overstate the count of software engineers in that firm.

inputs but IT labor expenditure would be under-reported. Various indicators have placed outsourcing at about 10-15% of total IT budgets Tambe and Hitt (2012b). In either case, this compares favorably with the 30-40% measurement error associated with using historical firm-level IT capital stock data. Moreover, if firms uniformly outsource IT labor it will not affect our estimates; if outsourcing is random it will generally understate measures of digital capital due to errors in variables bias. Tambe and Hitt (2012b) found that including measures of outsourcing (derived from firm-level surveys) does not affect production function estimates of the contribution IT labor.

Regardless of the measurement issues, IT outsourcing also requires its own set of intangible investments in efficient contracting and coordination, and many of the other intangible assets included as digital capital are still accumulated by the firm in the case that it outsources part of its digital capabilities. Cloud computing, for example, sends computational hardware assets outside of the firm, but the intangible investment in IT surrounding business processes, management practices, and applications of computational processes to business operations are still within the firm boundary (and accordingly, rents on this capital accrue to the firm and not their outsourcing partners). We therefore expect even outsourcing firms to employ meaningful quantities of the kinds of workers we use to track the digital capital stock.

Some of our specifications include salaries. As with survey data on employment counts, these salaries are also taken from the BLS-OES annual survey and matched to LinkedIn employment categories. The salaries reflect an average compensation rate by occupation-year, but salaries for the same types of roles can vary considerably by firm, region, seniority, and other factors. We expect averaging over many firms, occupations, and years to mitigate the variance in the salary-based measures, but it is still a component of the measure. There is also directional bias in the compensation measures to the extent that employers also compensate workers with equity or options. In that sense all of our salary-based measures are underestimates of the true employee compensation expenses. Coefficient estimates in these specifications should therefore be considered upper bounds, as variation from stock-based compensation is likely positively correlated with salary payments.

Our measures and estimation procedures also miss the social value of open source software (OSS). If a particular worker's human capital or a firm's digital capital does not generate rents and is otherwise costless to the market, then it is absent from stock market valuations. The value of open source tools is high and a driver of productivity, but not always appropriable (Greenstein and Nagle, 2014; Nagle, 2019). Our measures will miss the large social value of open source tools. However, they will capture the replacement cost and firm-specific rents associated with using otherwise free and standardized technologies.

4 Descriptive statistics and empirical analysis

4.1 Data, sample, and statistics

Applying Hall's Quantity Revelation Theorem to uncover firm-level digital capital quantities requires a firm-level time series of market and asset values that includes a correlate measure for digital capital as described above. To create a panel from these measures, we combined the IT labor series with firms' market values and assets collected from the Capital IQ database. Below, we report results from an analysis of two different panels, 1) a balanced panel consisting only of firms that appear in all years from 1987 to 2016 (n=277), and 2) an unbalanced panel over the same time period that can include firms that enter or exit the panel (n=5,215). Tests performed on the balanced panel have the advantage of consistency in sample composition over the duration of the panel, but they exclude firms that fail part way during the sample period as well as firms that only appeared in the late 1990's and after. For the balanced panel, the firms are disproportionately larger, older, and more established firms, so on a per-employee basis, our estimates of digital capital using this sample could be understated. Figure 4 illustrates the distribution of firms across NAICS industries for each of these two samples.

In Table 1, we report key statistics from the firm-level data in our panel. Statistics are reported from firms' 2005 values. The year 2005 was chosen because it is close to the midpoint of our panel but is not during the dot-com bust, which triggered large layoffs in IT sectors. The figures in the table indicate that firms in our sample are large, with an average market capitalization of about $35 billion and with roughly 57,000 employees. The average firm in our sample has about 3,200 IT employees, or a bit more than 5% of their workforce. This can be compared with BLS Occupational Employment Statistics that report that the IT workforce comprised slightly more than 2% of the overall US workforce in 2005.[9] This difference suggests that IT labor is employed in disproportionately larger numbers by the firms in our sample. Data Appendix A contains more detailed comparisons of occupational coverage in the LinkedIn data.

4.2 Digital capital market values, prices, and quantities

4.2.1 IT and market value

Table 2 reports results from regressions of market value on PP&E, IT, and other inputs using the Tobin's q framework shown in 8. The form of this regression is similar to that used in prior work Brynjolfsson et al. (2002) (hereafter, referred to as BHY), except that the IT investment data source for this research is different and the panel is longer. These regressions a) benchmark how the IT employment data perform in market value regressions and b) estimate contributions of IT investment to market value in a panel that post-dates the IT data used in existing firm-level research by almost two decades. We report estimates from Ordinary Least Squares (OLS) and fixed effects models as well as Least Absolute Deviation (LAD) models and all of the regressions in the table include year and two-digit industry fixed effects. In a market value context, LAD regressions reduce the impact of outliers (e.g. very large market value firms) on the estimates, although our point estimates do not change substantially when using these methods. For ease of interpretation, we scale the IT employment measures to millions of dollars by multiplying it by 0.10, which implies a cost of $100,000 per "unit" of IT labor. Below, we discuss how this wage assumption affects interpretation of our estimates. We first present results from the larger unbalanced panel and then from the balanced panel of firms for which data are available for all the years in our panel.

We first replicate results from earlier studies using our new IT employment dataset. Column (1) of Table 2 replicates the specification from BHY, using IT capital stock measures for the years from 1987 to 1998, constructed from the CITDB data, which is the same data set that was used in that paper.[10] For the unbalanced panel, estimates

[9] The BLS Occupational Employment Statistics reports that employment in Computer and Mathematical Occupations accounted for 2.27% of total US employment in 2005. Historical BLS occupational employment statistics can be found at https://www.bls.gov/oes/tables.htm. Last accessed on July 3, 2019.

[10] A detailed description of how these capital stock figures were created can be found in Brynjolfsson

on Property, Plant, and Equipment (PPE) and other assets (total assets minus physical capital) are close to their expected theoretical values of one, plus an increment for adjustment costs for that type of capital.

The coefficient estimate on IT capital indicates that a dollar of IT capital is associated with about twelve dollars of market value, although standard errors are large so we cannot reject the the hypothesis that the contribution of this type of capital to value is zero. However, it is notable that the point estimate is very close to the BHY estimates, which found that each dollar of computer investment was correlated with eleven dollars of market value.

In column (2), we substitute our new IT employment based measures for the IT capital measures, while using the same restricted time period used in column (1). The point estimate on IT employment suggests that a dollar of IT labor is associated with around nine dollars of market value (8.62, t=1.80). Unlike IT capital, IT labor is not owned by the firm's shareholders. Therefore, most of the dollar value implied by the coefficient estimate on IT labor should be assigned only to intangible correlates and their adjustment costs. That is, we do not need to subtract out the contribution of IT labor to the firm's market value.

Column (3) includes both IT measures in the same regression. The change in estimates on the IT labor and capital coefficients suggests that IT labor is a more precise measure of this type of investment, perhaps due to greater measurement error in the IT capital figures (Tambe and Hitt, 2012b) or as might be the case if spending on IT labor is a better indicator than IT capital of the development of new digital capital (17.70, t=1.83). This would be the case, for instance, if IT workers could build new digital intangibles on old servers or if the use of cloud computing is making measures of in-house IT capital a less informative measure of computing capacity

et al. (2002). Some of the description is replicated here for convenience, but the paper has additional details. "The measures of computer use were derived from the Computer Intelligence Infocorp (CII) installation database, which details IT spending by site for Fortune 1000 companies. Data from approximately 25,000 sites were aggregated to form the measures for the 1,000 companies that represent the total population in any given year. This database is compiled from telephone surveys that gather detailed information about the ownership of computer equipment and related products. Most sites are updated at least annually, with more frequent sampling for larger sites. The year-end state of the database for each year from 1987 to 1997 was used for the computer measures. From these data we obtained the total capital stock of computers (central processors, personal computers, and peripherals). The IT data do not include all types of information processing or communications equipment and are likely to miss some portion of computer equipment that is purchased by individuals or departments without the knowledge of information systems personnel."

relative to IT labor.[11]

Having found that these new data generate results largely consistent with prior results that had an overlapping (but considerably smaller) time period, we now report estimates from our full available sample that ranges from 1987 to 2016 (see Column (4)). In the extended panel, each IT employee is associated with about fourteen dollars of employer value (14.12, t=3.44). That is, the coefficient estimate on IT grows if we include the post-2001 years into the panel. Column (5) reports estimates from a similar analysis on all years but that uses a LAD regression to account for possible outliers. Here, the estimated magnitude on each IT employee is about the same as produced by the OLS regression in (3), indicating about twelve dollars of employee value per IT employee (11.98, t=12.92).

In column (6), which includes firm fixed-effects, the magnitude of the IT estimate falls to nine dollars of market value per dollar spent on IT labor (9.38, t=2.75). This indicates that about a third of the estimate on the IT variable from the prior column is due to cross-sectional heterogeneity in firms that invest in IT and those that do not. The theory suggests as well that q value is created by firm-specific capabilities. Firm fixed effects absorb time-invariant firm characteristics, such as fixed assets without ongoing investment requirements. The consistency of large positive and statistically significant coefficient estimates on IT labor across specifications with alternating fixed effect adjustments implies the presence of meaningful time-varying investment costs in digital capital.

Finally, the estimate in (7), which uses a somewhat more precise measure of IT wages, rises to about twelve dollars per IT employee (12.10, t=2.39). This more precise wage measure is computed by using the LinkedIn data to determine the mix of IT occupations within the firm, and using BLS occupation-level wages to compute the weighted IT wage number. In summary, the estimates in columns (3) through (7) suggest that in the last decade, the market continued to assign significant value to digital capital as estimated when using a proxy measure based on IT labor. The

[11] IT labor spending may also be a proxy for technology asset utilization. Although many firms may have a PC on every desk, the percent of time each PC is in use might vary from firm to firm with the percent of workers' jobs that use a PC. Moreover, even if worker use PC constantly, the CPU utilization of PCs might vary from firm to firm with the sophistication and computational intensity of the firm's software. Both of those utilization factors would presumably be correlated with IT labor and with returns. Returns on IT investments would be more strongly correlated to IT labor than IT asset because although the IT asset is necessary, it is not sufficient – idle PCs produce no returns.

higher estimate in column (7) than in column (6) is consistent with the idea that measurement error in IT wages in the prior column slightly reduces the estimate.

Table 3 shows results for the balanced panel, which is limited to firms that do not enter or exit the panel between 1987 and 2016. The pattern of results is similar to that reported in 2, although there is a fall in the magnitudes of the IT estimates, which suggests that the older firms in this sample have developed less digital capital and are less likely to be among the "superstar" firms that do especially well. Nevertheless, as with the unbalanced panel, the magnitude of the estimate on IT investment suggests the presence of significant correlated but unmeasured assets. Moreover, the estimated coefficients when using the IT employment measures suggest effects of similar size on the larger panels, and the estimate in column (3) is again consistent with the argument that IT labor is a more precise measure of the firm's IT assets than our IT capital measures.

4.2.2 Prices and quantities of digital capital

Recovering prices and quantities for digital capital requires a time-series of digital capital values. To estimate year-to-year changes in the value of firms' digital capital, we use the specification in equation 8 with the balanced panel, but we limit the sample for each regression to observations in a window around the focal year.[12] The IT coefficient estimates generated using this approach are shown in Figure 5. They rise and fall in a pattern that is consistent with some of the changes in levels of IT investment that occurred from 1995 through 2001 – the dot-com "boom" – and again in the years after 2010, which coincide with a recent wave of investment around data collection and mobile technologies as well as the data science and AI tools that are increasingly being used for data-driven decision-making.

Figure 6 plots the average of the imputed digital capital values for firms in the balanced panel over the course of our panel. The average digital capital value (i.e., the component of market value correlated with measures of IT employment) has fluctuated substantially, and the dot-com bubble and bust appears to have had a significant effect on the value of digital capital computed using this method. For the average firm in our panel, digital capital values rose to about $5 billion during the height of the dot-com boom, before falling to about half that number in the years following the bust. These values then continued to fall through the early and mid

[12] See Table 3 of Brynjolfsson et al. (2002) for use of a similar approach.

2000's, before beginning to rise again shortly after the Great Recession and onward through the most recent time period covered in the sample.

The fluctuations in digital capital value shown in Figure 6 suggest changes in the value of firms' digital capital over time. The conditions in equations 1 and 2 enable us to separate the role of changes in price and quantity in explaining these changes in value. Solving this system of equations, Figures 7 and 8 depict how the prices and quantities of digital capital have been changing for these firms during the years covered by our panel. The prices and quantities shown in the charts are average values for firms in the balanced panel.

The dominant feature for the price series in Figure 7 is its rise and subsequent fall corresponding to the late 1990's dot-com boom and bust. The theoretical prediction is that as Tobin's q for an asset rises above one, firms will have a greater incentive to invest into that asset. This theoretical relationship appears to be consistent with the two illustrations, as the rise in Tobin's q corresponds to accelerated investment by firms during the late 1990's. The dominant feature in Figure 8 is the rise in digital capital quantities in the panel, which occurs particularly rapidly during the dot-com boom years.

It is worth noting that, in contrast to the period around the dot-com boom, digital capital prices in the later part of the panel were quite stable despite a similar run-up in market values. Aside from the late 1990's and early 2000's, the price stays close to its theoretical value of one. Figure 8 implies a corresponding accumulation of digital capital quantities, except for a slowdown in the growth rate after the dot-com bust and a decline in quantities for a few years after. These figures suggest that the large market value changes around the dot-com boom and bust can be partly attributed to changes in digital capital price, not only to changes in the quantities of digital capital possessed by the firms in our sample. The bust removed the effects of possible investor mispricing from digital capital values, and the slowdown in the accumulation of digital capital stocks continued through the Great Recession. In contrast, there is little evidence that changes in price explain the more recent increases in digital capital value.

