Engineering Value: The Returns to Technological Talent and Investments in Artificial Intelligence

Daniel Rock[1]
MIT Sloan School of Management

This Version:
May 2019

Abstract

Engineers, as implementers of technology, are highly complementary to the intangible knowledge assets that firms accumulate. This paper seeks to address whether technical talent is a source of rents for corporate employers, both in general and in the specific case of the surprising open-source launch of TensorFlow, a deep learning software package, by Google. First, I present a simple model of how employers can use job design as a tool to exercise monopsony power by partially allocating employee time to firm-specific tasks. Then, using over 180 million position records and over 52 million skill records from LinkedIn, I build a panel of firm-level investment in technological human capital (information technology, research, and engineering talent quantities) to measure the market value of technological talent. I find that on average, an additional engineer at a firm is correlated with approximately $854,000 more market value. Firm fixed effects and instrumental variables analyses provide mixed evidence on the marginal causal value of engineers in general.

Specifically for AI talent, the value of engineering skills is clearer. AI skills are strongly correlated with market value, though variation in AI skills from 2014-2017 does not explain contemporaneous revenue productivity within firms. AI-intensive companies rapidly gained market value following the launch of TensorFlow, while companies with opportunities to automate relatively larger quantities of labor with machine learning did not. Using a difference-in-differences approach, I show that the TensorFlow launch is associated with an approximate market value increase of 4-7% for AI-using firms. Firms outside the top quintile of AI use (as measured by skill counts on LinkedIn) grow by approximately $3.56 million for a 1% increase in AI skill. AI superstar firms in the top quintile also appear to benefit, but show pre-trends in market value growth.

Acknowledgements: I thank my advisor, Erik Brynjolfsson, and thesis committee, Prasanna Tambe, Chad Syverson, and Andrew Lo, for extensive comments and guidance. I am grateful as well to Seth Benzell, Avi Collis, Guillaume Saint-Jacques, Sebastian Steffen, Lee Branstetter, Frank Nagle, Chiara Farronato, Danielle Li, Daniel Fehder, Daniela Scur, Joshua Krieger, Sam Zyontz, the LinkedIn EGR Team (particularly Di Mo, Guy Berger, and Jacqui Barrett), and numerous seminar participants for helpful suggestions. I received access to LinkedIn data as part of the LinkedIn Economic Graph Research Challenge. All errors are my own.

---

[1] Contact information: drock@mit.edu; 100 Main Street, Cambridge, MA 02142. Paper also available at http://drock.mit.edu/research.

"I was originally supposed to become an engineer but the thought of having to expend my creative energy on things that make practical everyday life even more refined, with a loathsome capital gain as the goal, was unbearable to me."

– Albert Einstein

1. Introduction

Technological labor is a well-established driver of corporate market value, innovation, and productivity (Hall 1993, 2006; Tambe and Hitt 2012; Tambe 2014). Technical knowledge is scarce as well. Engineers, research scientists, information technology workers, and other types of technically-skilled labor must invest for years in school and training to build their technical human capital. Their reward for devoting their creative energies to these pursuits is, in part, higher average wages. Still, the capital gains in applying highly specialized knowledge can be partially bargained away by employers in markets with competitive labor supply. There is enduring disagreement about how and to what extent firms can appropriate the human capital investments of their workers (Acemoglu and Pischke 1998; Acemoglu and Pischke 1999; Brynjolfsson et al. 2018). There are many ways in which the returns to worker investments in human capital might lead to employer value gains. Engineers, as implementers of technology, are highly complementary to the intangible knowledge assets that firms accumulate.

Firms hire engineers to build. The specifics of what technical workers build are subject to the discretion of their employers (presumably with capital gain as the goal). When workers prefer some tasks over others for non-pecuniary reasons, the firm gains monopsony power. In some cases, technological workers might sacrifice compensation for other perks or allowances (e.g. the right to publish scientific findings or working with cutting-edge technology) (Stern 2004; Mas and Pallais 2017). In labor market contexts where both supply and demand are competitive, there is little excess surplus to split. Each party earns the marginal product of what it provides. For employees with accumulated firm-specific knowledge and little competition (CEOs and high-level executives, for example), some of the surplus from firm-specific assets can be bargained away by the worker (Brynjolfsson 1994; Hart and Moore 1994). Some kinds of occupations, like managers and tech workers, have a mixture of both kinds of tasks. They do activities that require firm-specific capital that grow productivity and they also implement production processes in tasks that are more competitive.

This paper seeks to address whether technical talent is a source of rents for corporate employers. The answer to this question informs whether strategic technological labor resources are a direct source of sustainable competitive advantage or operate via other factors (Barney 1986; Crook et al. 2011). I approach this question in two ways: firstly by investigating the relationship between market value and aggregated measures of technological talent, and secondly by exploiting the unexpected open-sourcing of Google's TensorFlow, a machine learning software package particularly well-suited for deep learning, at the end of 2015. The overall engineering value estimation aggregates across many different technologies, whereas the TensorFlow analysis illustrates a case where engineering talent is highly scarce. As an emergent technology, I find that the marginal value net of wages of additional AI talent is still well above the breakeven point for the average publicly traded firm.

If technological human capital is highly complementary to the firm's asset base and both cause market value, for example, then failure to invest in worker retention may impair the value of the firm's non-human capital. Technological human capital should cause market value in the case that tech workers are employed building firm-specific assets (even if those assets remain in the workers' heads). In the first section of the empirical results, I use a firm-level panel of employment by worker type from LinkedIn merged to firm performance and value data to describe the component of market value attributable to technical talent. I estimate the causal effect of additional engineering investment using proximity to land-grant colleges (Moretti 2004; Bloom et al. 2018), changes in state-level covenant-to-not-compete (CNC) policy (Marx, Strumsky, and Fleming 2009; Marx 2011; Ewens and Marx 2017; Starr, Balasubramanian, and Sakakibara 2017; Balasubramanian et al. 2018; Jeffers 2017), and firm and state-level differences in the user cost of R&D capital (Lucking, Bloom, and Van Reenen 2017; Lucking 2018) as instruments for the engineering human capital in firms. I find that while engineering talent expenses and quantities are strongly correlated with market value the causal specifications using instrument variables and correlational estimates including firm fixed effects eliminate statistical significance. Point estimates for the value of engineers increase in specifications using land-grant and CNC instruments, though they are imprecise. Model specifications with the user cost of R&D instruments, to the contrary, indicate that the marginal engineer destroys market value adjusting for the total stock of R&D assets at the firm. This is indicative of the presence of firm-level intangible asset service flows complementing the labor of technological workers, consistent with Brynjolfsson, Hitt, and Yang (2002); Tambe, Hitt, and Brynjolfsson (2012).

Taken holistically, these results suggest that the causal effect of the marginal technological hire seems to add little to the firm's market value, but the average value of these workers is high in equilibrium. I find that each additional engineering worker is correlated with another $855,000 of market value for the firm, and in wage terms $1 of engineering wages is correlated with approximately $11.9 of market value. This suggests that the relationship between market value and technological talent is generated by firm-specific assets that are complementary to or perhaps embodied within generalized engineering labor. The lack of precision on the marginal effect of general engineering indicates the need for more skill-specific studies of labor value. Even aggregate measures of STEM employment fail to suggest easily earned rents.

In the current market environment, there are few, if any, technologies with the transformative potential of artificial intelligence and machine learning (Brynjolfsson, Rock, and Syverson 2018a; Brynjolfsson, Mitchell, and Rock 2018b; Cockburn, Henderson, and Stern 2018; Agrawal, Gans, and Goldfarb 2018). One of the primary obstacles to widespread adoption of artificial intelligence is the available labor supply, with top-tier scientists earning more than $1 million in some cases.[2] But the market is responding and the supply of machine learning engineers is increasing. Accordingly, to further understand the mechanism generating the large average correlational value of engineering talent and how it might be related to specific labor skills, the second part of the paper exploits the Google's open-source launch of TensorFlow in November 2015. TensorFlow has a Python-based application programming interface (API) which greatly facilitates the ease and efficiency in building (and learning to build) deep learning models. The TensorFlow launch serves as a shock to the fixed costs of learning how to build deep neural nets (DNNs) for software engineers and analysts. Prior to TensorFlow, the ability to train DNNs was rare and highly specialized. The launch of this tool both effectively commodified deep learning as a skill amongst those with Python ability and accelerated *expectations* for how soon deep learning would be easy to learn more generally.

Following the introduction of TensorFlow, I find a rapid increase in the rate and quantity of addition of Artificial Intelligence skills on LinkedIn. Mapping these increases to publicly-

---

[2] https://www.nytimes.com/2018/04/19/technology/artificial-intelligence-salaries-openai.html

traded firms, I find that the TensorFlow shock had differential effects on firm market value. The value of companies making investments in AI grew more following TensorFlow, even controlling for a wide variety of other complementary skills and including firm fixed effects. For firms in the third and fourth quintile of AI skills, each additional 1% in Artificial Intelligence (AI) skill record counts on LinkedIn is correlated with an increase in firm market value of nearly \$3.56 million following the introduction of TensorFlow. The TensorFlow launch provides evidence that talent scarcity can be an important bottleneck to the realization of returns on technological assets. Lowering the barriers to acquiring a formerly rare and valuable skill, as TensorFlow does, makes technological supply competitive. This increased competition of technology suppliers (i.e. engineers) enables their employers to earn returns on firm-specific assets. Additionally, I test whether it is firm-level opportunities to *apply* machine learning causing the increase in market value using the Suitability-for-Machine Learning (SML) measures in (Brynjolfsson, Mitchell, and Rock 2018a, 2018b). If anything, higher average firm SML scores are *negatively* correlated with market value. While it would be premature to assume that the stock market has fully priced in the automation potential of machine learning, this difference-in-differences result suggests AI-related repricing of corporate assets in 2016.

The set of mechanisms by which technology workers might generate market value is generally applicable to all kinds of human capital. However, technological skills can change or depreciate much faster than other kinds of human capital. What makes technology workers, and engineers in particular, useful for understanding the underlying value creation processes of workers in firms is this capacity for discrete changes in the competitive environment. TensorFlow's introduction is one such example among many. Technological shifts therefore supply outside researchers with a chance to study the outcomes of employment-related relational contracts. Analogous shifts to TensorFlow for managerial workers, for example, might be more challenging to find. Still, studying technological changes can supply insight into how companies and employees divide the gains from business activity. These conclusions, in some cases, can be applied to other kinds of employees. Ordinarily it is a substantial challenge to look within the firm with granular information about specific types of employed workers and the skills they have. This study is among the first to normalize and deploy detailed data on firm employment over time and how workers contribute to the value of their employers.

The paper is organized as follows: Section 2 describes the relevant literature in the economics and strategy of human capital, technology, and market value. Section 3 describes a theoretical model of how human capital can enter the valuation of firms. Section 4 details the construction of the datasets. Section 5 describes and analyzes the relationship between market value and aggregated engineering talent. Section 6 offers an empirical case study of the TensorFlow launch and discusses the market value effects of making AI talent more abundant. Section 7 concludes.

2. Related Research on Human Capital, Technology, and Market Value

This study fits within a tradition of human capital and technology studies that can be traced back to (Becker 1962). How human capital is accumulated and why firms are motivated to invest in it has long been a puzzle for social scientists. A key question is how easy it is for workers to apply their human capital across different employers. Transferrable skills and knowledge (general human capital) are subject to competitive bidding pressure from firms, while the returns to firm-specific investments are subject to bargaining arrangements in contracting. Labor market frictions might therefore create the right incentives for employers to invest in their workers' human capital (Acemoglu 1997; Acemoglu and Pischke 1998; Acemoglu and Pischke 1999). A related literature considers the market power aspect of these frictions, studying monopsony power (Bhaskar, Manning, and To 2002; Ashenfelter, Farber, and Ransom 2010). Firms insulate themselves from competitive pressure in myriad ways, including (but not limited to) regional concentration (Azar et al. 2018), organizational design and technology (Stole and Zwiebel 1996b), and incomplete contracts (Stole and Zwiebel 1996a; Simon 1991).

Match-specific value between employer and employee also leads to productivity-ability sorting, at which point larger surplus values are split between elite matches. Tervio (2008) offers an assignment model approach to understanding the value of CEOs that is instructive for the skilled worker context considered here. If part of the work activities bundled into a given technical job are investments in match-specific value and the remaining portion is assigned to competitive labor tasks, the firm pays wages at the competitive labor margin while appropriating the bundled match-specific value. This will not be the case, as in Tervio (2008), when much of the labor effort is devoted to match-specific tasks or, as in traditional neo-classical labor supply functions, the labor effort is competitive and commodified.

The context considered here is in-between those extremes: engineers, much like managers, spend part of their time building non-marketable firm-specific assets and part of their time maintaining or implementing production in competitive arenas. This generates an incentive for the firm to use the allocation of tasks (job design) as an instrument of monopsony power. As in Simon (1991), "the combination of uncertainty on the part of the employer (as to what will need to be done in the future) and broad acceptance of the employee (of what he or she will be ordered to do) makes the employment contract a very attractive bargain for both parties". Of course, firms can modify how workers *perceive* the firm-specificity of their human capital investments, wherein workers might be more willing to learn firm-specific skills if they believe them to be marketable (and conversely, less willing if the skills were perceived as unmarketable) (Coff and Raffiee 2015; Raffiee and Coff 2016).

This paper makes a simplifying assumption that workers and firms correctly understand the firm-specificity of their human capital, but that uncertainty at the time of contracting leads to ex-post rent sharing. Campbell, Coff, and Kryscynski (2012) argue for mobility constraints as a stronger influence than firm-specificity of human capital. Supply-side factors affecting worker bargaining power and the propensity for workers with firm-specific assets to start new firms (as in (Campbell et al. 2012; Eisfeldt and Papanikolaou 2014; Jovanovic 1979)) are an important component of the firm's share of created value. Additionally there is evidence that "scientists pay to be scientists" (Stern 2004) or, in other words, have preferences favoring academic rewards over monetary ones (Roach and Sauermann 2010). There is evidence this is also the case for technologists, who frequently prefer to work with cutting-edge technology (Tambe, Ye, and Cappelli 2019). More traditionally, monopsony power can be related to employer concentration (Azar et al. 2018; Benmelech, Bergman, and Kim 2018) and financial constraints also affect hiring decisions (Benmelech, Bergman, and Seru 2011). The mechanism for engineer value in this paper is the interaction of task allocation and firm-specific assets with engineering labor complements.

Fixed costs of capital investment apply to human capital as well, wherein quasi-rents can accrue to firms which have already sunk the necessary recruitment and training expenses required to make an employee productive (Hall 2001). The process for valuation of labor assets in these studies is functionally identical to the valuation of capital assets. Because the marginal adjustment costs of competitors set the price at which the asset is available, firms will hire capital until the marginal adjustment costs of competitors is equal to the marginal value created with that capital (Tobin's Q) (Hayashi 1982; Tobin et al. 1976; Kaldor 1966).[3] The difference between the firm's adjustment costs and those of its competitors pin down the excess profit of the firm in the short-run. These adjustment values for "vanilla" labor have been estimated at low values in the past (Hall 2004; Hall 2017).

Nevertheless, the firm's value share of some types of human capital has been estimated as large and meaningful. Specialized labor is highly sought-after by corporate employers. The H-1B visa program, which expands the talent pool for technically-savvy workers, is typically oversubscribed such that eligible talent from outside the U.S. must file for a lottery. This increased quantity of STEM workers from the H-1B program led to greater productivity (Peri, Shih, and Sparber 2015), within-firm employment (Kerr, Kerr, and Lincoln 2015), and rates of innovation and entrepreneurship (Kerr 2013). Technological talent is deployed not only in implementing the production function, but also in building the knowledge and business process assets of the firm which facilitate growth. Shifts in the availability of technological talent therefore cause valuation changes via many channels, including but not limited to: price effects on existing assets, appropriability of human capital, marginal labor productivity, and future innovation opportunities.

Research and development expenses, of which a substantial component is researcher salaries, are reliably strongly correlated with market value and drivers of patenting and other innovative activity (Hall 1993, 2006). The estimated average Q-value of R&D assets is nearly an order of magnitude above that of property, plant, and equipment in Compustat firms (Peters and Taylor 2017; Brynjolfsson, Rock, and Syverson 2018b). This is suggestive that R&D-intensive

[3] This leads to the main market value equation I apply in the empirical section.

firms tend to also accumulate hidden intangible complements – business processes, training, knowledge, and even firm culture – which contribute to market value in ways that are difficult to capitalize on a balance sheet. These factors, even with relatively coarse measures of organizational investment, have predictive power in the cross-section of stock returns (Eisfeldt and Papanikolaou 2013).

Intangible assets are complementary to and correlated with investment in technological human capital (Bresnahan, Brynjolfsson, and Hitt 2002; Brynjolfsson, Hitt, and Yang 2002; Saunders and Tambe 2015; Saunders and Brynjolfsson 2016). Further, the shift toward intangible assets in the digital age has opened up a research agenda into the productivity effects of IT capital, with technology diffusion serving as a leading explanation for the widening productivity differences between firms at the frontier and firms at the median productivity level (Syverson 2011; Lustig, Syverson, and Van Nieuwerburgh 2011; Andrews, Criscuolo, and Gal 2015). Intangible assets are inherently hard to measure and constitute an increasingly large component of the U.S. economy's asset stock. One explanation for the high market value of engineers in the empirical results of this paper is that firms with more engineers also tend to build up intangible assets which are left off of the corporate balance sheet. Firms often fail to capitalize software expenses, for example, causing correlational analyses to attribute market value to the observable complements (wage value, in this case). This quantity is the focal object of study in (Tambe et al. 2018). This study begins at the aggregated level of technological labor, and then studies the labor shock of TensorFlow in AI as a means of understanding specifically how scarce labor can serve as a bottleneck to firm value creation.

Similar studies of technological tool-based and technological knowledge-based exogenous events have addressed how such changes impact various performance measures for firms and other entities. Thompson (2017) studies the economic effects on firm productivity of the switch to multicore processing. Ewens, Nanda, and Rhodes-Kropf (2018) analyze the entrepreneurial effects of the launch of Amazon Web Services (AWS). AWS bundled a number of general-purpose technologies together and made computing infrastructure rentable.[4] This was a major reduction in the fixed costs of starting a new technology-oriented business. Teodoridis (2017) shows that the hack of the Microsoft Kinect made motion-sensing much cheaper, reducing the need for research teams to collaborate with specialists in motion-sensing technology. The hack democratized the technology, similar to the way in which the launch of TensorFlow has (partially) democratized deep learning. Technological advances need not have such an effect; burden-of-knowledge (Jones 2009) effects might dominate in the case that there is an exogenous increase in knowledge capital that is costly to process (Agrawal, Goldfarb, and Teodoridis 2016). This kind of change might necessitate the use of subject matter specialists in increased proportions. Most recently, (Zyontz 2018) has studied the team expertise structure of cell biology researchers following the advent of CRISPR, a precise gene editing tool.

These studies of how production changes in response to inventive activity have in common a theoretical underpinning in the value of "discovery information" (Hirshleifer 1978). Discovery information refers to "detection of properties of Nature that permit the development of new tools or the utilization of new techniques". In Hirshleifer's example, the prices of state-contingent claims adjust when the knowledge of the new state probabilities (e.g. the probability that deep learning will be made easier) will be obtained publicly before the close of trading. As with Eli Whitney's cotton gin, a "route to profit" other than patent protection for new intellectual property is in speculation on the business prospects of firms technologically exposed industries.

[4] Though not in 2006, AWS now includes AI services.

In AI, the pecuniary benefits of open-source innovation, by revealed preference, outstripped the benefits of private IP for Google in TensorFlow's case. AI shares characteristics of both types (tool and knowledge) of exogenous changes, though one of Google's stated aims in open-sourcing TensorFlow was to increase usability and accessibility of deep learning for engineers throughout the economy.[5] A primary pecuniary benefit of making AI models easier to build is an expected subsequent drop in marginal wage rates for AI-intensive human capital.

Like ICT, Artificial Intelligence-related assets are mostly intangible and the returns are mostly in the future at this point (Brynjolfsson, Rock, and Syverson 2018a). The recent progress in AI is mostly a result of advances in deep learning techniques, a specific kind of machine learning approach. Deep learning and neural net algorithms are decades old, but have only recently grown in popularity as large-scale datasets and cheap computational power have made them viable in new domains (White and Rosenblatt 1963; Rumelhart, Hinton, and Williams 1986; LeCun et al. 1998; LeCun, Bengio, and Hinton 2015). AI is an umbrella discipline, including machine learning (which itself includes deep learning) as well as rules-based or expert systems approaches to problem solving. As a new kind of software, however, deep learning and AI more broadly is a general purpose technology (Bresnahan and Trajtenberg 1995; Bresnahan 2010). It is potentially pervasive, improves over time as better and more data arrive, and can spawn complementary innovation. AI is a general purpose prediction technology (Agrawal, Gans, and Goldfarb 2017, 2018). Of course, deep learning is not the only prediction technology of its kind – similar problems might be solved by simpler methods like linear regression. Yet the performance of deep learning on formerly insurmountable tasks (e.g. image and speech recognition) has marked a watershed moment in the cost of prediction.

[5] https://www.wired.com/2015/11/google-open-sources-its-artificial-intelligence-engine/

Since prediction is pervasive throughout the economy, the promise of AI is that it will lead to business process innovation, job redesign, and new engineering advances across many domains in the economy (Furman and Seamans 2018; Brynjolfsson, Mitchell, and Rock 2018b). Critically, deep learning overcomes the obstacle of Polanyi's Paradox where "we know more than we can tell" (Polanyi 1966; Autor 2014). For deep learning models, we need only measure inputs and outputs. The map between them is learned by the algorithm. Artificial General Intelligence (AGI), where machine intelligence equals or surpasses human intelligence in all cognitive tasks, is technologically far away at the moment. But the relatively brittle, bespoke applications of deep learning could feasibly cause large shifts in labor demand and economic value creation processes (Brynjolfsson, Hui, and Liu 2018).[6]

Since the effects are mostly in the future, market value is one of a handful of measures which is sufficiently forward-looking to account for returns to investment activity in the present day. Further, as noted by (Raj and Seamans 2018), relatively little data on AI at the firm-level is available. This study builds and measures firm-level proxies for AI investment as of the end of 2017. As a GPT, the effects of machine learning on firms and labor markets will likely be diffuse across many industries (Brynjolfsson, Mitchell, and Rock 2018a; Felten, Raj, and Seamans 2018). Like prior waves of automation, machine learning will differentially impact tasks that are technologically and socially feasible (Autor, Levy, and Murnane 2003; Acemoglu and Autor 2011; Autor and Dorn 2013). Yet for engineering value, we can study now how the decision to make an advanced tool widely available lowered workers' entry costs and facilitated a shift in technological investment.