To place the growth of digital capital quantities in perspective, we can compare how digital capital growth compares to the growth of physical capital. Figure 9 makes this comparison explicit. By the end of our sample period, our imputed digital capital quantities had grown to approximately 25% of the magnitude of measured physical capital. It is worth noting that the sample for this comparison is the balanced panel, which does *not* include any firms such as Google or Facebook, which were founded later. In that respect, this comparison is likely a conservative one. It all suggests that the rise of digital intangible quantities is not only about the oft-discussed digital upstarts in pure-digital industries. Instead, we find that leading firms across other industries and over time have also built digital capabilities.

As described in Appendix B, the construction of these digital capital price and quantity measures is robust to a variety of depreciation assumptions on digital capital and technological labor. Two parameters are further required for the recursive methods used to recover prices and quantities from digital capital values: a) an adjustment cost parameter (α) which indicates the costs of installing new digital capital into the firm and b) initial quantities of digital capital owned by firms in the first year of the panel. Hall (2000) argues that using recursive methods to compute digital capital quantities in the way described earlier is not overly sensitive to either of these two assumptions. In Appendix C, we explicitly test the sensitivity of our estimates to these assumptions using the time series data and confirm that Hall's claims also apply in the case of our data.

4.3 Firm-level heterogeneity in digital capital

The construction of firm-level measures allows us to analyze how firms differ in terms of digital capital. In particular, there has been growing interest in the literature in the concentration of some types of capital in superstar firms in the US economy (Haskel and Westlake, 2018; Hall, 2018; Crouzet and Eberly, 2018; De Loecker et al., 2020; Autor et al., 2020). Using the firm-level data generated on digital capital quantities, we can compute distributional statistics on how digital capital has been accumulating in firms. Our approach to measuring digital capital is particularly useful here because we define superstar firms based on market value. By removing the effects of digital capital price and focusing solely on digital capital quantity, we eliminate the possibility that superstar firms appear to have more digital capital because they receive greater value for all their assets.

Figures 10a and 10b plot changes in the prices and quantities of digital capital when grouping firms in the sample into four quartiles according to their market values in the final year of the panel. In the two figures in the top panel, the different trend lines in each chart correspond to one of the four quartiles. The most salient feature in this set of charts is the relative concentration of digital capital quantities in the highest quartile group and that the gap in digital capital quantities between the top quartile and the bottom three quartiles has been growing larger over time. In particular, there appears to be less accumulation of digital capital in the lower quartiles and essentially no accumulation at all in the two bottom quartiles. Even within this limited sample of firms that appears in the balanced panel, therefore, there is evidence that the accumulation of digital capital assets is in line with a rise in the concentration of assets among US firms.

The charts in Figures 10c and 10d further subdivide firms in our sample into even narrower groups: into deciles in the bottom left figure and into ventiles in the bottom right figure. Both charts echo the finding of Figure 10b, which is that digital capital quantities are concentrated in firms at the top of the market value distribution. Digital capital appears to be especially concentrated in the top decile of firms, and this group seems to be pulling away from other firms during the years in the panel.

It should be again noted that this result is not driven by Internet-economy firms such as Google, Facebook and Amazon as they were founded after 1987; instead it reflects the differences in the willingness or ability of incumbent firms to accumulate digital capital. Indeed, when we perform a similar analysis on the unbalanced panel which includes firms founded later, the divergence is even greater (Figure 11). While these entrants are smaller (or enter during the panel) which lowers the average quantity of digital capital, the differences among firms in this analysis is even greater than those shown in Figure 10b, with essentially all the digital capital created since 2001 in just the top decile of firms.

These results are driven specifically by the accumulation of digital capital and not by differentials in growth rates across firms that would lead them to accumulate assets of all types. The charts in Figure 12 suggest that digital capital is unique in its pattern of accumulation in more valuable firms, at least when considered in relation to employment share. Figure 12a illustrates digital capital quantities per employee. In contrast, Figures 12b and 12c show changes in quantities of two other asset types normalized by employment: property, plant, and equipment as well as a measure of all other types of recorded assets (i.e., all assets excluding digital capital and PPE). These two figures indicate that while there is rising concentration within the deciles for these different asset classes, only digital capital is rising faster than employment share in higher value firms. In other words, after normalizing by employment, quantities of digital capital per worker were becoming particularly concentrated in the highest decile firms in terms of market value.

The Gini coefficient of the overall concentration in assets is plotted in Figure 13 and these findings echo those shown in Figure 12. Figure 13a shows that while concentration levels have been rising for most assets, they have been rising more rapidly for digital capital for those firms in our sample. This rise in concentration was most rapid in the years that immediately followed the dot-com bust, but the rising trend has continued in recent years. Similar and even more pronounced trends can be observed in Figure 13b, which plots changes in concentration of the employment normalized measures of digital capital quantities. In this plot, concentration rises rapidly after 2000 but plateaus after the Great Recession in 2008. Within the last decade, therefore, most of the rise in concentration in digital capital from Figure 13a has been moving in step with growth in firm size.

Figures 13c and 13d analyze the change in asset quantities and Gini coefficient concentration if firms are grouped by one-digit NAICS industry. In these figures, only the four one-digit industries with the highest numbers of firms are shown: NAICS sector "2" (Construction and Utilities), NAICS sector "3" (Manufacturing), NAICS sector "4" (Trade, Retail, and Transport), and NAICS sector "5" (Information, Finance, Insurance, Real Estate, and Professional Services). Figure 13c suggests that digital capital appears to be accumulating espeicially rapidly within firms in the NAICS "5" sector.

4.4 Digital capital and productivity

The productivity literature has found that the estimated output elasticity for IT capital in firm-level productivity regressions is typically two to three times larger than its cost share (see e.g., Brynjolfsson and Hitt (2003)). Researchers have suggested this gap is likely due to the output contribution of omitted but correlated factors of production such as the types of intangible assets that we call digital capital. Directly computing measures of digital capital quantities at the firm-level allows us to separately estimate the output contributions of traditional IT capital stock and digital capital stock. It is worth noting that production functions are technical relationships that require measures of quantities (i.e., capital stocks) to estimate inputs and output. The overall value of a firm's IT assets would be of little use here. Therefore, it is precisely the separation of IT market values into prices and quantities – the central contribution of this paper – which newly enables this type of analysis.

Table 4 reports results from productivity equations that include measures of labor, non-IT capital, IT capital, and digital capital in the same balanced panel of firms that was used to compute digital capital quantities. We assume production can be approximated by a Cobb-Douglas function, which has been the model of production most commonly used in the literature in this area, and the dependent variable we use is value-added, which is computed as output minus materials. The CITDB IT capital stock measures used in the regression are the same measures used in prior work on IT productivity at the firm-level unit of analysis (Brynjolfsson and Hitt, 2003). These data were collected through surveys, and they have the limitation that the data consistency begins to decline notably after the mid 1990's and particularly after 2000, so we limit our productivity estimates using these data to a shorter panel of years that ends in the late 1990's.

Column (1) includes IT capital, but not digital capital, as a factor of production and the estimated coefficient on the IT capital term suggests an output elasticity (0.037, t=2.31) that is similar to that found in other work that performs cross-sectional comparisons. Directly introducing digital capital measures in column (2) produces estimates which indicate that the contribution of digital capital to value-added is positive and about the same order of magnitude as IT capital (0.034, t=2.12). However, the relative magnitude of the coefficient estimate on digital capital stock is lower than what might be expected from survey-based estimates of the shares commanded by each of these inputs. For instance, Saunders and Brynjolfsson (2016) estimates that 70-80% of IT spending is on complementary organizational assets. One possible reason for the smaller than expected estimate is that the years that comprise most of the sample in our productivity regression are the late 1980's and early 1990's, which, as shown in our earlier figures, precede the larger run-up in digital capital stocks in this sample of firms.

We can directly test this assertion in column (3) by restricting our sample to observations that appear in the latter part of the sample used in column (2). The sample in column (3) is limited to the years 1997 to 2000, which were associated with a rapid increase in quantities of digital capital. The coefficient estimates on digital capital from these years suggest an output elasticity for digital capital that is more than double that of IT capital stock (0.093, t=1.90). In column (4), we introduce narrower, two-digit industry controls to remove some of the effects of industry-level heterogeneity on these estimates. This specification produces estimates of the output elasticity of digital capital that are similar in magnitude but because the standard errors rise, the estimates are no longer significantly different than zero. Column (5) adds R&D investment as a factor of production [13] that generates another type of intangible asset, intellectual property. R&D has been used extensively in prior work on the measurement of a firm's intangibles as a proxy for intellectual property. The digital capital estimates are of the same magnitude as the estimate in (2) which covers the same years, although in the limited sample, the estimate is no longer significantly different than zero. Regardless, this is likely a lower bound on digital capital contribution since these stocks have grown considerably (although we do not have data on IT capital in the longer sample to test this directly).

Errors in the measurement of digital capital quantities may not only exert a downward bias on the digital capital coefficient estimate, but can also transmit an upward bias on the IT capital stock measure because of the correlations between these two inputs. Fixed-effects estimates may remove some of this measurement error, but they may also remove some of the estimated effect of digital capital if the input varies slowly. Indeed, prior work has argued that IT organizational assets do change slowly relative to IT capital because of the high costs of organizational transformation and re-engineering (Bresnahan et al., 2002). Columns (6) and (7) report estimates after introducing firm fixed-effects. The magnitude of the coefficient estimates on IT and digital capital in column (6) (0.026, t=2.00) and column (7) (0.023, t=1.92) are similar in size although slightly smaller than in the cross-sectional regression results from columns (2) and (3) of the table.

Table 5 augments these analyses with measures of the firm's human capital computed from the LinkedIn data. The literature on skill-biased technical change links human capital to the productivity of digital enterprises. It argues that information technologies are complementary to higher levels of human capital in workers. Moreover, the notion that technology investments and skilled workers are key inputs into the production of digital capital is reflective of Hall's argument that college-educated workers are the key input into building "e-capital" (Hall, 2000), suggesting how these

[13] Not all firms, especially in the service sector, report R&D so we include a dummy variable if R&D is not available.

two are related in terms of a production relationship. In other words, technology and human capital are inputs into the production of digital capital, which, in turn, increases productivity levels.

Table 5 estimates a production function for digital capital stock that uses IT and human capital as inputs. The dependent variable in this table is digital capital intensity (i.e., computed on a per employee basis). Independent variables are capital intensity, IT intensity, and human capital intensity where all measures are computed as logs of the per-employee values. The regression output in column (2) indicates that higher levels of IT are correlated with the overall digital capital stock (0.350, t=2.97), but combinations of greater IT and human capital intensity are correlated with still greater levels of digital capital intensity (0.048, t=3.20). These estimates are consistent with the claim that these two factors are complementary in the production of digital capital.

Columns (3) and (4) reproduce these tests after including firm fixed-effects. The significant coefficient estimate on the interaction term persists, in column (4), after including firm fixed-effects (0.048, t=3.00), which indicates that the development of digital capital in firms coincides with simultaneously raising levels of IT and human capital.

The productivity results suggest that digital capital is a form of productive capital stock that is generated by combining technology and human capital (as well as other, possibly unmeasured, complements). However, it is often hypothesized that intangible capital of this type takes time to create, install, and become productive (Brynjolfsson and Hitt, 2000; Brynjolfsson et al., 2002) and that this is especially so in larger, higher value firms that require greater organizational changes, but for which the payoff to productivity may be higher (Tambe and Hitt, 2012a).

The next table tests the hypothesis that although digital capital is reflected in current market values, it impacts productivity only after a delay (Brynjolfsson et al., 2002). Table 6 tests timing for the baseline production regression, in which digital capital quantities are introduced with varying lags, such that moving across the table tests conditional correlations between levels of digital capital and future productivity levels. In each column, digital capital is lagged by the number of years specified in the column header. All columns in this table include firm fixed-effects. The estimates suggest that digital capital is correlated with future productivity as much as it is with contemporaneous productivity. The magnitude of the correlation peaks at around two years (0.024, t=2.40), which is consistent with the argument that digital capital requires installation time. Such a pattern, in which intangible capital contributes to the firm's current market value but does not contribute to productivity levels for a few years, is entirely consistent with the broader empirical literature on returns to IT investment (Brynjolfsson et al., 2018) broadly and digital capital specifically (Brynjolfsson et al., 2002).

A second way to test the argument that digital capital requires installation time is by estimating productivity in long-differences (Bartelsman et al., 1994; Brynjolfsson and Hitt, 2003). If digital capital requires time to install, we should observe returns in longer time differences that are larger than those in shorter time differences. The results of the long-differences tests are shown in Table 7. The first column of estimates is from the main productivity regression computed in long-differences, where changes in the difference length, which range from 1 year to 10 years, are shown on the left of the table, and only the coefficient estimates on digital capital are shown in the table. Moving down the table (longer difference lengths) corresponds with estimates that grow from .005 in one-year differences to .026 in 10-year differences (t=2.36), which is consistent with the argument that digital capital requires time to be adapted to the organization, and that its productivity effects appear only after a delay. The second column in the table adds two-digit NAICS industry into the regression, but the pattern of correlations moving down the table is very similar to that in the previous column. In sum, the estimates in Tables 6 and 7 are consistent with the argument that the productivity of digital capital assets takes several years to fully appear.

5 Conclusion

This paper uses new firm-level data on IT investment to develop panel measures of digital capital prices and quantities. In particular, building upon earlier work by Hall and others, we compute firm-level measures of intangible IT capital quantities that allow us to: 1) generate estimates of the annual growth of this asset, 2) to compare how these growth rates differ among firms of different value, and 3) analyze how the accumulation of these assets contributes to productivity differences among firms.

We find that the stock of digital capital has grown rapidly and by 2016 accounted for about 25% of total capital stock for firms in our sample. Changes in the value of digital capital in the years before and after the dot-com boom and bust appear to primarily be due to price fluctuations. However, since 2001, firms continued to accumulate significant amounts of digital capital while prices varied little. The most recent technology-related increases in the market value of firms appear to be due to changes in quantity, not price, as firms accumulate more and more digital capital.

We find striking firm-to-firm heterogeneity in digital capital value, with most of the value concentrated in a small group of superstar firms with market values in the top decile. Inequality in digital capital among firms is growing as the top firms pull further away from the rest. Moreover, per-capita digital capital stocks are substantially greater in firms with more educated workers. These findings are consistent with the emphasis that technology-intensive firms place on making investments in training and skills.

Our findings suggest that the higher values the financial markets have assigned to firms with large digital investments in recent years reflect greater digital capital quantities, rather than simply higher prices for existing assets. In other words, they reflect genuine improvements to firms' productive capacity. In fact, we find that digital capital, if included as a separate factor in firm-level production functions, predicts differences in output and productivity among firms. The firms that are pulling away in terms of digital capital also have an increasing advantage in productivity.

Our estimates of the output elasticity of digital capital suggest that it is several times greater than the output elasticity of IT capital. Our estimates are broadly consistent with earlier evidence (Saunders and Brynjolfsson, 2016) that indicates that IT hardware accounts for only about 10% of total digital investment, with investments in complementary intangibles–that is, digital capital–accounting for the rest.

One interpretation of our findings is that translating organizational innovations into productive capital requires significant investment in organizational re-engineering and skill development. Therefore, even if firms have the appropriate absorptive capacity, knowledge of how to construct digital assets will not automatically generate productive digital capital any more than access to the blueprints of a competitor's plant will directly lead to productive capacity. The ability to convert digital potential into realized value appears to be increasingly dominated by a small set of superstar firms.

References

Abis, S. and Veldkamp, L. (2020). The changing economics of knowledge production. *Available at SSRN 3570130*.

Autor, D., Dorn, D., Katz, L. F., Patterson, C., and Van Reenen, J. (2020). The fall of the labor share and the rise of superstar firms. *The Quarterly Journal of Economics*, 135(2):645–709.

Ayyagari, M., Demirguc-Kunt, A., and Maksimovic, V. (2019). *The Rise of Star Firms: Intangible Capital and Competition*. The World Bank.

Baily, M. N. (1981). Productivity and the services of capital and labor. *Brookings Papers on Economic Activity*, 12(1):1–66.

Bartelsman, E. J., Caballero, R. J., and Lyons, R. K. (1994). Customer-and supplier-driven externalities. *The American Economic Review*, 84(4):1075–1084.

Belo, F., Lin, X., and Bazdresch, S. (2014). Labor hiring, investment, and stock return predictability in the cross section. *Journal of Political Economy*, 122(1):129–177.

Bessen, J. (2020). Information technology and industry concentration. *Journal of Law and Economics*, 63(3):531–555.

Bessen, J. E., Denk, E., Kim, J., and Righi, C. (2020). Declining industrial disruption. *Boston Univ. School of Law, Law and Economics Research Paper Forthcoming*.

Bharadwaj, A. S., Bharadwaj, S. G., and Konsynski, B. R. (1999). Information technology effects on firm performance as measured by tobin's q. *Management science*, 45(7):1008–1024.

Bischoff, C. W. and Kokkelenberg, E. C. (1987). Capacity utilization and depreciation-in-use. *Applied Economics*, 19(8):995–1007.

Black, S. E. and Lynch, L. M. (2001). How to compete: the impact of workplace practices and information technology on productivity. *Review of Economics and statistics*, 83(3):434–445.

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).

Brynjolfsson, E. and Hitt, L. M. (2000). Beyond computation: Information technology, organizational transformation and business performance. *Journal of Economic perspectives*, 14(4):23–48.

Brynjolfsson, E. and Hitt, L. M. (2003). Computing productivity: Firm-level evidence. *Review of Economics and 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*, 2002(1):137–181.

Brynjolfsson, E., Rock, D., and Syverson, C. (2018). The productivity J-curve: How intangibles complement general purpose technologies.

Corrado, C., Haskel, J., Jona-Lasinio, C., and Iommi, M. (2012). Intangible capital and growth in advanced economies: Measurement methods and comparative results. Technical report, IZA Discussion Papers.

Corrado, C., Hulten, C., and Sichel, D. (2009a). *1. Measuring Capital and Technology: An Expanded Framework*. University of Chicago Press.

Corrado, C., Hulten, C., and Sichel, D. (2009b). Intangible capital and US economic growth. *Review of income and wealth*, 55(3):661–685.

Covarrubias, M., Gutiérrez, G., and Philippon, T. (2019). Explaining the rising concentration of us industries: Superstars, intangibles, globalization or barriers to entry? In *NBER Macroeconomics Annual 2019, volume 34*. University of Chicago Press.

Crouzet, N. and Eberly, J. (2018). Understanding weak capital investment: The role of market concentration and intangibles. In *prepared for the Jackson Hole Economic Policy Symposium*.

Cummins, J. G. (2005). A new approach to the valuation of intangible capital. In *Measuring capital in the new economy*, pages 47–72. University of Chicago Press.

De Loecker, J., Eeckhout, J., and Unger, G. (2020). The rise of market power and the macroeconomic implications. *The Quarterly Journal of Economics*, 135(2):561–644.

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

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

Eisfeldt, A. L. and Papanikolaou, D. (2013). Organization capital and the cross-section of expected returns. *The Journal of Finance*, 68(4):1365–1406.

Epstein, L. and Denny, M. (1980). Endogenous capital utilization in a short-run production model: Theory and an empiral application. *Journal of Econometrics*, 12(2):189–207.

Farboodi, M., Mihet, R., Philippon, T., and Veldkamp, L. (2019). Big data and firm dynamics. In *AEA papers and proceedings*, volume 109, pages 38–42.

Greenstein, S. and Nagle, F. (2014). Digital dark matter and the economic contribution of apache. *Research Policy*, 43(4):623–631.

Hall, B. H. (1993). The stock market's valuation of R&D investment during the 1980's. *The American Economic Review*, 83(2):259–264.

Hall, R. E. (2000). E-capital: The link between the stock market and the labor market in the 1990s. *Brookings Papers on Economic Activity*, 2000(2):73–102.

Hall, R. E. (2001). The stock market and capital accumulation. *American Economic Review*, 91(5):1185–1202.

Hall, R. E. (2018). New evidence on the markup of prices over marginal costs and the role of mega-firms in the us economy. Technical report, National Bureau of Economic Research.

Hamermesh, D. S. and Pfann, G. A. (1996). Adjustment costs in factor demand. *Journal of Economic literature*, 34(3):1264–1292.

Haskel, J. and Westlake, S. (2018). *Capitalism without capital: the rise of intangible economy*. Oxford.

Hayashi, F. (1982). Tobin's marginal q and average q: a neoclassical interpretation. *Econometrica*, (50):213–224.

Holt, C. C., Modigliani, F., Muth, J., and Simon, H. (1960). Planning production, inventories, and work force.

Hubmer, J. and Restrepo, P. (2021). Not a typical firm: The joint dynamics of firms, labor shares, and capital–labor substitution. Technical report, National Bureau of Economic Research.

Kogan, L. and Papanikolaou, D. (2019). Technological innovation, intangible capital, and asset prices. *Annual Review of Financial Economics*, 11:221–242.

Lev, B. (2000). *Intangibles: Management, measurement, and reporting*. Brookings institution press.

Lichtenberg, F. R. (1995). The output contributions of computer equipment and personnel: A firm-level analysis. *Economics of innovation and new technology*, 3(3-4):201–218.

McGrattan, E. R. (2020). Intangible capital and measured productivity. *Review of Economic Dynamics*.

McGrattan, E. R. and Prescott, E. C. (2010). Technology capital and the us current account. *American Economic Review*, 100(4):1493–1522.

Milgrom, P. and Roberts, J. (1990). The economics of modern manufacturing: Technology, strategy, and organization. *The American Economic Review*, pages 511–528.

Nadiri, M. I. and Prucha, I. R. (1996). Estimation of the depreciation rate of physical and R&D capital in the us total manufacturing sector. *Economic Inquiry*, 34(1):43–56.

Nagle, F. (2019). Open source software and firm productivity. *Management Science*, 65(3):1191–1215.

Pakes, A. and Schankerman, M. (1984). The rate of obsolescence of patents, research gestation lags, and the private rate of return to research resources. In *R&D, patents, and productivity*, pages 73–88. University of Chicago Press.

Rock, D. (2019). Engineering value: The returns to technological talent and investments in artificial intelligence. *Unpublished Working Paper*.

Saunders, A. and Brynjolfsson, E. (2016). Valuing information technology related intangible assets. *Mis Quarterly*, 40(1).

Sculley, D., Holt, G., Golovin, D., Davydov, E., Phillips, T., Ebner, D., Chaudhary, V., and Young, M. (2014). Machine learning: The high interest credit card of technical debt.

Solow, R. M. et al. (1960). Investment and technical progress. *Mathematical methods in the social sciences*, 1:48–93.

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

Tambe, P. and Hitt, L. M. (2012a). Now IT's personal: Offshoring and the shifting skill composition of the US information technology workforce. *Management Science*, 58(4):678–695.

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

Van Reenen, J. (2018). Increasing differences between firms: market power and the macro-economy.

Wu, L., Hitt, L., and Lou, B. (2020). Data analytics, innovation, and firm productivity. *Management Science*, 66(5):2017–2039.

Figure 1: Market capitalization of top 5 firms in the S&P 500

Figure Notes: This chart illustrates changes in the market capitalization of the top 5 firms in the S&P 500 as a percent of the total market value of the S&P 500. The chart was created by Goldman Sachs Global Investment Research and was reproduced from https://markets.businessinsider.com/news/stocks/sp500-concentration-large-cap-bad-sign-future-returns-effect-market-2020-4-1029133505, accessed on May 14, 2020.

Figure 2: Price and quantity of capital

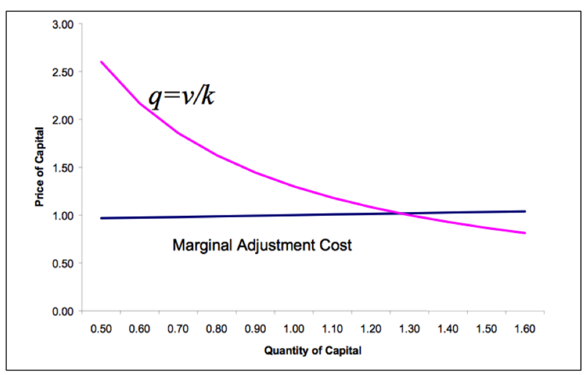


Figure Notes: This chart is reproduced from (Hall, 2001). It graphs the solution to Hall's Quantity Revelation Theorem, which solves for the equilibrium price and quantity of capital.

Figure 3: Year-to-year measures of firms' IT employment

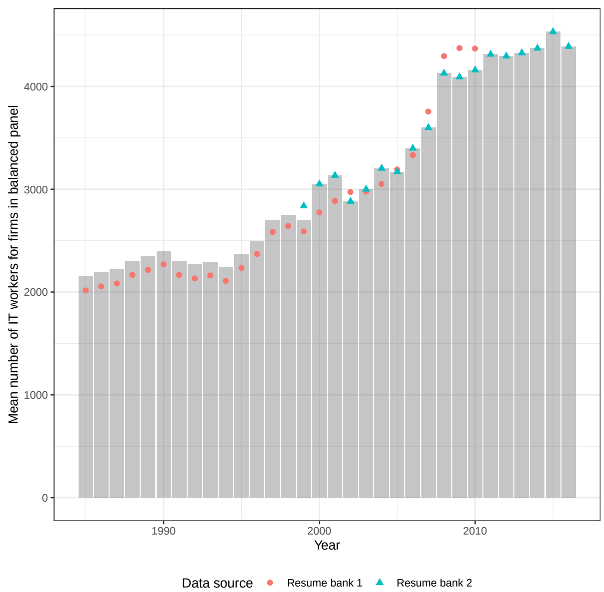
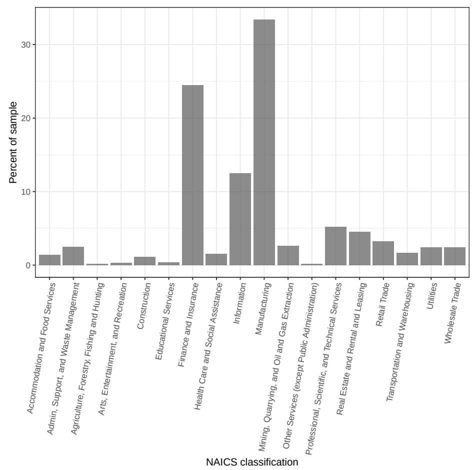
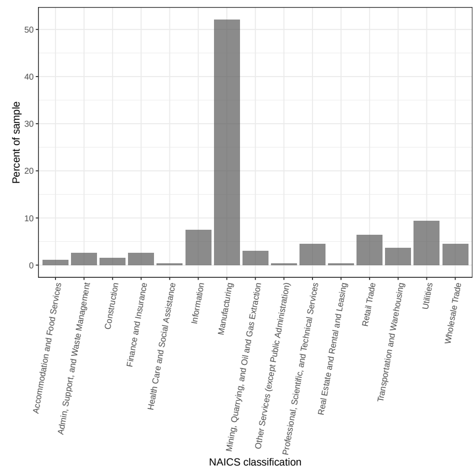


Figure Notes: This chart illustrates changes in measures of IT employment using data from the two different resumé banks where normalized quantities from the two different resumé banks are denoted with the two sequences of plotted points. The height of the gray bar indicates a synthetic measure that combines the two different data sources into a single IT employment sequence.