[6] Applications so far include: Self-driving cars, call center automation, insurance claims processing, materials discovery, drug discovery, and language translation

3. Theoretical Framework for Valuing the Firm's Share of Human Capital Investment

Standard investment theory requires little modification for human capital to enter into market valuations of firms (Tervio 2008). Here I follow a setup common to (Lucas 1967; Hayashi 1982; Wildasin 1984; Hayashi and Inoue 1991; Yang and Brynjolfsson 2001; Brynjolfsson, Rock, and Syverson 2018b). The firm must choose the right investment and labor quantities to maximize profits:

$$\max_{I,L} V(0) = \int_0^{\infty} \pi(t)u(t)dt; \quad \pi(t) = pf(K,I,L,t) - w^t L - z^t I - r^t K \quad (1)$$

Profits are denoted by $\pi(t)$; $u(t)$ denotes the compound discount factor and p is the price of output. F is assumed nondecreasing and concave in capital (K) and labor (L), and nonincreasing and convex in investment I. W and Z refer to the price vectors at time t for wages and investment, respectively (superscript t is dropped from now on). Assume F is homogeneous in the first degree. We have the following growth constraint on capital stocks of varieties indexed by j:

$$\frac{dK_j}{dt} = I_j - \delta_j K_j \quad \forall j = 1,2,\dots,J$$

Then the firm's Hamiltonian to maximize is:

$$H(K,I,L,t) = (pf(K,I,L,t) - wL - zI)u(t) + \sum_{j=1}^{J} \lambda_j \left(I_j - \delta_j K_j\right) \quad (2)$$

With the following constraints[7]:

$$\frac{\partial H}{\partial \lambda_j} = \dot{K}_j = I_j - \delta_j K_j \quad \forall j \text{ and } \forall t \in [0,\infty]$$

[7] $\dot{x} \equiv \frac{dx(t)}{dt} \quad \forall x(t)$

$$\frac{\partial H}{\partial K_j} = -\dot{\lambda}_j = pF_k u - \lambda_j \delta_j \qquad \forall j, t$$

$$\frac{\partial H}{\partial I_j} = 0 = \left(pF_{I_j} - z_j\right) u + \lambda_j \qquad \forall j, t$$

$$\frac{\partial H}{\partial L_i} = 0 = \left(pF_{L_i} - w_i\right) u \qquad \forall i = 1,2,3, \dots, L \; and \; \forall t$$

$$\lim_{t \to \infty} \boldsymbol{\lambda}(t) \boldsymbol{K}(t) = 0$$

This implies that the firm's market value is the sum of the quantities of its capital assets multiplied by their replacement value prices added to the installed value (where λ is the installed asset price). The solution for the firm's market value under the condition that marginal and average wages are equivalent in all cases is then[8]:

$$V(0) = \sum_{j=1}^{J} \lambda_j(0) K_j(0) \qquad (3)$$

We can relax the assumption that wages are equal to marginal and average products of labor.

$$\int_0^{\infty} \left[\sum_{j=1}^{J} \left(pF_{K_j} K_j + pF_{I_j} I - z_j I_j \right) + \sum_{i=1}^{L} \left(pF_{L_i} L_i - w_i \right) \right] u(t) dt = V(0) = V_K + V_L \qquad (4)$$

And now we are primarily concerned with the value of labor term V_L, and labor varieties are indexed by i.

$$V_L = \int_{t=0}^{\infty} V_L^t u(t) dt = \int_{t=0}^{\infty} \sum_{i=1}^{L} \left(pF_{L_i} L_i - w_i \right) u(t) dt \qquad (5)$$

In the case that wages and marginal products of labor are equivalent for all labor types, at all employment quantities, and in all time periods, this term in equation (5) is zero. Since wages are

[8] $V(0) = \sum_{j=1}^{J} \lambda_j(0) K_j(0) = \sum_{j=1}^{J} (\lambda_j(0) K_j(0) - \lim_{t \to \infty} \lambda_j(t) K_j(t)) = \sum_{j=1}^{J} \int_0^{\infty} (-\dot{\lambda}_j K_j - \lambda_j \dot{K}_j) dt = \int_0^{\infty} (\sum_{j=1}^{J} \left(pF_{K_j} K_j + pF_{I_j} I - z_j I_j \right) + \sum_{i=1}^{L} (pF_{L_i} L_i - w_i)) u(t) dt = V(0)$

set competitively by the market and the asset holdings of different firms vary, it is unlikely to be the case that all firms individually face the same marginal productivity of labor as the aggregate. Wages, on the other hand, are more likely to be consistent given ability or skills in the same region and industry, though there is evidence to suggest wage inequality can also be driven by firm characteristics (Song et al. 2015).

Equation (5) describes the potential surplus that an employer receives from the aggregated marginal products of its employees. The worker problem is deliberately simple: workers seek to maximize their wage subject to a constraint that it be above their reservation wage. Workers have one divisible unit of labor to supply. Assume now that the production function can be decomposed as follows as a function of the inputs in (1), but now firms can choose to assign workers to firm-specific labor (H) or general labor (L).

$$\max_{I,H,L} V(0) = \int_0^\infty \pi(t)u(t)dt\,;\ \pi(t) = pA^tF(K,I,H,L,t) - w_H^tH - w_L^tL - z^tI - r^tK \qquad (6)$$

Subject to additional constraints for workers indexed by *l:*

$$H_l + L_l = 1\ \forall l\ \in 1,2,3,\dots,N$$

$$H_l \geq 0, L_l \geq 0\quad \forall l\ \in 1,2,3,\dots,N$$

$$w_{H_l} \geq w_0, w_{L_l} \geq w_0\ \forall l\ \in 1,2,3,\dots,N\ \forall t$$

So that all workers have only one unit to supply, that negative labor is not possible, and that the reservation wage for all workers w_0 is met, guaranteeing participation. Of course, every employer has *some* tasks that are firm-specific and that value, rather ironically, will be compensated as part of a general wage because workers can foresee some of what their employers will have them do. The firm-specific value I consider here is "extra" and arises from an incomplete ex-ante contracting problem. The worker does not really know everything their employer might ask of them. Realistically extra firm-specificity might have its roots in job search frictions, worker preferences for more specialized tasks, unique business strategies and production functions, or even programs where workers are allowed to spend some proportion of their time as they choose (Stern 2004; Tambe, Ye, and Cappelli 2019). Given that this firm-specificity of tasks exists, however, the job design is a channel for the exercise of monopsony power in the labor market.

The firm will maximize the present value of all future discounted profits (productivity term A is included to note that productivity can change over time). $F(K, I, H, L, t)$ serves to transform capital, investment, firm-specific labor, and general labor into production output. In addition to the same set of assumptions on equation (1) to get the solution in equation (4), I assume that F is nondecreasing and concave in H and L. Firms therefore know the ability of the workers they hire, and choose for them whether they work on "H" tasks or on "L" tasks. I also assume that the surplus from firm-specific tasks H are more appropriable to the firm than non-specific tasks, and that there are strictly increasing differences in the marginal value of H and L tasks (this guarantees a single-crossing of the marginal values of firm-specific and general labor task marginal products). Assume the firm's share of firm-specific task marginal products of labor is $\beta \in [0,1]$. This parameter is assumed to be exogenous, to avoid focusing on a specific bargaining process. Formally, if Q = H + L,

$$\exists Q^* = H^* + L^* \ s.t.$$

$$\beta \frac{\partial F}{\partial H}(Q, K, I) > \frac{\partial F}{\partial L}(Q, K, I) \quad \forall Q < Q^*$$

and

$$\beta \frac{\partial F}{\partial H}(Q, K, I) \leq \frac{\partial F}{\partial L}(Q, K, I) \quad \forall Q \geq Q^*$$

$$\because \frac{\partial^2 F}{\partial H \partial L}(H, L, K, I) > 0 \quad \forall (H, L, K, I) \quad (7)$$

This assumption characterizes firm-specific and general labor as complements, both required as part of the production function. Though I assume this for tractability purposes, future work could investigate relaxing this assumption. The firm now observes its hired labor pool and assigns each worker l to some proportion α_l of firm-specific tasks, with the remainder of the worker's time spent on non-specific tasks (or general tasks).[9] The employers make these decisions perhaps with the workers' skills or abilities in mind. The assignment function for the employer seeks simply to maximize the marginal product of each worker's labor given the other inputs factor vectors. $\boldsymbol{\alpha}$ denotes the Nx1 vector of assignments to firm-specific tasks and each worker has 1 unit of labor to supply:

$$\boldsymbol{\alpha}^* = \underset{\boldsymbol{\alpha}}{\operatorname{argmax}}\{\beta \frac{\partial F}{\partial H}(H_{\boldsymbol{\alpha}}, L_{\boldsymbol{\alpha}}, K, I) + \frac{\partial F}{\partial L}(H_{\boldsymbol{\alpha}}, L_{\boldsymbol{\alpha}}, K, I)\} \quad (8)$$

$$H_{\boldsymbol{\alpha}} = \int_{l=0}^{N} \alpha_l dl$$

$$L_{\boldsymbol{\alpha}} = \int_{l=0}^{N} (1 - \alpha_l) dl = N - H_{\boldsymbol{\alpha}}$$

$$s.t.\ \alpha_l \in [0,1]\ \forall l$$

Here the firm only receives β units of marginal product for each unit of H tasks; workers bargain away the remaining proportion. The employer's share of the marginal product of each worker is a sum of the task-specific marginal products. We guarantee participation of each of the workers with the participation constraint in (6). Sorting workers in order of their proportion of work in the firm-specific task (index the highest proportion worker at 0, lowest at N), we get the period-specific employers' and workers' returns to labor in equations (9) and (10) (given t). The workers in general tasks earn a wage w_L for which the firm is a price-taker. While the share is

[9] Assume, for example, that each worker has an endowed positive productivity level in H and L tasks (ω_H, ω_L), drawn from a stable distribution, that the firm can observe.

temporarily fixed, workers bargain away surplus proportional to the marginal product of firm-specific tasks at their rank *l*. The workers' share of firm-specific surplus would be lower if all workers accepted the same share of the aggregate marginal value of firm-specific labor.[10]

$$V_L^t = \int_{l=0}^{N} \left(\alpha_l^* \beta \frac{\partial F}{\partial H}(H_l, L_l, K, I) + (1-\alpha_l^*) \left(\frac{\partial F}{\partial L}(H_l, L_l, K, I) - w_L \right) \right) dl \quad (9)$$

$$w_l = \alpha_l^*(1-\beta) \frac{\partial F}{\partial H}(H, L, K, I) + (1-\alpha_l^*) w_L \qquad (10)$$

where

$$X_l = \int_{j=0}^{l} \frac{\partial X}{\partial j} dj \;\; for\; X = H, L$$

And at $l = N$ *where N is chosen to maximize profits*[11]:

$$\alpha_N \beta \frac{\partial F}{\partial H}(H_N, L_N, K, I) + (1-\alpha_N) \left(\frac{\partial F}{\partial L}(H_N, L_N, K, I) - w_L \right) \leq 0$$

The firm's returns to labor investments come from two sources: the share of total surplus they recover from firm-specific labor H and the difference between the marginal product of general-task labor and the general labor wage set by the market. Under the standard neoclassical assumptions that general-task labor is competitive and the demand curve for L is perfectly elastic, this second term under the integral in equation (9) goes to zero. If employers do not invest in firm-specific labor tasks, the market value of labor will be zero. Otherwise, the first term in the integrand in (9) represents the employer share of firm-specific labor. Maintaining the assumption that the firm surplus for general task labor is close to zero, the correlation between human capital measures and market value is affected by 1) changes in the firm's share of surplus

[10] Again for tractability, I assume the rate of change in labor of both types to be differentiable in the index of the workers *l*. This could easily be discretized and summed instead.
[11] If the share of general tasks is 1 at *l*=N, this reduces to difference between the marginal product of general labor and the prevailing general wage in the market (which we expect is close to zero). High frictions are also important, though not directly included in the model (they can be thought of as entering through the firm's bargained share).

β and 2) The quantity of firm-specific labor at the firm, and 3) complementarities (or substitution effects) between capital and firm-specific labor. Market value regressions without firm-specific labor measures may therefore be sensitive to omitted variable bias.

On the labor side, earnings come from two sources: general-purpose labor marginal wages and the workers' bargained shares of *marginal* firm-specific labor value. Differences between average and marginal wages accrue to the firm. Incidentally this model may partly explain the recent separation between labor productivity growth and wage growth (Brynjolfsson and McAfee 2014; Bivens and Mishel 2015; Stansbury and Summers 2017). If general labor wages are falling, perhaps due to outsourcing or increased labor supply, wage growth must come from either increased bargaining for firm-specific surplus or larger overall firm-specific surplus from labor. If employers, perhaps because of specialization or insulation from competition in labor markets, shift their share of work to firm-specific tasks, this could also put downward pressure on wages.[12][13] The single-crossing property in (7) suggests too that as employers increase total employment past Q^*, general-task labor will increase as a share of total labor in the firm.[14]

The model can be applied to specific occupations as well. Hiring larger quantities of a given occupations will mean that more of the available firm-specific surplus is captured. As bottlenecks of expensive firm-specific labor are alleviated, the firm comes closer to realizing the maximal value of complementary assets. Sample inverse labor demand curves are displayed

---

[12] The bargained firm-specific wage must be as least as large as the general wage, or workers can move to a firm that will pay them for entirely general labor.

[13] Lippman and Rumelt (1982) suggest another possible explanation in business complexity. Competitive pressure is ameliorated when business processes are especially difficult to reverse-engineer. Technology has expanded the combinatorial space of viable business models. In the long-run profits are competed away, but in the short-run a curse of economic dimensionality with scarce productive inputs can make competitive effects slow to act. I leave the exploration of the cardinality of competition to future work.

below in the figure. The figure shows wage as a function of quantity demanded (Q_D) for general-task labor (blue), firm-specific task labor less the baseline wage for general labor (red), and the aggregate attainable by varying alpha in a linearly decreasing manner (purple). The single-crossing point is where the red and blue curves intersect.

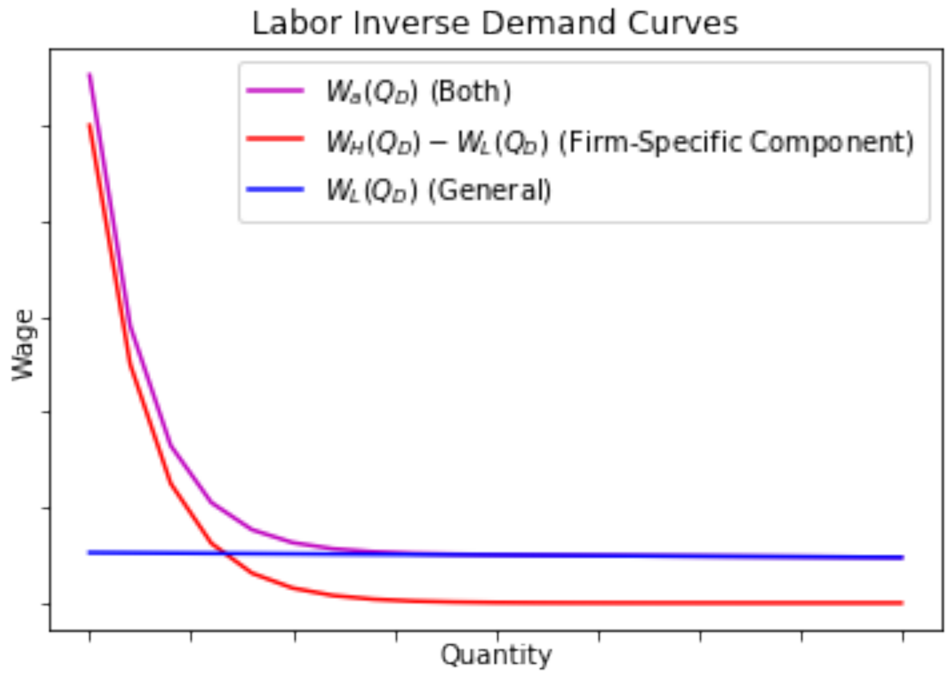


Figure 1 – Labor Inverse Demand Curves with Firm-Specific Tasks

These dynamics provide a motivation for why the labor buyer surplus in (5) might vary as a function of the productivity of firm-specific assets and labor. Since wages are a function of supply of labor conditional on skills and abilities, a technological shock increasing supply and driving wages for general-task labor down (or expected future wages) would change the firm's valuation. This mechanism works both through the direct effect of lower wages, but also through the marginal productivity of firm-specific capital and installed capital assets, i.e. changes in $\boldsymbol{\lambda}$, at the new equilibrium. TensorFlow is the expected future labor supply-increasing shock I now turn to for the case of AI. Machine learning talent is an expensive complement to machine learning capital assets like data and computational power. TensorFlow leads to more abundant machine learning labor, alleviating a bottleneck preventing employers with machine learning assets from realizing larger returns. For the case of machine learning talent, the expectation is that TensorFlow or similar software cause a future increase in the available labor supply. The figure below displays a simplified version of the intended supply shift from the green supply curve to the orange one. This moves the labor market equilibrium from the intersection of the purple inverse demand function (as above) and the green supply curve to the intersection of the inverse demand function and the orange supply curve. Wages are set by the intersection of the supply curves with the purple demand curve, and therefore the shift from the green supply to the yellow supply increases the firm's surplus as the buyer of labor (in red). In the last part of equation (9), the firm sets N such that the marginal surplus of an additional worker is less than or equal to zero. The alleviation of large search frictions and/or inelastic supply (omitted from the model, but a major concern for AI employers) would cause the equilibrium hiring breakeven point to increase. In Section 6 I provide evidence that this is what happens following the launch of TensorFlow.

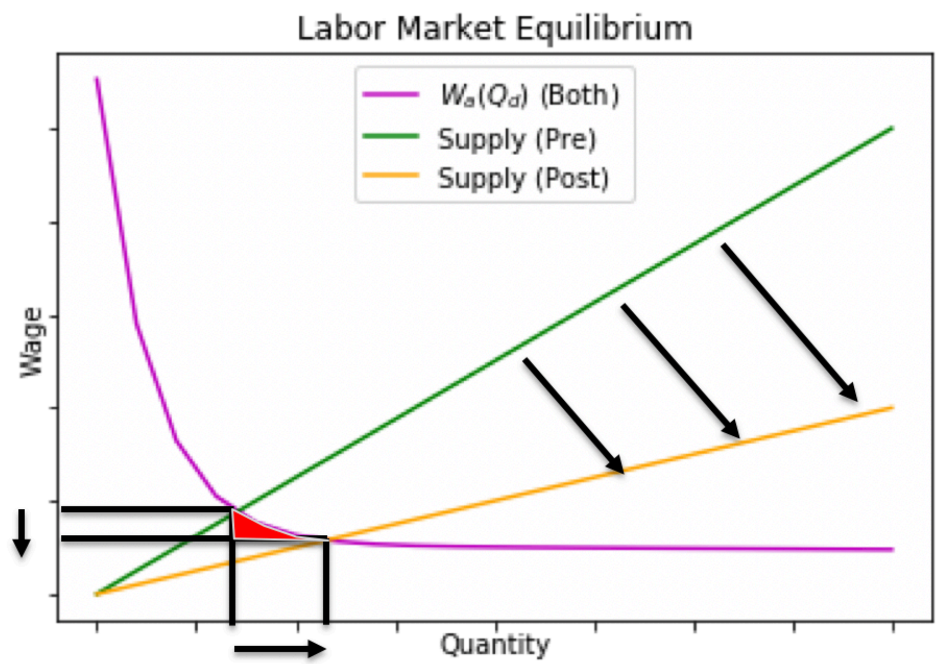


Figure 2 – Labor Market Equilibria with Firm-Specific Tasks and Increasing Skill Supply

4. Dataset Construction

Member profile information from LinkedIn serves as my main data source. Part of the reason the firm value of engineering and technological talent has been difficult to measure in the past is because of a relative paucity of granular data in this area. Online platforms like LinkedIn present an opportunity to tie organizations to the skills, education, career histories, and professional networks of their staff. Outside of governmental and administrative datasets, data at this scale and level of detail is unusual. LinkedIn has over 575 million members in over 200 countries and territories (more than 150 million U.S. members, 15 million in Canada, and 25 million in the U.K.). Additionally, over 26 million companies, 60 thousand schools, and 35 thousand skills are represented on LinkedIn.[15] The LinkedIn platform has become a standard tool for job seekers in many labor markets. With over 180 million individual position records spanning from 2000 to 2017, I build firm-level aggregates of worker years of education as well as counts and total wage bills (employee counts multiplied by Bureau of Labor Statistics average wages) of specific varieties of worker. Engineering, information technology, and research worker talent counts and wage bills are some of the aggregated firm variables I construct. The process is similar to the variable construction in (Brynjolfsson et al. 2018; Benzell, Lagarda, and Rock 2018). Figures 3a-c show a representative LinkedIn profile.

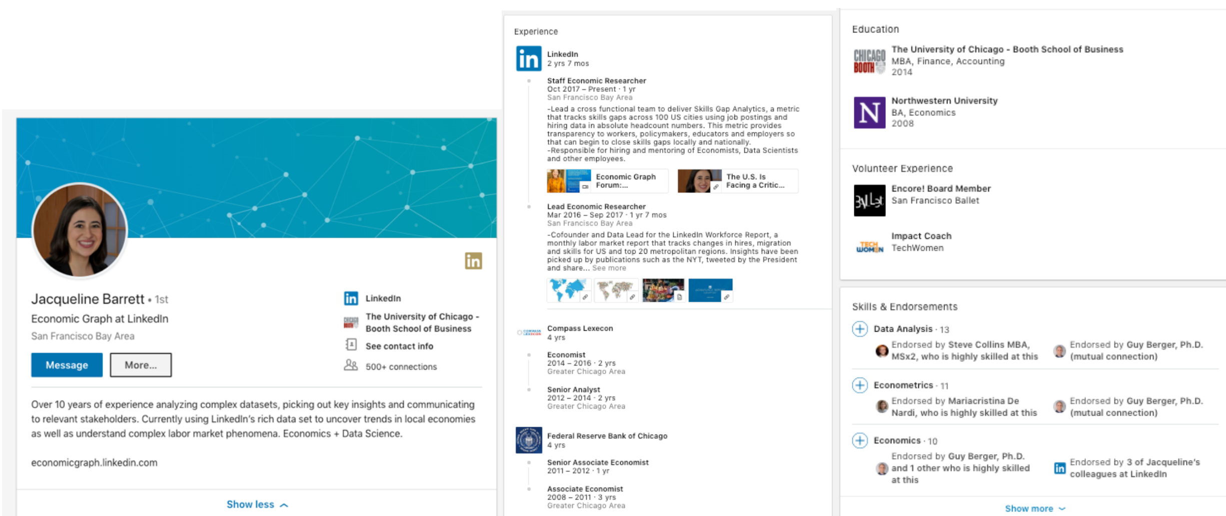


Figure 3a, 3b, 3c – LinkedIn Profile, Experience, Education and Skills

The LinkedIn data covers a substantial portion of the global knowledge and human capital-intensive worker population. The representativeness of the LinkedIn panel is imperfect, with predictably sparser coverage of smaller (non-public) organizations, less educated workers, blue-collar workers, and non-U.S. firms. Further the sample quality varies by year as LinkedIn's

[15] Source: The LinkedIn Economic Graph Research team. About 70% of platform membership is outside the U.S. The growth rate of membership is approximately 2 members per second as of July 2018.

adoption diffused through the workforce. While there are data going back substantially farther than 2000, the coverage at that point relies upon members populating their pages with highly detailed work histories. Additionally, the incentives governing whether to post certain information differ across workers. The selection of workers observed on LinkedIn is likely to differ in meaningful ways from the underlying employee population. Workers seeking employment, for example, are more likely to have updated employment history and skills information on their profiles.