Figure 4: NAICS distributions of balanced and unbalanced panels



(a) NAICS industries, unbalanced panel



(b) NAICS industries, balanced panel

Figure Notes: These charts illustrate the distributions of firms in the 2005 cross sections of the unbalanced and balanced panels over major NAICS industry codes (2 digit level). There are a total of 5,215 firms in the unbalanced panel cross section and 277 firms in the balanced panel cross section.

Table 1: Summary statistics for firm-level measures

Statistic	Mean	St. Dev.	Min	Max	N
Market value	35,042.34	64,248.57	143.06	471,424.40	264
PPE	6,158.58	12,808.84	3.74	107,010.00	264
Employment	56.74	137.39	0.32	1,800.00	231
Other assets	14,857.04	38,089.26	55.66	397,677.00	264
IT employment	3,197.72	7,631.98	12.71	93,652.23	264
Value added	4,582.25	7,947.51	21.88	58,131.00	231
Capital	10,262.37	22,297.88	11.89	223,252.00	231
Labor	2,396.43	4,323.38	23.99	39,186.00	231
Materials	12,408.07	28,907.59	40.94	244,953.70	231
Log wage bill	18.03	1.49	13.66	22.54	264

Table Notes: This table reports the summary values of key statistics for firms from the 2005 sample year of the panel. Values are collected from the Capital IQ database, except IT employment measures, which are generated using the LinkedIn database, as described in the paper and Data appendix. Wage bill measures are computed using BLS occupational wages and the occupational mix of the firm as computed from the LinkedIn database. Employment is measured in thousands, and IT employment figures are not scaled. All other variables are measured in millions.

Table 2: Regressions of assets on market value using the unbalanced panel

	DV: Market Value						
	OLS	OLS	OLS	OLS	LAD	FE	OLS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
PPE	1.667***	1.833***	1.606***	1.527***	1.428***	1.303***	1.531***
	(0.133)	(0.157)	(0.141)	(0.119)	(0.018)	(0.134)	(0.126)
Other assets	1.084***	1.057***	1.049***	1.004***	1.016***	0.998***	1.004***
	(0.128)	(0.018)	(0.131)	(0.006)	(0.002)	(0.006)	(0.006)
IT capital	12.226		−1.851				
	(8.040)		(10.816)				
IT labor		8.623*	17.702*	14.116***	11.975***	9.378***	
		(4.789)	(9.694)	(4.102)	(0.927)	(3.408)	
IT wage bill							12.103**
							(5.066)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	No	No	No	No	No	Yes	No
Industry FE	Yes	Yes	Yes	Yes	Yes	No	Yes
Observations	2,957	16,722	2,922	84,045	84,045	84,045	65,238
R^2	0.736	0.939	0.739	0.990		0.995	0.990

Table Notes: This table reports results from the regression $MV_{it} = \beta_{PPE}PPE_{it} + \beta_{OA}OASSET_{it} + \beta_{IT}IT_{it} + \gamma_{it} + \epsilon_{it}$. MV is market value, PPE is property, plant, and equipment, IT is IT capital, $OASSET$ is all other assets, i indexes the firm, t indexes the year, and γ is a vector of fixed effects that can include year, industry, or firm as indicated in the table. It uses data from the unbalanced panel. Column (1) is an OLS regression using the IT capital measures with the sample restricted to the years 1987-1998. Column (2) is an OLS regression using the IT employment measures with the sample restricted to the years 1987-1998. Column (3) is an OLS regression using the IT capital and IT employment measures with the full sample for the years 1987-1998 for which both measures are available. Column (4) uses the IT employment measures for the full sample of years. Column (5) is an LAD regression on the same sample. Column (6) adds firm fixed effects to the specification used in column (4). Column (7) uses firm-level IT occupational mix from LinkedIn along with BLS salary data to translate IT labor to a wage bill. Where included, industry fixed-effects are included at the two-digit NAICS level. Standard errors are clustered on firm and shown in parentheses, with *, **, and *** denoting significance at the 10%, 5%, and 1% level, respectively.

Table 3: Regressions of assets on market value using the balanced panel

	DV: Market Value					
	OLS	OLS	OLS	OLS	FE	OLS
	(1)	(2)	(3)	(4)	(5)	(6)
PPE	1.651***	1.453***	1.630***	1.495***	1.272***	1.529***
	(0.128)	(0.152)	(0.126)	(0.232)	(0.207)	(0.238)
Other assets	0.990***	1.120***	0.952***	1.258***	1.328***	1.269***
	(0.174)	(0.173)	(0.177)	(0.189)	(0.254)	(0.189)
IT capital	8.580		−0.265			
	(8.825)		(10.409)			
IT labor		3.533**	3.170	9.099***	10.381***	
		(1.489)	(3.139)	(2.357)	(3.736)	
IT wage bill						12.215***
						(3.397)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	No	No	No	No	Yes	No
Industry FE	Yes	Yes	Yes	Yes	No	Yes
Observations	1,604	3,017	1,603	8,521	8,521	5,540
R^2	0.724	0.760	0.725	0.814	0.902	0.808

Table notes: This table reports results from the regression $MV_{it} = \beta_{PPE}PPE_{it} + \beta_{OA}OASSET_{it} + \beta_{IT}IT_{it} + \epsilon_{it}$. MV is market value, PPE is property, plant, and equipment, IT is IT capital, and $OASSET$ is all other assets. It uses data from the balanced panel. Column (1) is an OLS regression using the IT capital measures with the sample restricted to the years 1987-1998. Column (2) is an OLS regression using the IT employment measures with the sample restricted to the years 1987-1998. Column (3) is an OLS regression using the IT capital and IT employment measures with the full sample for the years 1987-1998 for which both measures are available. Column (4) uses the IT employment measures for the full sample of years. Column (5) adds firm fixed effects to the specification used in column (4). Column (6) uses firm-level IT occupational mix from LinkedIn along with BLS salary data to translate IT labor to a wage bill. Industry fixed-effects are included at the two-digit NAICS level. Standard errors are clustered on firm and shown in parentheses, with *, **, and *** denoting significance at the 10%, 5%, and 1% level, respectively.

Figure 5: IT coefficients from year-to-year market value regressions

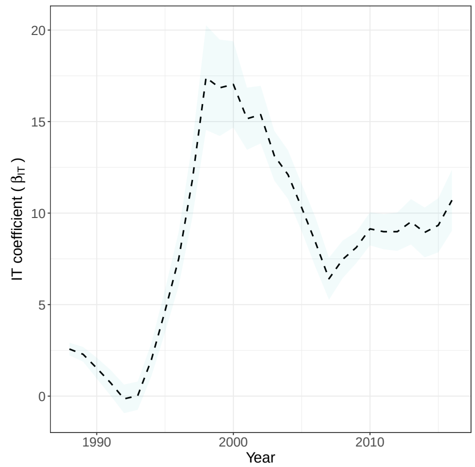


Figure Notes: This chart plots the estimated IT coefficient (β_{IT}) from the regression $MV_{it} = \beta_{PPE}PPE_{it} + \beta_{OA}OA_{it} + \beta_{IT}IT_{it} + \gamma_{it} + \epsilon_{it}$ using a rolling window around each focal year [-1, +1], and where i indexes the firm and t indexes the year. γ is a vector of fixed effects that include industry and year. Standard errors are clustered at the firm level. The shaded region represents the standard error band.

Figure 6: Digital capital market value

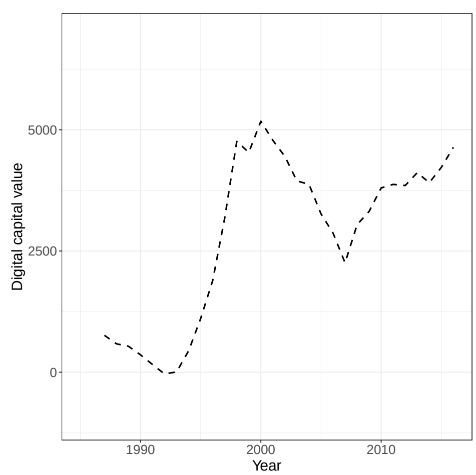


Figure Notes: This chart illustrates average values of digital capital for the firms in our balanced panel. The market value of digital capital is computed in two steps. First, we estimate the regression $MV_{it} = \beta_{PPE}PPE_{it} + \beta_{OA}OA_{it} + \beta_{IT}IT_{it} + \epsilon_{it}$ using a rolling window around each focal year. Then, the coefficient β_{IT} for each firm is multiplied by the firm's IT employment to compute the component of its market value that is correlated with the firm's IT assets.

Figure 7: Digital capital prices

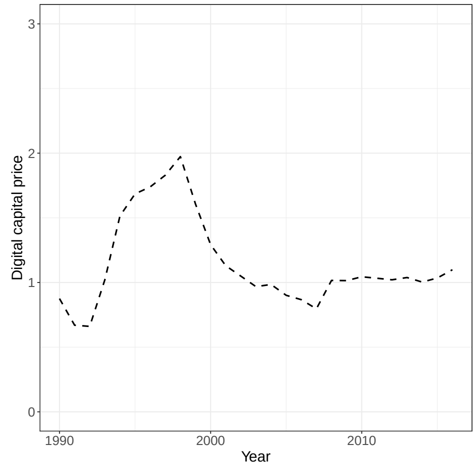


Figure Notes: This chart illustrates average prices for digital capital for firms in the balanced panel. The price of digital capital for each firm-year is computed by solving equations 1 and 2 in the main text, and the index in the chart above is computed by averaging across firms in each year.

Figure 8: Digital capital quantities

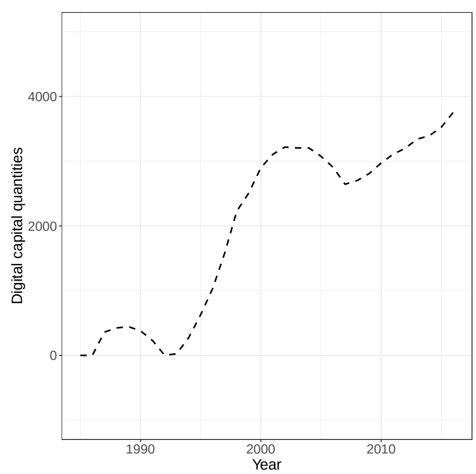


Figure Notes: This chart illustrates average quantities of digital capital for firms in the balanced panel. The average quantity of digital capital for each firm-year is computed by solving equations 1 and 2 in the main text, and the index in the chart above is computed by averaging across firms in each year.

Figure 9: Comparison of changes in PPE and digital capital quantities

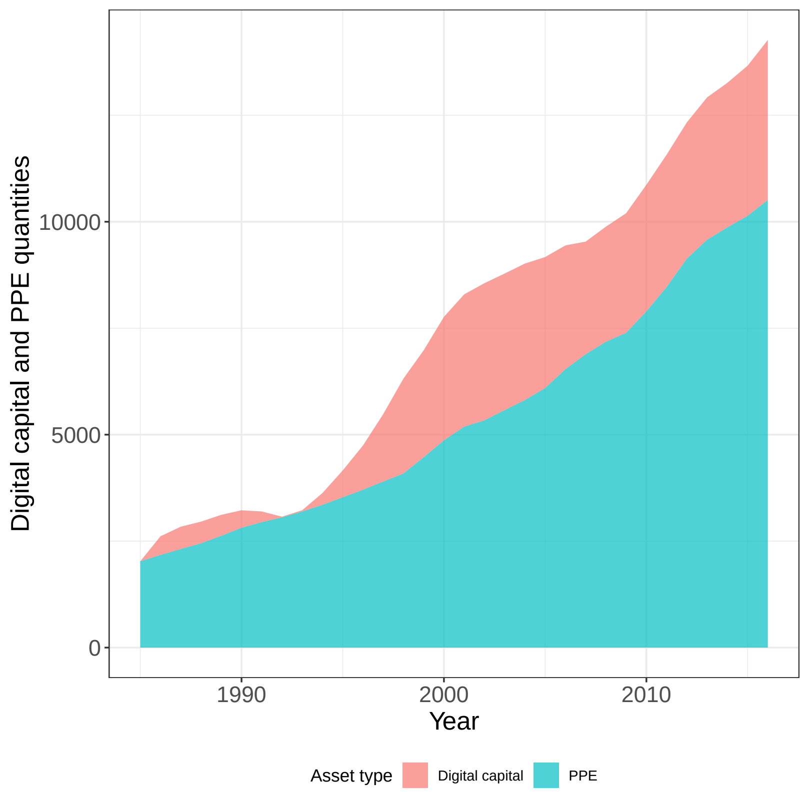
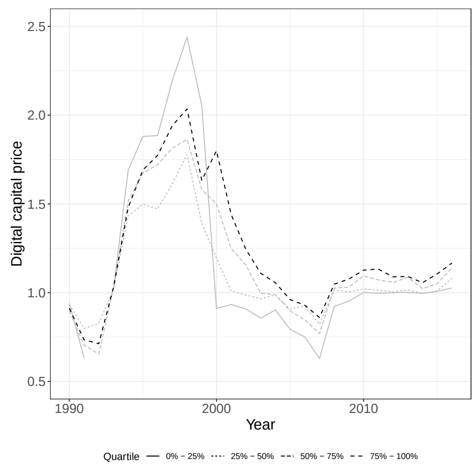
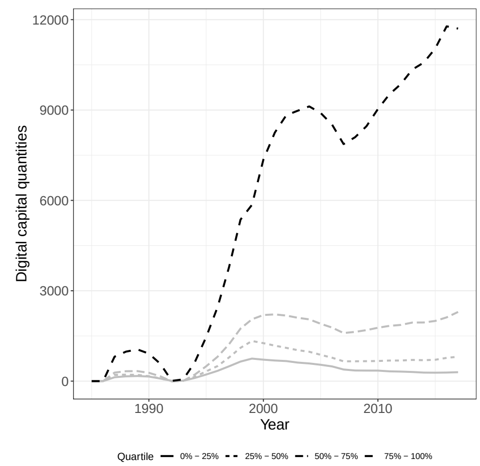


Figure Notes: This chart compares year-on-year accumulation of quantities of digital capital with that of Property, Plant, and Equipment (PPE) for the balanced panel. The average quantity of digital capital for all firms in a year is computed by solving equations 1 and 2 in the main text. Average PPE quantities for firms are computed from data from Capital IQ.