I pursue a number of strategies to mitigate these potential sources of bias. The simplest is the inclusion of combinations of firm, industry-time, and time fixed effects in all regression specifications. In all specifications, however, I correct for occupation, year, and firm-based discrepancies between LinkedIn and administrative labor datasets from the Bureau of Labor Statistics Occupational Employment Survey (BLS-OES). The BLS-OES survey provides detailed industry-level measures of occupational employment and wage. As in Brynjolfsson et al. (2018) and Benzell, Lagarda, and Rock (2018), I build a crosswalk between LinkedIn's internal occupational classification system and the BLS-OES Standard Occupational Classification (SOC) Code by year. For firm-level aggregate employment data, I use the Compustat/Capital IQ North America database value of EMP. In the case that the EMP value is missing or erroneous, I substitute the predicted value of EMP from a linear model trained on known EMP values of the following form[16]:

$$\widehat{EMP}_{it} = \alpha + LI_{it}'\beta + \gamma_{jt} + \epsilon_{it} \qquad (10)$$

The predicted EMP for firm *i* in year *t* is a function of the intercept, the LinkedIn total count for that firm in that year, a fixed effect for that industry-year combination, and an error

[16] Prediction accuracy gains from models with higher complexity (e.g. tree-based models or support vector machines) were relatively small

term. With knowledge of the total firm-year varying employment, the industry classification (3-Digit NAICS Code), the LinkedIn employment counts by LinkedIn occupational category, and the industry-level employment composition according to the BLS-OES, I build a firm-year-occupation-level coverage ratio for all of the publicly traded firms in Compustat/Capital IQ. Whereas omitting the occupational coverage differences within firm implicitly assumes all workers in the same firm face the same incentives to post information to their profile, this adjustment assumes that all workers with the same occupation in the same firm in the same year are subject to similar data supply incentives. Firm-level differences and year-level differences in coverage are even more substantial, and handled by this procedure. Meanwhile this adjustment does make a potentially significant assumption that workers employed by U.S. publicly traded firms but working elsewhere are employed in similar proportions to the BLS-OES industrial occupational employment shares. The appendix has the regression results for equation (10) in Table A1. Typically firms have about 1.9 times as many employees as are available on LinkedIn, controlling for the asset base size and industry-year.

In detail, first I take the occupational employment shares by industry-year from the BLS-OES. I then calculate the industry-employment shares by industry from Compustat using either EMP or predicted EMP from (10). Re-weighting the BLS-OES occupation-industry-year shares by the Compustat industry-year shares and summing by occupation yields the Compustat occupation-year shares. These Compustat occupation-year shares are multiplied by total Compustat employment (emp or predicted emp) to get the total Compustat employment by occupation-year. The total employment by occupation in publicly traded firms on LinkedIn is compared to this Compustat employment by occupation value to get a job-year-level coverage value λ_{jt} for the proportion of Compustat employment in job j and year t captured on LinkedIn.

The total LinkedIn count in year t at the firm i is then divided by the total Compustat employment in that firm to get θ_{it}, the firm-year coverage ratio. Multiplying these two factors is analogous to flipping two biased coins – one for if the worker in firm i is captured by Compustat and LinkedIn, and another for if the worker with job j is on Compustat and LinkedIn. Since these coverages will double-count the employment weighted average coverage ratio by firm $\overline{\theta_{Jt}}$, we divide that out such that total adjusted LinkedIn employment is equal to total Compustat employment. The relatively simple normalization function to convert observed LinkedIn occupation-firm-year counts into BLS-OES-Compustat standard occupation-firm-year counts is as follows:

$$LI_{ijt} = \frac{(\theta_{it}\lambda_{jt})}{\overline{\theta_{Jt}}} Compustat_{it} \qquad (11)$$

The end result is Compustat-BLS-OES-consistent firm-year-occupation employment coverage ratios. Occupations like software engineer, unsurprisingly, have high fidelity and near complete coverage for U.S. firms. A few other titles, like dentist or transportation specialists, have lower baseline levels of coverage but are adjusted to BLS-OES consistency with this process. Nevertheless, the occupations and firms for which LinkedIn membership is relatively sparse will have noisier adjusted employment shares as well. To handle these issues as well as to effectively tackle the research questions in the paper, fixed effects at the industry, year, and firm-level are included in regression specifications. For engineering, research, and IT worker stocks, the relative presence on LinkedIn is higher in comparison to other occupations.[17] LinkedIn defines Engineering, Information Technology, and Research as separate functional areas within a

[17] IT, Research, and Engineering are all defined as functional areas by LinkedIn. Workers of specific job titles in specific job functions are mapped into these functional areas. I aggregate counts, wage bills, and human capital after applying the normalization procedure detailed above.

firm. When members submit their profile information, they are additionally classified into a given functional area. Occupations are distributed across these different domains, not always into the same functional area of the company. Software engineers are most frequently included in the Engineering category (as are most occupations with "engineer" in the title), but may also be categorized in Information Technology. I calculate the total employee counts in each of these different categories. The normalized counts of workers are taken as the output of the adjustment represented in (11). Those employee counts are multiplied by their BLS-OES wage in the relevant respective year to construct the wage bill variables.

I also construct a total education years variable for each firm in each year as a control for the overall level of human capital at each firm. For this variable, following Brynjolfsson et al. (2018), I aggregate the educational records of the workers according to the years of education required to achieve each listed degree.[18] That is, an Associate's degree counts as two years, a Bachelor's degree counts for four years, a Master's degree counts for two years, a research doctorate or medical doctor degree counts as six years. High school, for an alternative measure of education years, is counted as 12 years. These values are adjusted for coverage in the procedure above, and summed by firm-year to generate a total education years control.[19] Descriptive statistics for the LinkedIn human capital measures for 2006 and 2016 can be found in Table 1 below.

[18] The normalization of education years to adjust for coverage is an identical process to the count normalization process.
[19] Can be considered as the answer to the question "how many years has the firm gone to school?"

2016 Summary Stats	Market Value (MM USD)	Total Assets (MM USD)	IT Employment	Research Employment	Engineering Employment	IT Wage Bill	Research Wage Bill	Engineering Wage	Education Years
Count	7,365	7,365	1,828	1,828	1,828	1,828	1,828	1,828	2,249
Mean	17,081.06	14,133.41	972.40	198.57	1,556.18	96,903,181.60	27,533,898.31	120,046,634.46	35,189.00
Standard Deviation	112,933.43	109,319.50	4,024.43	692.30	4,759.50	405,197,401.23	93,265,923.08	387,763,106.94	115,419.34
0.25	70.24	40.51	35.01	5.21	42.96	3,374,545.19	827,245.57	3,142,992.88	1,053.44
0.50	694.88	436.15	140.77	22.91	239.57	13,708,210.73	3,350,099.03	16,838,521.82	5,495.35
0.75	4,260.37	2,712.24	539.67	109.90	977.20	52,722,407.66	15,822,372.62	70,194,131.19	23,867.57

2006 Summary Stats	Market Value (MM USD)	Total Assets (MM USD)	IT Employment	Research Employment	Engineering Employment	IT Wage Bill	Research Wage Bill	Engineering Wage	Education Years
Count	8,453	8,453	2,211	2,211	2,211	2,211	2,211	2,211	2,994
Mean	11,290.06	9,080.16	517.95	128.16	1,089.52	35,424,886.64	9,797,113.28	63,189,456.84	24,032.11
Standard Deviation	85,117.95	80,507.76	2,033.99	567.22	3,935.39	141,847,510.53	34,689,741.29	242,870,484.33	90,667.33
0.25	77.53	35.89	16.87	2.55	26.54	1,126,353.06	286,148.23	1,584,233.32	600.57
0.50	441.37	246.90	69.90	11.11	128.77	4,666,854.58	1,162,623.08	7,112,400.35	2,647.09
0.75	2,282.32	1,394.58	278.49	51.24	543.20	18,859,136.17	4,921,754.10	29,885,631.95	13,434.37

Table 1: Descriptive Statistics of Employment Measures 2006 and 2016

The final set of LinkedIn-derived values come from the relatively recently constructed panel of detailed skills detail. LinkedIn first rolled out the skills product in 2011, though collection of high-fidelity records of member additions of skills began in 2014. Recently, LinkedIn has categorized and standardized the over 35,000 unique skills on its standard platform into a set of skills clusters using nonlinear embedding spaces.[20][21] These clusters are seeded by humans and subsequently applied to co-occurrences of skills on profiles across the entire platform. Skills are related by distance in "skill space" as a result of this machine learning-driven encoding. Skills that tend to be closer in this space are more likely to be associated together and tagged with a common human-curated cluster name. Likewise, skills that co-occur less frequently are classified in separate clusters. I make use of the production neural skills embeddings supplied by the LinkedIn engineering team.

The result is a series of aggregated counts of skills additions in different categories which I then aggregate, accumulate, and normalize at the firm-year-occupation and firm-year levels. I also extract specific skill counts for deep learning, machine learning, R, SPSS, and a handful of other data science skills. All of these measures are then joined to Compustat measures of

[20] Clusters including Agronomy, Artificial Intelligence, People Management, and Digital Literacy (amongst others) and rely upon user-supplied data. Because the user-supplied data is highly variable, all skills go through a standardization algorithm before being made available for analysis.
[21] See https://en.wikipedia.org/wiki/Nonlinear_dimensionality_reduction for a set of useful embedding algorithms. TensorFlow can be used to build some of these models.

financial performance by firm and year. Figures 4A-D show the aggregate skill additions for AI-related skills and advertising. There is some seasonality in the data, with more skills getting added in the beginning of the year. The table below shows some example skills for different aggregated categories.

Skills	ArtificialIntelligence	DigitalLiteracy	DataScience	Advertising	BusinessManagement	DataStorageTechnologies
1	artificialintelligence	microsoftexcel	forecasting	advertising	management	microsoftsqlserver
2	machinelearning	microsoftword	modeling	campaigns	strategy	mysql
3	classification	microsoftpowerpoint	statistics	collateral	strategicbusiness	sql
4	informationretrieval	microsoftoffice	analytics	sponsorship	smallbusiness	forecasting
5	computervision	microsoftaccess	dataintegrity	directmarketing	strategicplanning	databases
6	neuralnetworks	microsoftoutlook	statisticaltools	searchenginemarketing(sem)	changemanagement	datacenter
7	speechrecognition	email	dataanalysis	branddevelopment	executivemanagement	storage
8	semanticweb	spreadsheets	sas	mediaplanning	serviceproviders	datawarehousing
9	parsing	mac	datamining	emailmarketing	outsourcing	hpproducts
10	patternrecognition	lotusnotes	sampling	media&entertainment	businessplanning	pl/sql

Example Skills Table

Note: Example skills are not exhaustive – some categories have hundreds of skills.

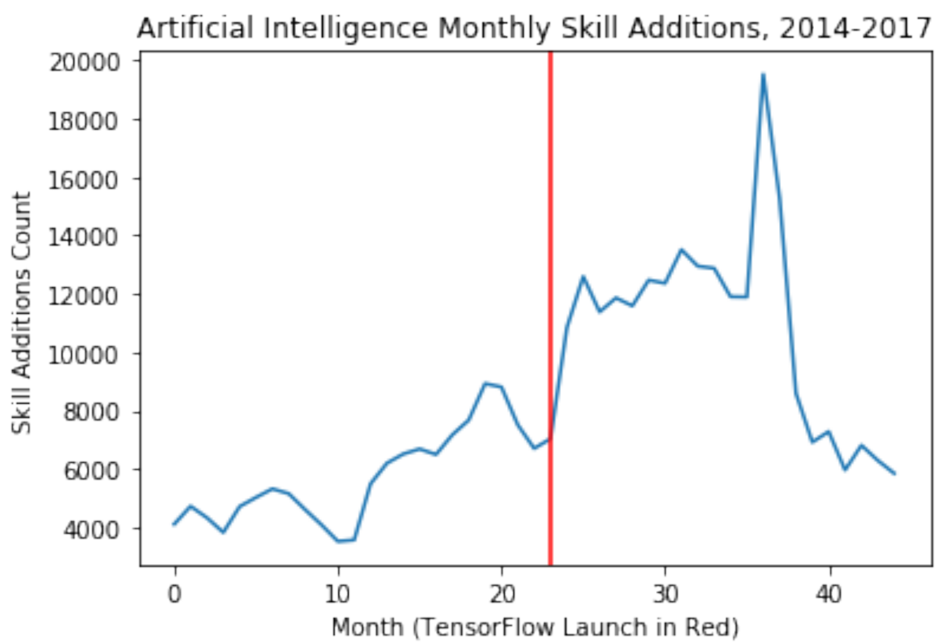


Figure 4A – Artificial Intelligence Skills 2014-2017

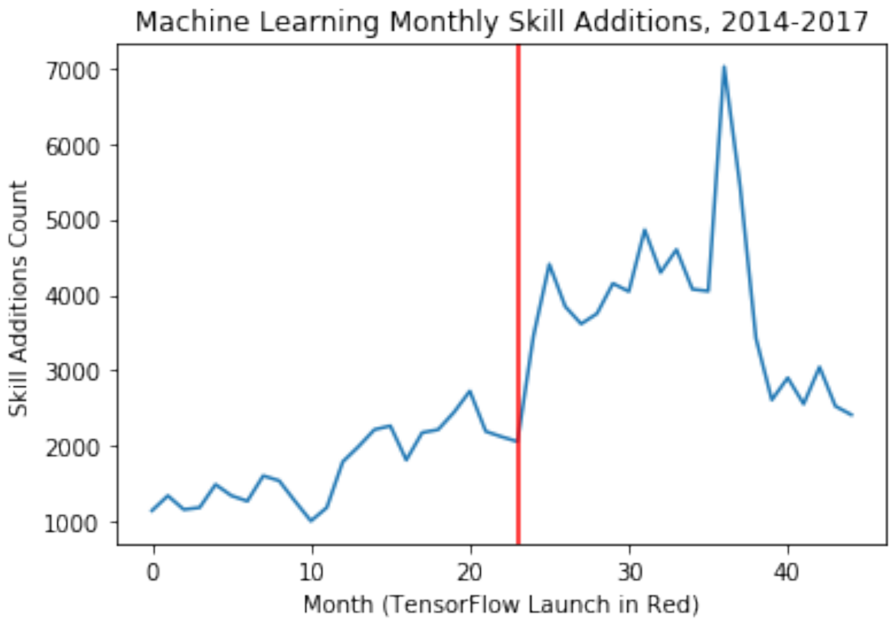


Figure 4B – Machine Learning Skills 2014-2017

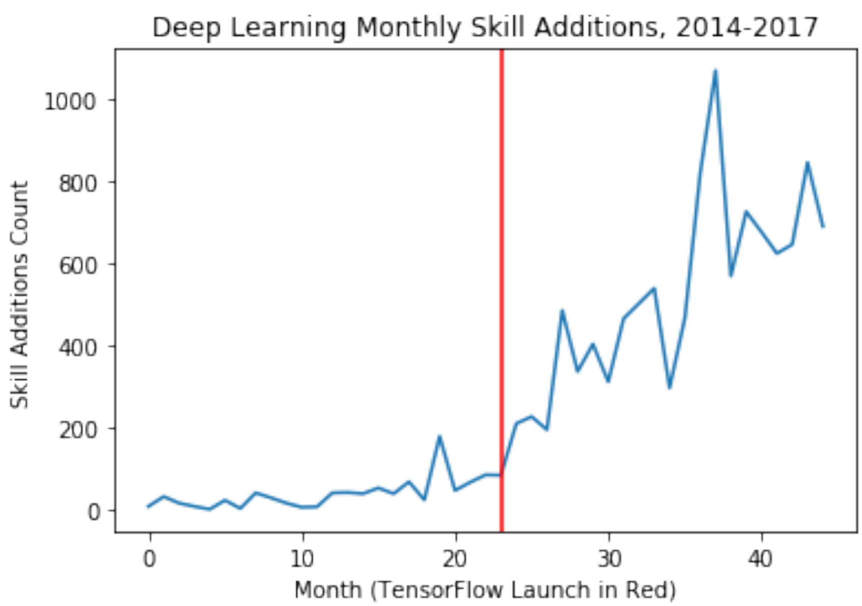


Figure 4C – Deep Learning Skills 2014-2017

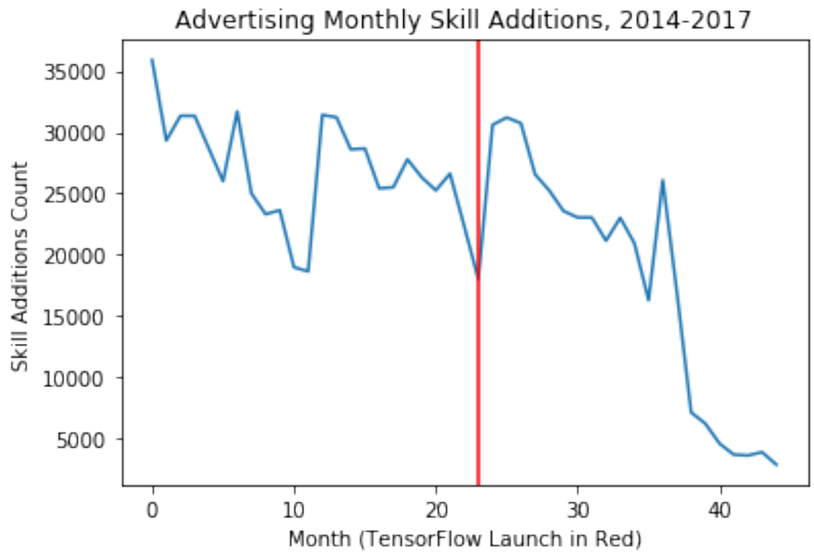


Figure 4D – Advertising Skills 2014-2017

Compustat also indicates the state, zip code, and county for the corporate headquarters. I link these states to changes in covenant-to-not-compete (CNC) policy changes at the state-level using the set of states in (Ewens and Marx 2017) and (Jeffers 2017) separately. These changes in CNC policy are known to impact knowledge workers with greater propensity than other kinds of employees (Balasubramanian et al. 2018), and they constitute a first-order supply-side shock as mobility restrictions on human capital when the enforceability of these contracts increases. CNC enforceability changes should therefore serve as an instrument for the technical talent hired by the firm. Following Ewens and Marx (2017) and Jeffers (2017), I code increased enforceability by state courts as a 1, unchanged values as a 0, and weakened enforceability as -1.

I also merge the corporate headquarters zip and county codes to the zip and county codes of the land-grant universities. Land-grant universities, established by the Morrill Acts of 1862 and 1890, provided for the creation of colleges in each state following the sale of federal lands.

As in Moretti (2004), worker proximity to land-grant institutions predicts higher likelihood of human capital accumulation. To attempt to recover a causal estimate of the market value of technological workers controlling for overall human capital, I include a dummy variable in instrumental variables specifications for proximity to land-grant institutions of the corporate address at the county level.

For the final instrument set, following (Lucking, Bloom, and Van Reenen 2017; Lucking 2018), I match each firm-year to the logged user cost of R&D capital value for both the firm and the best matching state. R&D investment subsidies for each state-year and firm-year lower the cost of corporate R&D investments. Specifications logging these rates and including time fixed effects leave only the state or firm-level variation net of changes in overall capital costs. Since R&D stocks are a known channel for these cost decreases to increase market value, I also include R&D stocks at the firm-level calculated from Peters and Taylor (2017). The conditional effect estimated is therefore adjusted for the innovative activities of the firm, leaving only non-innovative engineering.

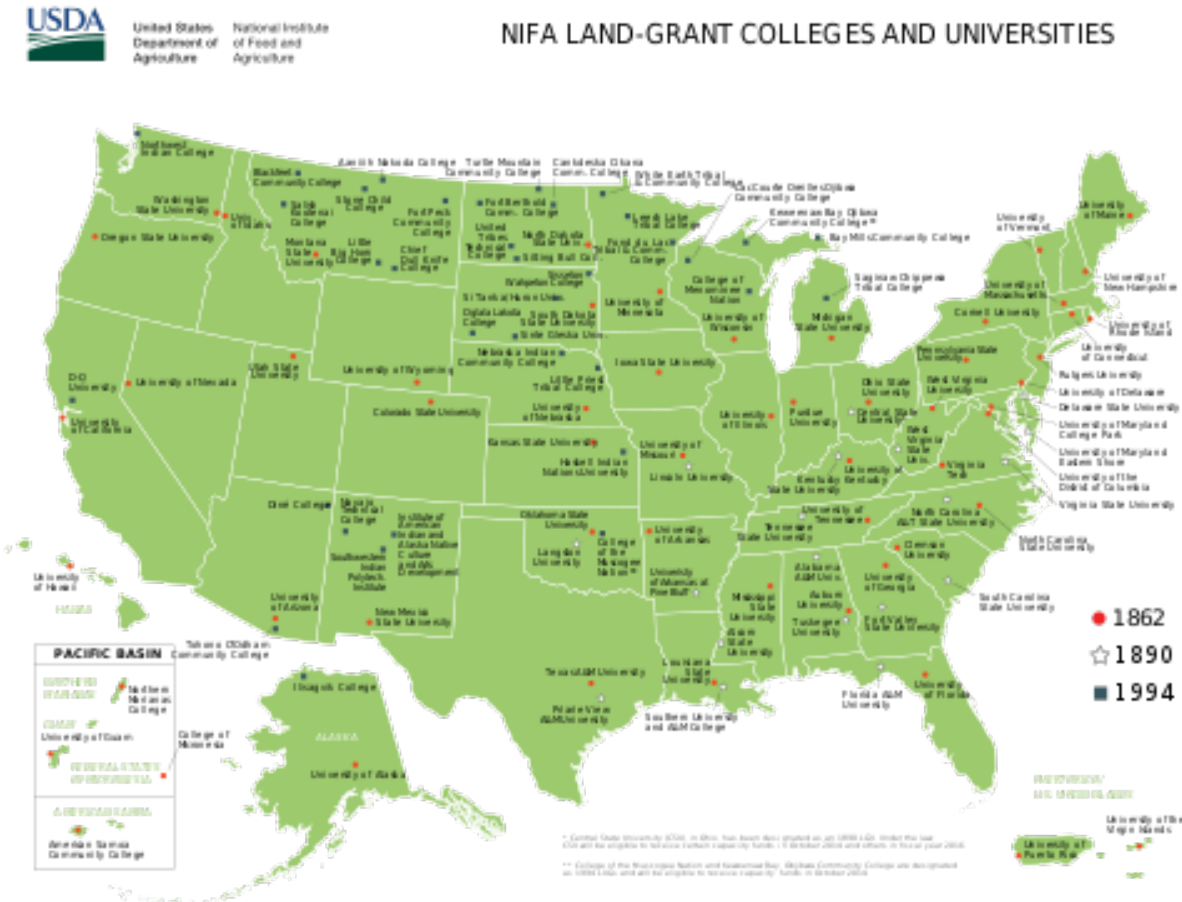


Figure 5 – Land-grant Colleges

The Compustat component of the dataset is mostly taken as-is, with market value (MV) constructed as the total book value of the firm plus the market value of equity at the end of the fiscal year less the book value of common equity. Total Assets (TA) is included in most regression specifications as a control for the capital size of the firm.