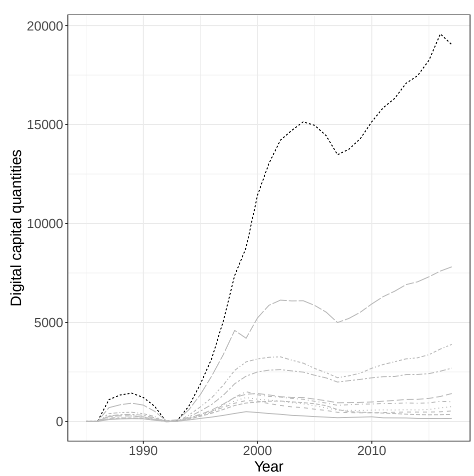
Figure 10: Digital capital prices and quantities computed by market value quantiles



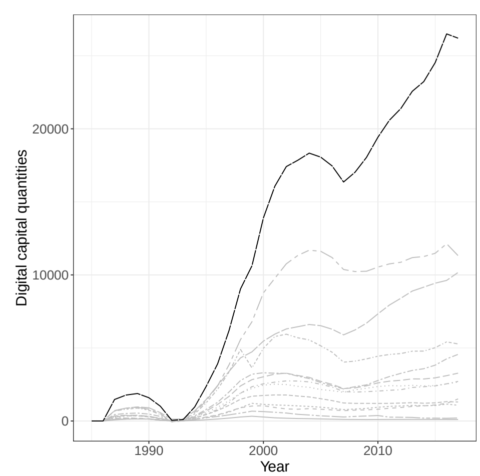
(a) Digital capital prices by market value



(b) Digital capital quantities by market value (25%)



(c) Digital capital quantities by market value (10%)



(d) Digital capital quantities by market value (5%)

Figure Notes: These charts illustrate prices and quantities of digital capital where firms are separated into quantiles according to their market values at the end of the panel. Annual quantities are computed by solving equations 1 and 2. Subfigure (a) shows prices by quartile and (b) shows quantities by quartile. Subfigure (c) shows quantities with firms separated into ten quantiles. Subfigure (d) shows quantities with firms separated into twenty quantiles. The darker line in each chart is the top quantile of firms in terms of market value.

Figure 11: Digital capital quantities for the unbalanced panel

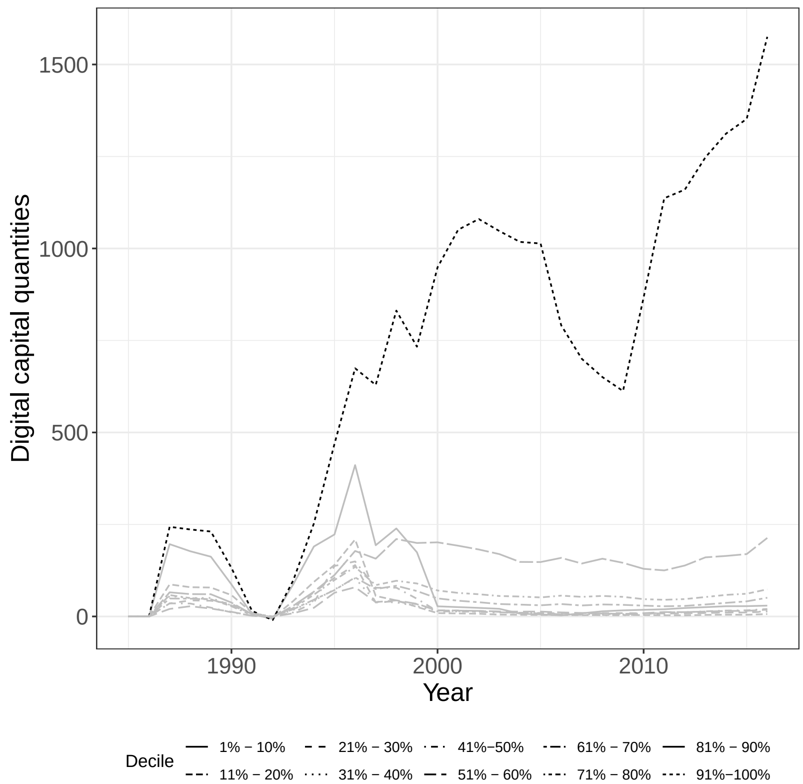
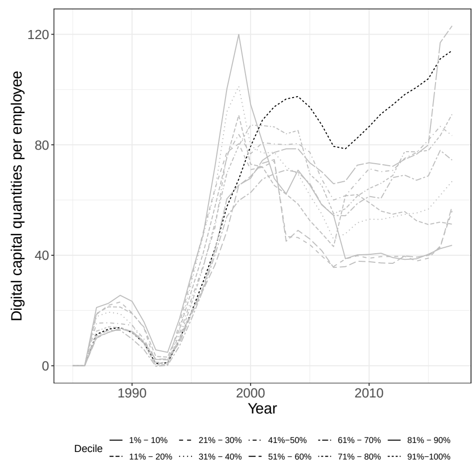
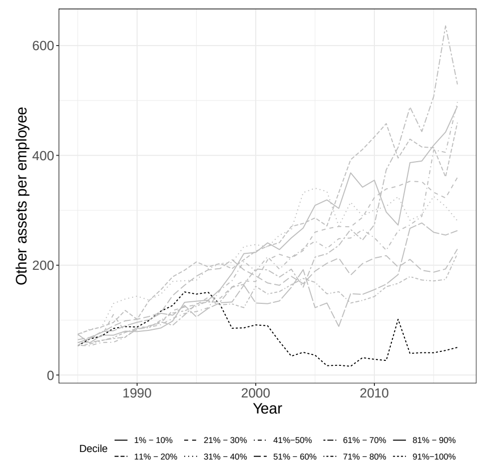


Figure Notes: This chart illustrates average quantities of digital capital for firms by market value decile in the unbalanced panel. The average quantity of digital capital for each firm-year is computed by solving equations 1 and 2 in the main text and then averaged across firms in each year. The darker line is the top decile group in terms of market value.

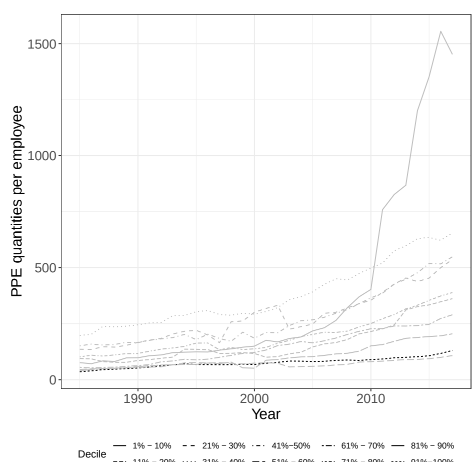
Figure 12: Per capita values of factors separated by market value deciles



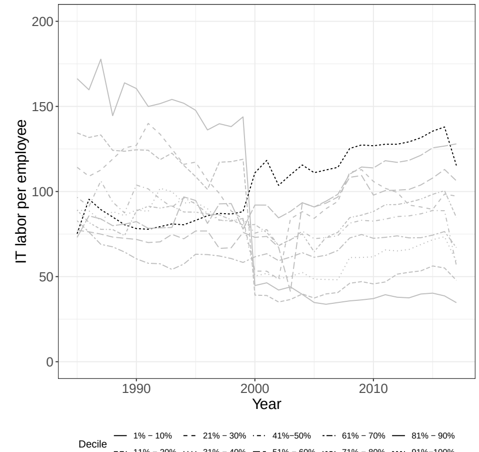
(a) Digital capital



(b) Other assets



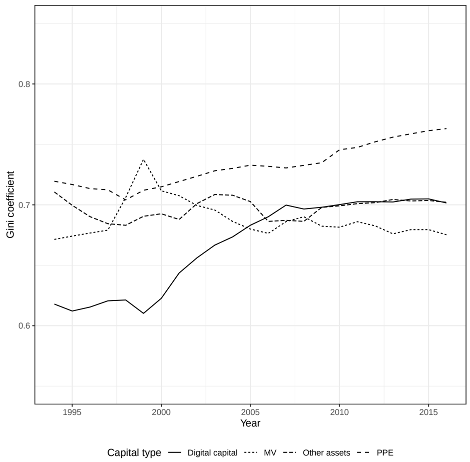
(c) PPE



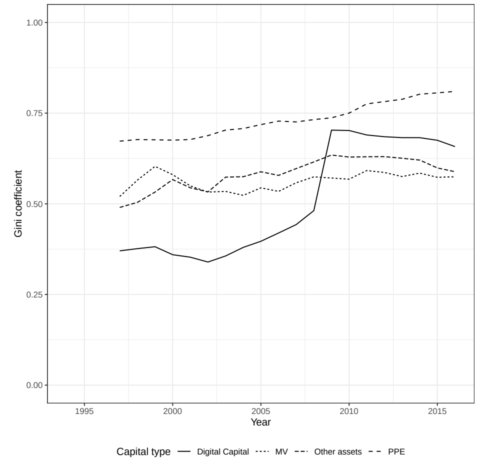
(d) IT labor

Figure Notes: These charts illustrate per capita quantities of different assets where firms are separated into deciles according to their 2016 market values. Digital capital quantities for each year are computed by solving equations 1 and 2. Subfigure (a) illustrates digital capital quantities, subfigure (b) shows other assets (total assets minus physical capital, including receivables, inventories, cash, and other accounting assets), subfigure (c) shows Property, Plant, and Equipment, and subfigure (d) shows IT labor. The darker line in each chart is the top quantile of firms in terms of market value.

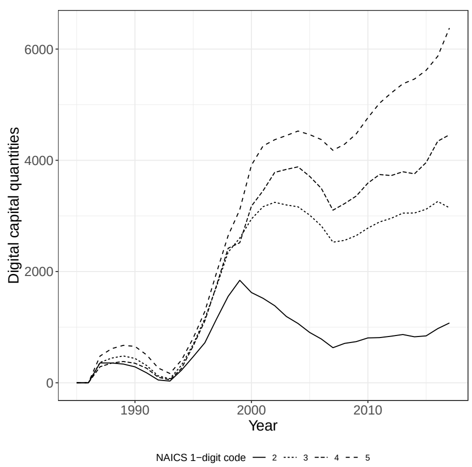
Figure 13: Digital capital concentration by capital type and industry



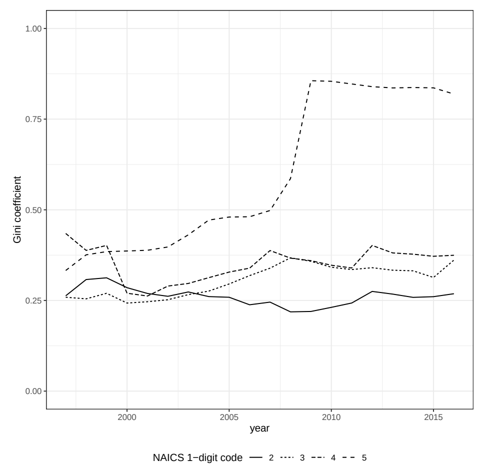
(a) Digital capital concentration



(b) Digital capital per capita concentration



(c) Digital capital quantities by industry



(d) Digital capital concentration by industry

Figure Notes: These charts illustrate changes in digital capital concentration over time. Subfigure (a) shows digital capital concentration and (b) shows concentration in digital capital per capita. Subfigure (c) charts digital capital by industry, where firms are placed in industries at the 1 digit NAICS level, and only quantities from the four industries with the largest number of firms in the sample are shown. NAICS industries in category "2" include Mining, Quarrying, Oil and Gas Extraction, Utilities, and Construction, those in "3" include Manufacturing, those in "4" include Wholesale Trade, Retail Trade, and Transportation and Warehousing, and those in "5" include Information, Finance and Insurance, Real Estate and Rental and Leasing, Professional, Scientific, and Technical Services, Management of Companies and Enterprises, and Administrative and Support and Waste Management and Remediation Services. Subfigure (d) shows capital concentration by the same industry categories used in (c).

Table 4: Regressions of digital capital on productivity

	DV: Log(Value Added)						
	OLS	OLS	OLS	OLS	OLS	FE	FE
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log(Non-IT capital)	0.259***	0.250***	0.259***	0.290***	0.242***	0.402***	0.394***
	(0.027)	(0.025)	(0.035)	(0.045)	(0.025)	(0.113)	(0.109)
Log(Labor)	0.672***	0.662***	0.587***	0.592***	0.646***	0.435***	0.429***
	(0.044)	(0.047)	(0.097)	(0.112)	(0.048)	(0.139)	(0.138)
Log(IT capital)	0.037**	0.031**	0.031	0.018	0.015	0.026**	0.027**
	(0.017)	(0.015)	(0.026)	(0.023)	(0.014)	(0.013)	(0.013)
Log(DC)		0.034**	0.093*	0.077	0.022		0.023**
		(0.016)	(0.049)	(0.047)	(0.016)		(0.012)
Log(R&D)					0.064***		
					(0.017)		
R&D reported?					−0.346***		
					(0.101)		
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	No	No
Firm FE	No	No	No	No	No	Yes	Yes
Observations	1,827	1,827	610	610	1,827	1,827	1,827
R^2	0.951	0.952	0.926	0.932	0.954	0.985	0.985

Table Notes: This table reports results from the regression $Log(VA)_{it} = \beta_K Log(K)_{it} + \beta_L Log(L)_{it} + \beta_{IT} Log(IT)_{it} + \beta_{DC} Log(DC)_{it} + \epsilon_{it}$. VA is value added (output minus materials), K is capital, L is labor, IT is IT capital, and DC is digital capital. Column (1) is the baseline regression on the full set of observations for which the CITDB IT capital data are available on a consistent basis (through 2000). Column (2) includes the computed digital capital stock measures. Column (3) restricts the sample to the years 1997 to 2000 and column (4) uses finer industry controls (two-digit instead of one-digit NAICS industry controls). Column (5) includes measures of R&D as well as measures of whether or not R&D is reported. Columns (6) and (7) include firm fixed-effects. Standard errors are shown in parentheses and are clustered on firm with *, **, and *** denoting significance at the 10%, 5%, and 1% level, respectively.