I also join in measures of *Suitability for Machine Learning* (SML) at the occupation level. SML measures are from Brynjolfsson, Mitchell, and Rock (2018a), wherein they use the crowdsourced evaluation of thousands of rubric surveys to construct measures of task SML and measurability for each of the detailed work activities (DWAs) supplied in the O*NET database (Brynjolfsson and Mitchell 2017). High relative values of SML indicate an opportunity to use machine learning (and deep learning in particular) to automate aspects of a task. These scores are aggregated across tasks to the occupational level, and I subsequently aggregate the occupation-level SML scores to the firm-year by generating wage bill-weighted averages of SML scores. If TensorFlow is a shock to the availability of talent, it is also a shock to the opportunities for automation (but not necessarily at the same firm). With measures of which firms have an opportunity to deploy deep learning talent and which firms have an opportunity to use what deep learning engineers build, I can address whether it is the firms who create or the firms who consume technology who capture the rents to technology's effects on labor.

5. Overall Engineering: Empirical Results and Discussion

Technological human capital assets, if they contribute to the market value of the firm, can be priced following an equilibrium relationship that the asset's marginal Q-value above the asset replacement cost must equal the marginal adjustment costs of competitors (Wildasin 1984; Hayashi and Inoue 1991; Hall 2001). Summing over assets within firm and year, we get that the market value of the firm is equal to the sum of the market value of its constituent assets priced at Q. In other words, a regression of market value on measures of the assets of the firm will recover as coefficients the value per unit (dollars) for the asset in equilibrium. To find the value of different types of technological labor, I estimate the coefficient vector for the following regression:

$$MV_{it} = \beta * \text{Total Assets}_{\text{it}} + \gamma * HK_{total_{it}} + \sum_{m=1}^{M} \delta_m TechHK_{mit} + \boldsymbol{Z}'_{\boldsymbol{it}}\boldsymbol{\lambda} + \epsilon_{it} \qquad (12)$$

In this regression, *i* indexes the firm, *t* indexes the year, and *m* indexes the variety of technological labor. The equation therefore describes the decomposition of the panel of firm market values on the total book value of assets (Total Assets), the total education years in the firm at year t (HK_total), and each type of technological labor TechHK.[22] A vector **Z** of controls including industry-year fixed effects and firm fixed effects in some of the specifications are also included. This is a standard market value regression of the sort in (Hall 1993; Brynjolfsson, Hitt, and Yang 2002; Brynjolfsson et al. 2018). However, as suggested by the results in equations (4) and (9), the market value of the firm is also a function of the integrated differences between worker marginal products and worker wages. Equation (12) therefore sets a specification by which we can decompose market value into the value of observable assets and the value of labor-correlated inputs. The coefficient vector for technological talent recovers the "installed" average price of the talent itself and the value of omitted yet talent-correlated assets.

The results for the set of OLS regressions from (12) are in Tables 2 (counts) and 3 (wage bills), revealing strong correlations between market value and engineering talent where firm fixed effects are left out. While IT workers and Research workers appear not to be correlated with market value after controlling for the total education years attained by the firm and the total book value of assets, along with the appropriate fixed effects, an additional engineer is correlated with an increase of approximately $855,000 of market value (Table 2, column 5), and each dollar of engineering wage bill is correlated with an additional $11-12 of market value (Table 3, column 5). However, when including firm-level fixed effects, the correlation between all types of technological talent and market value is no longer statistically significant and the point estimates on engineering talent drop. They are not statistically different than the industry-year estimates. Firm fixed effects effectively adjust for the aspects of the firm that allow it to successfully generate value. Including firm fixed effects, in effect, results in an estimate of market value as though the firm were competitive with itself. That is, the variation which generates value is

[22] Depending on specification, wage bill or counts might be included here

cross-sectional (and over time). The within-firm estimates might be thought of as a lower bound if corporate asset bases are relatively static.

This is consistent with an explanation of firm-specific assets driving the returns to investments in technological labor. That is, there is either a valuable (priced) intangible correlate asset for technological labor that is firm-specific generating a correlation between market value and engineering labor as an omitted variable, or the component of engineering labor that causes market value is firm-specific and time-invariant.[23] That firm fixed effects drive the coefficient on engineering talent to lose statistical significance is consistent with either omitted firm-specific assets being the primary source of the correlation between market value and engineering talent or the marginal value of firm-specific labor being close to zero. The former explanation suggests off-balance sheet capital is the source of the empirical relationship. This latter explanation would describe a marginal causal effect of engineering talent on market value. Both explanations are consistent with monopsony power coming from the firm-specificity of human capital.

The wage regressions (Table 3) compare the flow of wages to workers of a given type to the market value stock outcome variable, equal to the present value of all future cash flows. Since the wages are what the worker earns, the coefficient on the wage is the present value per \$1 of wages paid in a given year of *all future flows* to the firm. That is, the wage is what workers earn in a year; the table coefficient is the discounted value of what the firm will get in all future years. This inflates the coefficient value because each year's wages must represent all future wages to be comparable to the firm's value. Either representing wages as a stock or representing market value as a one-year flow fixes the comparison problem. Unfortunately, the former approach requires knowledge of the initial stock of firm capital in wages and the depreciation

[23] As a sanity check, the correlations between total assets and market value are close to replacement cost (\$1) and the market value of human capital is positive and statistically significant.

rate (to calculate a perpetual inventory) and the latter requires knowledge of the appropriate discount rate for engineering capital's share of market value. As a back-of-the-envelope calculation, the stock of assets in wages paid to technological talent is 14.3 to 20 times the wage value.[24]. This would imply that engineer wages as a stock are worth about 59 to 83 cents per dollar net of what the worker is paid if there are no omitted variables. In the case that marginal and average products are equal to wages, the 59 to 83 cents estimate is the value of off-balance sheet assets correlated with the presence of engineering talent.[25]

Table 2: Market Value – Worker Count Regressions	(1) No Tech	(2) IT Value	(3) Engineering Value	(4) Research Value	(5) All Tech Value	(6) All Tech Value (Firm FE)	(7) Tech w/o HK Value
Total Assets	1.012***	1.012***	1.013***	1.014***	1.012***	1.003***	1.004***
	(0.00768)	(0.00691)	(0.00627)	(0.00751)	(0.00615)	(0.0122)	(0.0119)
Total Years of Education	0.0187***	0.0138***	0.0125***	0.0125**	0.00622	0.00873	
	(0.00451)	(0.00429)	(0.00362)	(0.00524)	(0.00448)	(0.00988)	
IT Employees		1.301**			0.496	-0.111	0.0737
		(0.498)			(0.475)	(0.302)	(0.273)
Engineering Employees			1.129***		0.855***	0.567	0.817*
			(0.289)		(0.276)	(0.422)	(0.478)
Research Employees				7.164*	6.745	0.843	2.208
				(4.231)	(4.428)	(2.288)	(2.391)
Observations	50,501	37,813	37,813	37,813	37,813	37,825	37,825
R-squared	0.984	0.984	0.984	0.984	0.984	0.994	0.994
Industry-Year FE	Yes	Yes	Yes	Yes	Yes	No	No
Firm and Year FE	No	No	No	No	No	Yes	Yes

Table Notes: Robust standard errors in parentheses, Standard errors clustered at the industry (3-Digit NAICS) for columns 1-5, firm for columns 6-7. Market value is in millions USD.

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

12

[24] Assuming an aggregated human capital depreciation rate of between 5% and 7%, the approximate rate of interest for acquiring human capital: https://studentaid.ed.gov/sa/types/loans/interest-rates

[25] Actual discount rates for this kind of talent vary by firm. Some firms will have much higher (lower) discount rates, in which case the employer share of wage value would be substantially higher (lower).

Table 2 – OLS Market Value Regressions on Worker Counts

Table 3: Market Value – Wage Bill Regressions	(1) IT Value	(2) Engineering Value	(3) Research Value	(4) All Tech Value	(5) All Tech Value	(6) Tech w/o HK Value
Total Assets	1.012***	1.013***	1.011***	1.010***	1.003***	1.004***
	(0.00693)	(0.00638)	(0.00690)	(0.00609)	(0.0122)	(0.0119)
Total Years of Education	0.0148***	0.0134***	0.0136***	0.00985***	0.00912	
	(0.00431)	(0.00366)	(0.00424)	(0.00359)	(0.00931)	
IT Wage Bill	1.30e-05**			1.68e-06	-5.13e-06	-3.63e-06
	(5.10e-06)			(3.92e-06)	(3.38e-06)	(2.89e-06)
Engineering Wage Bill		1.57e-05***		1.19e-05**	1.02e-05	1.33e-05*
		(4.75e-06)		(4.93e-06)	(6.62e-06)	(7.57e-06)
Research Wage Bill			7.32e-05**	6.04e-05	9.91e-06	1.51e-05
			(3.20e-05)	(4.00e-05)	(1.17e-05)	(1.13e-05)
Observations	37,813	37,813	37,813	37,813	37,825	37,825
R-squared	0.984	0.984	0.984	0.984	0.994	0.994
Industry-Year FE	Yes	Yes	Yes	Yes	No	No
Firm and Year FE	No	No	No	No	Yes	Yes

Table Notes: Robust standard errors in parentheses

Standard errors clustered at the industry (3-Digit NAICS) for columns 1-4, firm for columns 5-6. Market value is in millions USD. The wage bill is equal to the prevailing wage for a given occupation-year grouping, where the occupational wage is the employment-weighted average wage of all BLS-OES occupation categories within a given LinkedIn occupational category, matched many-to-one. The wage bill is then a *flow* measure and the market value is a *stock* equal to the present value of all future flows.

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

Table 3 – OLS Market Value Regressions on Worker Wage Bills

In the case that the human capital value is firm-specific, a causal shift in employee supply driven by CNC enforceability tightening, for example, would suggest market values should increase. In this scenario, the policy change forecloses on employment alternatives for employed workers. On the margin, newly lower opportunity costs for employees might make further capital accumulation attractive. If the instrumented engineering labor does not cause market value, it is suggestive evidence that (for compliers), the correlation between market value and engineering talent is driven by hidden intangible assets owned by the firm. If this is instead the case, the valuation of off-balance sheet assets correlated with engineering talent values is less sensitive to employee opportunity cost changes. Monopsonists will hire less than the socially optimal amount if they have to pay all employees the same wage. They receive a mark-down on wages. Supply-increasing shocks will increase both the market value of the firm and the quantity of labor hired if wages for existing workers are renegotiable. Wages might be sticky though. The fact that CNCs also affect existing workers by shifting the optimal balance of firm-specific and general tasks means that even if the firms keep their hiring the same, the market value of firms will respond to CNC policy changes.

As described above, I will use the proximity of the corporate headquarters to land-grant universities and changes in state-level covenant-to-not-compete (CNC) policy as instruments for the firm's stock of human capital (in the land-grant case) and as supply shocks for engineering talent wherein increased (decreased) enforceability reduces (increases) the quality of workers' outside options (in the CNC case). The nature of the instruments does not allow inclusion of firm fixed effects in the specification as the covariate matrix will be of deficient rank. But interacting the CNC policy changes with the land-grant proximity dummies permits an overidentified model. In all specifications, the weak identification test F statistic is well above 20 and the overidentification test (a joint test of local average treatment effect homogeneity and correlation of residuals and instruments) fails to reject the null. This is reassuring, though not dispositive, with respect to endogeneity concerns of some subset of the instruments. The second stage results for engineering employee counts and wage bills are in Tables 4A and 4B (respectively). The first stage results are reported in Tables 5A and 5B.

	(1)	(2)	(3)	(4)	(5)

Table 4A: IV MV Regressions: Engineering Labor	Land-grant	CNC-Jeffers	CNC-Ewens-Marx	LG+CNC-Ewens-Marx	LG+CNC-Ewens-Marx
Engineering Employees	0.965	4.297	0.404	1.067	1.484
	(1.973)	(33.29)	(5.717)	(1.844)	(1.692)
Total Assets	1.035***	1.023***	1.027***	1.035***	1.032***
	(0.0113)	(0.0339)	(0.0117)	(0.0110)	(0.0104)
Total Years of Education	0.0139	-0.00137	0.0162	0.0134	0.00232
	(0.0102)	(0.151)	(0.0263)	(0.00947)	(0.00423)
IT Employees					1.052
					(1.684)
Research Employees					6.740
					(4.447)
Observations	31,822	32,622	32,622	31,375	31,375
R-squared	0.973	0.976	0.977	0.973	0.974
Industry-Year FE	Yes	Yes	Yes	Yes	Yes

Table Notes: Robust standard errors in parentheses, SEs Clustered by 3-Digit NAICS. Market value is represented in millions USD.

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

Table 4A – IV Market Value Regressions on Engineering Worker Counts

	(1)	(2)	(3)	(4)	(5)
Table 4B: IV MV Regressions: Engineering Wage Bill	Land-grant	CNC-Jeffers	CNC-Ewens-Marx	LG+CNC-Ewens-Marx	LG+CNC-Ewens-Marx
Engineering Wage Bill	1.51e-05	-5.60e-05	7.74e-06	1.68e-05	2.12e-05
	(3.10e-05)	(0.000524)	(0.000109)	(2.90e-05)	(2.87e-05)
Total Assets	1.035***	1.032***	1.027***	1.035***	1.029***

	(0.0112)	(0.0448)	(0.0130)	(0.0109)	(0.00974)
Total Years of Education	0.0141	0.0332	0.0159	0.0137	0.00452
	(0.00972)	(0.142)	(0.0300)	(0.00907)	(0.00339)
IT Wage Bill					7.27e-06
					(1.98e-05)
Research Wage Bill					9.88e-05*
					(5.55e-05)
Observations	31,822	32,622	32,622	31,375	31,375
R-squared	0.973	0.963	0.977	0.973	0.975
Industry-Year FE	Yes	Yes	Yes	Yes	Yes

Table Notes: Robust standard errors in parentheses, SEs clustered by Industry (1-3), Firm (4-5). Wage bill is calculated as in Table 3 (see table notes).

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

Table 4B – IV Market Value Regressions on Engineering Wage Bill

Table 4C: R&D User Cost IV (Engineer Counts)	(1)	(2)	(3)	(4)
Log Market Value	User Cost: State	User Cost: All	User Cost: State	User Cost: All
Log Engineers	-1.251	-1.072**	-0.165	-0.579**
	(0.845)	(0.513)	(0.367)	(0.260)
Log(Assets)	0.979***	0.939***	0.932***	0.904***
	(0.0527)	(0.0487)	(0.0231)	(0.0448)
Log(Years of Edu.)	0.676	0.656**	0.119	0.420**
	(0.428)	(0.289)	(0.198)	(0.163)
Log(R&D Capital)	0.109***	0.128*	0.0557***	0.100***
	(0.0397)	(0.0638)	(0.0161)	(0.0291)
Log(IT Workers)	0.467	0.351	0.0279	0.139
	(0.357)	(0.212)	(0.148)	(0.0873)
Log(Researchers)	0.0551	0.0143	0.0361*	0.0313
	(0.0477)	(0.0454)	(0.0203)	(0.0247)
Observations	16,252	12,450	16,246	12,441
Industry-Year FE	Yes	Yes	Yes	Yes
City FE	No	No	Yes	Yes

Standard errors clustered by industry (3-Digit NAICS) in parentheses

$^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.1$

Table 4C – IV Market Value Regressions on Engineer Counts Instrumented by R&D User Costs

The causal estimates for the land-grant and CNC local average treatment effects (LATEs) of engineering human capital on market value are imprecise and statistically indistinguishable from zero. While the point estimates are larger ($7.7 to $21.2 per dollar of wages for specifications including the Ewens-Marx CNC data), they are also not statistically significantly different than the fixed effect OLS estimates.

Firm fixed effects soak up the variation generating a statistically significant relationship between market value and engineering talent measures. Broadly, this indicates that the average product of engineering talent in firms is weakly larger than the wages paid over all employed engineers.[26] Nevertheless, there is little case that employers can extract freely available rents by hiring more engineers. The imprecisely estimated zero marginal value of an engineer net of wages is consistent with value at or below the opportunity cost of recruiting search frictions, training. and adjustment costs for employers. That engineers are so highly correlated on average with market value is suggestive of complementarities between engineers and firm-specific assets. The land-grant instrument is designed to approximate (statically) a measure of the available supply of engineers to the firm. Ideally there would be an experiment randomly assigning assets and workers of different types to different companies. Under the assumption that sharing county locations with land-grant universities (as opposed to other universities) is otherwise excluded from market value, the land-grant IV detects contribution of wage changes to firm market value. The first stage suggests that, controlling for worker education and the other inputs, firms near a

[26] And strictly larger for some engineers

land-grant college are less likely to hire more engineers. This could be, for example, because human capital-intensive firms far from land-grant colleges hire relatively more engineers. This could be the case if the first high human capital employees are engineers for most firms, and on the extensive margin other types of educated workers are hired.

CNC policy changes, however, shift the outside options of incumbent workers, making it potentially more attractive for firms to invest in firm-specific training for those workers as departure is less likely. At the same time, hiring prospects are somewhat diminished. So CNC changes primarily operate through the price of human capital assets. Since neither of these instruments reveal a conclusive causal effect of engineering employment on market value, even if the coefficients are larger than they are in the OLS case. As a caveat, this is to be expected if equilibria are insufficiently changed by the instruments. Further, the inclusion of firm fixed effects and more precisely estimated lack of statistically significant market value correlation with engineering talent is evidence against a large *within-firm* value of engineering labor investment, but not on *average* in the cross-section. The firm fixed effects control for time-invariant omitted variables. The correlation between market value and engineering talent diminishes within firm. The difference between specifications with firm fixed effects and without them is suggestive then of a possible omitted asset generating rents. A large, statistically significant effect in the IV regressions would suggest the talent itself as the source of rents on the hiring margin. The R&D user cost instrument shows effects in the negative direction, adjusting for the innovative activities of the employers. Table 4C below shows the R&D user cost instrument specifications. These regression specifications have logged market value and logged covariates because of the skewed nature of R&D stocks. Table 5C has the corresponding first stage.

The construction of the R&D user cost instruments is to take the log of the firm and state-level user costs by firm-year.[27] The firm-level user cost varies as a result of federal subsidies for R&D at the firm level, but does not apply to all companies as the policy does not apply to newer firms. The state-level subsidy is applicable only to R&D expenditures within a given state. These state-level costs are matched to the firm locations. Additionally included is an interaction of these two variables. Firms that take up the subsidy will have lower costs of capital for R&D investment and ideally will accumulate larger R&D stocks. Adjusting for these R&D stocks (and other types of assets), the conditional IV estimator represents the causal effect of hiring another engineer *not* engaged in R&D activity. Since R&D expenses have a large salary component, this adjusted estimate represents the value of marginal engineering activity. It is unlikely, however, that any engineer is engaged completely in non-innovative activity.

The estimate value from the R&D user cost specifications suggest that the marginal engineer engaged in non-innovative activity (somewhat surprisingly) destroys nearly $580,000 in value with a standard error of $260,000, adjusting for the corporate city location, innovative activity of the firm, and other assets. This could be due to the conditions under which such an engineer would be hired. Tight labor markets might mean that hiring an engineer under problematic conditions is more an indication that market value is declining. Since location fixed effects do attenuate the coefficient estimate toward zero, the availability of talent is likely an important component of the incentives facing employers and their employees. It may also be the case that more engineers does not always make for easier problem solving and on the margin, removing innovative activity, this is a net drain on the firm's value for the complier companies. Another possibility is that if the value of the engineer is a bundled combination of maintenance

[27] Lucking (2018) has greater detail on the construction of these instruments.

activities which have negative value and innovation activities with positive value, then each marginal engineer could have zero value but some activities create a cross-subsidy. Firms facing fixed job designs would therefore have an incentive to redesign the bundle of tasks in engineering jobs, disentangling innovative from non-innovative job tasks.

The CNC, land-grant, and R&D user cost IV regressions fail to provide a conclusive story. The first stage coefficients on CNC policy changes are not statistically significant, but they are for land-grant university proximity. The technological talent markets seem to be in equilibrium with respect to CNCs. The value of engineering talent is high on average, but nearly zero on the margin of the local average treatment effects given by the CNC and land-grant instruments (as suggested by firm fixed effects and both sets of IV regressions). If land-grant proximity does predict hiring, but shows no statistically significant estimated causal effect of hiring on market value, then on the margin we fail to reject the hypothesis that wages are equal to worker product. Finally, the R&D user cost regressions suggest that the marginal engineer may destroy value once the value of innovative activities is priced.

In the case that labor is relatively abundant, we might expect that the marginal hire is devoted principally to general tasks and contributes little to market value net of wages. In the case of engineers, the average value is high are relatively precise but the marginal value is noisy. On balance, these IV specifications rule out little more than large deviations from zero marginal value. There is little evidence of abundant rents for firms seeking to hire more technological talent *in general.* This is possibly because the firm-specific tasks have already reached the point at which general tasks are more valuable to the employer. This suggests the need to look at more granular cases to understand where employers might be able to hire more and generate value at the same time.

Some talent is bottlenecked, i.e. finding and recruiting workers of that type is very costly. For these workers, it is more likely that they work on firm-specific tasks which are higher marginal value. Workers for which there is a bottleneck will have high marginal and average contribution to market value as the frictions to find and/or train them will create a wedge between their wages and the value they create for their employers. AI talent offers a recent technological case study for what can happen when the market expects a previously bottlenecked talent to become much more abundant.

Table 5A: IV Regression First Stage (Worker Counts)	(1) Land-grant	(2) CNC-Jeffers	(3) CNC-Ewens-Marx	(4) LG+CNC-Ewens-Marx	(5) LG+CNC-Ewens-Marx
Total Assets	0.00178	0.000999	0.00100	0.00165	-0.000599
	(0.00209)	(0.00180)	(0.00180)	(0.00194)	(0.000848)
Total Years of Education	0.00453***	0.00452***	0.00452***	0.00447***	0.00114
	(0.00138)	(0.00137)	(0.00137)	(0.00135)	(0.000931)
CNC (Ewens-Marx)			87.11	64.02	-64.22
			(88.37)	(289.3)	(203.1)
Land-grant County Dummy	-485.4***			-460.9***	-324.6**
	(174.5)			(168.6)	(127.5)
CNC (Ewens-Marx) X Land-grant				36.13	163.4
				(271.1)	(196.6)
CNC (Jeffers)		21.59			
		(112.7)			
IT Employees					0.983***
					(0.175)
Research Employees					0.0738
					(0.262)
Observations	31,822	32,622	32,622	31,375	31,375
Industry-Year FE	Yes	Yes	Yes	Yes	Yes
Firm and Year FE	No	No	No	No	No

Robust standard errors in parentheses, SEs Clustered by 3-Digit NAICS

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

Table 5A – Worker Counts First Stage Regression (Companion Table to Table 4A)

Table 5B: IV Regression First Stage (Engineering Wage Bill)	(1) Land-grant	(2) CNC-Jeffers	(3) CNC-Ewens-Marx	(4) LG+CNC-Ewens-Marx	(5) LG+CNC-Ewens-Marx
Total Assets	130.6	82.06	82.34	122.9	-26.08
	(136.6)	(118.0)	(118.1)	(127.6)	(47.90)
Total Years of Education	272.0***	271.4***	271.5***	268.3***	67.41
	(93.71)	(93.16)	(93.12)	(91.67)	(58.16)
CNC (Ewens-Marx)			4.548e+06	909,957	-7.630e+06
			(6.071e+06)	(1.872e+07)	(1.390e+07)
Land-grant County Dummy	-3.092e+07***			-2.938e+07***	-2.010e+07**
	(1.125e+07)			(1.091e+07)	(9.199e+06)
CNC (Ewens-Marx) X Land-grant				4.899e+06	1.397e+07
				(1.739e+07)	(1.340e+07)
CNC (Jeffers)		-1.658e+06			
		(7.834e+06)			
IT Employees					0.712***
					(0.0776)
Research Employees					0.210**
					(0.0891)
Observations	31,822	32,622	32,622	31,375	31,375
Industry-Year FE	Yes	Yes	Yes	Yes	Yes
Firm and Year FE	No	No	No	No	No

Table Notes: Robust standard errors in parentheses, SEs Clustered by 3-Digit NAICS. Outcome for first stage is the dollar wage bill (wage times worker counts).

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

Table 5B – Worker Wage Bill First Stage Regression (Companion Table to Table 4B)

	(1)	(2)	(3)	(4)

Table 5C: IV Regression First Stage (Engineering Counts)	User Cost: State	User Cost: All	User Cost: State	User Cost: All
Log(Assets)	0.0447**	0.0500*	0.0431**	0.0230
	(0.0215)	(0.0266)	(0.0208)	(0.0184)
Log(Years of Edu.)	0.539***	0.602***	0.539***	0.622***
	(0.0697)	(0.0811)	(0.0678)	(0.0542)
Log(R&D Capital)	0.0378	0.0186	0.0331	0.0351
	(0.0272)	(0.0743)	(0.0202)	(0.0624)
Log(IT Workers)	0.388***	0.359***	0.369***	0.310***
	(0.0396)	(0.0440)	(0.0496)	(0.0503)
Log(Researchers)	0.0154	-0.0200	0.0426	0.0276
	(0.0337)	(0.0489)	(0.0355)	(0.0435)
Log(Firm User Cost)		1.840*		1.339*
		(0.952)		(0.707)
Log(State User Cost)	-0.0314	-0.0208	-0.0636	-0.0751
	(0.0247)	(0.0359)	(0.0634)	(0.0861)
Log(Firm*State)		1.723		1.186*
		(1.140)		(0.654)
Observations	16,252	12,450	16,246	12,441
Industry-Year FE	Yes	Yes	Yes	Yes
City FE	No	No	Yes	Yes

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 5C – Engineering Counts First Stage Regression (Companion Table to Table 4C)

6. TensorFlow, the Deep Learning Toolkit, and Bottlenecked Talent

The open-source launch of Google Brain's TensorFlow machine learning toolkit on November 9, 2015 was a departure from expectations that Google would try to safeguard all of its AI-related intellectual property.[28] The project grew out of a 2011 Google Brain initiative

[28] As noted in Wired (https://www.wired.com/2015/11/google-open-sources-its-artificial-intelligence-engine/): "With TensorFlow, however, the company has changed tack, freely sharing some of its newest—and, indeed, most important—software. Yes, Google open sources parts of its Android mobile operating system and so many other smaller software projects. But this is different. In releasing TensorFlow, Google is open sourcing software that sits at the heart of its empire. 'It's a pretty big shift,' says Dean, who helped build so much of the company's groundbreaking data center software…"

called DistBelief to build and train deep neural nets for research and commercial applications (Abadi et al. 2016).[29] TensorFlow was unique in that it was designed to serve as a single system that could run on a variety of platforms, ranging from mobile devices to "large-scale training systems running on hundreds of specialized machines with thousands of GPUs". Its release meant the wide availability of production-level software packages with greater stability and simplicity than other popular packages at the time (e.g. Theano, Caffe, and Torch). TensorFlow can be installed as a Python module or in C++, taking advantage of popular programming languages to make deep learning available to as many people as possible.

The package also includes a set of software pipelining tools such as TensorBoard, which helps machine learning engineers visualize the computational graph they have built, and performance tracing which helps track threads as they are processed. At the time, few of the comparable systems (Caffe, Chainer, Theano, and Torch) simultaneously supported symbolic differentiation, was written C++ to facilitate high performance production code, and could easily be mapped to many machines at once. Further, the Python interface and training documentation provided a baseline on which the open-source community could improve. What had been an experts' game was, at least in the near future, going to be something any reasonably talented coder could implement. Soon after, additional abstraction layers like Keras (Chollet and others 2015) and PyTorch (Paszke et al. 2017), a Pythonized version of the popular Torch software developed by employees at Facebook, would enter as competitors for TensorFlow.[30]

[29] Usually called "deep" when a standard neural net architecture has 4 or more layers.
[30] Keras and TensorFlow are now implemented for R as well. PyTorch and TensorFlow both have another abstraction layer module called fast.ai which is gaining popularity. Its creators, Rachel Thomas and Jeremy Howard, frequently implement new technical advances into the fast.ai module. PyTorch remains a favorite package in the research community.

But was the TensorFlow launch decision about talent? Oren Etzioni, a machine learning expert and executive director of the Allen Institute for Artificial Intelligence, at the time stated that Google was trying to "attract developers and new hires to its technology".[31] With new technologies, especially open-source software packages, adoption dynamics and value creation can be highly sensitive to network effects (Von Hippel and Krogh 2003). One interpretation then is that the TensorFlow open-source strategy meant Google could capture more of the rents in the economic applications of machine learning. Another is that their software platform would improve with the benefit of a community of contributors. Indeed, to date TensorFlow's GitHub project has over 40,000 code commits, 20 branches, 68 releases, and over 1,600 contributors.

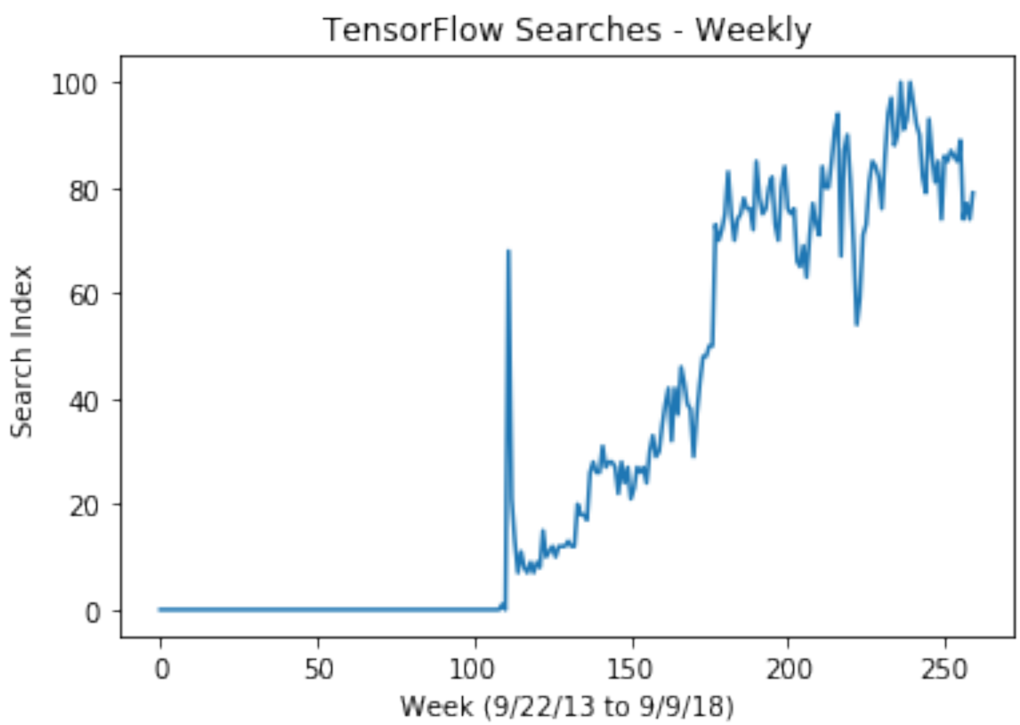


Figure 5 – TensorFlow Searches

[31] https://bits.blogs.nytimes.com/2015/11/09/google-offers-free-software-in-bid-to-gain-an-edge-in-machine-learning/?mtrref=undefined

Figure 5 shows the Google Trends searches for TensorFlow. Figures 6A and 7 show TensorFlow code and TensorBoard output, respectively. Given that TensorFlow constitutes a surprising reduction in the barriers to learn how to deep learn, it serves as an opportunity to understand an important talent bottleneck to the diffusion of the machine learning general-purpose technology. Figure 6B shows a code snippet for autograd, one of the core functions performed by TensorFlow. This is one of the necessary steps (and hundreds of lines as a module) for computing optimal parameter vectors in deep learning models, reduced to only a couple lines in TensorFlow. Google is now working to automate more of the deep learning model construction process with AutoML. The goal of AutoML, according to Google CEO Sundar Pichai, is to "take an ability that a few PhDs have today and will make it possible in three to five years for hundreds of thousands of developers to design new neural nets for their particular needs".[32] Building and launching TensorFlow was a first step in service of that goal.

```python
import tensorflow as tf
mnist = tf.keras.datasets.mnist

(x_train, y_train),(x_test, y_test) = mnist.load_data()
x_train, x_test = x_train / 255.0, x_test / 255.0

model = tf.keras.models.Sequential([
  tf.keras.layers.Flatten(),
  tf.keras.layers.Dense(512, activation=tf.nn.relu),
  tf.keras.layers.Dropout(0.2),
  tf.keras.layers.Dense(10, activation=tf.nn.softmax)
])
model.compile(optimizer='adam',
              loss='sparse_categorical_crossentropy',
              metrics=['accuracy'])

model.fit(x_train, y_train, epochs=5)
model.evaluate(x_test, y_test)
```

333 lines (287 sloc) | 11.9 KB   Raw Blame History

```python
from itertools import count
from functools import reduce
from .tracer import trace, primitive, toposort, Node, Box, isbox, getval
from .util import func, subval

# -------------------- reverse mode --------------------

def make_vjp(fun, x):
    start_node = VJPNode.new_root()
    end_value, end_node =  trace(start_node, fun, x)
    if end_node is None:
        def vjp(g): return vspace(x).zeros()
    else:
        def vjp(g): return backward_pass(g, end_node)
    return vjp, end_value

def backward_pass(g, end_node):
    outgrads = {end_node : (g, False)}
    for node in toposort(end_node):
        outgrad = outgrads.pop(node)
        ingrads = node.vjp(outgrad[0])
        for parent, ingrad in zip(node.parents, ingrads):
            outgrads[parent] = add_outgrads(outgrads.get(parent), ingrad)
    return outgrad[0]

class VJPNode(Node):
    __slots__ = ['parents', 'vjp']
    def __init__(self, value, fun, args, kwargs, parent_argnums, parents):
        self.parents = parents
        try:
            vjpmaker = primitive_vjps[fun]
        except KeyError:
            fun_name = getattr(fun, '__name__', fun)
            raise NotImplementedError("VJP of {} wrt argnums {} not defined"
                                      .format(fun_name, parent_argnums))
        self.vjp = vjpmaker(parent_argnums, value, args, kwargs)

    def initialize_root(self):
        self.parents = []
        self.vjp = lambda g: ()

primitive_vjps = {}
def defvjp_argnums(fun, vjpmaker):
    primitive_vjps[fun] = vjpmaker

def defvjp_argnum(fun, vjpmaker):
    def vjp_argnums(argnums, *args):
        vjps = [vjpmaker(argnum, *args) for argnum in argnums]
        return lambda g: (vjp(g) for vjp in vjps)
    defvjp_argnums(fun, vjp_argnums)

def defvjp(fun, *vjpmakers, **kwargs):
    argnums = kwargs.get('argnums', count())
    vjps_dict = {argnum : translate_vjp(vjpmaker, fun, argnum)
                 for argnum, vjpmaker in zip(argnums, vjpmakers)}
    def vjp_argnums(argnums, ans, args, kwargs):
        L = len(argnums)
        # These first two cases are just optimizations
        if L == 1:
            argnum = argnums[0]
            try:
                vjpfun = vjps_dict[argnum]
            except KeyError:
                raise NotImplementedError(
                    "VJP of {} wrt argnum 0 not defined".format(fun.__name__))
            vjp = vjpfun(ans, *args, **kwargs)
            return lambda g: (vjp(g),)
        elif L == 2:
            argnum_0, argnum_1 = argnums
            try:
                vjp_0_fun = vjps_dict[argnum_0]
                vjp_1_fun = vjps_dict[argnum_1]
            except KeyError:
                raise NotImplementedError(
                    "VJP of {} wrt argnums 0, 1 not defined".format(fun.__name__))
            vjp_0 = vjp_0_fun(ans, *args, **kwargs)
            vjp_1 = vjp_1_fun(ans, *args, **kwargs)
            return lambda g: (vjp_0(g), vjp_1(g))
        else:
```

[32] https://blog.google/technology/ai/making-ai-work-for-everyone/

Figure 6A, 6B – TensorFlow Code (source: TensorFlow.org) and Autograd Code (source: https://github.com/HIPS/autograd/blob/master/autograd/core.py)

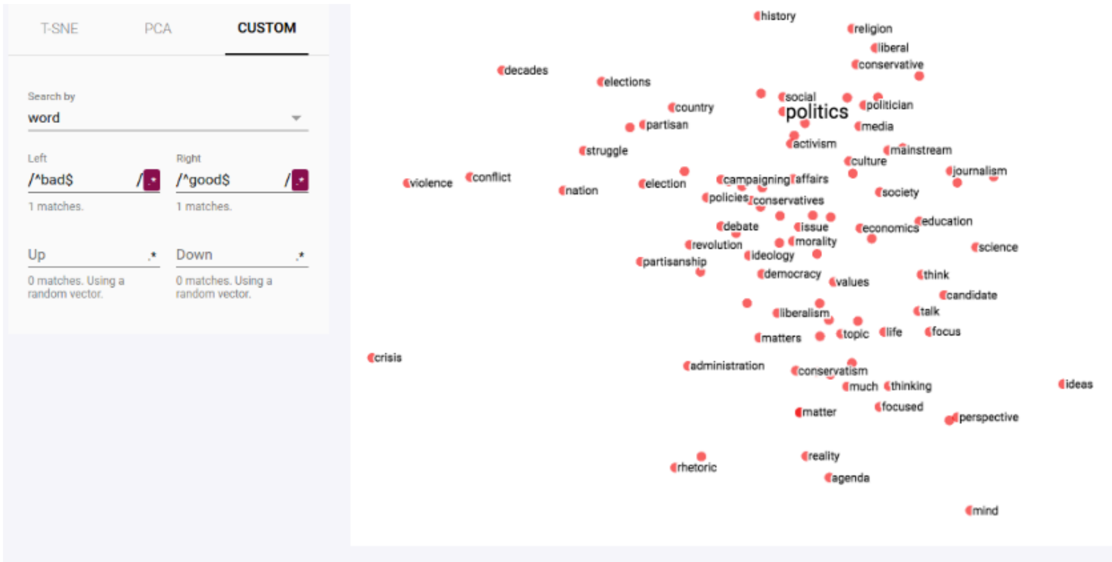


Figure 7 – TensorBoard Embedding Interface

This explicit objective on Google's part is evidence of the talent bottleneck to realizing the economic possibilities of machine learning. Though TensorFlow is the most popular package in use today, Google was and is not the only organization trying to make deep learning more accessible. Many other deep learning packages emerged around the same time as competition. PyTorch, developed by Facebook's AI research group, is a favorite amongst researchers. PyTorch supports a number of common optimization algorithms and a convenient neural net module for building more complicated networks. Fast.ai, a module built and maintained by researchers Rachel Thomas and Jeremy Howard, is explicitly dedicated to "making the power of deep learning accessible to all" and implements many state-of-the-art techniques as they are published.[33] These packages emerged as improvements to the existing toolkits between 2015 and

[33] Coders with basic python knowledge can build and train a high quality image classifier in only a couple hours of training using the fast.ai tutorial. See: http://www.fast.ai/about/

2018 in response to the growing disparity between the demand for AI talent and the difficulty in building and implementing deep learning models.

A natural starting question is whether or not AI talent is correlated with market value or productivity at all. Table 6A details some suggestive evidence that this is indeed the case. Using a balanced panel of publicly traded firms from 2014-2017 and a variety of skills indices, a regression of excess market value (total market value less the book value of assets) on these logged skills indices returns a coefficient of $9.28 million (standard error $2.485 million) per 1% increase in the AI skills index for specification (1) with firm and year fixed effects, and $63.84 million (standard error $12.6 million) for specification (2) with industry-year fixed effects. If market prices are perfectly efficient, this reflects a pricing of AI talent at the firm level. Notably most of the variation is between firm, not within firm. The firm fixed effects reduce the estimated coefficient by a factor of 7.

In specifications (3) and (4), the predicted output is revenue instead. A 1% increase in the AI skills index is correlated with an approximate $33 million (standard error $6.6 million) increase in revenues for specification (4), but including firm fixed effects reduces this point estimate considerably and leads to a loss of statistical significance. Again, the variation in firms drives the variation in revenues and adoption of AI. Firms with greater revenues tend to invest more in AI talent, adjusting for firm size. This has implications for the productivity effects of AI in the short-run. For a simple Cobb-Douglas production function with these skills indices as inputs, for example, the residual can be related to an index of total factor productivity. If the AI skills index variance cannot explain any of the variation in the revenue residual after adjusting for firm labor and capital inputs, then it is difficult to make a case that AI is driving contemporaneous productivity improvements for adopter firms.

That the AI skills index correlation persists for market value, but not for revenues, suggests two major possibilities. The first is that AI skills have been mispriced and that the market prices for AI talent are irrationally high. But assuming that the prices are rational, this evidence is suggestive that there are implementation lags to earning the productivity returns to AI investments, a phenomenon commonly associated with new GPTs and information technology (Hornstein and Krusell 1996; Greenwood and Yorukoglu 1997; Brynjolfsson, Rock, and Syverson 2018a). That is, the benefits are to come in the future. This is consistent also with the hypothesis that new general-purpose technologies induce a "J-Curve" for productivity (mis)measurement: in early stages, productivity growth can be slower as measurable resources are devoted to producing unmeasured intangible asset outputs (Brynjolfsson, Rock, and Syverson 2018b). Later, in harvest periods, these intangible assets will contribute a capital service flow.

Now I return to the case that a skill goes from being specialized to generalized in expectation, extending the results in Table 6A. This is the context for TensorFlow, when returns to firm-specific assets are otherwise bottlenecked by a lack of available talent at attractive prices. In spite of the rapid growth in the deep learning-skilled community, it still may be the case that deep learning talent is capacity-constrained. Yet the change that the TensorFlow API, PyTorch, Keras, fast.ai, and similar packages introduce is the expectation that in the future these skills will be generalized instead of specialized. To test the market value effects of the democratization of deep learning via TensorFlow, I use a differences-in-differences continuous treatment specification of the following familiar form:

$$MV_{it} = \beta * \text{Total Assets}_{\text{it}} + \gamma * HK_{total_{it}} + \lambda_t + \eta_i + \nu * (POST_{TF} * AISKILL_{it}) + \boldsymbol{Z}'_{\boldsymbol{it}}\boldsymbol{\delta} + \epsilon_{it} \qquad (13)$$

Market value for firm *i* at time *t* is a function of total assets, human capital (as total education years), a time fixed effect, a firm fixed effect (which absorbs the firm's average AI skill level), and an interaction between the post-TensorFlow launch dummy and the cumulative

AI skill counts for the firm in period *t*. **Z** denotes the vector of additional cumulative skills indices and the AI skills index which might otherwise confound the analysis. I consider a number of related skills, and a handful of seemingly unrelated ones (e.g. advertising) to test the relationship between AI skills and market value following the launch of TensorFlow.[34] These regressions are pooled over the course of the year, testing for a structural break related to the acquisition of AI-related skills.

Table 6B has the results of the difference-in-difference analysis. Other than total assets and lagged market value, the index values from LinkedIn are all logged. This allows for an easier interpretation of the coefficients in percentage terms. These difference-in-difference estimates explain the Q-value created by firms with different levels of AI skills. Interestingly, the AI skills index on its own negatively predicts market value in the pre-TensorFlow period. AI skills have an economically and statistically significant relationship with market value. Each additional order of magnitude increase in AI skill for the firm is associated with an increase in market value of $1.9 to $2.2 billion in the post-TensorFlow period, with standard errors of approximately $520 million. Including lagged market value (Table 6B, Column 7) and fixing the sample to be a balanced panel from 2014-2017 increases the coefficient values by a meaningful factor to the upper end of estimates. Digital Literacy, a category including Microsoft Office and other standard computer skills, and Advertising skills indices are negatively correlated with market value as well. Business Management skills tend to be positively (albeit imprecisely) correlated with market value. Cloud computing skills are negatively correlated with market value, suggesting that lower market-to-book value firms may be investing more in cloud skills at higher rates adjusting for the other skills indices. As the Compustat sample overrepresents

[34] Google is excluded from the analysis out of concern for possible endogeneity issues. Including Google in the sample generally increases the correlation between AI skills and market value.

manufacturing firms relative to the entire U.S. economy, this is suggestive of smaller market-to-book manufacturing firms investing more in the cloud.

Multiplying the index value by a 1-year lead (Table 7) suggests that the market prices the value of commoditized AI talent at the end of 2015 without earlier anticipation. The 1 year lead coefficient is not statistically significant. In other words, a year prior to the launch of TensorFlow, AI skills increases predicted *no change* in market value. The coefficient is however attenuated, suggesting a need for more detailed analysis of pre-trends.