Table 5: Relationship between IT, human capital, and digital capital

	Log(DC/e)			
	OLS	OLS	FE	FE
	(1)	(2)	(3)	(4)
Log(K/e)	0.200***	0.187***	0.255	0.259
	(0.034)	(0.034)	(0.171)	(0.172)
Log(IT/e)	0.703***	0.350***	0.519***	0.194
	(0.035)	(0.118)	(0.050)	(0.118)
Log(HK/e)	−0.103	−0.236***	−0.399***	−0.397***
	(0.080)	(0.058)	(0.082)	(0.076)
Log(IT/e)×Log(HK/e)		0.048***		0.048***
		(0.015)		(0.016)
Year FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	No	No
Firm FE	No	No	Yes	Yes
Observations	1,805	1,805	1,805	1,805
R^2	0.656	0.661	0.852	0.853

Table notes: This table reports results from the regression $Log(DC/e)_{it} = \beta_k Log(K/e)_{it} + \beta_{hk} Log(HK/e)_{it} + \beta_{IT_e} Log(IT/e)_{it} + \beta_{ITxHK}(Log(HK/e)_{it} \times Log(IT/e)_{it}) + \epsilon_{it}$. All columns test the per capita relationships between HK and IT investment and digital capital quantities and include year fixed-effects. Columns (1) and (2) also include two-digit NAICS industry controls but they do not include firm-fixed effects. Columns (3) and (4) add firm fixed-effects to the regressions. Standard errors are clustered on firm and shown in parentheses with *, **, and *** denoting significance at the 10%, 5%, and 1% level, respectively.

Table 6: Productivity regressions with lagged digital capital

	DV: Log(Value Added)			
	0 years	1 year	2 years	3 years
	(1)	(2)	(3)	(4)
Log(Capital)	0.245***	0.248***	0.249***	0.243***
	(0.055)	(0.056)	(0.057)	(0.060)
Log(Labor)	0.657***	0.660***	0.661***	0.659***
	(0.076)	(0.076)	(0.076)	(0.080)
Log(DC)	0.019**	0.013**	0.010*	0.010**
	(0.009)	(0.006)	(0.005)	(0.005)
Year FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Observations	4,167	4,166	4,166	3,976
R^2	0.987	0.987	0.987	0.987

Table Notes: This table reports results from the regression $Log(VA)_{it} = \beta_K Log(K)_{it} + \beta_L Log(L)_{it} + \beta_{DC} Log(DC)_{i,t-n} + \epsilon_{it}$ where n varies between 0 and 3 years. VA is value added (output minus materials), K is capital, L is labor, and DC is digital capital lagged by the number of years specified in the column header. The first column is a regression of value added on the production factors in that year, the second column is a regression of value added on contemporaneous capital and labor and the digital capital stocks in the prior year, the third column is a regression of value added on contemporaneous capital and labor and digital capital stocks from two years prior, and the last column is a regression of value added on contemporaneous capital and labor and the digital capital stocks from three years prior. All four columns include fixed effects for firm and year. Standard errors are shown in parentheses and are clustered on firm with *, **, and *** denoting significance at the 10%, 5%, and 1% level, respectively.

A Online Appendix: Data for Digital Capital and Superstar Firms

A.1 Data Descriptions

Three of the datasets used in this paper are 1) online resume data from LinkedIn provided via the Economic Graph Research (EGR) team, 2) Standard and Poor's Compustat-Capital IQ Database (Compustat) for publicly traded corporate financial data, and 3) the Bureau of Labor Statistics Occupational Employment Survey (BLS OES). Access to LinkedIn's internal databases and taxonomies was arranged as part of an ongoing research partnership between the authors and the EGR team. Many of the taxonomies and standardization tools, in addition to the data, are proprietary to LinkedIn. We are grateful to the LinkedIn EGR team for supporting this work, providing valuable feedback, and making available the data to make our analyses possible.

Accessing Compustat is possible through Wharton Research Data Services (WRDS). WRDS is available here: https://wrds-www.wharton.upenn.edu/. To get access to WRDS data, you will need to register for an account through your institution or otherwise.[14]

The BLS OES data can be accessed here: https://www.bls.gov/oes/tables.htm. These datasets detail employment counts and salary estimates by occupation. Estimates are available for the entire U.S. economy, by industry, and by region.

A.1.1 Standard and Poor's Compustat-Capital IQ Database

Variables Used and Construction The Compustat items used in this analysis include:

- *gvkey* (firm identifier)
- *ticker symbol* (firm identifier)
- *NAICS* (NAICS industry code)
- *SIC* (SIC industry code - used for years where NAICS was not available in BLS OES)
- *fyear* (reporting year)
- *at* (total assets)
- *prcc_ c* (share price)
- *csho* (common shares outstanding)
- *che* (cash and equivalents)
- *ceq* (book value of common equity)

[14] Account registration is available at: https://wrds-web.wharton.upenn.edu/wrds/?register.

Table 7: Long-differenced estimates of digital capital on productivity

Year diffs	Year controls	Ind+Year controls	Sample size
1	0.005	0.005	3790
	(0.004)	(0.004)	
2	0.008	0.008	3543
	(0.005)	(0.005)	
3	0.013**	0.012**	3312
	(0.006)	(0.006)	
4	0.014*	0.012*	3078
	(0.007)	(0.007)	
5	0.014*	0.013	2852
	(0.008)	(0.008)	
6	0.017*	0.015*	2727
	(0.009)	(0.009)	
7	0.021*	0.019	2507
	(0.012)	(0.012)	
8	0.028**	0.026**	2298
	(0.012)	(0.012)	
9	0.024**	0.021**	2083
	(0.01)	(0.01)	
10	0.026**	0.022**	1869
	(0.011)	(0.011)	

Table Notes: This table reports estimates from long-differenced regressions of digital capital and other factors on productivity. The estimated equation is $(log(VA)_{i,t} - log(VA)_{i,t-n}) = \beta_K(log(K)_{i,t} - log(K)_{i,t-n}) + \beta_L(log(L)_{i,t} - log(L)_{i,t-n}) + \beta_{DC}(log(DC)_{i,t} - log(DC)_{i,t-n}) + \gamma_{it} + \epsilon_{it}$. VA is value added (output minus materials), K is capital, L is labor, and DC is digital capital, n is the difference length, which varies from 1 to 10 years, and γ is a vector of dummy variables. The first column indicates the difference length in years. The second column reports the coefficient estimate on the differenced log digital capital measure when the regression equation includes dummy variables for year. The third column reports results when the regression equation includes both year and two-digit NAICS industry variables. The fourth column reports how sample size changes as the difference lengths are increased. Standard errors are clustered on firm with *, **, and *** denoting significance at the 10%, 5%, and 1% level, respectively.

- *ppegt* (gross property, plant, and equipment)
- *intan* (book value of intangible assets)
- *gdwl* (goodwill)
- *emp* (employee count in thousands)

These financial reporting items are used to construct measures of market value and to inform our measures of employment in Compustat firms (including by industry and occupation).

A.1.2 Bureau of Labor Statistics Occupational Employment Survey (BLS OES)

Variables Used and Construction First we combine and standardize BLS OES data across years. The BLS OES data we use for this come from the industry-level estimates of employment and salary by 2-digit SIC code (1999 and 2000) and by 3-digit NAICS code (2001-2017) for each of the 6-digit Standard Occupational Classification System (SOC) Codes. These data constitute a panel of employment (TOT_EMP) and average annual salary (A_MEAN) by detailed occupation from 2001-2017 for the entire U.S. economy.

Not all occupations are represented in all industries. Further, the sum of employment in "detailed" categories adds up to less than 100% of the employment in a given 3-digit NAICS group. We estimate the residual percentage by proportionally allocating workers to detailed employment categories, applying the employment mixture of the 2-digit NAICS group to which the 3-digit industry belongs. For example, if 96.36% of the employment in NAICS 237 (heavy and civil engineering construction) is represented by occupations with a "detailed" designation in the BLS OES, the residual 3.64% of total employment will be allocated proportionately to detailed occupational groupings from NAICS 23 (construction). We then add these additional counts to our estimates by detailed SOC code in NAICS 237 such that the total adds to 100%. This gives us an estimate of the U.S. economy's occupational composition in all major industries.

Adjustment to match Compustat The industrial composition of the entire U.S. economy differs from that of publicly listed corporations. Estimating employment mixes in a specific company using the industrial composition data from the BLS OES is somewhat noisy. However, by applying the employment-weighted Compustat industrial composition (using EMP and NAICS/SIC) to the BLS OES, we build measures of the employment counts and average salaries of all Compustat firms *in aggregate* as a subset of the economy. Note that this assumes the international labor composition of firms in Compustat is roughly similar to their employment mix in the U.S. We have made this occupation composition constancy assumption, because detailed international data is difficult to obtain and link to our other data sources.

This aggregated Compustat employment and salary data by occupation-year serves as our main benchmark for evaluating the overall coverage and consistency of the LinkedIn profile data. The SOC Codes are crosswalked directly to a LinkedIn employment group, then we calculate total employment and weighted average salaries by LinkedIn Occupation type. These total employment and weighted average salaries by LinkedIn Occupation type represent the entirety of Compustat firms, but not the entire economy. This is how we build our measure of the expected total employment of a given LinkedIn Occupation type for a given year in all of Compustat. We compare this measure by occupation-year to counts of profiles reported on LinkedIn within an occupation-year to build a picture of LinkedIn's reach for different types of workers.

Note also that some firms do not report EMP or instead there are coding errors in the EMP field. In the case that EMP is unreported or a firm reports fewer than 100 employees, we replace EMP with a predicted value from a regression procedure described in a section below.

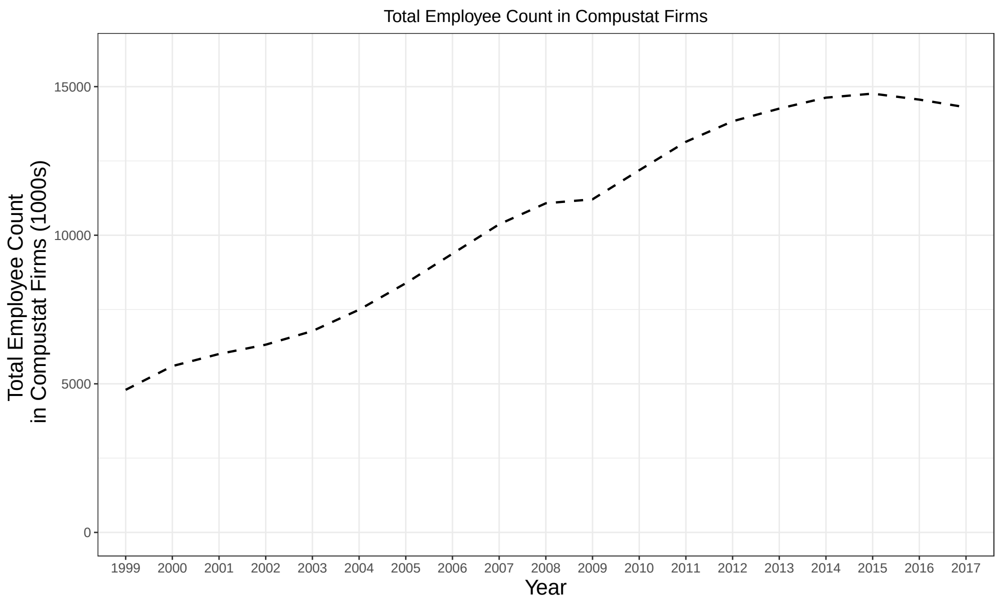
A.1.3 LinkedIn

Variables Used and Construction LinkedIn's profile database has over 180 million profile records. Records consist of a job title and employer and time of employment at a minimum, though many profiles have data on skills, educational achievements, qualifications, and other resume data. We build a panel of firm-occupation-year tuples using a set of standardized occupations and firms available via our research partners on the LinkedIn Economic Graph Research (EGR) Team.

Standardizing job titles and firms The raw input text job titles for occupations throughout the global economy vary considerably, even if the work within some groupings of titles is relatively similar. We apply LinkedIn's internal taxonomy to aggregate profile-level employment classifications into standardized groups. This taxonomy is hierarchical; some levels have as many as 20,000 types of occupations, while higher levels (e.g., organizational functional groups) are fewer in number. We map the level of the taxonomy with the greatest similarity to the 6-Digit SOC Code to a set of aggregated categories such that each SOC Code can be directly crosswalked to an aggregated category of LinkedIn occupations (many-to-one). In our final analysis, we have 136 manually constructed unique aggregated LinkedIn categories. The many-to-one construction makes adjusting for LinkedIn's dynamic coverage of the economy over time more straightforward. Prevalence of some of these aggregated LinkedIn categories is detailed below.

For standardizing firms to link Compustat-Capital IQ to LinkedIn records, we apply the internal standardization tools created by LinkedIn matching corporate records to ticker symbols. For additional matches where a standardized linkage was not present,

Figure A.1: LinkedIn Employment Records for Compustat Firms



we employed a mixture of fuzzy string matching on firm names and corporate weburl identifiers combined with a manual checking process. This combination of techniques, while subject to some matching error, enabled us to capture mergers and acquisitions and increase the relevant sample size. The number of publicly traded firms is declining over time for our sample, but LinkedIn's coverage of the overall economy is growing faster than this decline. This growth partially offsets downward trends in overall employment from the Great Recession in 2008-2009. Our match count somewhat declines from 2015 to 2017, leading to a modest drop in employee counts. These time-varying macroeconomic changes impacting all firms are one of the primary reasons we include time-period fixed effects in most specifications.