Table 6A: Excess Market Value (Market less Book) and Revenue Regressions Balanced Panel - 2014-2017	(1) Excess Market Value (Millions USD)	(2) Excess Market Value (Millions USD)	(3) Revenue (Millions USD)	(4) Revenue (Millions USD)
Total Assets			0.0843***	0.0359***
			(0.0215)	(0.00650)
Log(Education Years)	1,362*	3,456***	950.3***	3,321***
	(801.6)	(662.6)	(305.4)	(662.1)
Log(AI Skills Index)	928.8***	6,384***	-71.35	3,332***
	(248.5)	(1,260)	(151.8)	(673.2)
Log(Data Science Skills Index)	767.0*	-1,461*	90.72	-1,220*
	(392.4)	(795.8)	(80.55)	(623.7)
Log(Cloud Computing Skills Index)	189.7	1,075	31.52	437.5
	(236.0)	(730.7)	(179.6)	(546.3)
Log(Data Storage Skills Index)	-5,651	-103,178***	-2,940	-44,439*
	(14,102)	(36,148)	(9,127)	(22,590)
Log(Digital Literacy Skills Index)	127.0	454.5	125.6	205.1
	(360.2)	(1,624)	(157.5)	(651.8)
Log(Management Skills Index)	186.9	-2,112	143.6	-595.4
	(631.8)	(1,507)	(411.3)	(702.9)
Log(Advertising Skills Index)	293.9	2,966**	111.1	572.0
	(223.6)	(1,294)	(183.9)	(638.3)
Observations	5,288	5,076	5,288	5,076
R-squared	0.939	0.340	0.985	0.539
Firm and Year FE	Yes	No	Yes	No
Industry-Year FE	No	Yes	No	Yes

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

Notes: Standard errors are clustered at the firm level for (1) and (3), industry level (4-Digit

NAICS Code) for (2) and (4). Data skills indices are computed as the logged cumulative total skills from the LinkedIn resume records within each skills group, aggregated by firm and year. Excess market value is defined as the publicly traded value of firm less the book value of total assets from Compustat. This is equivalent to assuming the coefficient on total assets is 1. Firms must be present in all years 2014-2017. Google is included in these specifications, as this table serves only to provide correlational evidence on the relationship between AI skills, market value, and productivity.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Table 6B: Market Value - AI Skills Difference-in-Difference (MM USD)	AI Skills	+Data Science	+Cloud Computing	+Data Storage	+Digital Literacy	+Bus.Mgmt and Advertising	+Lagged MV	+Balanced Panel
Lagged Market Value							0.149*	
							(0.0761)	
Total Assets	1.099***	1.099***	1.099***	1.099***	1.099***	1.099***	0.957***	1.122***
	(0.0478)	(0.0478)	(0.0478)	(0.0478)	(0.0478)	(0.0477)	(0.0619)	(0.0556)
Log(Edu. Years)	932.4	938.7	958.6	964.2	983.4	993.3	828.5	1,224*
	(639.0)	(633.2)	(637.2)	(638.2)	(629.2)	(635.7)	(526.6)	(738.7)
Log(AI Index)	-2,058***	-2,045***	-1,988***	-1,985***	-1,980***	-1,970***	-2,071***	-2,370***
	(685.1)	(691.6)	(677.1)	(676.4)	(677.7)	(672.8)	(643.2)	(777.3)
Log(AI Index x Post TF)	1,900***	1,903***	1,916***	1,917***	1,921***	1,926***	1,998***	2,202***
	(460.7)	(459.5)	(462.5)	(462.7)	(462.0)	(464.7)	(442.9)	(520.0)
Log(Data Science Index)		-124.7	-18.81	-6.796	146.7	67.07	-135.6	-105.2
		(281.7)	(290.8)	(291.4)	(259.5)	(270.7)	(278.0)	(319.3)
Log(Cloud Computing Index)			-469.5**	-458.1**	-440.5*	-465.5*	-519.9**	-535.2*
			(231.8)	(230.7)	(231.2)	(243.4)	(242.5)	(277.6)
Log(Data Storage Technology Index)				-11,102	-9,944	-10,023	-9,750	-20,331
				(10,326)	(10,546)	(10,438)	(9,877)	(14,296)
Log(Digital Literacy Index)					-332.7	-435.8*	-461.1*	-555.8*
					(293.8)	(254.5)	(275.1)	(319.0)
Log(Bus. Management Index)						559.4	832.2	883.7

						(537.5)	(676.5)	(713.5)
Log(Advertising Index)						-160.4	-195.9	-117.3
						(182.5)	(188.3)	(218.2)
Observations	6,440	6,440	6,440	6,440	6,440	6,440	6,393	5,284
R-squared	0.998	0.998	0.998	0.998	0.998	0.998	0.999	0.998
Firm and Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 6B: Market Value Difference-in-Difference on AI Skills During TensorFlow Launch

Table 7: AI Difference-in-Difference Leads Robustness Check	(1) AI Cluster	(2) +Data Science	(3) +Other Indices
Total Assets	0.987***	0.987***	0.988***
	(0.0154)	(0.0154)	(0.0154)
Log(Edu. Years)	432.0	426.1	485.9
	(554.9)	(554.4)	(569.5)
Log(AI Index)	-987.1	-996.9	-963.1
	(785.8)	(781.3)	(768.0)
Log(AI Index x Post TF+1 Year Lead)	9.425	6.701	17.49
	(440.5)	(441.8)	(447.0)
Log(AI Index x Post TF)	895.5***	893.5***	900.0***
	(185.4)	(186.2)	(188.4)
Log(Data Science Index)		120.0	260.5
		(208.6)	(212.7)
Log(Cloud Computing Index)			-339.2
			(223.7)
Log(Data Storage Technology Index)			-5,548
			(9,019)
Log(Digital Literacy Index)			-269.7
			(184.0)
Log(Bus. Management Index)			74.10
			(428.4)
Log(Advertising Index)			97.97
			(161.6)
Observations	4,587	4,587	4,587

Firm and Year FE	Yes	Yes	Yes

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Table 7: Lead Check for AI Skill Difference-in-Difference

Further Robustness Checks

The launch of TensorFlow and other deep learning packages might have coincided with other conditions or endogenous firm decisions which limit any chance to make causal claims. So far, a causal interpretation of the difference-in-differences coefficients in the above tables assumes that market value does not cause adoption of AI, that there are parallel trends in the AI using and non-AI using firms, and the stable unit treatment value assumption (SUTVA) holds.[35] The first condition is mitigated by including lagged market value as a control, and the latter condition can be investigated partially with a balanced panel assuming no spillovers between firms. The parallel trends assumption is trickier and requires a more granular time series analysis. I create new group variables "AI Quintile" and "Suitability for Machine Learning (SML) Quintile" for the quintile groups in which each firm falls in AI skill employment and SML ranking (respectively) as of the fourth quarter of 2015. The median firm has no listed AI skills at that time. In fact, up to the 59th percentile of AI use the LinkedIn skill count for AI is zero.[36] In these specifications, I calculate the log indices of the input skills to recover a "percentage increase per quantile change" interpretation of the coefficients.[37] Using a balanced panel, I estimate the following specification, where total assets and education years are included in the **Z** vector:

$$Log(MV_{it}) = \lambda_t + \eta_i + \nu_t * (D_t * AIQuantile_i) + \boldsymbol{Z}'_{it}\boldsymbol{\delta} + \epsilon_{it} \qquad (14)$$

[35] This assumes that spillovers and compositional changes in the sample are not confounding.
[36] The appendix has a histogram of AI group percentages. The groups are uneven because of ties in AI skill counts.
[37] Technically each index is shifted by 1, so the log of market value variable is log(MV+1).

Each year is represented as a dummy D_t and interacted with the AI Quantile dummy variable. The results are in the Figure 8 below, showing an increase in market value of approximately 4-7% for all AI-using firms relative to non-AI firms (bottom, second, and part of the third quintile) in the quarter of the TensorFlow launch. This is the only quarter for which the change in market value by AI quintile is statistically significant for the second-highest AI-intensive quintile, and one of two for the third quintile. The quarters themselves are not statistically significantly different, but the pooled estimates from Table 8 suggest a structural break. The top AI using firms appear to have a non-parallel trend relative to non-AI firms in the pre-launch period. As seen in green, this high end of AI-using firms has a statistically significantly higher growth rate. This would invalidate the parallel trends assumption *for a quintile group-based* difference-in-difference estimate above and motivates the continuous skill value regression with firm fixed effects. The Appendix has the corresponding table of estimates, including interactions for SML quintile. Interestingly, the SML quintiles interacted with time dummies fail to reflect an effect of Tensorflow. Instead, the market value for those firms decreases later in the same year. This calls into question whether the TensorFlow (and AI talent) shock expectation is driving the downward trend in high SML firms. It is possible that this high SML revaluation occurs with a lag, or that another effect is driving the negative correlation between SML value and market value. Figure 9 shows the continuous treatment version of the effect on deep learning skills in particular. The effect precision narrows as more LinkedIn members post deep learning skills, but here again there is an increase in the market value of firms following the launch of TensorFlow in Q4 2015. This launch is statistically significant from zero AI use (the baseline), while all of the previous quarters are not statistically significantly different from the zero baseline. We fail to rule out parallel trends for the continuous treatment version, though it may be that the statistically significant market value pop due to deep learning skills was by chance. It is convincing, however, that the effects diminish to an imprecisely measured zero in the quarters following Q4 2015. Figures 10A-D have data science skill, linear regression, management, and advertising skill index specifications (respectively) for comparison. All specifications include the full set of skill index covariates, firm fixed effects, industry-time fixed effects, lagged market value, education years, and total asset. The error bars are the 95% confidence interval using standard errors clustered by firm. The pooled and time-series regressions suggest strong market value effects on AI-using companies from the launch of TensorFlow and related packages. Of course, any shock affecting AI-using companies in the same quarter will also show up in the coefficient estimates. These concurrent unobserved shocks are a threat to any causal interpretation. At a minimum, however, it appears that there was a strong upward repricing of all companies employing AI talent at the same time that Google made the decision to make TensorFlow open-source. The evidence that this launch caused a decrease in the value of companies with high SML tasks is weaker unless the effects occurred with nearly a one-year lag. At the same time, the value of other types of skills do not seem to have responded to the TensorFlow launch.

These results would seem at first glance to stand in stark contrast to the previous section's results on the general value of engineering talent, but instead may indicate that the firm-specific assets already accumulated by companies are strongly complementary to the general deep learning talent pool. The $1.9 to $2.2 billion per LinkedIn AI skill order of magnitude increase predicted by the difference-in-difference analysis in Table 6 is unexpectedly large given that even high salaries for AI workers are typically less than $2 million at the moment. Noticing this pre-trend in the top quintile of AI-using firms, I adjust the difference-in-difference analysis below to exclude the top quintile in Tables 8 and 9.

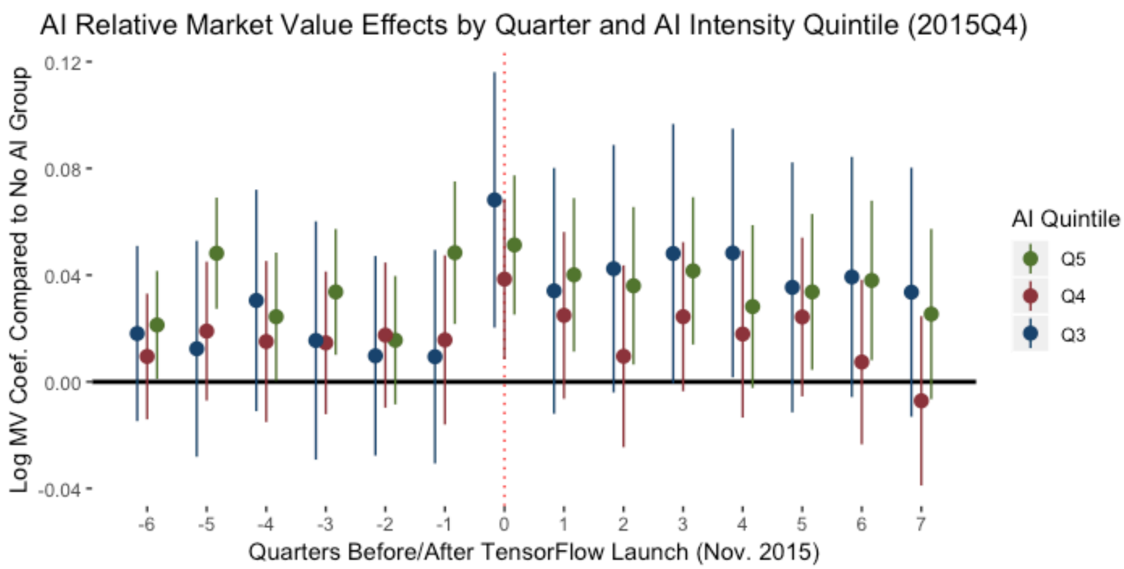


Figure 8 - AI Quintile Effects

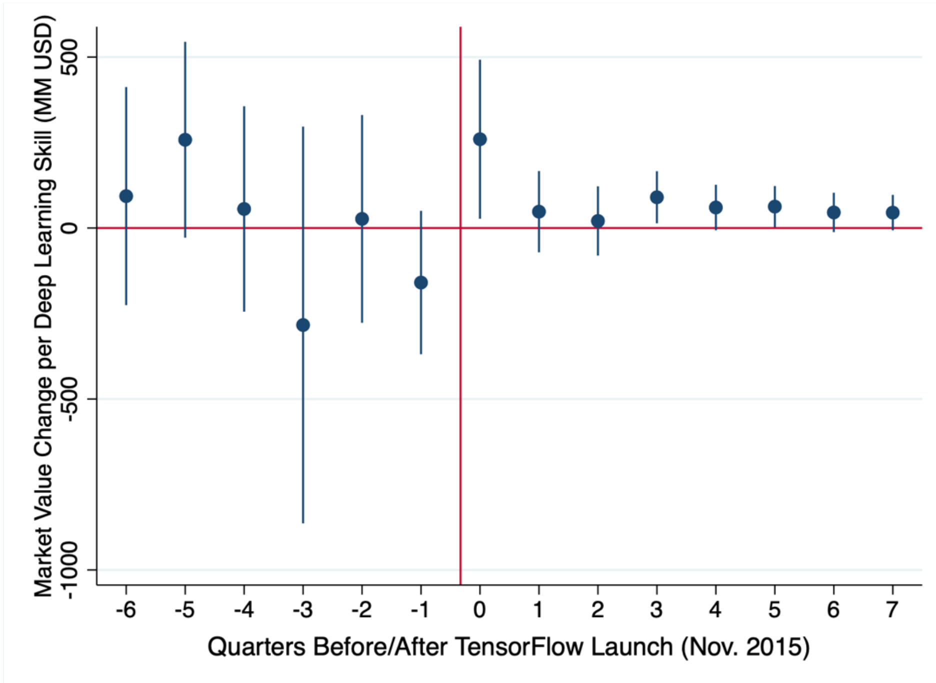


Figure 9 – Market Value Change per Deep Learning Skill

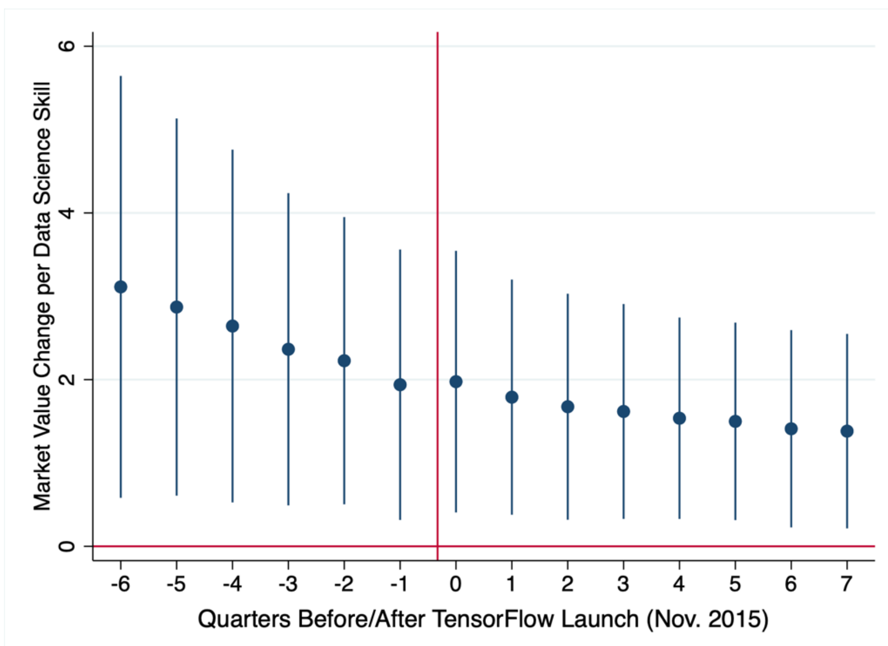


Figure 10A - Data Science Skills

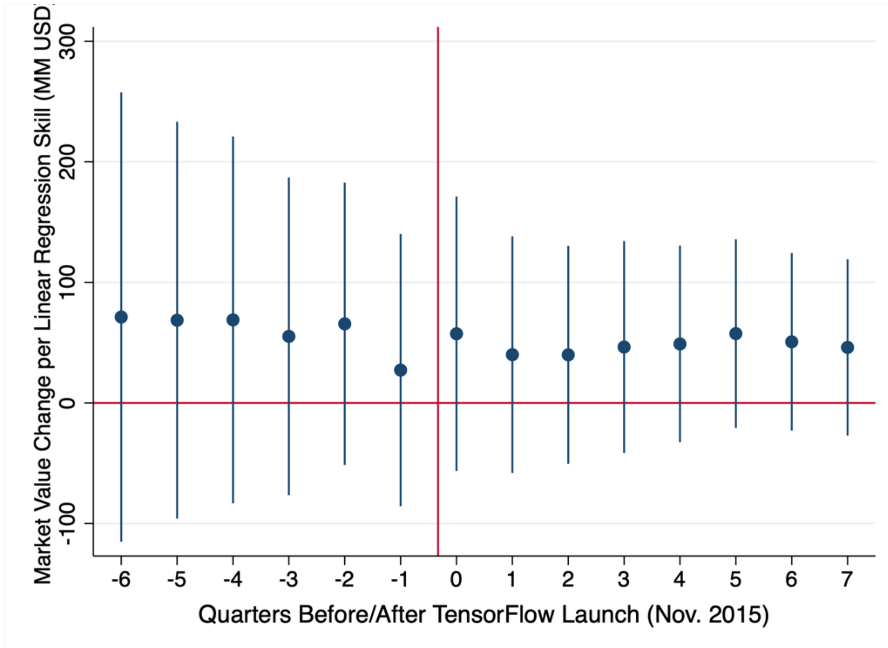


Figure 10B – Linear Regression Skills

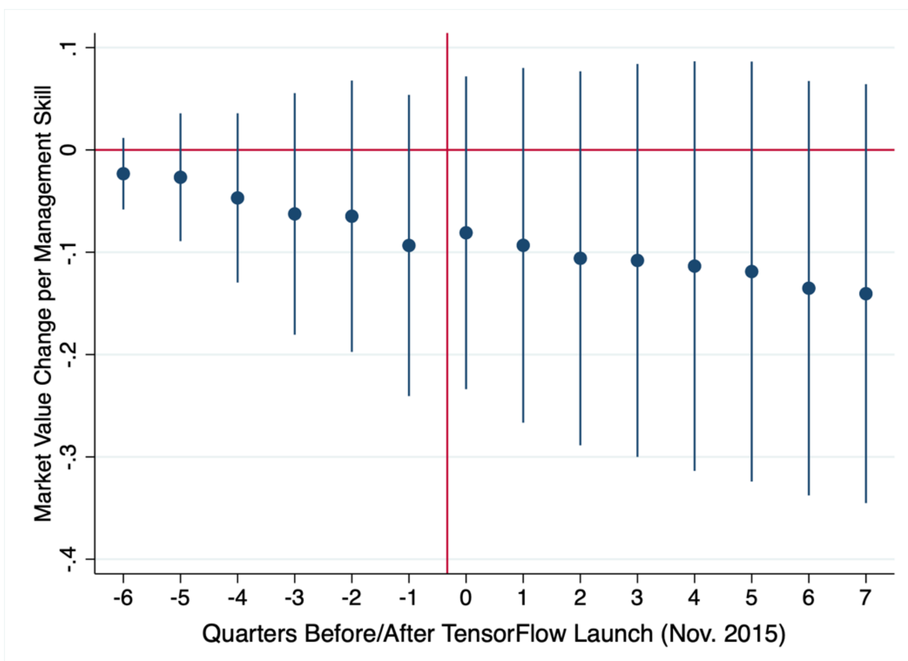


Figure 10C – Management Skills

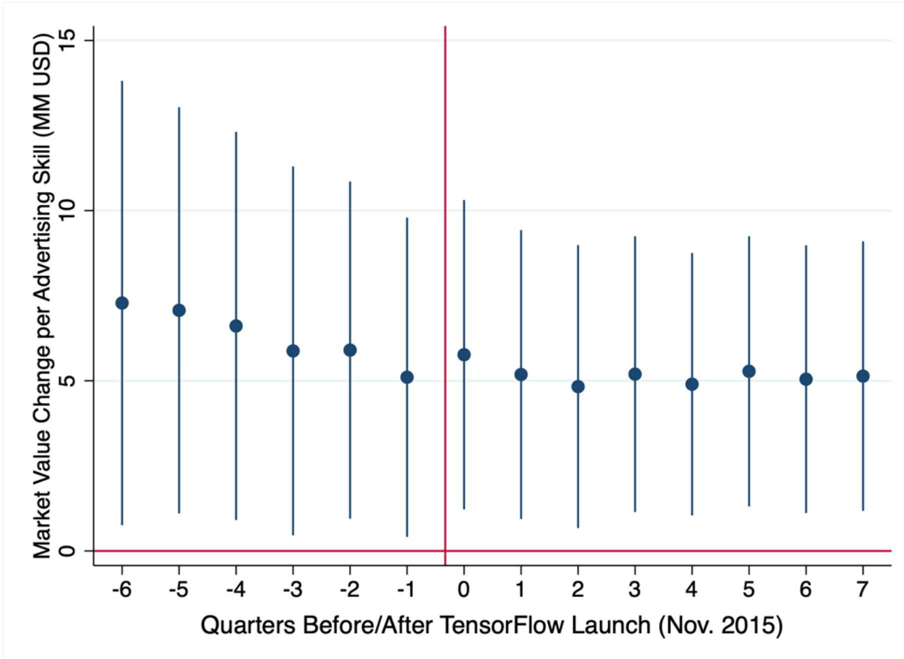


Figure 10D – Advertising Skills

Table 8: AI Difference-in-Difference w/o Top Quintile Market Value (MM USD)	(1) AI Cluster	(2) +Data Science	(3) +Cloud Computing	(4) +Data Storage	(5) +Digital Literacy	(6) +Bus.Mgmt and Advertising	(7) +Lagged MV	(8) +Balanced Panel
Lagged Market Value							0.00473	
							(0.0240)	
Total Assets	0.990***	0.990***	0.990***	0.990***	0.990***	0.990***	0.986***	1.000***
	(0.0180)	(0.0180)	(0.0180)	(0.0180)	(0.0180)	(0.0180)	(0.0209)	(0.0164)
Log(Edu. Years)	299.0**	306.1**	309.1**	313.2**	322.7**	318.1**	323.0**	360.0**
	(130.0)	(130.9)	(131.1)	(131.0)	(133.3)	(132.7)	(136.3)	(151.9)
Log(AI	-293.0	-278.5	-272.6	-270.1	-266.9	-269.7	-278.4	-277.7

Index)								
	(196.4)	(197.6)	(198.8)	(198.9)	(198.5)	(198.7)	(199.4)	(222.5)
Log(AI Index x Post TF)	352.3***	357.4***	360.3***	360.3***	361.6***	358.3***	366.1***	356.7**
	(136.2)	(136.5)	(136.5)	(136.5)	(136.5)	(136.6)	(138.6)	(144.3)
Log(Data Science Index)		-152.9	-139.3	-129.9	-51.85	-8.269	2.779	-99.55
		(130.0)	(126.4)	(125.4)	(119.0)	(132.0)	(139.0)	(150.4)
Log(Cloud Computing Index)			-60.82	-51.47	-42.69	-28.63	-39.55	-57.00
			(137.7)	(138.7)	(138.0)	(140.7)	(140.9)	(158.1)
Log(Data Storage Technology Index)				-8,939*	-8,350	-8,373	-8,481	-14,033*
				(5,230)	(5,309)	(5,245)	(5,245)	(7,491)
Log(Digital Literacy Index)					-171.4	-117.9	-129.1	-217.3
					(136.7)	(135.4)	(139.0)	(169.4)
Log(Bus. Management Index)						-273.9	-271.5	-224.4
						(207.5)	(223.2)	(255.4)
Log(Advertising Index)						56.82	51.24	149.4
						(126.2)	(126.2)	(137.2)
Observations	5,864	5,864	5,864	5,864	5,864	5,864	5,819	4,784
Firm and Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 8: Adjusted Difference-in-Differences AI Skill Valuation Post-TensorFlow Ex. Top Quintile of AI-Using Firms

The coefficients excluding the top AI-using quintile are substantially attenuated. Within the remaining AI-using firms as of the end of Q4 2015, the effect of the TensorFlow launch is that a 100% increase in AI skills corresponds to a $352 to $367 million dollar increase in market value with standard errors of approximately $140 million. This is still a substantial amount, and qualitatively similar given that many of the top AI-using firms are the largest firms. The effect of the TensorFlow launch survives excluding the AI superstar firms with pre-trends. Similarly, Table 9 has the leads analysis corresponding to Table 7. The 1 year leads fail to indicate any kind of anticipation, though in this set of specifications the ordinary effects are not detected either.

Table 9: AI Difference-in-Difference Leads w/o Top AI Quintile	(1) AI Cluster	(2) +Data Science	(3) +Other Indices
Total Assets	0.974***	0.974***	0.974***
	(0.0144)	(0.0144)	(0.0144)
Log(Edu. Years)	-26.97	-28.39	-15.62
	(190.1)	(191.0)	(190.9)
Log(AI Index)	-367.4	-369.5	-366.1
	(269.2)	(268.4)	(269.0)
Log(AI Index x Post TF+1 Year Lead)	102.2	100.9	99.71
	(156.0)	(156.2)	(155.8)
Log(AI Index x Post TF)	51.93	51.20	50.53
	(125.5)	(125.9)	(126.4)
Log(Data Science Index)		33.97	133.3
		(121.6)	(108.4)
Log(Cloud Computing Index)			0.879
			(146.2)
Log(Data Storage Technology Index)			-3,625
			(5,751)
Log(Digital Literacy Index)			-119.5
			(142.9)
Log(Bus. Management Index)			-202.3
			(198.3)
Log(Advertising Index)			81.55
			(138.6)
Observations	4,187	4,187	4,187
Firm and Year FE	Yes	Yes	Yes

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Taken together, these results suggest story like that of Hirshleifer's pecuniary benefits of technological change. Having AI talent in 2015 is a signal that there are other assets at the firm that are complementary to AI talent. TensorFlow makes a specialist skill into a generalist skill. As a result, more workers can build engineering value using commodity deep learning packages, and firms receive the capital service flow from their assets (for example, large scale databases or automation projects) formerly bottlenecked by the scarcity of available talent. TensorFlow, by making a scarce complement cheaper and more abundant, increases the value of firms positioned to invest in AI. Importantly, these effects are democratizing for at least one quarter. AI-using firms without pre-trends catch up to their AI superstar competitors for the fourth quarter of 2015. Afterward, their market value performance is statistically similar to peer companies without AI talent (as in Figure 9). Further, no similar effects can be observed in other types of skills where we might expect to see similar changes. Data Science, Linear Regression, Advertising, and Management fail to demonstrate any of the effects present for Deep Learning and AI skills more broadly. This suggests that the repricing of AI companies was indeed generated by the launch of TensorFlow.

Suitability for Machine Learning Analysis

Independent of AI talent, what happened to companies where the employed workers might be expected to be impacted by machine learning technology? Tables 10 and 11 show the results of the same analysis run for Suitability for Machine Learning (SML) scores aggregated to the firm level (wage bill-weighted averages) and in logged terms (respectively). Notably the launch of TensorFlow has a statistically significant negative association with the value of firms with higher potential to automate tasks with machine learning. A 1% increase in the overall SML of a firm is correlated with a 0.5% decrease in the firm's market value post-2016. This is consistent with the idea that the assets complementary to machine learning engineering are valuable, but potentially productivity-enhancing (not profitability) innovations without a source of rents might force firms to invest more rapidly than they would have otherwise done. Convex ex-ante fixed costs of investment therefore might drive down market value for firms that have to change their business models. Figure 11 shows the correlation between market value and SML for 2016, which is very close to zero.

Asset managers need not have the granular detail of the SML scores for these price effects to be incorporated. Since high SML tasks tend to be clerical and data-intensive routine work, the firms employing lots of these types of workers with intangible assets optimized for these purposes might newly be vulnerable to ML and AI-powered competition. However, these effects are likely not driven by TensorFlow. Interacting the quarter dummies with the SML quintile reveals that the structural change is indeed in the same year, but the repricing does not occur in the same quarter as the TensorFlow launch. Under the assumption that markets efficiently price the effects of TensorFlow at the time of release, it is implausible to conclude that SML-intensive companies lost market value as a result of the open sourcing event.

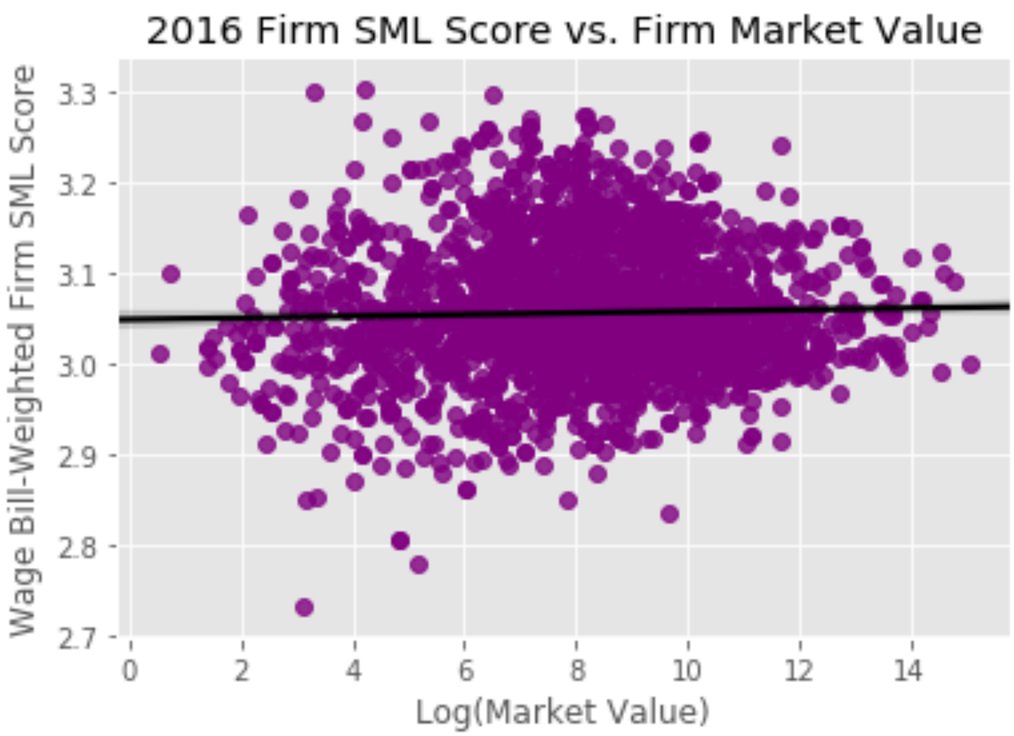


Figure 11: Firm-Level Suitability for Machine Learning (SML) vs. Log Market Value

Table 10: Market Value (SML) Difference-in-Differences	(1) SML	(2) +Data Science	(3) +Cloud Computing	(4) +Data Storage	(5) +Digital Literacy	(6) +Bus.Mgmt and Advertising
Total Assets	1.062***	1.099***	1.100***	1.100***	1.078***	1.043***
	(0.0287)	(0.0446)	(0.0447)	(0.0447)	(0.0384)	(0.0302)
Total Education Years	0.0185	0.0190	0.0192	0.0192	0.0210	0.0137
	(0.0183)	(0.0177)	(0.0179)	(0.0179)	(0.0176)	(0.00933)
SML X Post-TensorFlow	-4,336*	-3,684	-3,667	-3,661	-6,477**	-5,920**
	(2,257)	(2,791)	(2,803)	(2,790)	(2,821)	(2,476)
Data Science Skill Index		0.503	0.365	0.365	-0.870*	1.584**
		(0.310)	(0.402)	(0.402)	(0.483)	(0.769)
Cloud Computing Skill Index			0.261	0.261	0.213	0.0218
			(0.761)	(0.761)	(0.690)	(0.432)
Data Storage Skill Index				1,534	6,843	-1,527
				(10,438)	(10,318)	(8,721)
Digital Literacy Skill Index					0.498***	0.360***
					(0.0858)	(0.130)
Business Mgmt. Skill Index						-0.717***
						(0.198)

Advertising Skill Index						4.523***
						(1.586)
Observations	8,764	6,437	6,437	6,437	6,437	6,437
Firm and Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parentheses, SEs clustered by Firm
*** p<0.01, ** p<0.05, * p<0.1
Table 10: Suitability for Machine Learning (SML) Difference-in-Differences

Table 11: Log Market Value – SML Difference-in-Differences	(1) SML	(2) +Data Science	(3) +Cloud Computing	(4) +Data Storage	(5) +Digital Literacy	(6) +Bus.Mgmt and Advertising
Log(Total Assets)	0.607***	0.584***	0.583***	0.584***	0.582***	0.582***
	(0.0660)	(0.0849)	(0.0849)	(0.0850)	(0.0851)	(0.0851)
Log(Education Years)	0.0232	0.0307	0.0316	0.0311	0.0296	0.0290
	(0.0153)	(0.0207)	(0.0207)	(0.0207)	(0.0206)	(0.0206)
Log(SMLxPost-TF)	-0.484**	-0.520**	-0.514**	-0.506**	-0.525**	-0.499*
	(0.235)	(0.257)	(0.256)	(0.256)	(0.255)	(0.255)
Log(Data Science)		-0.00454	0.000376	-0.000913	-0.0150	-0.0123
		(0.0187)	(0.0190)	(0.0193)	(0.0188)	(0.0196)
Log(Cloud Computing)			-0.0154	-0.0166	-0.0187*	-0.0183*
			(0.0109)	(0.0108)	(0.0106)	(0.0105)
Log(Data Storage Tech.)				1.036	0.925	0.945
				(0.871)	(0.857)	(0.868)
Log(Digital Literacy)					0.0299*	0.0344**
					(0.0160)	(0.0156)
Log(Business Management)						-0.0292
						(0.0237)
Log(Advertising)						0.0136
						(0.0165)
Observations	8,764	6,437	6,437	6,437	6,437	6,437
Firm and Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parentheses, SEs Clustered by Firm
*** p<0.01, ** p<0.05, * p<0.1
Table 11: Market Value and Log(SML) Difference-in-Difference

7. Conclusion

It remains an open and context-specific question whether the market value of firms is driven by appropriation of the value of technological human capital. Not all varieties of technological human capital are correlated with market value after controlling for generalized education level and the firm's asset base. Engineering talent, however, is. Using a panel of corporate fixed assets and human capital, I have measured the average and marginal returns to investments in technological labor. I find that on average, engineering talent is strongly correlated with market value, but the marginal causal effect of hiring more engineers as estimated by instrumental variables analyses and controlling for time-invariant firm-specific factors is indistinguishable from zero. Nevertheless, exciting growth in market value can occur when formerly specialized skillsets are converted into general ones, as is the case with Google's launch of TensorFlow. The reduced barriers to entry in AI led to growth of roughly $3.56 million per 1% increase in AI skill excluding top quintile AI firms. This corresponds to a 4-7% contemporaneous increase in market value for firms outside the top quintile of AI skills with the launch of TensorFlow. This is the case even controlling for growth in other kinds of related skillsets, like Cloud Computing or Data Science. Further, Suitability for Machine Learning (SML) is unlikely to be a pathway via which TensorFlow positively impacted market value. If higher SML companies increased in value following the launch of TensorFlow, it would indicate that in expectation companies currently employing lots of high SML labor might appropriate the productivity gains for automating tasks with machine learning in the future. There is little evidence that the SML scores are positively related to market value following the launch of TensorFlow.

This kind of discrete technological change is informative about the processes which limit diffusion of new general-purpose technologies before they become generally-applied technologies. Namely, the tools or training have to be available to the engineers who are to build the assets that generate the technology's value. The choice to make specialized technology with large potential more widely available is sometimes one that can be made by corporate actors in industry. Talent is not always a bottleneck, but when it is firms may be more likely to designate firm-specific tasks with high marginal value to workers with scarce skills. This makes it more difficult for competitors to bid up the wages of those types of workers, but at the same time more abundant skills are likely to be competitively priced. This combination of firm-specific assets with complementary applications of engineering skills means that firms can appropriate some of their employees' investments in human capital. The paradox that, controlling for non-human capital assets, technological talent can have high value on average but marginally low value is resolved when firms can assign tasks which their competitors do not value (on average) but do value (on the margin).

Managers expecting to pay all incoming workers the same amount as their incumbent staff are faced with a challenge when talent is scarce. Do they want to give everyone a raise just to hire one more person? When competition on the bases of wages is difficult, the assignment of firm-specific tasks is a potential mechanism to bargain away part of the employees' talent value. Managers therefore force workers to compete with each other inside the firm while insulating their employer from outside competition. It is therefore potentially lucrative to expand the available talent pool to capture the full possible value of firm-specific tasks. TensorFlow made expectations of future deep learning talent, a previously scarce skillset, much higher. This suggests that all AI employers expected to find more AI talent in the coming years, permitting assignment of a greater range of firm-specific tasks. This suggests, for example, that open-source production decisions may generate rents for adopting firms via the talent channel. Many previous studies of employer power in the labor market have focused on employer-occupation-market concentration or policy changes changing the bargaining power of employees. Yet if firms are benefitting from exercise of labor market monopsony power, it should show up in their valuations. This paper shows that for a specific type of technological talent – engineers – it can be the case that firm value is in part driven by employer appropriation of employee human capital. Companies can do this by allocating their employees to firm-specific job tasks and finding an employment niche. Meanwhile, workers should carefully consider their contractual arrangements and how their employers are engineering value. In the case that work tasks are not competitively decided, the distribution of value might not reward the employee for their full contribution.

References

Abadi, Martin, Paul Barham, Jianmin Chen, Zhifeng Chen, Andy Davis, Jeffrey Dean, Matthieu Devin, et al. 2016. "Tensorflow: A System for Large-Scale Machine Learning." In *OSDI*, 16:265–83.

Acemoglu, D., and J.-S. Pischke. 1998. "Why Do Firms Train? Theory and Evidence." *The Quarterly Journal of Economics* 113 (1):79–119.

Acemoglu, Daron. 1997. "Training and Innovation in an Imperfect Labour Market." *The Review of Economic Studies* 64 (3). Wiley-Blackwell:445–64.

Acemoglu, Daron, and David Autor. 2011. *Skills, Tasks and Technologies: Implications for Employment and Earnings*. *Handbook of Labor Economics*. Vol. 4.

Acemoglu, Daron, and Jorn-Steffen Pischke. 1999. "Beyond Becker: Training in Imperfect Labour Markets." *The Economic Journal* 109 (453):112–42.

Agrawal, Ajay, Joshua S Gans, and Avi Goldfarb. 2017. "Prediction, Judgment, and Uncertainty." In *Economics of Artificial Intelligence*. University of Chicago Press.

———. 2018. *Prediction Machines: The Simple Economics of Artificial Intelligence*. Boston: Harvard Business Review Press.

Agrawal, Ajay, Avi Goldfarb, and Florenta Teodoridis. 2016. "Understanding the Changing Structure of Scientific Inquiry." *American Economic Journal: Applied Economics* 8 (1):100–128.

Andrews, Dan, Chiara Criscuolo, and Peter N Gal. 2015. "Frontier Firms, Technology Diffusion and Public Policy: Micro Evidence from OECD Countries." *Working Paper*, 39.

Ashenfelter, Orley C, Henry Farber, and Michael R Ransom. 2010. "Labor Market Monopsony." *Journal of Labor Economics* 28 (2). The University of Chicago Press:203–10.

Autor, David H. 2014. “Polanyi’s Paradox and the Shape of Employment Growth.” *Working Paper*, 129–78.
Autor, David H., and David Dorn. 2013. “The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market.” *American Economic Review* 103 (5):1553–97.
Autor, David, Frank Levy, and Richard J. Murnane. 2003. “The Skill Content of Recent Technological Change: An Empirical Exploration.” *The Quarterly Journal of Economics* 118 (4):1279–1333.
Azar, José A, Ioana Marinescu, Marshall I Steinbaum, and Bledi Taska. 2018. “Concentration in US Labor Markets: Evidence from Online Vacancy Data.”
Balasubramanian, Natarajan, Jin Woo Chang, Mariko Sakakibara, Jagadeesh Sivadasan, and Evan Starr. 2018. “Locked In? The Enforceability of Covenants Not to Compete and the Careers of High-Tech Workers.”
Barney, Jay B. 1986. “Strategic Factor Markets: Expectations, Luck, and Business Strategy.” *Management Science* 32 (10). INFORMS:1231–41.
Becker, Gary S. 1962. “Investment in Human Capital: A Theoretical Analysis.” *Journal of Political Economy* 70 (5, Part 2):9–49.
Benmelech, Efraim, Nittai K Bergman, and Amit Seru. 2011. “Financing Labor.”
Benmelech, Efraim, Nittai Bergman, and Hyunseob Kim. 2018. “Strong Employers and Weak Employees: How Does Employer Concentration Affect Wages?”
Benzell, Seth, Guillermo Lagarda, and Daniel Rock. 2018. “Skill-Biased Technological Change Within and Across Firms: 2000-2016.”
Bhaskar, Venkataraman, Alan Manning, and Ted To. 2002. “Oligopsony and Monopsonistic Competition in Labor Markets.” *Journal of Economic Perspectives* 16 (2):155–74.
Bivens, Josh, and Lawrence Mishel. 2015. “Understanding the Historic Divergence between Productivity and a Typical Worker“ s Pay: Why It Matters and Why It“ s Real.” *Washington DC: Economic Policy Institute (Instituto de Pol{\'\i}ticas Económicas). Http://Www. Epi. Org/Publication/Understanding-the-Historic-Divergence-Betweenproductivity-and-a-Typical-Workers-Pay-Why-It-Matters-and-Why-Its-Real*.
Bloom, Nicholas, Erik Brynjolfsson, Lucia Foster, Ron Jarmin, Megha Patnaik, Itay Saporta-Eksten, and John Van Reenen. 2018. “What Drives Differences in Management Practices?” *American Economic Review*.
Bresnahan, Timothy. 2010. “General Purpose Technologies.” *Handbook of the Economics of Innovation* 2 (1):761–91.
Bresnahan, Timothy F., and M. Trajtenberg. 1995. “General Purpose Technologies ‘Engines of Growth’?” *Journal of Econometrics* 65 (1):83–108.
Bresnahan, Timothy F, Erik Brynjolfsson, and Lorin M Hitt. 2002. “Information Technology, Workplace Organization, and the Demand for Skilled Labor: Firm-Level Evidence.” *The Quarterly Journal of Economics* 117 (1). MIT Press:339–76.
Brynjolfsson, Erik. 1994. “Information Assets, Technology and Organization.” *Management Science* 40 (12). INFORMS:1645–62.
Brynjolfsson, Erik, Lorin M. (Lorin Moultrie) Hitt, and Shinkyu Yang. 2002. “Intangible Assets: Computers and Organizational Capital.” *Brookings Papers on Economic Activity* 2002 (1):137–98.
Brynjolfsson, Erik, Xiang Hui, and Meng Liu. 2018. “Does Machine Translation Affect International Trade? Evidence from a Large Digital Platform.”
Brynjolfsson, Erik, and Andrew McAfee. 2014. *The Second Machine Age: Work, Progress, and*

Prosperity in a Time of Brilliant Technologies. WW Norton & Company.
Brynjolfsson, Erik, and Tom Mitchell. 2017. “What Can Machine Learning Do? Workforce Implications.” *Science* 358 (6370):1530–34.
Brynjolfsson, Erik, Tom Mitchell, and Daniel Rock. 2018a. “Machine Learning and Occupational Change.” *Unpublished Working Paper. MIT*.
———. 2018b. “What Can Machines Learn, and What Does It Mean for Occupations and the Economy?” *AEA Papers and Proceedings*, 43–47.
Brynjolfsson, Erik, Daniel Rock, and Chad Syverson. 2018a. “Artificial Intelligence and the Modern Productivity Paradox: A Clash of Expectations and Statistics.” In *Economics of Artificial Intelligence*. University of Chicago Press.
———. 2018b. “The Productivity J-Curve: How Intangibles Complement General Purpose Technologies.”
Brynjolfsson, Erik, Daniel Rock, Prasanna Tambe, and Jacqueline Barrett. 2018. “People Are Our Greatest Asset: Measuring the Market Value of Human Capital.” MIT Working Paper.
Campbell, Benjamin A, Russell Coff, and David Kryscynski. 2012. “Rethinking Sustained Competitive Advantage from Human Capital.” *Academy of Management Review* 37 (3). Academy of Management Briarcliff Manor, NY:376–95.
Campbell, Benjamin A, Martin Ganco, April M Franco, and Rajshree Agarwal. 2012. “Who Leaves, Where to, and Why Worry? Employee Mobility, Entrepreneurship and Effects on Source Firm Performance.” *Strategic Management Journal* 33 (1). Wiley Online Library:65–87.
Chollet, François, and others. 2015. “Keras: Deep Learning Library for Theano and Tensorflow.” *URL: Https://Keras. Io/K* 7 (8).
Cockburn, Iain M, Rebecca Henderson, and Scott Stern. 2018. “The Impact of Artificial Intelligence on Innovation: An Exploratory Analysis.” *National Bureau of Economic Research* No. w24449 (September).
Coff, Russell, and Joseph Raffiee. 2015. “Toward a Theory of Perceived Firm-Specific Human Capital.” *Academy of Management Perspectives* 29 (3). Academy of Management Briarcliff Manor, NY:326–41.
Crook, T Russell, Samuel Y Todd, James G Combs, David J Woehr, and David J Ketchen Jr. 2011. “Does Human Capital Matter? A Meta-Analysis of the Relationship between Human Capital and Firm Performance.” *Journal of Applied Psychology* 96 (3). American Psychological Association:443.
Eisfeldt, Andrea L., and Dimitris Papanikolaou. 2013. “Organization Capital and the Cross-Section of Expected Returns.” *Journal of Finance* 68 (4):1365–1406.
———. 2014. “The Value and Ownership of Intangible Capital.” *American Economic Review* 104 (5):189–94.
Ewens, Michael, and Matt Marx. 2017. “Founder Replacement and Startup Performance.” *The Review of Financial Studies* 31 (4). Oxford University Press:1532–65.
Ewens, Michael, Ramana Nanda, and Matthew Rhodes-Kropf. 2018. “Cost of Experimentation and the Evolution of Venture Capital.” *Journal of Financial Economics* 128 (3). Elsevier:422–42.
Felten, Edward W., Manav Raj, and Robert Seamans. 2018. “A Method to Link Advances in Artificial Intelligence to Occupational Abilities.” *AEA Papers and Proceedings* 108:54–57.
Furman, Jason, and Robert Seamans. 2018. “AI and the Economy,” 1–33. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3186591.