Occupation counts We have restricted our analysis to firms present in Standard and Poor's Compustat-Capital IQ database. These firms are well represented on LinkedIn, particularly in the years from 1999 to the present. While LinkedIn did not exist as early as 1999, many users have populated their profile data going back into the late 1990s. Position record counts on LinkedIn within year have climbed steadily over the sample period. Our LinkedIn sample goes until 2017.

Coverage on LinkedIn is not uniform across occupations, firms, or years. Since this variation in coverage and incentives to report employment poses an empirical challenge for our results, our analyses deploy calculated "derived" counts of workers in different occupation groupings (especially IT-related occupations). "Derived" counts refer to coverage-normalized counts given the propensity for a worker in a particular occupation group at a specific firm in a given year to report their employment on the LinkedIn platform.

Inferring firm-level employment and coverage LinkedIn's coverage of Compustat firms varies across a number of factors in ways that might affect our analysis. We therefore undertake a number of adjustments to build more accurate measures. LinkedIn position records differ markedly in their presence on the platform conditional on the time they were active, the firm or organization linked to the worker, and the worker's occupation. We build a firm-occupation-year coverage ratio to build a mapping from the observed LinkedIn profile record counts at each employer in each year of our Compustat sample. The first step is on average checking what an additional LinkedIn profile would suggest about the overall employment in a firm-year. We regress the Compustat EMP variable on LinkedIn profile records for a few different simple specifications in the table below.

Table A.1: Prediction of Compustat EMP from LinkedIn Profile Records

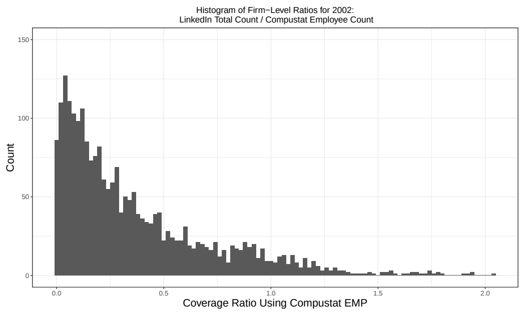
	(1)	(2)	(3)	(4)
Total Assets	0.00004	0.00004 * *	0.00004 * **	0.00003 * **
	(0.0000278)	(0.0000187)	(0.0000146)	(0.0000083)
LinkedIn Worker Count	0.00200 * **	0.00192 * **	0.00190 * **	0.00145 * **
	(0.0005717)	(0.0004699)	(0.0004729)	(0.0003746)
NAICS2-Year FE	✓			
NAICS3-Year FE		✓		
NAICS4-Year FE			✓	
Firm and Year FE				✓
R^2	.3243493	.3897643	.4217879	.9414679
N	53,699	53,607	52,767	53,657

Note: *p<0.1; **p<0.05; ***p<0.01

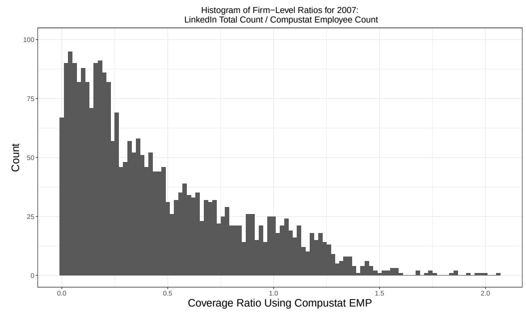
Table Notes: Results are from regressions of the EMP variable (measured in thousands) in Compustat-CapitalIQ's North America Annual Firm Financials database on the Total Assets variable and author-constructed aggregate worker counts from LinkedIn. Specifications (1)-(3) include 2, 3, and 4 digit NAICS code fixed effects by firm interacted with a set of year fixed effects. Specification (4) has firm and year fixed effects. The estimates are qualitatively similar: adjusting for total assets and these fixed effects, an additional LinkedIn profile in a Compustat firm predicts an additional 1.45 to 2.0 employees reported in Compustat-CapitalIQ (coefficients are in thousands of employees). Our adjustments in the main results use specification (3). Occasionally EMP is not well-populated in Compustat. In the case that EMP is missing or less than 100 employees, we replace EMP with the predicted EMP from this regression.

Since not all Compustat firms have EMP populated, we can fill in the missing firms' data with a predictive model. We select specification (3) from the table above because it allows for granular industry-year variation. On average, each additional LinkedIn position record in a firm over the entire sample period is correlated with another 1.9 employees recorded in the Compustat EMP field. This coverage, however, hides considerable heterogeneity across years, firms, and occupations. LinkedIn's representation

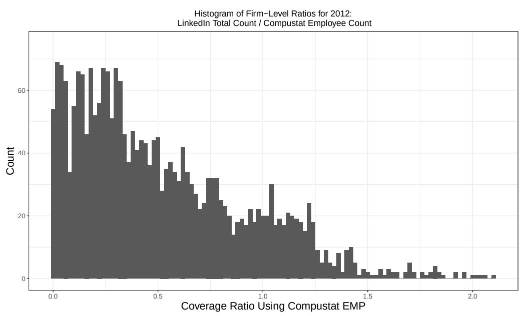
Figure A.2: Comparing LinkedIn and Compustat Employment



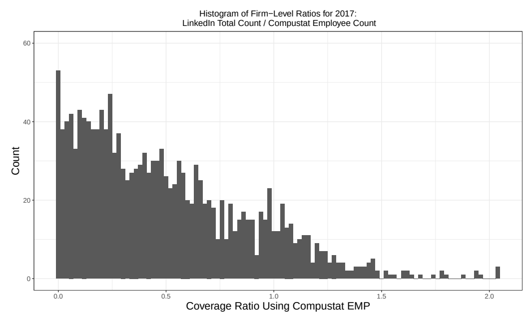
(a) 2002



(b) 2007



(c) 2012



(d) 2017

in Compustat firms improves considerably from 2000 to 2017. Below are histograms of firm-level average coverage (LinkedIn profile counts divided by Compustat EMP where populated) for 2002, 2007, 2012, and 2017. The mean coverage expands each year.

Some firms have more employees recorded in LinkedIn than reported in Compustat. Part of this reflects occupational mixes at the firm level. Many Compustat employers hire large quantities of contractors who will report working at their client firm, especially if the client firm is more prestigious than the contracting employer. On the other hand, worker composition can affect the overall coverage of LinkedIn by firm. Some occupations are less likely to populate profile information than subothers.

This mixture of issues poses a challenge that the BLS OES survey can help solve. Following the adjustment of the BLS OES occupation composition by industry to match Compustat's industry-level composition, we know the approximate composition of workers by occupation-year in aggregate across Compustat (maintaining the assumptions described above as well). With the Compustat EMP variable and regression prediction of firm-level employment for missing EMP entries, we then have the approximate total employment count in Compustat firms. Applying the BLS OES-Compustat normalized composition by occupation to the total employment count by year, we derive an approximate measure of the count of each workers employed in each 6-digit SOC Code across all of Compustat. We crosswalk these 6-digit SOC employment counts to the LinkedIn internal taxonomy categories of occupations. Comparing the count on LinkedIn within a category-year to the overall computed BLS-OES-Compustat count yields an occupation-year coverage rate.

Let this occupation-year coverage rate be λ_{jt} for occupation category j in year t. The firm-year coverage rate is calculated simply as the total profile record count within a year on LinkedIn divided by Compustat EMP or, in the case that there are fewer than 100 employees or data missing for EMP, the predicted value from specification (3) above. This estimate of EMP for firm i in year t is $FirmEmp_{it}$. Let the firm-year coverage rate in firm i for year t be θ_{it}. To get the LinkedIn occupation category coverage rate in firm i, occupation j, and year t, we multiply the firm-level coverage by the occupation-level coverage and divide out the common factor such that the sum over all derived occupation counts matches $FirmEmp_{it}$. Specifically,

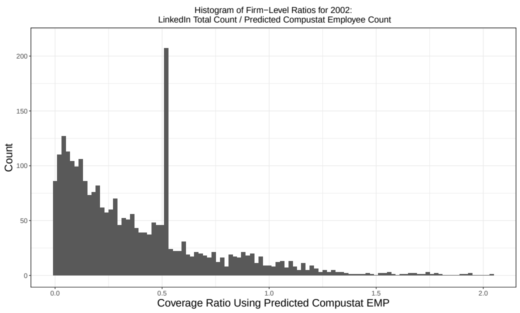
$$\frac{LI_{ijt}}{FirmEmp_{it}} = \frac{\theta_{it}\lambda_{jt}}{\overline{\theta_{jt}}} \tag{9}$$

where $\overline{\theta_{jt}}$ is the employment weighted-average firm-year coverage rate. The left-hand side of the equation is the proportion of employment in occupation j for a given firm-year, as the numerator is the count of profile records in occupation j for a given firm-year. Summing across all occupations yields 100%, and summing across all occupation employment in that firm-year yields the firm's approximate employment. The coverage ratio is represented on the right-hand side, and is our means of calculating the derived occupation counts in the numerator on the left-hand side that serve as our primary occupation count measure. A similar procedure is followed in Rock (2019).

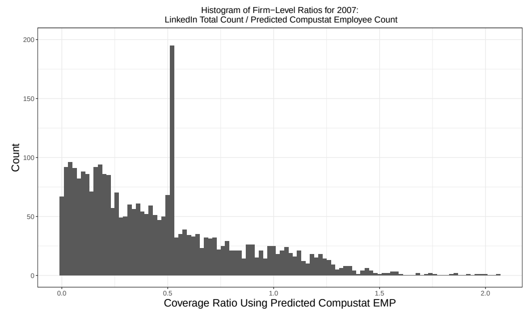
The EMP prediction measure mostly preserves the firm coverage histograms above, but for the firms wherein EMP is replaced by predicted EMP, the coverage rates are shifted toward the model prediction as shown in the figures below. This has little effect on the analysis, but does allow for the expansion of the sample size. The spike of additional firms near 50% coverage reflects the model coefficient in our regression specification.

Inferring occupation-level employment and coverage Our focus for the majority of this work is the IT sector, where the occupation-level coverage rates tend to be close to the calculated total amounts across Compustat. LinkedIn reporting rates for knowledge workers are favorably high. We use the LinkedIn taxonomy's hierarchical categorization of occupations belonging to Information Technology functions at firms. The following occupations constitute the set of worker types categorized as IT workers for our analysis:

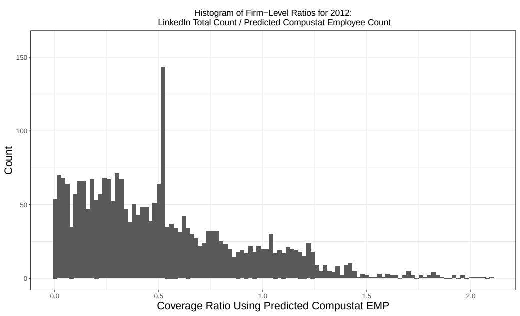
Figure A.3: Comparing LinkedIn Counts With Predicted Compustat Employment



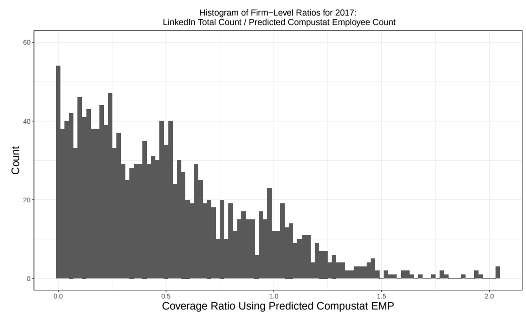
(a) 2002



(b) 2007



(c) 2012



(d) 2017

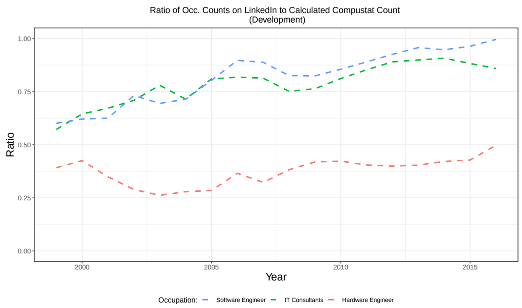
Actuaries, Audio-Visual Specialists, Business Analysis Professionals, Computer-Aided Designers, Control Systems Engineers, Data Analysts (includes Data Scientists), Data Entry Clerks, Database and Network Engineers, Economists, Hardware Engineers, IT Consultants, IT Security and Audit Specialists, IT Support Specialists, IT System Administrators, Market Research Professionals, Multimedia Specialists, Quality Assurance Testers, Research Fellows, Software Engineers, Surveyors and GIS Specialists, Telecommunications Specialists, and Web Designers.

Some actuaries, for example, might not belong to the IT function. But we include the actuaries who are designated as IT workers in our sample. For a few occupations, we also include workers outside the IT function because they are likely IT workers in other departments. Those occupation groups are:

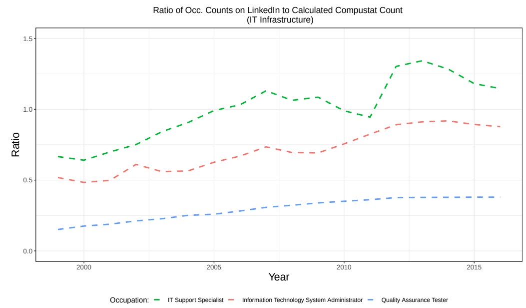
Computer Aided Designers, Data Analyst, IT Consultants, IT Security and Audit Specialist, IT Support Specialists, IT System Administrator, Software Engineers, and Web Designers

Coverage rates for selected occupations in aggregate across Compustat are reported in the plots below. These coverage rates are calculated using derived count for the denominator according to the procedure above.