Greenwood, Jeremy, and Mehmet Yorukoglu. 1997. “1974.” In *Carnegie-Rochester Conference Series on Public Policy*, 46:49–95.
Hall, Bronwyn H. 1993. “The Stock Market’s Valuation of R&D Investment during the 1980’s.” *The American Economic Review* 83 (2):259–64.
———. 2006. “R&D, Productivity, and Market Value.” *IFS Working Papers, Institute for Fiscal Studies (IFS) 06/23*.
Hall, R. E. 2004. “Measuring Factor Adjustment Costs.” *The Quarterly Journal of Economics* 119 (3):899–927.
Hall, Robert E. 2001. “The Stock Market and Capital Accumulation.” *The American Economic Review* 91 (5):1185–1202. http://www.jstor.org/stable/2677921%5Cnhttp://www.jstor.org/stable/pdfplus/2677921.pdf?acceptTC=true.
Hall, Robert E. 2017. “High Discounts and High Unemployment.” *American Economic Review* 107 (2):305–30.
Hart, Oliver, and John Moore. 1994. “A Theory of Debt Based on the Inalienability of Human Capital.” *The Quarterly Journal of Economics* 109 (4). MIT Press:841–79.
Hayashi, Fumio. 1982. “Tobin’s Marginal q and Average q: A Neoclassical Interpretation.” *Econometrica* 50 (1):213–24.
Hayashi, Fumio, and Tohru Inoue. 1991. “The Relation Between Firm Growth and Q with Multiple Capital Goods: Theory and Evidence from Panel Data on Japanese Firms.” *Econometrica* 59 (3):731–53.
Hippel, Eric von, and Georg von Krogh. 2003. “Open Source Software and the ‘Private-Collective’ Innovation Model: Issues for Organization Science.” *Organization Science* 14 (2). INFORMS:209–23.
Hirshleifer, Jack. 1978. “The Private and Social Value of Information and the Reward to Inventive Activity.” In *Uncertainty in Economics*, 541–56. Elsevier.
Hornstein, Andreas, and Per Krusell. 1996. “Can Technology Improvements Cause Productivity Slowdowns?” *NBER Macroeconomics Annual* 11. MIT press:209–59.
Jeffers, Jessica S. 2017. “The Impact of Restricting Labor Mobility on Corporate Investment and Entrepreneurship.” *Job Market Paper*.
Jones, Benjamin F. 2009. “The Burden of Knowledge and the ‘Death of the Renaissance Man’: Is Innovation Getting Harder?” *The Review of Economic Studies* 76 (1). Wiley-Blackwell:283–317.
Jovanovic, Boyan. 1979. “Firm-Specific Capital and Turnover.” *Journal of Political Economy* 87 (6). The University of Chicago Press:1246–60.
Kaldor, Nicholas. 1966. “Marginal Productivity and the Macro-Economic Theories of Distribution: Comment on Samuelson and Modigliani.” *The Review of Economic Studies* 33 (4). JSTOR:309–19.
Kerr, Sari Pekkala, William R Kerr, and William F Lincoln. 2015. “Skilled Immigration and the Employment Structures of US Firms.” *Journal of Labor Economics* 33 (S1). University of Chicago Press Chicago, IL:S147--S186.
Kerr, William R. 2013. “U.S. High-Skilled Immigration, Innovation, and Entrepreneurship: Empirical Approaches and Evidence.” *Nber Working Paper Series* 53:1689–99.
LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. 2015. “Deep Learning.” *Nature* 521 (7553):436–44.
LeCun, Yann, Léon Bottou, Yoshua Bengio, and Patrick Haffner. 1998. “Gradient-Based

Learning Applied to Document Recognition." *Proceedings of the IEEE* 86 (11):2278–2323.
Lippman, Steven A, and Richard P Rumelt. 1982. "Uncertain Imitability: An Analysis of Interfirm Differences in Efficiency under Competition." *The Bell Journal of Economics*. JSTOR, 418–38.
Lucas, Robert E. 1967. "Adjustment Costs and the Theory of Supply." *Journal of Political Economy* 75 (4, Part 1):321–34.
Lucking, Brian. 2018. "Do R&D Tax Credits Create Jobs?"
Lucking, Brian, Nicholas Bloom, and John Van Reenen. 2017. "Have R&D Spillovers Changed?" *SSRN*.
Lustig, Hanno, Chad Syverson, and Stijn Van Nieuwerburgh. 2011. "Technological Change and the Growing Inequality in Managerial Compensation." *Journal of Financial Economics* 99 (3). Elsevier:601–27.
Marx, Matt. 2011. "The Firm Strikes Back: Non-Compete Agreements and the Mobility of Technical Professionals." *American Sociological Review* 76 (5). Sage Publications Sage CA: Los Angeles, CA:695–712.
Marx, Matt, Deborah Strumsky, and Lee Fleming. 2009. "Mobility, Skills, and the Michigan Non-Compete Experiment." *Management Science* 55 (6):875–89.
Mas, Alexandre, and Amanda Pallais. 2017. "Valuing Alternative Work Arrangements." *American Economic Review* 107 (12):3722–59.
Moretti, Enrico. 2004. "Estimating the Social Return to Higher Education: Evidence from Longitudinal and Repeated Cross-Sectional Data." *Journal of Econometrics*.
Paszke, Adam, Sam Gross, Soumith Chintala, Gregory Chanan, Edward Yang, Zachary DeVito, Zeming Lin, Alban Desmaison, Luca Antiga, and Adam Lerer. 2017. "Automatic Differentiation in Pytorch."
Peri, Giovanni, Kevin Shih, and Chad Sparber. 2015. "STEM Workers, H-1B Visas, and Productivity in US Cities." *Journal of Labor Economics* 33 (S1). University of Chicago Press Chicago, IL:S225--S255.
Peters, Ryan H., and Lucian A. Taylor. 2017. "Intangible Capital and the Investment-q Relation." *Journal of Financial Economics* 123 (2):251–72.
Polanyi, Michael. 1966. "The Logic of Tacit Inference." *Philosophy* 41 (155):1–18.
Raffiee, Joseph, and Russell Coff. 2016. "Micro-Foundations of Firm-Specific Human Capital: When Do Employees Perceive Their Skills to Be Firm-Specific?" *Academy of Management Journal* 59 (3). Academy of Management Briarcliff Manor, NY:766–90.
Raj, Manav, and Robert Seamans. 2018. "AI, Labor, Productivity, and the Need for Firm-Level Data." In *Economics of Artificial Intelligence*. University of Chicago Press.
Roach, Michael, and Henry Sauermann. 2010. "A Taste for Science? PhD Scientists' Academic Orientation and Self-Selection into Research Careers in Industry." *Research Policy*.
Rumelhart, David E., Geoffrey E. Hinton, and Ronald J. Williams. 1986. "Learning Representations by Back-Propagating Errors." *Nature* 323 (6088):533–36.
Saunders, Adam, and Erik Brynjolfsson. 2016. "Valuing Information Technology Related Intangible Assets." *Mis Quarterly* 40 (1).
Saunders, Adam, and Prasanna Tambe. 2015. "Data Assets and Industry Competition: Evidence from 10-K Filings," no. September.
Simon, Herbert A. 1991. "Organizations and Markets." *The Journal of Economic Perspectives* 5 (2):25–44.
Song, Jae, David J. Price, Fatih Guvenen, Nicholas Bloom, and Till von Watcher. 2015.

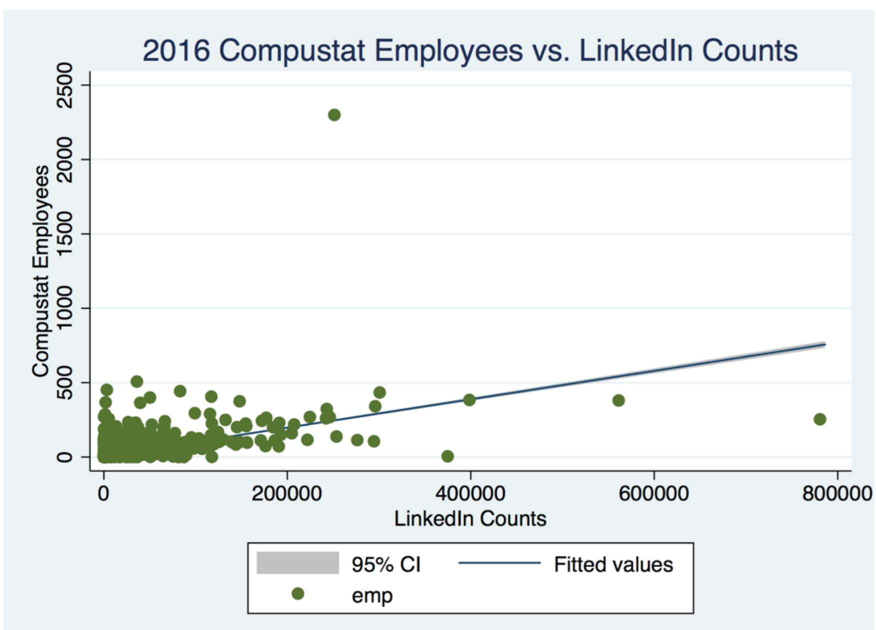
“Firming Up Inequality.” *NBER Working Paper*.
Stansbury, Anna M, and Lawrence H Summers. 2017. “Productivity and Pay: Is the Link Broken?”
Starr, Evan, Natarajan Balasubramanian, and Mariko Sakakibara. 2017. “Screening Spinouts? How Noncompete Enforceability Affects the Creation, Growth, and Survival of New Firms.” *Management Science* 64 (2). INFORMS:552–72.
Stern, Scott. 2004. “Do Scientists Pay to Be Scientists?” *Management Science* 50 (6). INFORMS:835–53.
Stole, Lars A, and Jeffrey Zwiebel. 1996a. “Intra-Firm Bargaining under Non-Binding Contracts.” *The Review of Economic Studies* 63 (3). Wiley-Blackwell:375–410.
———. 1996b. “Organizational Design and Technology Choice under Intrafirm Bargaining.” *The American Economic Review*. JSTOR, 195–222.
Syverson, Chad. 2011. “What Determines Productivity?” *Journal of Economic Literature* 49 (2):326–65.
Tambe, Prasanna. 2014. “Big Data Investment, Skills, and Firm Value.” *Management Science* 60 (6):1452–69.
Tambe, Prasanna, Lorin and Hitt, Erik Brynjolfsson, and Daniel Rock. 2018. “The Price and Quantity of IT-Related Intangible Capital.”
Tambe, Prasanna, and Lorin M. Hitt. 2012. “The Productivity of Information Technology Investments: New Evidence from IT Labor Data.” *Information Systems Research*.
Tambe, Prasanna, Lorin M. Hitt, and Erik Brynjolfsson. 2012. “The Extroverted Firm: How External Information Practices Affect Innovation and Productivity.” *Management Science* 58 (5):843–59.
Tambe, Prasanna, Xuan Ye, and Peter Cappelli. 2019. “Paying to Program? Engineering Brand and High-Tech Wages.” 25552. National Bureau of Economic Research Working Paper Series. https://www.nber.org/papers/w25552.pdf.
Teodoridis, Florenta. 2017. “Understanding Team Knowledge Production: The Interrelated Roles of Technology and Expertise.” *Management Science*. INFORMS.
Tervio, Marko. 2008. “The Difference That CEOs Make: An Assignment Model Approach.” *American Economic Review* 98 (3):642–68.
Thompson, Neil. 2017. “The Economic Impact of Moore’s Law: Evidence from When It Faltered.”
Tobin, James, William C Brainard, and others. 1976. “Asset Markets and the Cost of Capital.”
White, B. W., and Frank Rosenblatt. 1963. “Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms.” *The American Journal of Psychology* 76 (4):705.
Wildasin, David E. 1984. “The q Theory of Investment with Many Capital Goods.” *American Economic Review*.
Yang, S, and E Brynjolfsson. 2001. “Intangible Assets and Growth Accounting: Evidence from Computer Investments.” *Unpublished Paper. MIT* 85 (61):28.
Zyontz, Samantha. 2018. “Running with (CRISPR) Scissors: Tool Specialists and Innovation with New Technologies.”

Appendix: Additional Regression Results on Coverage and Robustness

LinkedIn Coverage	(1) Compustat Count (Thousands)
LinkedIn Count	0.00190***
	(1.59e-05)
Total Assets	4.27e-05***
	(1.70e-06)
Observations	52,767
R-squared	0.422
Industry-Time FE	Yes

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Coefficients represent the predicted per LinkedIn user employee count (in thousands) for Compustat firms with employee count data populated. This regression is used to predict the employee count in the case that Compustat is missing data.



Log(Market Value)	(1) SML Quintiles	(2) AI Quintiles
Log(Lagged Market Value)	0.646***	0.645***
	(0.0169)	(0.0170)
Log(Total Assets)	0.305***	0.307***

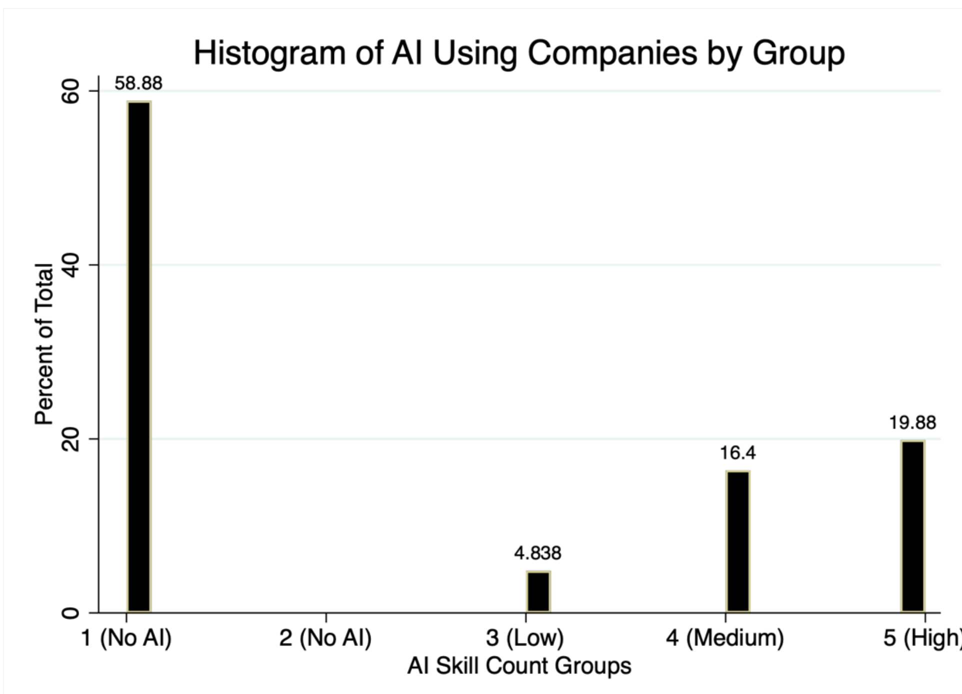
	(0.0181)	(0.0181)
Log(Education Years)	-0.00521	-0.00500
	(0.00840)	(0.00840)
Log(Business Mgmt.)	-0.0128	-0.0128
	(0.00823)	(0.00831)
Log(Cloud Computing)	0.00300	0.00114
	(0.00454)	(0.00460)
Log(Data Science)	-0.00590	-0.00662
	(0.00755)	(0.00753)
Log(Digital Literacy)	0.00977	0.00870
	(0.00643)	(0.00649)
Log(Data Storage)	-0.00268	-0.00266
	(0.00513)	(0.00509)
Log(Big Data)	-0.00111	-0.00276
	(0.00326)	(0.00353)
Quintile 2x6	-0.00201	
	(0.0149)	
Quintile 2x7	-0.00638	
	(0.0143)	
Quintile 2x8	-0.0240	
	(0.0158)	
Quintile 2x9	-0.0224	
	(0.0160)	
Quintile 2x10	-0.0110	
	(0.0209)	
Quintile 2x11	-0.0255	
	(0.0179)	
Quintile 2x12	-0.00180	
	(0.0183)	
Quintile 2x13	-0.00327	
	(0.0207)	
Quintile 2x14	-0.0440**	
	(0.0206)	
Quintile 2x15	-0.0169	
	(0.0166)	
Quintile 2x16	-0.0484**	
	(0.0194)	
Quintile 2x17	-0.0454**	
	(0.0179)	
Quintile 2x18	-0.0187	
	(0.0181)	
Quintile 2x19	-0.0330*	
	(0.0189)	
Quintile 3x6	-0.0156	0.0181
	(0.0174)	(0.0167)
Quintile 3x7	-0.00405	0.0124
	(0.0166)	(0.0207)
Quintile 3x8	0.000366	0.0305
	(0.0207)	(0.0211)
Quintile 3x9	-0.0340*	0.0155
	(0.0198)	(0.0228)
Quintile 3x10	-0.00357	0.00974
	(0.0208)	(0.0191)
Quintile 3x11	-0.00515	0.00940
	(0.0207)	(0.0206)
Quintile 3x12	-0.00215	0.0682***

	(0.0209)	(0.0244)
Quintile 3x13	-0.00881	0.0341
	(0.0210)	(0.0236)
Quintile 3x14	-0.0146	0.0424*
	(0.0211)	(0.0236)
Quintile 3x15	-0.0232	0.0481*
	(0.0214)	(0.0248)
Quintile 3x16	-0.0339	0.0483**
	(0.0223)	(0.0237)
Quintile 3x17	-0.0249	0.0354
	(0.0210)	(0.0239)
Quintile 3x18	0.00932	0.0393*
	(0.0197)	(0.0230)
Quintile 3x19	-0.0306	0.0336
	(0.0222)	(0.0238)
Quintile 4x6	-0.0190	0.00950
	(0.0167)	(0.0120)
Quintile 4x7	-0.0159	0.0190
	(0.0159)	(0.0132)
Quintile 4x8	-0.00286	0.0151
	(0.0190)	(0.0153)
Quintile 4x9	-0.0203	0.0146
	(0.0186)	(0.0137)
Quintile 4x10	-0.00866	0.0175
	(0.0190)	(0.0138)
Quintile 4x11	0.0180	0.0157
	(0.0183)	(0.0162)
Quintile 4x12	-0.00516	0.0385**
	(0.0199)	(0.0153)
Quintile 4x13	-0.00899	0.0249
	(0.0194)	(0.0160)
Quintile 4x14	-0.0311	0.00957
	(0.0200)	(0.0174)
Quintile 4x15	-0.0131	0.0244*
	(0.0204)	(0.0143)
Quintile 4x16	-0.0416*	0.0179
	(0.0231)	(0.0160)
Quintile 4x17	-0.0301	0.0243
	(0.0197)	(0.0152)
Quintile 4x18	-0.00189	0.00740
	(0.0197)	(0.0157)
Quintile 4x19	-0.0357*	-0.00713
	(0.0209)	(0.0162)
Quintile 5x6	-0.0394*	0.0213**
	(0.0209)	(0.0103)
Quintile 5x7	-0.0365	0.0482***
	(0.0223)	(0.0107)
Quintile 5x8	-0.0356	0.0244**
	(0.0264)	(0.0123)
Quintile 5x9	-0.0338	0.0337***
	(0.0229)	(0.0120)
Quintile 5x10	-0.0280	0.0156
	(0.0211)	(0.0122)
Quintile 5x11	-0.0290	0.0484***
	(0.0247)	(0.0136)
Quintile 5x12	-0.0366	0.0513***

	(1)	(2)
	(0.0232)	(0.0133)
Quintile 5x13	-0.0502*	0.0401***
	(0.0264)	(0.0147)
Quintile 5x14	-0.0437*	0.0360**
	(0.0253)	(0.0151)
Quintile 5x15	-0.0272	0.0416***
	(0.0229)	(0.0141)
Quintile 5x16	-0.0613**	0.0282*
	(0.0259)	(0.0156)
Quintile 5x17	-0.0388	0.0337**
	(0.0252)	(0.0149)
Quintile 5x18	-0.0107	0.0380**
	(0.0262)	(0.0153)
Quintile 5x19	-0.0433*	0.0254
	(0.0247)	(0.0163)
Observations	20,522	20,526
R-squared	0.998	0.998

Table Note: Robust standard errors in parentheses, SEs Clustered by Firm. TensorFlow Launch corresponds to year-quarter 12. Quintile 5x12 is the coefficient on the fourth quarter of 2015 dummy interacted with the highest quintile for SML (1) and AI (2). None of the second quintile data are available for AI because the second quintile still does not use AI skills. Other values are relative differences to the first quintile.

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



Percentage of Firms by AI Skill Group (Quintiles). Note: Counts are uneven because of ties in the skill counts.