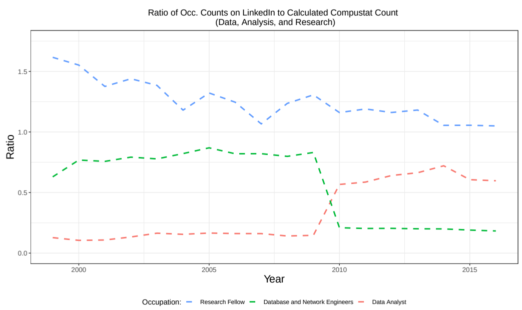
Figure A.4: LinkedIn Occupational Coverage



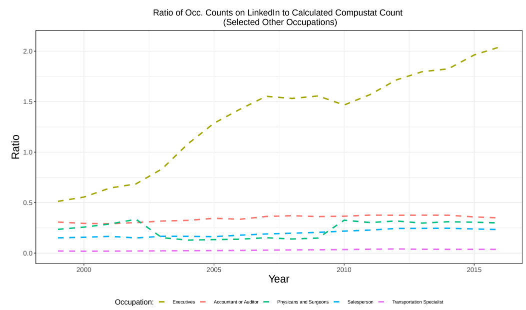
(a) IT Development



(b) IT Infrastructure



(c) Analysis



(d) Selected Other

Software Engineers and IT Consultants, by the end of the sample period, are approaching nearly the full quantity of records that we expect to be represented. Hardware engineers have lower derived coverage rates overall, but are still well represented.

Other types of IT support staff are also well covered by the LinkedIn data. IT Support is somewhat overrepresented, perhaps because these occupations are often found in higher proportions abroad than they are domestically.

Research and analysis staff are also well covered by the data. Interestingly, the creation of the "Data Scientist" job title shifts the composition of the data and underscores the importance of time-varying occupational categories and coverage. Many workers who would have previously been categorized as Database Engineers (often Data Engineers and Data Scientists as we might refer to them today) start to reclassify their occupational titles between 2009 and into 2012. Our ratios capture these changes and adjust accordingly.

The above plot describes a selected group of other occupations. Executives, while found in many different functions, tend to be somewhat overrepresented on the LinkedIn platform. Many workers use their online resumes to describe themselves as executives. While some of these are from establishments abroad, it is likely the case that some of these titles are aspirational. A number of other occupations tend to be less well-represented via profile records. Salespeople, physicians, and transportation specialists have lower coverage rates than workers in the IT and knowledge functions of Compustat firms.

Human capital measures In addition to position records, a large number of profiles also contain education records. LinkedIn creates standardized ids for schools and degree types, allowing us to aggregate counts of degrees by type within firms by adding the degrees held by a given company's employees in a given year. We aggregate these counts for Associate's degrees, Bachelor's degrees, Master's degrees, research doctorates, medical doctorates, and law degrees. We then create a worker-level measure of equivalent years of education by assigning 2 years for Associate's and Master's degrees, 3 years for law degrees (J.D.), 4 years for Bachelor's degrees, and 5 years for research and medical doctorates. Summing these counts within our occupational classification system allows us to build a firm-level measure of total equivalent years of education in the workforce. We use this aggregated equivalent education-year count as a proxy for general human capital in some of our regression specifications. Equivalent years of education are inflated or deflated by worker category according to the firm-occupation-year coverage rates that we calculate for the derived counts of workers by occupation-firm-year.

B Online Appendix: Digital capital depreciation

This section examines the construction of measures of digital capital stocks under different assumptions from those used in the main analysis. Rather than treating IT labor as a proxy measure of the stock of digital capital available to a firm, we can alternatively treat it as a flow of investment into this capital. By assuming a depreciation rate for digital capital (δ), the stock of digital capital in any year can be recovered from the time series of the investments into IT labor by using the perpetual inventory equation shown in equation 10.

$$DC_t = (1 - \delta)DC_{t-1} + I_t \tag{10}$$

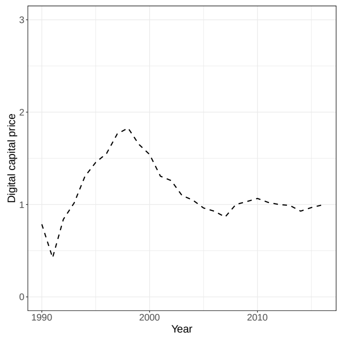
To the best of our knowledge, no prior work has explicitly estimated the depreciation rate of what we call digital capital. Studies on the depreciation of a broader set of intangibles assign fairly high rates of depreciation to different categories of this capital, ranging from 20-60% (Corrado et al., 2009a,b). For the categories closest to ours, they assume depreciation rates of 33% (computerized information, including custom software and data) and 40% (firm-specific resources, including firm-specific human capital and organizational structure). By comparison, studies that estimate the depreciation of physical capital produce estimates that are lower, ranging from .030 to .12 (Epstein and Denny, 1980; Bischoff and Kokkelenberg, 1987; Nadiri and Prucha, 1996), and studies estimating the depreciation of R&D capital produce estimates that fall somewhere in between physical capital and computerized intangible capital, ranging from 0.12 to 0.40 (Pakes and Schankerman, 1984; Nadiri and Prucha, 1996).

We assume that the starting values of digital capital are zero for all firms in the panel. To the extent that depreciation rates are fairly high for this type of capital, this assumption should have a relatively small effect on our analysis.

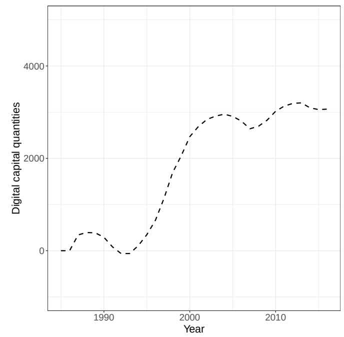
Using this alternative approach to utilize our IT labor measures to develop a series of the firm's stock of digital capital, we can re-compute changes in prices and quantities of digital capital for the firms in our panel. These price and quantity series are shown for three different depreciation rates in Figure B.1. The top panel illustrates changes in prices and quantities assuming a relatively low depreciation rate of .15 for digital capital. The middle panel illustrates changes in prices and quantities assuming a higher depreciation rate of .40 for digital capital, which is close to the value used for organizational structure and firm-specific human capital in Corrado et al. (2009b). The bottom panel assumes instant depreciation for digital capital, which is equivalent to the assumption used in the main text, because only current investments in IT labor matter for imputing quantities of digital capital stock. The prices and quantities computed in the bottom panel of the Figure are, therefore, the same as those reported in the main text and are shown for reference.

If we treat IT-related labor as perfectly flexible such that the labor service flows in these specifications are mere proxies for hidden digital capital stocks (i.e. full depreciation of IT labor every period), then the IT labor measures are a reflection of the digital capital stock complements that the firm owns. On the other hand, if IT labor does not depreciate, we can consider the total IT employment of the firm as a form of capital that is also correlated with corporate digital capital stocks. We do not separate the component of digital capital embodied by employees in their human capital from assets held separately by the company; our estimation procedures reveal the sum of both capital varieties in the case that depreciation of labor services is nonzero. The figure below shows that the overall trends in digital capital prices and quantities are robust to either assumption.

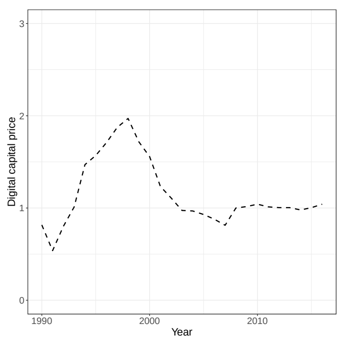
Figure B.1: Prices and quantities of digital capital with different depreciation



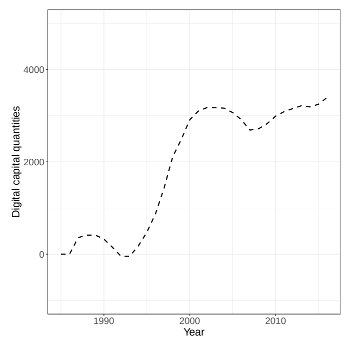
(a) Price, $\delta = .15$



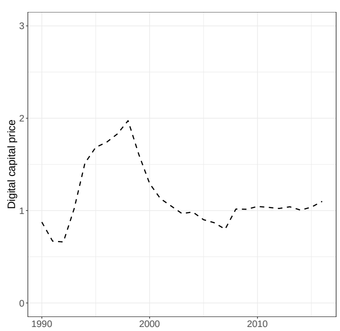
(b) Quantities, $\delta = .15$



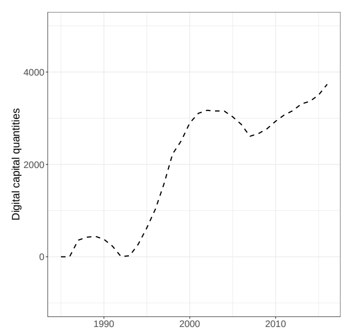
(c) Price, $\delta = .4$



(d) Quantities, $\delta = .4$



(e) Price, $\delta = 1$



(f) Quantities, $\delta = 1$

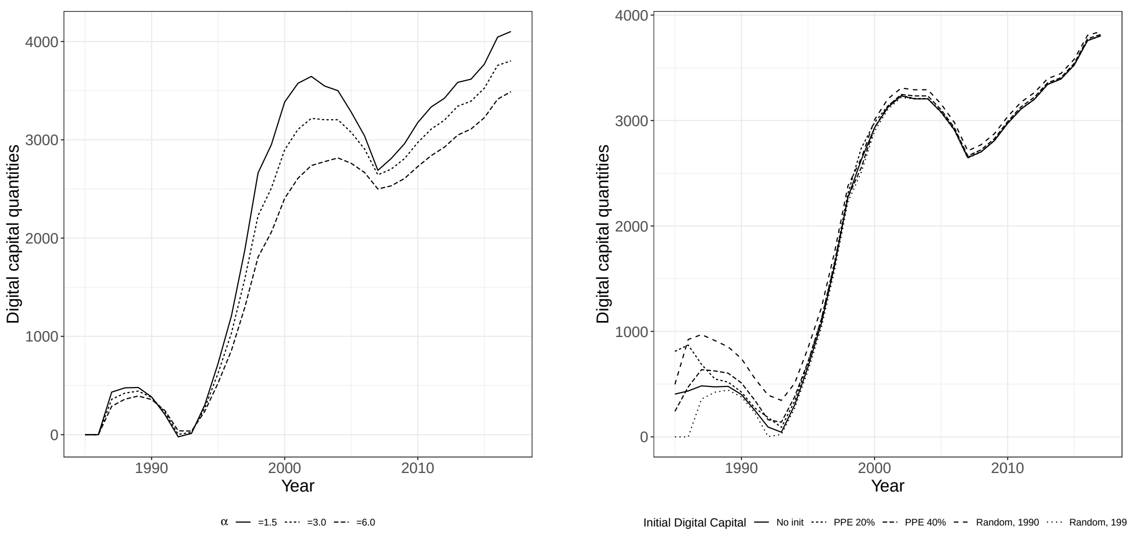
C Online Appendix: Sensitivity to parameter value assumptions

Figure C.1a illustrates how varying values of the adjustment cost parameter α affect the estimated trajectory of digital capital quantities. Our adjustment cost parameters are estimated at the firm-level. In Figure C.1a, along with the computed quantity series that was already shown in Figure 8 (which uses $\alpha = 3.0$), we also present digital capital quantities generated by halving and doubling the α parameter (setting it equal to 1.5 and 6.0, respectively). Using these different values implies larger and smaller costs of installing new capital in the firm, respectively, and it shifts the computed levels of estimated digital capital quantities up and down, but does not impact the overall trajectory of the curve. In other words, digital capital quantities follow the same accumulation path, but assumptions about higher (lower) adjustment costs attenuate (amplify) the increase in price that occurs around the late 1990's as well as the size of the fall in quantities of digital capital that occurred after the 2001 technology recession.

In Figure C.1b, we test how setting initial quantities of digital capital stock to zero affected the digital capital quantities we computed in later years. As an alternative, we change initial digital capital values such that instead of being zero, they are proportional to firms' levels of physical capital, 20% and 40%, respectively. As another alternative, we can seed firms with an initial digital capital quantity that is randomly drawn from a range between zero and twice the mean digital capital quantity computed for firms in 1990 and then 1995. All of the data series that result from using these alternative starting points are supportive of the notion that assumptions about initial levels of intangible capital do not significantly impact the inferences drawn from subsequent values due to the convergent nature of the process.

In sum, although it is difficult to pin down precise values for firms' adjustment costs and digital capital quantities in the years preceding the years in our panel, Figures C.1a and C.1b support the argument that estimated digital capital quantities are not particularly sensitive to these two choices in the recursion model. When these parameters are fixed, the quantity of intangible capital is revealed by its market value if firms make capital investments in a manner consistent with financial investment theory. Hall shows that over a long enough panel, this assumption was a reasonable one for the last half of the twentieth century. It may not describe firms' behaviors on a year-to-year basis, but it is a reasonable approximation over a longer time period.

Figure C.1: Sensitivity to assumptions about model parameter values



(a) Adjustment costs (α) **(b) Firms' initial digital capital levels**

Figure Notes: These charts perturb two parameters of the model used to compute digital capital quantities. Panel (a) alters values of α, the adjustment cost parameter used in equation 7. In the figure, the α parameter is set to three different values: 1.5, 3.0, and 6.0. Panel (b) alters firms' starting values of digital capital. The chart plots quantities with initial digital capital values set to five different levels: i) zero, ii) 20% of PPE, iii) 40% of PPE, iv) a value that was randomly chosen from between zero and twice the value of the firm mean quantities in 1990, and v) a value randomly chosen from between zero and twice the value of firms' mean quantities in 1995.